upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Departmento de Estadística, Informática y Matemáticas
Escuela de Doctorado de Navarra (EDONA)
Grupo de Inteligencia Artificial y Razonamiento Aproximado (GIARA)

Doctoral Thesis

Dealing with uncertainty

a human perception approach in image processing

Cédric Marco Detchart

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Dealing with uncertainty

a human perception approach in image processing

Cédric Marco Detchart

September, 2019

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Universidad Pública de Navarra

upna

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Departmento de Estadística, Informática y Matemáticas
Escuela de Doctorado de Navarra (EDONA)
Grupo de Inteligencia Artificial y Razonamiento Aproximado (GIARA)

Doctoral Thesis

Dealing with uncertainty

a human perception approach in image processing

Cédric Marco Detchart

Supervisors Carlos López Molina and
Francisco Javier Fernández Fernández

September, 2019

Cédric Marco Detchart

Dealing with uncertainty: a human perception approach in image processing

Doctoral Thesis, September, 2019

Supervisors: Carlos López Molina and Francisco Javier Fernández Fernández

Universidad Pública de Navarra

Grupo de Inteligencia Artificial y Razonamiento Aproximado (GIARA)

Escuela de Doctorado de Navarra (EDONA)

Departmento de Estadística, Informática y Matemáticas

Campus Arrosadía

31006 Pamplona

Autorización

Dr. Carlos López Molina: Profesor de la Universidad Publica de Navarra en el Área de Conocimiento de Ciencias de la Computación e Inteligencia Artificial.

Dr. Francisco Javier Fernández Fernández: Profesor de la Universidad Pública de Navarra en el Área de Conocimiento de Ciencias de la Computación e Inteligencia Artificial.

HACEN CONSTAR que el presente trabajo titulado *"Dealing with uncertainty: a human perception approach in image processing"* ha sido realizado bajo su dirección por D. Cédric Marco Detchart. Autorizándole a presentarlo como Memoria para optar al grado de Doctor por la Universidad Pública de Navarra.

El Doctorando

Fdo.: Cédric Marco Detchart

Los Directores

Fdo.: Carlos López Molina

Fdo.: Francisco Javier Fernández Fernández

Pamplona, September, 2019

A Giulia, a Thierry
y a mis padres.

Agradecimientos

En primer lugar me gustaría agradecer a mis directores por aceptar la dirección de esta tesis y ayudarme en este largo camino. A Carlos por su apoyo en todo momento, sus consejos y conversaciones. Gracias por las ideas, las lecturas recomendadas, la paciencia y las discusiones mantenidas a lo largo de estos años. A Javier por su guía y conocimiento. Por la comprensión en momentos difíciles y por el optimismo, arrojando siempre algo de luz hacia el final del camino. Sin vosotros esta tesis no hubiera salido adelante.

A Humberto, por haberme abierto las puertas del grupo de investigación, por sus amplios conocimientos, su ayuda y por haberme introducido en el mundo de la investigación.

Además a aquellas personas con las que he estado durante estos años compartiendo horas y horas. Gracias a Laura, por sus consejos, momentos de escucha y conversaciones en todo momento. A Mikel por las charlas y los cafés. A Edurne, por su ayuda y por haber sido mi primer contacto con el grupo. A los demás miembros del grupo de investigación GIARA y a la Universidad Pública de Navarra por haber hecho posible este trabajo.

A Graçaliz por su acogida en Brasil, su ayuda y sus innumerables lecciones de historia y cultura. Por los buenos ratos pasados tan lejos de casa. A Giancarlo, por su amistad, sus conversaciones y por su guía y conocimientos.

A mi familia, en especial a mis padres, por animarme a conseguir mis objetivos y sin los que no hubiera llegado hasta aquí. A mi hermano, Thierry, por estar siempre a mi lado.

Finalmente, a Giulia, por su constante apoyo y ánimo, que me han acompañado desde la distancia en todo este viaje que ha sido esta tesis y cuya conclusión nos reunirá en otros viajes.

¡Muchas gracias a todos!

Abstract

An effective way of comparing objects is through analysing their similarity (comparing their common attributes) or dissimilarity (comparing their differences). Similarity has been studied in a wide variety of ways, from different points of view and different disciplines such as psychology, neuroscience or mathematics. A topic where distance metrics and similarities are used is through fuzzy logic which permits a human approach to comparison measures, providing a tool that captures the uncertainty inherent to the perception of distance. In the context of comparison and similarity, we find an important related concept, that is, *feature matching*, which is one of the most used technique in detecting patterns or measuring algorithm performance.

In this Ph.D. thesis we have studied similarity in the context of fuzzy logic along with different ways it can be used; among other we found it useful in pattern extraction in the context of fingerprint analysis, which in this thesis has come to the development of new concepts, named Restricted Radial Equivalence Functions (REF) and Radial Similarity Measure (RSM), which models the perceived similarity between scalar and vectorial pieces of radial data. Moreover, as a more pure distance analysis in the context of fuzzy logic, we have studied distance measures, similarity measures and entropies in interval-valued fuzzy sets. To do so, as a novelty, we have included the width of the interval in the computation of the measure. This add-on has permitted to be able to connect the uncertainty of the outputs with the inputs.

Finally, as a complement we have studied an application of our theoretical results in image processing. In particular, we have used ordered directionally monotone functions in edge detection problems and consensus techniques to build a new edge detector. In addition, we have conducted a study about all the different method presented in the literature for feature matching evaluation quality, resulting in a novel taxonomy of the different methods and an analysis of their behaviour, showing that none of the method is better than the others and being equivalent.

Resumen

Una de las maneras de abordar la comparación entre objetos es mediante el uso de la similitud (comparando sus atributos comunes) o la disimilitud (comparando sus diferencias). La similitud ha sido estudiada de maneras muy diversas, desde muchos puntos de vista y en una variedad de campos tales que la psicología, la neurociencia o las matemáticas. Uno de los principales temas en los que las distancias y las similitudes han sido abordadas es a través de la lógica difusa, que permite implementar una perspectiva humana en las medidas de comparación, aportando una herramienta que permite capturar la incertidumbre inherente en la percepción de la distancia. Dentro de este contexto de la comparación y la similitud, encontramos un concepto relacionado, como es la correspondencia de características. Este concepto es una de las técnicas más utilizada en la detección de patrones o la evaluación del rendimiento de un algoritmo.

En esta tesis doctoral hemos estudiado la similitud en el contexto de la lógica difusa junto con una serie de propuestas en las que tiene cabida. Entre otras, podemos citar la extracción de patrones en el análisis de huellas dactilares, cuyo estudio ha llevado en esta tesis al desarrollo de nuevos conceptos como las Funciones de Equivalencia Restringidas Radiales y las Medidas de Similitud Radiales, que modelan la similitud percibida entre datos radiales, tanto escalares como vectoriales. Además, como un estudio más explícito de las distancias en la lógica difusa, hemos abordado el estudio de las distancias, las medidas de similitud y la entropía en conjuntos difusos intervalo-valorados. Para ello, hemos incluido la amplitud del intervalo en el cálculo de las medidas. Esta condición adicional nos ha permitido conectar la incertidumbre contenida en el intervalo final con la del intervalo inicial.

Finalmente, de manera complementaria hemos estudiado la aplicación de nuestros resultados teóricos en tareas de procesamiento de imagen. Hemos desarrollado un detector de bordes mediante el uso de funciones monótonas direccionalmente ordenadas y técnicas de toma de decisión por consenso. Además, hemos realizado un estudio que recoge las diferentes técnicas de análisis de calidad para los métodos de extracción de bordes, resultando en una nueva taxonomía de los diferentes métodos y un análisis de su comportamiento, mostrando que ninguno de los métodos es más adecuado que otro, siendo equivalentes entre si.

Contents

1 Introduction **1**

1.1 Motivation and Problem Statement 16
1.2 Objectives 18
1.3 Thesis Structure 19

2 Phd thesis report **21**

2.1 A framework for radial data comparison and its application to fingerprint analysis 21
2.2 Width-based interval-valued distances and entropies 29
2.3 Similarity between interval-valued fuzzy sets taking into account the width of the intervals 35
2.4 Ordered directional monotonicity in the construction of edge detectors 41
2.5 A survey on matching algorithms for boundary image comparison and evaluation 45
2.6 Optical images-based edge detection in Synthetic Aperture Radar images 49
2.7 Conclusions 51
2.8 Future research lines 52
2.9 Conclusiones (versión en español) 53

3 Publications: published, accepted and submitted **55**

3.1 A framework for radial data comparison and its application to fingerprint analysis 55
3.2 Width-based interval-valued distances and entropies 70
3.3 Similarity between interval-valued fuzzy sets taking into account the width of the intervals 85
3.4 Ordered directional monotonicity in the construction of edge detectors 111
3.5 A survey on matching algorithms for boundary image comparison and evaluation 139
3.6 Optical images-based edge detection in Synthetic Aperture Radar images 154

Bibliography **165**

1 Introduction

> *Exploring the unknown requires tolerating uncertainty.*
>
> — **Brian Greene**
> (Theoretical physicist, mathematician)

HUMAN learning is a process carried out ever since early stages of life. However, far from being an evident process, it has been a matter of study by many researchers through the ages. These researches have taken multiple paths and approaches in an attempt to model human learning, in all its different shapes and times. As of now, we can find a myriad of factors being involved, including the environment [1], the pre-acquired knowledge (past experience) [2] and the ability to put it all together to make decisions [3]. The complexity of the topic has made it necessary to involve many different disciplines of research in order to explain even the most basic learning-oriented tasks. In fact, and although the ideal result of all these research efforts would be a global, computational model for human learning, we can currently only aspire to explain task-localized human learning.

At a broad level, the human being constantly receives an infinity of stimuli that permit to learn and acquire experience. Hence, an evident taxonomy for learning processes can focus on the type of stimulus (or stimuli) that abilitated (or fired up) the learning process. These stimuli come in different ways, from sensations to sounds and touch, being visual stimuli one of the learning cornerstones [4]. This dissertation focuses on learning based on visual information, specifically on producing computational models that attempt to explain one or many of the tasks involved in vision-based learning.

Learning based on visual information has characteristics which are akin to any other type of learning, while also having its own special features. As for the former, we

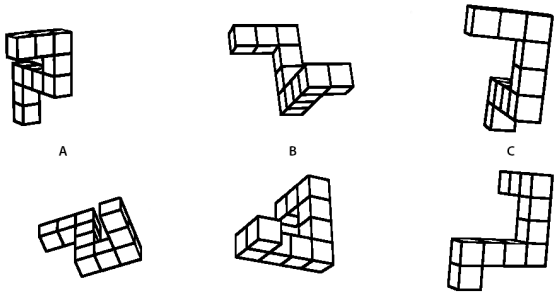


Fig. 1.1: Figures from the Shepard and Metzler's experiment on mental rotation. The experiment shows two equal figures rotated and mirrored (A and B) and a third one which is impossible to obtain by either rotation or mirroring (C).

find the need of representing information in context-agnostic, invariant manners. This is referred to by Zadeh as *tolerance for imprecision*. Humans do not need to keep a myriad of different images of a singular object to understand that it is the same object. In the experiment by Shepard and Metzler [5] the mental perception of objects is analysed with a series of figures (Fig. 1.1) where rotation and reflection has been applied. The experiment consisted in recovering the time spent by a person to conclude the equality of objects and how the rotation and mirroring where obtained, mentally trying to find the same object by modifying it based on the mental model.

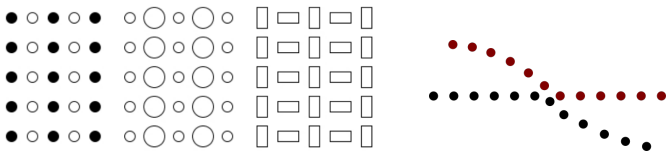
This effort in modelling human visual recognition has not stayed in theoretical or mildly applied fields. In fact, in recent years, authors have even tried to replicate early cortical connections in human brains by implementing Convolutional Neural Networks (CNNs). These works, majorly influenced by the Mental Rotation Tests [6], aim at modelling human recognition abilities in the Human Visual System (HVS), specially regarding rotation-, scale- and eccentricity-invariance [7], [8]. These works mostly use line-based draws (e.g., in [9], Korean characters). CNNs properly replicating human recognition abilities are really appealing, since they might connect physiological and theoretical understanding of the HVS. However, would also yield questions due to the hardly-understandable behaviour of CNNs. For example, as reported by Nguyen, Yosinski and Clune [10], state-of-the-art CNNs yield aberrant results in object recognition. Such errors are, moreover, yield in almost-full ($> 99\%$) confidence. As the authors state, their findings raise *questions about the true generalization capabilities of DNNs* [10], which are of paramount importance for the present task. Humans are extremely good at generalization,

and hence CNNs might find here a limitation in their aim to replicate human behaviour.

Nevertheless, there are aspects of vision-based learning that make the process rather different from other learning processes. For example, we have the huge rate of information a human is able to capture, way superior to that by any other sense. This has led to extremely complex systems of selective focus and image simplification, which allow humans to understand scenes through simplification and abstraction. Another characteristic which makes vision-based learning so special is the low quality of the information received, and the extreme proficiency of humans in overcoming such problems. Humans often capture poorly-presented visual information, yet being able to make a full comprehension out of it. Reasons for low quality information might relate to physiological problems, but also to characteristics of the scene (low luminosity, fog, shades, occlusions), and to other factors. As a result, visual information often contains errors and doubty items, *e.g.*, the same scene can be seen differently by two persons. In addition, the memorization process suffers from imperfection as well as the patterns and models *stored* in our brain (acquired knowledge), needing a certain tolerance for imprecision [11], which is to be exploited to better achieve robustness and low solution cost.

Our visual system is capable from tiny points (natural images) to perceive complete scenes forming groups (shapes as lines, squares or any other form). Many studies have attempted to propose a framework to explain the *enigma of perception* (how we interpret our world), being one of the first and major approximations the Gestalt Theory [12], [13]. This theory states that we tend to organize what we see into groups from the most detailed component to larger objects, according to laws that describe components relations and conflicts, and identify them to some type of abstraction (model) we already have. For example, some laws of this theory deal with similarity (colour, shape, texture, etc.) or continuity of direction (Fig. 1.2b), *i.e.*, looking in a particular direction makes the path we follow go in that direction even in case of occlusions or when the perception of a shape goes beyond its limits. An example of how these laws work is depicted in Fig. 1.2a where we see a set of drawings with different shapes, colours and orientations, instantly differentiating each of them by grouping their characteristics, and Fig. 1.2b showing how the perception is influenced by the direction or the colour of the objects.

Other researchers, *e.g.* Gibson, propose a different model for human vision, clearly oriented towards a computational model. This line of research states that the



(a) Perception of grouping by similarity of colour, size and orientation.

(b) Perception of continuity in direction.

Fig. 1.2: Illustration showing some laws from the Gestalt Theory [13].

surrounding environment, or context, play an essential role in our perception and can be divided and analysed at different levels of detail, defining a hierarchy in the elements. His approach is not centred on understanding the visual system and how humans obtain the images, but rather focused on the different task of perception, trying to understand the properties of the environment and what is perceived. For Gibson, perception implies to detect the information that best describes the main environmental properties [1]. One of the main questions drawn up by Gibson is to determine how humans perceive stimulus and build a solid knowledge from a constant evolving and changing world. This led further on to one of the key topics in understanding the human visual systems: the extraction of characteristics or features that represent objects. These characteristics are supposed to be captured by humans and learning, then used to discern whether an object is the type of object we think it is. In this way, objects would not be compared from their full representation, but from the comparison of specific features extracted from them.

In a line similar to Gibson's, another important theory is that of Marr, where vision is seen as a process producing a useful description of real world images, removing all unnecessary information [14]. The vision process is presented as a computational model referred to as *Primal Sketch*, which involves four different phases: image representation, features, information extraction and recognition. In his proposal Marr states that the primary representation of a raw image is some sort of *sketch* where intensity changes, local relations and geometric structures are analysed. In Marr's approach to human visual system, the framework proposed present vision as series of representations, where images (a set of points of intensities) are decomposed firstly as a primitive sketch (*e.g.*, edges), then as a more advanced sketch (*e.g.*, orientations) and finally as a 3D model (hierarchy of spatial organization).

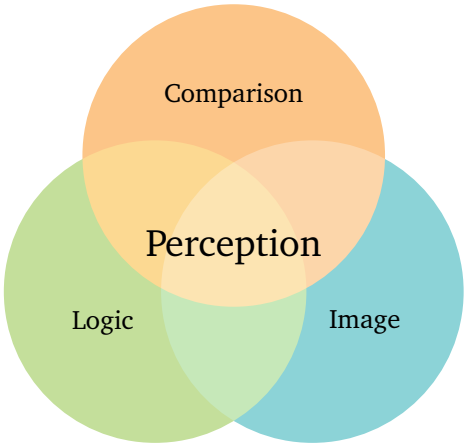


Fig. 1.3: General scheme of the thesis, representing the three main pillars analysed, Comparison, Logic and Image, which serves to build the idea of Perception.

This dissertation inspires from previous theories and focuses on one of the subtasks with greater importance in them: visual comparison. The ability to compare information is key for any learning process, since humans are generally unable to keep in mind perfect copies of informations. Their knowledge is messy, imprecise and vague, as masterfully presented by Zadeh when advocating for imprecision-tolerant systems. Humans are able to naturally lead with imprecisions and inexact information. And the human visual system is not an exception.

In this dissertation we focus our efforts on intensively analysing the idea of human visual information comparison. It considers its very core, as it is the models that could embody *how humans compare*, as well as peripheral issues, as the extraction of characteristics that are used in such comparisons, or the ability to produce logical consequences out of the comparison process. All of it heavily oriented towards the management of uncertainty, which is an unavoidable factor to be taken into account when understanding human learning (in general) and the human visual system (in particular). This dissertation is supported by three basic pillars: comparison, logic and image, as illustrated in Fig. 1.3. We tackle the theory of comparison from a human based approach and how researchers have tried to mimic the way humans compare. Then, coupled to comparison and as a tool to represent information, first in the human reasoning and then in computing, we find logic, and more precisely human-like approaches. Finally, we put comparison, uncertainty and logic all together in image analysis, proposing new ways of managing uncertainty in comparison tasks for computer vision.

Certainly, ours is not the first effort focusing on the psychological or physiopsychological fundamentals of comparison, since many others share or have shared interest in understanding how humans compare (in terms of vision, or not). Comparison is a fundamental task for human knowledge, being a basic tool for leaning and producing a model of the surrounding environment, improving and updating their knowledge by means of comparing objects, situations, persons, etc. Comparison has been approached in many research areas as psychology, neuroscience and mathematics. As appointed out by Attneave [15], the key question might be as simple as: *What makes things seem alike or seem different?* There is not an exact and unique answer, as comparison is very dependent of the context where it is made as well as of the stimulus it depends of. Moreover, in order to compare things we do not compare them as a whole, but rather focus on certain details or features, based on our knowledge or our intuition, that characterize them. As a simple example, cars are not compared as cars. The comparison is made as a composite of parts that represent the idea of car: number of seats, type of engine, colour, capabilities, etc. This example leads to secondary questions, *e.g. a better car is a faster one?* How can we answer this questions? And, considering humans could provide quick and efficient answers to those questions, how is such conclusion reached?

It is acceptable to believe that comparison is based on a dual representation of information: a new object/feeling/idea to a pre-stored one, hypothetically involving a list of items to which a new object is to be compared. In order to understand how humans compare, researchers have always tried to analyse the psychological relations of closeness, as understood by humans. As said by Shepard [16], humans are able to say that *two words are...closely associated* like when we usually associate the word *butter* to the word *bread* although apparently they have nothing in common or when *one of two similar colours is said to be very near to the other*. So, how is this *closeness* measured? Shepard introduce the notion of *psychological closeness* or *degree of proximity*, that tries to encompass how concepts, attitudes and stimulus are related.

Shepard's approach to closeness has led researchers to develop mathematical concepts for a quantitative analysis of comparison. As a first approach and as a basic operation with data, we find equality, which indicates whether two objects are exactly the same. The main problem in equality-based comparison is that if the elements compared are not equal the result is negative, and does not recover the idea of degree of proximity from human perception. Humans might see different objects that look different, yet understand that they are the *same* object. It is true

that when humans (since childhood) first compare objects they tend to see if they are the same (*e.g.*, children try to fit objects with shapes in a box with the shape of the objects) but they progressively try to learn to identify characteristics that differentiate them and use them in order to evaluate their similarity. So, how can we consider the existent variations to obtain how two things are with respect to each other?

Taking the idea that similarity can be interpreted as a relation of proximity a spatial representation of the data being analysed is needed. The most common approach in the literature has been to use a metric space [16], generally an Euclidean space, to handle object comparison. In this approach, an object is as similar to another as the value distance between them; hence, comparison reduces to the application and understanding of the value yield by a metric. A metric is a function that defines a distance between a pair of objects, and is able to capture the earlier mentioned proximity in a given universe. This type of functions are characterized by three axioms, such as:

- *Minimality*: states that if the distance between two objects is zero this means that they are the same object;
- *Symmetry*: demands that the distance between two objects or points is the same whether we measure on one direction or the contrary;
- *Triangle inequality*: requires that the shortest distance between two objects is the straight one.

These properties significantly ease the computation of a similarity value as they define boundaries to what can be done.

Metrics have been a precious tool to the comparison theory, and very specifically for computer-based applications [17]. From a mathematical point of view, the axiomatic representation, which gives some rigidity and mathematical consistency, provides us a manner to interpret the notion we have of proximity between objects in a quantitative way, easing the computational process. But from the human perception point of view the usual axioms are not so evident. Many criticism has been made to similarity approached in these terms. Starting from the assumption of an Euclidean space, as stated by Attneave [15], assuming that *psychological space is Euclidean in its character* is *exceedingly precarious*. To other authors that question

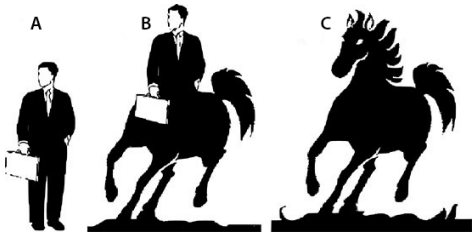


Fig. 1.4: Illustration used by Tversky and Gati [19] to show that in terms of human perception the triangle inequality does not always hold. **A** and **B** are totally different, but the transition from **A** to **B** and from **B** to **C**, is perceived as shorter.

that human perception follows metric axioms [18] and doubt about the necessity of metric conditions for measuring similarity.

The opposition to metric axioms has been backed by fields like psychology and neuroscience, being one of the main opponents Tversky [20]. Tversky argues that similarity is an asymmetric relation and comparison is made in a directional way, and also defend that the triangle inequality not always holds. The symmetry states that the choice of the subject to be compared conditions the measure. As presented by Tversky, we use to say "the portrait resembles the person" not "the person resembles the portrait". The subject of comparison tend to be the strongest object and the referent the weakest and the chosen direction in the comparison can derive in different meanings. In terms of the triangle inequality, the unfulfillment of this axiom can be clearly seen in the example from Tversky and Gati [19] in Fig. 1.4, where the different images are used to illustrate that the human perception does not always follows this property. We can barely relate a man to a horse, but the relationship through a Minotaur seems rather smooth. Put to raw mathematical terms, and being the man (a), the Minotaur (b) and the horse (c) the objects to compare, it might seem that $sim(a, c) \ll sim(a, b) + sim(b, c)$, with sim representing the similarity. In order to recover Tversky ideas about the axioms required for comparison, while maintaining a geometric approach, Krumhansl [21] propose to add to the measurement of similarity the idea of spatial density. With the assumption that spatial density affects similarity the violation of certain axioms is admitted, as with equal distance two points may have a different similarity depending on the density of the region where they are located.

As a result of these criticisms, other measures inspired by metrics have arisen, often driven by the relaxation of the metric axioms. For example we found *pseudo-metrics* which follows the axioms defined for a metric, except that the identity

axiom has been modified. In this way, the pseudo-metric will account 0 only if the elements compared are the same, being possible to have a distance of 0 with different elements. Or another type of measure known as *quasi-metrics* which drop the symmetry axiom. For example, if we consider the time for going from a point A to a point B crossing mountains, the time spent to cross up hill is not the same as the one for going down.

The deepest form of relaxation of the metric axioms has given birth to the concept of Similarity Measure (SM), which is able to encompass different ways of comparing objects. In fact, as there is not any restriction as compared to distance metrics, the similarity measure definition is quite vague and usually defined *ad-hoc* for the problem to be solved [22]–[24]. Moreover, most of the literature expects no other restriction to SM than obtaining maximal values when comparing identical objects.

All the previous forms of comparison affect to 1-vs-1 comparison. However, there are other physiological tasks that heavily relate to comparison, and do not properly involve 1-vs-1 confrontation. A clear example of such task is the evaluation of the heterogeneity within a group of objects. An effort in modelling such situations is that of informational entropy, focused on the amount of chaos in a series of events. Entropy is defined by Shannon [25] to measure the *missing information*, or the *measure of randomness* within data; Alternatively, De Luca and Termini [26], define entropy as a *measure of a quantity of information which is not necessarily related to random experiments* giving a global measure of the uncertainty inherent to a situation. Another type of measures in groups of data are *dispersion measures* which are able to capture how the distribution of the data is, *e.g.*, variation, standard deviation, mean absolute difference, etc. In fact, as pointed out in [27] dispersion measures are related to multidistances [28], where the concept of distance measured between two elements is generalized to work with collections of elements, *i.e.*, to be able to measure how much separated are not only two elements of a set but any finite list.

Not included in all the previous types of measures for comparison we found other operators in the extensive analysis of the ordering procedure from Barnett [29], where comparison plays a key role. He states that *the ordering principle is clear and unambiguous* for univariate samples, while ordering multivariate ones in a clear way is not always possible and some *restricted form of ordering...is feasible and advantageous*, like relative order comparison or partial ordering. Also a more

specific type operator is studied in Eckert and Klamler work [30], namely distance-based aggregation. This operator is proposed in order to solve the problem of aggregating multiple objects into one that represents them all, capturing the intuition of consensus.

To sum up, all these operators depend in fact on the representation of the data being compared. It is clear that most of the approaches are based on geometric models and multidimensional scaling [31], as it is convenient from a mathematical point of view, providing a quantitative description and a parametric representation of the objects being analysed. Connected to the representation model we also have what we represent, that is, the features representing objects. These features are what we need to compare or match in order to measure similarity and their extraction is one of the comparison's cornerstone as well as a problematic process [20], because features can be any property describing an object.

Even if it has driven a large effort on theoretical studies, comparison has an essential role to play in applications, and its mostly driven by its usefulness. It is part of a process where learning is the result of comparing previous stored knowledge with a new stimuli, resulting in decisions. These decisions are at the same time used over some data to obtain knowledge and hence to learn, but is this new knowledge? Or it is something underlying that was already known? An example of application where comparison is a cornerstone is pattern-based decision problems, and in general, any classification problem or clustering, where data needs to be separated in groups. Any application which uses real, and hence, imperfect data, is subject to the needs of comparing in order to generate some type of decision.

It is a clear thought that comparison is at the hearth of decision making and to go trough, humans use mental tools to represent their thoughts [32], [33]. The understanding and object representation needs for a way to channel how decision are made by humans, and since early times researchers have been seeking for an explanation analysing the models involved and the type of reasoning made in our brain [34]–[36].

Within all the previous maremagnum of object representations, and axioms tailored to capture human comparison, there is a centric spot for fuzzy set theory and related concepts. When describing the world, humans express feelings or interests with words. When appreciating a piece of art we say that it is a *marvellous* piece of art or that *we like it a lot*. In the case of making decisions, a similar situation

occurs, for example, if it is raining we take an umbrella, and when it is sunny we do not. But what do we decide when it drizzles? Do we take the umbrella or not? Clearly drizzling is not the same as raining, but maybe we are getting wet, but not as much as if it were *really raining*. The decision to be taken regarding an umbrella might be depending upon whether it drizzles "more like it does not rain" or more "like it rains". This, obviously, relates to partial truths. And, hence, to multivalued or real-valued truth values. So, it might seem that the natural output of a human comparison is in fact a partial truth, and hence the very comparison process might resemble a logic-based procedure, instead of a topology-inspired operator. This can be further analysed from the point of view of human reasoning.

Many research has been done in order to represent how decisions are made, from the logical point of view. As a primer approach we found classical logic, which is somehow a static representation of the truth, where two possible values for a sentence can be chosen, *i.e.*, true or false, being possible only one of them at the same time. Then, for a stimuli received, a fact, a sentence, etc. we will be able to evaluate them with one of the two possible values, obtaining in this way a degree of truth. The truth values of the classical logic can also take a different representation in form of numerical values, being the truth value 1 and the false value 0. This logic leads us to propositional logic going a step forward and permitting to express natural language situations, which is possible due to a series of statements coupled to a set of operators. Advanced statements based on individual and indivisible propositions can be made, mimicking at a basic level the human logic. A possible example could be:

Premise 1: If it's raining then it's cloudy.

Premise 2: It's raining.

Conclusion: It's cloudy.

In a first approach classical logic could seem a good approximation but it lacks a possible value in order to represent the existence of uncertainty. To cope with this inconvenient, in one of the three-valued logics, Łukasiewicz [37] introduced a third possible truth value as the *undecided* term. Also in the context of three-valued logic, partial logic presented by Blamey [38] propose to use, in addition to true and false values, some sort of *partially defined* predicates to work with.

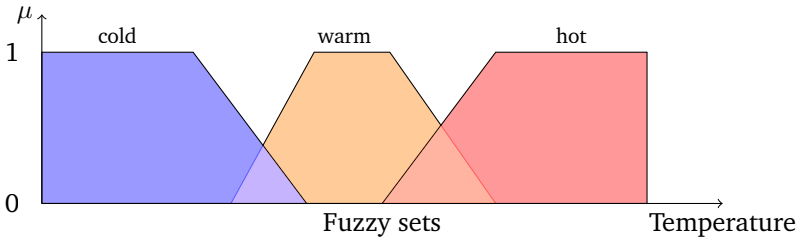


Fig. 1.5: Fuzzy sets representing the transitions between human concepts of temperature sensations.

Even these types of logic, which consider an uncertainty term, are not able to capture the human-like, logic-based reasoning. The vagueness of mental representation carries not well defined linguistic terms, nor boundaries in the definition of a class of object. An alternative representation would be to use real numbers to obtain a certain precision when constructing a model. Extending the possible truth values that can be taken into account we arrive at multi-valued logic where a finite possible values are considered to represent the degree of truth or even use an infinite set of values, that is, fuzzy logic.

In fuzzy logic, instead of considering truth values representing the degree of truth on a statement, membership degree is used. This type of logic comes as a consequence of the introduction of fuzzy sets by Zadeh [39]. Zadeh states that "classes of objects...do not have precisely defined criteria of membership", and that this imprecision in the definition of classes plays a crucial role in human thinking. Then, a *fuzzy set* is a class that can take a membership value in a continuous set of real numbers in the interval $[0, 1]$. In Fig. 1.5 we can see the fuzzy sets representing the sensations of temperature, having *soft transitions* between changes. Further on fuzzy sets have been extended to consider variations in the specific manner in which membership (for sets) and truth (for logic) is represented. Zadeh introduced in [40] interval-valued fuzzy sets, which considers the definition of the set by representing the membership degree as an interval in the $[0, 1]$ range instead of only one number in that same range. Even more, Zadeh introduced Type-2 Fuzzy Sets where both the lower and upper membership degree are represented by intervals. Many more generalizations of the idea of fuzzy set have raised from this theory, among others, Intuitionistic Fuzzy Sets [41], for which Attanasov introduce, coupled to the membership degree, the concept of non-membership degree, being able to capture more uncertainty as it quantifies how many of a variable does not correspond to a given set.

From a philosophical vision, fuzzy sets and its generalizations are a tool that permit a human approach to quantify phenomenons. As humans use words to express their reasoning and conclusions, from the premises extracted or learned. In this sense, Zadeh [42] present Fuzzy Logic as a tool to be able to compute with words and express propositions and object features in natural language, just like humans do.

In the particular context of fuzzy logic, comparison and, specifically, similarity have been defined by Zadeh as a generalization of the notion of equivalence. Many other authors have since studied its properties or its relation with perception [18]. As indicated by Zadeh, considering fuzzy sets as a linguistic representation, permits to interpret similarity with a semantic meaning, being able to mimic the human logic decision process in comparison tasks.

Literature contains different proposals to construct similarity measures in the context of fuzzy logic, mostly based on metrics adapted to fuzzy concepts. Among other, those based on the *Minkowsky r-metric* [43], like the $(d_2)^2$ proposed by Kacprzyk [44], or the dissemblance index by Kaufman and Gupta [45] (a deeper analysis and further examples are shown in [46], [47]). A particular way of constructing SM is by means of Restricted Equivalence Functions (REF) [48] which permits the comparison of membership degrees, *i.e.*, working with fuzzy information and hence with uncertainty.

Theory of comparison, in the context of fuzzy set theory, also relates to tasks other than pure comparison. For example, it relates to data ordering through a hypothetical triangular inequality, or to entropy, trough a multiple comparison. In fact there is a relation between all this concepts, and usually the notion of metric, due to its mathematical applicability, has been used as a similarity/dissimilarity measure. All of these concepts have been used in a variety of application topics, as image understanding [49], data clustering [50] or analysis of proximities [16].

As we have said earlier, humans receive a large quantity of information, mainly trough visual stimuli, forming an image of the real world. The acquisition process of the human visual system (HVS) consist in an optical transformation made in the eye, that makes the optical flow to be projected on the retina and then passed to the visual cortex [51]. From this optical flow the HVS is capable of extracting simple sketches and patterns building a mental model, used further on in comparison tasks and hence in the learning process. In contrast to the HVS, the digital model

of the natural vision, mimicked by images captured through camera sensors, are a discretization of the real word represented as a matrix where each point (pixel) contains colour information. Due to the inherent uncertainty of the discretization process, which can transmit errors, generating artefacts and deformation or capture partial information, interpreting pixel values is a complex task. There is a need for new tools being able to consider this uncertainty and that permit to process information as humans do.

We are in a great momentum, where digital imagery has acquired a great importance, as images are everywhere and are at the very centre of knowledge. Even if mostly all the imagery produced in the world has a social and recreation purpose, there exist other core subjects in the society such as industrial, security [52], [53] or medical applications [54], where image analysis plays an important role. In these specific fields, the use of images has been usually oriented to task supervised by humans, where their knowledge plays an essential role; until now, when the great increase in image quantity and quality, comes paired with unmanageable amounts of information and hence with the need of an automatic or semi-automatic processing. For this reason, we focus our analysis in image processing.

As we want computers/machines to see like humans, both comparison and human-like logical processing are tools to achieve our goal. Being able to perceive like humans implies understanding the wide variety of images that exists, applying the right tools to the right images. Images goes from the usual and well known grey-scale or colour images to binary ones (used in matching phases, *e.g.*, contour groundtruth images), or more specific types as the ones found in the medical context [55] (magnetic resonance, ultrasound, tomography, etc.), hyper-spectral imaging, *e.g.*in agriculture [56], [57], and very specific cases as orientation maps [58] used, *e.g.* in fingerprint analysis. All this variety of images lead us to ask ourselves what is an image? And, what makes and image to be an image?

On one hand we have the raw information, captured by sensors or generated by somebody, on the other hand what can be extracted, the features/characteristics that permit further processing, comparison and understanding [59]–[62]. The crucial point about images is not the image by itself or how it is represented but what it contains, and the availability of the correct comparison tools to extract and use the knowledge enclosed in it. In fact, a large variety of descriptors can be extracted from images to use them in comparison task and image interpretation, among other, we found texture [63] information as well as edges and contours [64]. This

is why there is a need of new approaches for image processing, more specifically, different ways of managing the inherent uncertainty, specific to each type of image, and being able to use it in comparison tasks.

Uncertainty in images can come in a variety of forms [65], [66], as patterns are not clearly defined; images can be corrupted during the acquisition process, objects in the image or the knowledge about them can be doubtful, or presented in ambiguous terms, etc. Joining the very nature of images with comparison and fuzzy logic, which is meant to deal with uncertainty, has been proved to provide an interesting approach that has lead many researchers to investigate this topic [67]. Different applications can be considered using fuzzy sets, from image enhancement [68] to edge detection and segmentation [69]. Moreover, comparison within the image context cover a extent of applications, from low-end to high-end processing, such as face recognition [70], image classification [71], boundary matching [72], image clustering [73] or deep convolutional neural networks for image labelling [74].

1.1 Motivation and Problem Statement

Digital image data has become an important piece of information in our world and needs an automatic processing to extract knowledge. The main problem with images is that the data they contain does not mean anything by itself, and when comparing data we have to manage the uncertainty inherent to this specific type of data. As a way to works as close as possible to human perception and how the human visual system works and manage the uncertainty, we want to join and work together with different topics, namely comparison, fuzzy logic and computer vision. We have tackled some aspects of digital processing, where fuzzy logic has an important role to play. From the most basic extraction of features of an image to the quality evaluation of a solution done by an automatic technique we need to manage uncertainty and comparison.

In the context of fuzzy sets, a common way to compare is through Similarity Measures (SM) [75], which, among other methods, can be constructed by means of aggregation functions [76] and Restricted Equivalence Functions (REF) [48]. These approaches perform well in scalar data and normal conditions. But, to the best of our knowledge, when considering other data types, *e.g.*, radial data, there are no studies quantifying the similarity in a radial universe. Moreover there exist no framework established to work with this type of data and only some proposed solutions adapted to a specific problem [77]–[80], but that in any case are interpretable in relation to adapt the human perception. To capture the essence of radial data, where the perception of increasing the distance between points makes the similarity to be smaller we adapted the concepts of REF and SM to radial data and analysed its behaviour. Also in the SM and REF context we propose another approach to capture the uncertainty in interval-valued data. Until now, for interval-valued data, in most of the cases only a partial order is considered, not permitting to recover the fundamental notions of fuzzy set theory, such as aggregation functions, implications, etc. This particular situation may lead to incomparable intervals and hence situations where monotonicity needs to be weakened. Moreover, to the best of our knowledge, there is no proposal considering the width of the interval in order to be able to capture the uncertainty between the input and output interval.

While analysing comparison aspects in fuzzy logic and images, to close the complete image processing framework, we finish our work with a new way of extracting

features in the context of edge detection and the construction of a taxonomy for different feature matching methods in edge extraction quality evaluation.

The main difficulty in edge detection arises from the inherent process of capturing real word information. An edge can be regarded as a location in which a big enough jump between neighbour pixel intensities happen. However, even this basic, loose definition can be criticized, since it would not consider situations as textures (where edges should not be detected) or hallucinated boundaries [81] (where boundaries appear with little or no intensity contrast). Many attempts have been done to define a framework for edge detection, being one of the most important the one presented by Bezdek [82], who propose four phases: *conditioning*, *feature extraction*, *blending* and *scaling*. Our efforts in edge detection techniques are mainly focused on the *feature extraction* step, that can be seen as an operation where the information received from each pixel is converted into problem-specific information. Even if this information is usually interpreted with gradient magnitudes, we also consider gradient direction in our proposal. We study the viability of using Ordered Directionally Monotone functions to fuse pixel neighbourhood information and build a feature map. The main advantage of this type of functions is that they are monotonic along different directions over the decreasingly ordered input vector, recovering the importance of the intensity differences.

Finally to close the analysis of the role of comparison in the image context we review the most relevant boundary matching techniques for edge quality evaluation, comparing their performance and seeking for the existence of a real impact over the evaluation depending on the approach used.

1.2 Objectives

The main objective of this thesis is to explore new ways of working with uncertainty for different types of data, considered from the point of view of human perception, in comparison tasks with a particular analysis in image processing.

In order to gain this objective we present some sub-objectives:

- Develop new concepts of similarity measures for different data types
- Define the concept of Restricted Equivalence Function and Similarity Measure to deal with radial data.
- Construct new distance measures and entropies for interval-valued data considering the with of the interval, understood as a measure of the uncertainty linked to that data, and using admissible and total orders.
- Present a new methodology to use the new concepts for radial data in the context of fingerprint analysis.
- Develop a new edge detector method based on ordered directionally monotone functions and consensus techniques.
- Build a taxonomy and compare the different quality evaluation matching methods for boundary detection.

1.3 Thesis Structure

This thesis is structured in two main parts. The first one (Chapter 2) is devoted to present the different concepts developed trough the Ph.D. thesis work analysing each one of the papers done for the research. We present each of the new concepts developed along with applications in the context of computer vision, comparing the results obtained with other methods from the literature.

Following the different applications and data types, we summarize all the theoretical concepts developed throughout this thesis and present a series of conclusions along with some future research lines.

The last part (Chapter 3) present all the papers developed during this thesis. Each one of the publications is accompanied by its status, the journal where it has been submitted, its impact factor and the text of the paper. The following papers are the ones associated with this thesis:

- A framework for radial data comparison and its application to fingerprint analysis [83]
- Width-based interval-valued distances and entropies [84]
- Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders [85]
- Ordered directional monotonicity in the construction of edge detectors [86]
- A survey on matching algorithms for boundary image comparison and evaluation [87]
- Optical images-based edge detection in Synthetic Aperture Radar images [88]

2 Phd thesis report

> *Information is the resolution of uncertainty.*
>
> — **Claude Shannon**
> (Mathematician)

In this chapter we expose the different concepts developed throughout this Ph.D. thesis detailing a series of preliminary concepts, an in depth analysis and some conclusions of each of the papers that conform this thesis. Then some general conclusions and future research lines are outlined.

2.1 A framework for radial data comparison and its application to fingerprint analysis

In this work we address the problem of working with radial data in the context of comparison. We study the similarity of scalar and non-scalar radial data and present the new concepts of Radial Restricted Equivalence Functions (RREF) and Radial Similarity Measures (RSM) as a new way of measuring distances for radial data.

Radial data is very often treated in computer vision problems, in a wide variety of forms, *e.g.*, angular, vector or tensorial data [89]. The utility of radial data requires well-defined tools to work with, *e.g.*, when comparing data.

Due to the specificity of this type of data and to be able to do some type of processing over it, there is a need for some conventions that we have used through this work.

In order to tackle the fact that two angles indicate the same direction, *e.g.*, $\frac{\pi}{2}$ and $\frac{5\pi}{2}$, we propose an equivalence relation R. Let $a, b \in \mathbb{R}$. Then, aRb if and only if $a = b + 2k\pi$, where $k \in \mathbb{Z}$. In this way, the equivalence class $[a] = \{b \mid bRa\}$ is a set that contains all the data sharing the same direction. Then, a semiopen interval having a width of 2π (*i.e.*, whose interval is $[\omega, \omega + 2\pi[$) is the quotient set (*i.e.*, it contains only one element of each equivalence class). A particular case of quotient set, which is the most frequent in radial data are $[0, 2\pi[$ and $[-\pi, \pi[$

The quotient that we work with in this thesis is the following, $\Omega = [0, 2\pi[$ and we define on Ω the classical operations sum, $a \oplus b = [a + b]$, and difference, $a \ominus b = [a - b]$, where $[t]$ denotes the only element $z \in \Omega$ such that zRt. Then under these conditions, we refer to *mirroring* as the mapping: $m : \Omega \longrightarrow \Omega$ such that $m(a) = a \oplus \pi$.

For the sake of our proposal to compare radial data on the quotient set Ω, we need the general definition of metric:

Definition 1 *A function $d : U \to \mathbb{R}^+$ is called metric on U if it satisfies the following:*

(D1) $d(a, b) = 0$ *if and only if* $a = b$*;*

(D2) $d(a, b) = d(b, a)$ *for all* $a, b \in U$*;*

(D3) $d(a, c) \leq d(a, b) + d(b, c)$ *for all* $a, b, c \in U$.

Radial data come with interesting characteristics for the field of similarity. In fact, while we increase de farness of two elements it can occur that we are making them closer. This apparent contradiction, which only occurs in linear data, is inherent to radial data.

The main contributions of this work are the definition of the concepts od RREF and SM and a way of construct them from automorphisms in the unit interval.

Definition 2 *[48] A mapping $r : [0, 1]^2 \to [0, 1]$ is said to be a Restricted Equivalence Function (REF) associated with the strong negation n if it satisfies the following:*

(R1) $r(x, y) = r(y, x)$ *for all* $x, y \in [0, 1]$*;*

(R2) $r(x,y) = 1$ *if and only if* $x = y$*;*

(R3) $r(x,y) = 0$ *if and only if* $\{x,y\} = \{0,1\}$*;*

(R4) $r(x,y) = r(n(x), n(y))$ *for all* $x, y \in [0,1]$*;*

(R5) For all $x, y, z, t \in [0,1]$*, such that* $x \leq y \leq z \leq t$ *then* $r(y,z) \geq r(x,t)$.

Note that (R5) means that, for all $x, y, z \in [0,1]$, if $x \leq y \leq z$ then $r(x,y) \geq r(x,z)$ and $r(y,z) \geq r(x,z)$.

Definition 3 *[75] A mapping* $s : [0,1]^k \times [0,1]^k \to \mathbb{R}^+$ *is called* k*-ary Similarity Measure (SM) associated with the strong negation* n *if it satisfies the following:*

(S1) $s(\boldsymbol{x},\boldsymbol{y}) = s(\boldsymbol{y},\boldsymbol{x})$ *for all* $\boldsymbol{x},\boldsymbol{y} \in [0,1]^k$*;*

(S2) $s(\boldsymbol{x}, n(\boldsymbol{x})) = 0$ *if and only if* $x_i \in \{0,1\}$ *for all* $i \in \{1,\ldots,k\}$ *and* $n(\boldsymbol{x}) = (n(x_1),\ldots,n(x_k))$*;*

(S3) $s(\boldsymbol{z},\boldsymbol{z}) = \max_{(\boldsymbol{x},\boldsymbol{y} \in [0,1]^k)} s(\boldsymbol{x},\boldsymbol{y})$ *for all* $z \in [0,1]^k$*;*

(S4) For all $\boldsymbol{x},\boldsymbol{y},\boldsymbol{z},\boldsymbol{t} \in [0,1]^k$*, if* $\boldsymbol{x} \leq \boldsymbol{y} \leq \boldsymbol{z} \leq \boldsymbol{t}$ *then* $s(\boldsymbol{y},\boldsymbol{z}) \geq s(\boldsymbol{x},\boldsymbol{t})$ *where* $\boldsymbol{x} \leq \boldsymbol{y}$ *implies that* $x_i \leq y_i$ *for all* $i \in \{1,\ldots,k\}$.

Similarly to what happens for REFs, the property (S4) is equivalent to: For all $\mathbf{x},\mathbf{y},\mathbf{z} \in [0,1]^k$, if $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z}$ then $s(\mathbf{x},\mathbf{y}) \geq s(\mathbf{x},\mathbf{z})$ and $s(\mathbf{y},\mathbf{z}) \geq s(\mathbf{x},\mathbf{z})$.

Similarity Measures (SMs) can be constructed in different ways, although the most popular method originates from the combination of REFs and aggregation functions [76].

Definition 4 *A mapping* $r_\theta : \Omega^2 \to [0,1]$ *is called a Restricted Radial Equivalence Function (RREF) associated with the metric* d *if it satisfies the following:*

(RR1) $r_\theta(a,b) = r_\theta(b,a)$ *for all* $a, b \in \Omega$*;*

(RR2) $r_\theta(a,b) = 1$ *if and only if* $d(a,b) = 0$*;*

(RR3) $r_\theta(a,b) = 0$ *if and only if* $d(a,b)$ *is maximum;*

(RR4) $r_\theta(a,b) = r_\theta(m(a), m(b))$ *for all* $a, b \in \Omega$*;*

(RR5) For all $a, b, c, d \in \Omega$*, if* $d(b,c) \leq d(a,d)$*, then* $r_\theta(b,c) \geq r_\theta(a,d)$*.*

Definition 4 is not a direct extension of Definition 2 to radial data. The differences arise from the use of distances in (RR5) instead of orders (as in (R5)). Even considering this change, which is due to the inherent interpretation of orders for radial data, we understand that the semantics and sense of REFs are preserved.

Proposition 1 *Let* r_θ *be a RREF associated with the metric* d^**. For all* $a_1, b_1, a_2, b_2 \in \Omega$*, if* $d^*(a_1, b_1) = d^*(a_2, b_2)$ *then* $r_\theta(a_1, b_1) = r_\theta(a_2, b_2)$*.*

In this work, we only consider RREFs associated with the angular metric $d^*(a,b) = \min(|b-a|, 2\pi - |b-a|)$.

Following the construction of REFs, we also define a new possible construction method for RREFs, which is also based on automorphisms.

Proposition 2 *Let* φ *and* ψ *be automorphisms of the intervals* $[0,1]$ *and* $[0,\pi]$*, respectively. The mapping* $t : \Omega^2 \to [0,1]$ *given by*

$$t(a,b) = \varphi^{-1}\left(1 - \left(\frac{1}{\pi}\psi\left(d^*(a,b)\right)\right)\right) \tag{2.1}$$

is a RREF.

Definition 5 *A mapping* $s_\theta : \Omega^k \times \Omega^k \to \mathbb{R}^+$ *is said to be a* k*-ary Radial Similarity Measure (RSM) associated with the metric* d^* *if it satisfies the following:*

(SR1) $s_\theta(\boldsymbol{a}, \boldsymbol{b}) = s_\theta(\boldsymbol{b}, \boldsymbol{a})$ *for all* $\boldsymbol{a}, \boldsymbol{b} \in \Omega^k$*;*

(SR2) $s_\theta(\boldsymbol{a}, \boldsymbol{b}) = 0$ *if and only if* $d^*(a_i, b_i) = \pi$ *for all* $i \in \{1, \ldots, k\}$*;*

(SR3) $s_\theta(\boldsymbol{c},\boldsymbol{c}) = Max_{\boldsymbol{a},\boldsymbol{b}\in\Omega^k} s_\theta(\boldsymbol{a},\boldsymbol{b})$ *for all* $c \in \Omega^k$*;*

(SR4) For all $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d} \in \Omega^k$*, if* $d^*(\boldsymbol{a},\boldsymbol{d}) \geq d^*(\boldsymbol{b},\boldsymbol{c})$ *then* $s_\theta(\boldsymbol{a},\boldsymbol{d}) \leq s_\theta(\boldsymbol{b},\boldsymbol{c})$*, where* $d^*(\boldsymbol{a},\boldsymbol{d}) \geq d^*(\boldsymbol{b},\boldsymbol{c})$ *implies that* $d^*(a_i,d_i) \geq d^*(b_i,c_i)$ *for all* $i \in \{1,\ldots,k\}$*.*

Then, using the same methodology as in SM, to construct RSMs we use RREFs, to aggregate their results.

Proposition 3 *Let* r_θ *be a RREF and let* f *be a* k*-ary aggregation function such that* $f(\boldsymbol{x}) = 0$ *if and only if* $x_i = 0$ *for all* $i \in \{1,\ldots,k\}$ *and* $f(\boldsymbol{x}) = 1$ *if and only if* $x_i = 1$ *for all* $i \in \{1,\ldots,k\}$*. The function* $s_{\theta_{[f,r_\theta]}} : \Omega^k \times \Omega^k$*, given by*

$$s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{b}) = f(r_\theta(a_1,b_1),\ldots,r_\theta(a_k,b_k)) \tag{2.2}$$

is a k*-ary radial similarity measure that satisfies*

- $s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{b}) = s_{\theta_{[f,r_\theta]}}(\boldsymbol{b},\boldsymbol{a})$ *for all* $\boldsymbol{a},\boldsymbol{b} \in \Omega^k$*;*
- $s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{b}) = 0$ *if and only if* $d^*(a_i,b_i) = \pi$ *for all* $i \in \{1,\ldots,k\}$*;*
- $s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{b}) = 1$ *if and only if* $a_i = b_i$ *for all* $i \in \{1,\ldots,k\}$*;*
- *For all* $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c},\boldsymbol{d} \in \Omega^k$*, if* $d^*(\boldsymbol{a},\boldsymbol{d}) \geq d^*(\boldsymbol{b},\boldsymbol{c})$ *then* $s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{d}) \leq s_{\theta_{[f,r_\theta]}}(\boldsymbol{b},\boldsymbol{c})$*, where* $d^*(\boldsymbol{a},\boldsymbol{d}) \geq d^*(\boldsymbol{b},\boldsymbol{c})$ *implies that* $d^*(a_i,d_i) \geq d^*(b_i,c_i)$ *for all* $i \in \{1,\ldots,k\}$*.*
- $s_{\theta_{[f,r_\theta]}}(\boldsymbol{a},\boldsymbol{b}) = s_{\theta_{[f,r_\theta]}}(m(\boldsymbol{a}),m(\boldsymbol{b}))$ *for all* $\boldsymbol{a},\boldsymbol{b} \in \Omega^k$ *where* $m(\boldsymbol{a}) = (m(a_1),\ldots,m(a_k))$*.*

In addition to the theoretical advances presented in this work, we present a possible application where the new concepts are used. Particularly we use RREF and RSM in singular point detection for fingerprint analysis. Fingerprint-based authentication consists in analysing ridges present in our fingertips in order to find patterns. These local patterns are known as *minutiae*, and are usually bifurcations, discontinuities or ridge breaks [90], [91]. There exist two main type of local patterns in our fingertips that automated systems look for: cores and deltas. Then the process of finding those anomalies in the fingertips of an individual being and compare it with another one is known as matching [92].

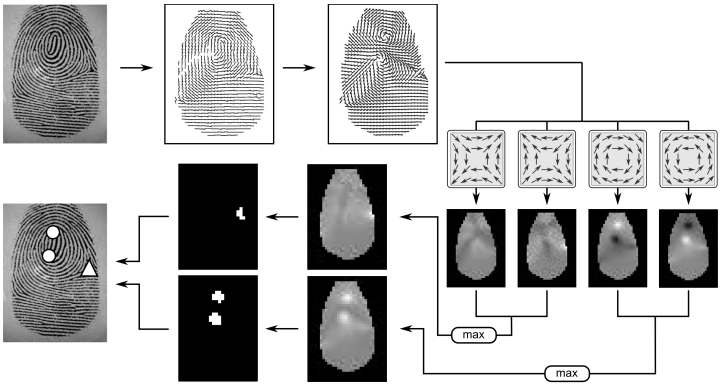


Fig. 2.1: Schematic representation of the proposed framework for singular point detection using orientation templates and radial similarity measures, namely Template-based singular point detection (TSPD).

The complete process of our proposal is represented in Fig. 2.1. The different steps of the TSPD are the following:

1. Dividing the image into non-overlapping blocks.

2. Segmenting the image using the previous calculated blocks.

    a) Normalizing the image to a desired mean and variance, 100 and 1000, respectively [91].

    b) Segmenting the fingerprint and the background, by assigning to the latter those blocks for which the variance of the pixel intensities is greater than 30 [91].

3. Calculating the orientation map over the segmented image.

    a) Computing the gradient at each pixel of the image (e.g., using Sobel masks) [93].

    b) Since the gradients are computed for each pixel and the result of a single pixel may not be reliable enough, the OM is smoothed to get more accurate orientations. In order to do so, the technique by Kass and Witkin [94] is used.

c) Creating the OM from the regularized orientations.

4. Creating the SqOM by multiplying by two the values in the OM.

5. Detecting singular points.

   a) Computing the similarity map for each template. This is done by comparing the elements of the SqOM and those in each templates using the RSMs.

   b) Fusing the similarity maps corresponding to each type of SP. This is done by obtaining, at each block, the maximum response for the cores and, in parallel, for the deltas.

   c) Selecting cores and deltas. This is done by taking the two points with the highest local response for core and delta similarity map in parallel.

We have tested our proposal with well-known state of the art approaches as Poincaré and Liu.

With this work we arrive to two main conclusions. We have made an important contribution, adapting the concepts of REF and SM to radial data. The new concepts that arose, namely RREF and RSM, are able to capture the nature of radial data while retaining the behaviour and semantics of the original operators. Moreover, we have proved the usefulness of the new concepts and their validity in a complex context as fingerprint analysis

This section of the thesis is associated with the following publication:

- C. Marco-Detchart, J. Cerron, L. De Miguel, C. Lopez-Molina, H. Bustince, and M. Galar, "A framework for radial data comparison and its application to fingerprint analysis", *Applied Soft Computing Journal*, vol. 46, pp. 246–259, 2016

2.2 Width-based interval-valued distances and entropies

In this paper, we study the behaviour of entropies and distance measures in the context of interval-valued fuzzy sets. These two concepts are well-known approaches in the comparison theory and widely used notions in the fuzzy sets theory. Our proposal in this work has mainly two novelties. The first one consists in using the width of intervals in order to connect the uncertainty of the output with the uncertainty of the inputs. And the second one lies in bringing in total orders when working between intervals, instead of partial ones, being able to recover for interval-valued set-ups some of the advantages of the notions of monotonicity. The construction of distance measures and entropies is done trough the aggregation of interval-valued restricted dissimilarity functions and interval-valued normal E_N-functions.

The main concepts developed in this work are the following.

Proposition 4 *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function and let $d : [0,1]^2 \to [0,1]$ be a restricted dissimilarity function. Then, the function $d_{IV} : L([0,1])^2 \to L([0,1])$ given by*

$$d_{IV}(X,Y) = \big[\min\big(d\left(K_\alpha(X), K_\alpha(Y)\right), 1 - M(w(X), w(Y))\big), \\ \min\big(1, d\left(K_\alpha(X), K_\alpha(Y)\right) + M(w(X), w(Y))\big)\big] \quad (2.3)$$

is an IV restricted dissimilarity function w.r.t. any admissible order $\leq_{TL}$. Moreover, d_{IV} is w-preserving.

The construction method given in Proposition (4) can be simplified with some additional assumptions on M and d.

Corollary 1 *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function such that $M(x,y) \leq \min\big((1-\alpha)x + \alpha y, \alpha x + (1-\alpha)y\big)$ for all $x, y \in [0,1]$ and let $d : [0,1]^2 \to [0,1]$ be a restricted dissimilarity function such*

that $d(x,y) \leq |x-y|$ for all $x \in [0,1]$. Then, the function $d_{IV} : L([0,1])^2 \to L([0,1])$ given by

$$d_{IV}(X,Y) = [d\left(K_\alpha(X), K_\alpha(Y)\right), d\left(K_\alpha(X), K_\alpha(Y)\right) + M(w(X), w(Y))] \quad (2.4)$$

is an IV restricted dissimilarity function w.r.t. any admissible order $\leq_{TL}$. Moreover, d_{IV} is w-preserving.

Proposition 5 *Let $n : [0,1] \to [0,1]$ be a strong negation with equilibrium point e. Let $\alpha, \beta \in]0,1[$, $\beta \neq \alpha$ and $N : L([0,1]) \to L([0,1])$ be a strong IV negation w.r.t. $\leq_{\alpha,\beta}$ with equilibrium point ε and such that $K_\alpha(\varepsilon) = e$. Let $E_N : [0,1]^2 \to [0,1]$ be a normal E_N-function w.r.t. n. Then, the function $EN_{IV} : L([0,1]) \to L([0,1])$ given by*

$$\begin{aligned} EN_{IV}(X) \quad = \quad & \left[\max\left(0, E_N\left(K_\alpha(X)\right) - w(X)\right),\right. \\ & \left.\max\left(E_N\left(K_\alpha(X)\right), w(X)\right)\right] \end{aligned} \quad (2.5)$$

is an IV normal E_N-function w.r.t. N. Moreover, EN_{IV} is w-preserving.

Corollary 2 *Let $\leq_{XY}$ be the Xu and Yager order and $N : L([0,1]) \to L([0,1])$ be an IV strong negation w.r.t. $\leq_{XY}$ with the equilibrium point ε such that $\underline{\varepsilon} + \overline{\varepsilon} = 1$. Let $E_N : [0,1]^2 \to [0,1]$ be a normal E_N-function w.r.t. a strong negation n with the equilibrium point $e = 1/2$ such that $E_N(x) \geq 1 - |2x-1|$ for all $x \in [0,1]$. Then, the function $EN_{IV} : L([0,1]) \to L([0,1])$ given by*

$$EN_{IV}(X) = \left[E_N\left(\frac{\underline{X} + \overline{X}}{2}\right) - w(X), E_N\left(\frac{\underline{X} + \overline{X}}{2}\right)\right] \quad (2.6)$$

is an IV normal E_N-function w.r.t. N. Moreover, EN_{IV} is w-preserving.

Then using w-preserving IV aggregation functions w.r.t. $\leq_{\alpha,\beta}$ we can construct and IV entropy.

Proposition 6 *Let $U = \{u_1, \ldots, u_n\}$ and let $N : L([0,1]) \to L([0,1])$ be a strong IV negation w.r.t. a total order $\leq_{TL}$. Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be an idempotent IV aggregation function w.r.t. $\leq_{TL}$ satisfying $M_{IV}(X_1, \ldots, X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Let $EN_{IV} : L([0,1]) \to L([0,1])$ be an IV normal E_N-function*

*w.r.t. N (given by Definition **??**). Then, the function $E : IVFS(U) \to L([0,1])$, defined by:*

$$E(A) = M_{IV}\left(EN_{IV}\left(A(u_1)\right), \ldots, EN_{IV}\left(A(u_n)\right)\right)$$

for all $A \in IVFS(U)$, is an IV entropy on $IVFS(U)$ with respect to the strong IV negation N.

We study the conditions under which the function EN_{IV} given by Equation (2.5) can be used in the previous proposition to obtain an IV entropy.

Corollary 3 *Let $U = \{u_1, \ldots, u_n\}$ and $\alpha, \beta \in]0,1[$ with $\beta \neq \alpha$. Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be an IV aggregation function w.r.t. $\leq_{\alpha,\beta}$ defined by two aggregation functions M_1, M_2, as in Proposition 2. Let $EN_{IV} : L([0,1]) \to L([0,1])$ be an IV normal E_N-function given in terms of a normal E_N-function E_N, as in Proposition 5, with N a strong IV negation w.r.t. $\leq_{\alpha,\beta}$ with an equilibrium point ε. Let $E : IVFS(U) \to L([0,1])$ be a function defined by:*

$$E(A) = M_{IV}\left(EN_{IV}\left(A(u_1)\right), \ldots, EN_{IV}\left(A(u_n)\right)\right)$$

for all $A \in IVFS(U)$. Then

(i) E satisfies axiom (E1), if

$$M_1(x_1, \ldots, x_n) = 0 \text{ if and only if } x_1 = \ldots = x_n = 0.$$

(ii) E satisfies axiom (E2), if M_1 and M_2 are idempotent.

(iii) E satisfies axiom (E3) w.r.t. $\leq_{\alpha,\beta}$.

(iv) Let M_2 be idempotent. Then, for all $A \in IVFS(U)$, $w(A(u_1)) = \ldots = w(A(u_n))$ implies $w(E(A)) = w(A(u_1))$.

With the approach presented in this work we are getting an interval-valued output, which permits to recover the uncertainty of the input interval and maintain the interval-valued representation.

Once we have defined the new expressions to work with both distance measures and entropy preserving the width of the interval we focus our attention on the

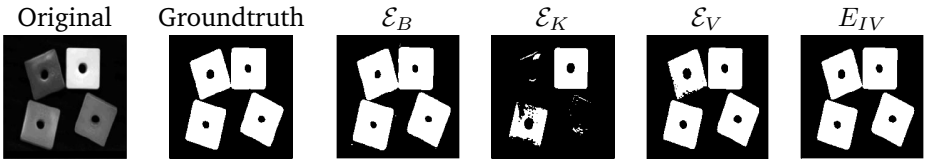


Fig. 2.2: Comparison of the different approaches for thresholding an image compared with the groundtruth provided by the dataset.

application part. One of the well-studied processes in the literature of computer vision is image segmentation, which consists in the extraction of partitions from an image (regions sharing common characteristics) that represent the objects in it [95]. Segmenting an image, is done by assigning the same label to all the pixels (with the same properties) representing the possible object to whom they belong. Segmentation can be as fine or coarse as the number of labels we want to assign to the image.

To prove the validity of the theoretical concept presented in this work we put it to the test by using it in the context of image thresholding, as being one of the usual and most basic techniques to label an image [96]. The thresholding process consists in finding the grey intensity value that best separates two regions (background and object) of an image.

We obtain these two regions by using a modified version of the algorithm presented by Huang and Wang [97] for working with interval-valued data, which is illustrated in Algorithm 1.

Algorithm 1 Algorithm for thresholding an image using entropy.

Input: Image I with L intensity values.
Output: Image threshold t.

1: **for** each level of intensity t, $(t = 0, t = 1, \ldots, L-1)$ **do**
2: Build k fuzzy sets $Q_t^1 \ldots Q_t^k$;
3: Build an IVFS $\tilde{Q}_t$ from the fuzzy sets $Q_t^1 \ldots Q_t^k$;
4: **for** each $q \in \{0, \ldots, L-1\}$ **do**
5: $\mu_{\tilde{Q}_t}(q) = \left[T\left(\mu_{Q_t^1}(q),\ \ldots,\ \mu_{Q_t^k}(q)\right), S\left(\mu_{Q_t^1}(q), \ldots, \mu_{Q_t^k}(q)\right)\right]$, with T a t-norm and S a t-conorm.
6: **end for**
7: Compute the entropy of each of the L interval valued fuzzy sets $\tilde{Q}_t$;
8: **end for**
9: Select the threshold t with the smallest entropy.;

In the different experiments made in this work, where we vary the construction method for the fuzzy sets of step 2, we can see that our new entropy expression behave quite similarly to other expressions in the literature. When only using REF membership function our new entropy is comparable, while when combining REF and triangular membership functions, our expression does not perform well and remain just under other expressions. Finally, when adding to the membership function construction S-Z functions, our entropy is the best performer.

For comparison purposes we test our new entropy definition versus well-known expressions of the literature as Sambuc's indetermination index [98], Kacpryzk and Smidzt expression [99] and Vlachos and Sergiaidis expression [100].

Moreover we have put to the test our approach with classical methods for thresholding as *Otsu* [101], *Area-based* [102] and *Tizhoosh* [103] (the specific parameters for each one of the approaches are detailed in the paper in Section 2.2). A visual example of the results obtained with the different approaches is shown in Fig 2.2.

With this work we have approached a different type of data and managed to define an expression for interval-valued dissimilarity functions and interval-valued E_N-functions. One of the important achievements of this work is that the two concepts introduced have been defined taking into account a total order between intervals, which have lead to the construction of interval-valued distances and entropies.

This section of the thesis is associated with the following publication:

- Z. Takáč, H. Bustince, J. M. Pintor, C. Marco-Detchart, and I. Couso, "Width-based interval-valued distances and fuzzy entropies", *IEEE Access*, vol. 7, pp. 14 044–14 057, 2019

2.3 Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders

In this study we present a new type of similarity measures between interval-valued data. The new construction method is based on two novelties. On one hand, we construct the new similarity expression with respect to a total order for intervals. On the other hand, we consider the width of the interval in order to manage the uncertainty of the output linked to the one from the input.

In order to be able to define the new similarity measure for interval-valued fuzzy sets we need to define restricted equivalence functions and aggregation functions in the interval-valued setting.

Definition 6 *Let $\leq_L$ be an order on $L([0,1])$. An interval-valued restricted equivalence function w.r.t. the order $\leq_L$ is a function $R_{IV} : L([0,1])^2 \to L([0,1])$ such that:*

1. *$R_{IV}(X,Y) = 0_L$ if and only if $\{X,Y\} = \{0_L, 1_L\}$;*

2. *$R_{IV}(X,X) = [1 - w(X), 1]$ for all $X \in L([0,1])$;*

3. *$R_{IV}(X,Y) = R_{IV}(Y,X)$ for all $X, Y \in L([0,1])$;*

4. *If $X, Y, Z \in L([0,1])$ are such that $X \leq_L Y \leq_L Z$ and $w(X) = w(Y) = w(Z)$, then $R_{IV}(X,Z) \leq_L R_{IV}(X,Y)$ and $R_{IV}(X,Z) \leq_L R_{IV}(Y,Z)$.*

Note that the previous definition is does not preserve the with of the intervals but is a preliminary step to define a w-preserving IVREF.

Theorem 1 *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function and let $R : [0,1]^2 \to [0,1]$ be a restricted equivalence function. Then, the function $R_{IV} : L([0,1])^2 \to L([0,1])$ given by*

$$R_{IV}(X,Y) = \big[\max\big(0, R\left(K_\alpha(X), K_\alpha(Y)\right) - M(w(X), w(Y))\big), \\ \max\left(R\left(K_\alpha(X), K_\alpha(Y)\right), M(w(X), w(Y))\right)\big] \quad (2.7)$$

is an IV restricted equivalence function w.r.t. any admissible order $\leq_{TL}$. Moreover, R_{IV} is w-preserving.

When constructing an IVREF any REF R and any idempotent symmetric aggregation function M can be used in Equation (2.7). Then if additional assumptions are imposed to R and M Equation (2.7) can be simplified.

Corollary 4 *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function such that $M(x,y) \leq \min\big((1-\alpha)x + \alpha y, \alpha x + (1-\alpha)y\big)$ for all $x, y \in [0,1]$, and let $R : [0,1]^2 \to [0,1]$ be a restricted equivalence function such that $R(x,y) \geq 1 - |x-y|$ for all $x, y \in [0,1]$. Then, the function $R_{IV} : L([0,1])^2 \to L([0,1])$ given by*

$$R_{IV}(X,Y) = \big[R\left(K_\alpha(X), K_\alpha(Y)\right) - M(w(X), w(Y)), R\left(K_\alpha(X), K_\alpha(Y)\right)\big] \quad (2.8)$$

is an IV restricted equivalence function w.r.t. any admissible order $\leq_{TL}$. Moreover, R_{IV} is w-preserving.

In order to construct IV-aggregation functions preserving the width of the interval, w.r.t. $\leq_{\alpha,\beta}$, the following two properties are needed:

Given an aggregation function $M : [0,1]^n \to [0,1]$,

(P1) $M(cx_1, \ldots, cx_n) \geq cM(x_1, \ldots, x_n)$ for all $c \in [0,1]$, $x_1, \ldots, x_n \in [0,1]$.

(P2) $M(x_1, \ldots, x_n) \leq 1 - M(1-x_1, \ldots, 1-x_n)$ for all $x_1, \ldots, x_n \in [0,1]$.

Theorem 2 *Let $\alpha, \beta \in [0,1]$, $\beta \neq \alpha$. Let $M_1, M_2 : [0,1]^n \to [0,1]$ be aggregation functions such that M_1 is strictly increasing, $M_1(x_1, \ldots, x_n) \geq M_2(x_1, \ldots, x_n)$ for all*

$x_1, \ldots, x_n \in [0,1]$, *$M_1$ or M_2 satisfies property (P1) and M_1 or M_2 satisfies property (P2). Then $M_{IV} : (L([0,1]))^n \to L([0,1])$ defined by:*

$$M_{IV}(X_1, \ldots, X_n) = Y, \quad \text{where} \quad \begin{cases} K_\alpha(Y) = M_1\left(K_\alpha(X_1), \ldots, K_\alpha(X_n)\right), \\ w(Y) = M_2\left(w(X_1), \ldots, w(X_n)\right), \end{cases}$$

for all $X_1, \ldots, X_n \in L([0,1])$, is an IV aggregation function with respect to $\leq_{\alpha,\beta}$.

Moreover, if M_2 is idempotent, then M_{IV} is w-preserving.

Once we have defined an w-preserving IV-aggregation we can construct an IV-similarity preserving the width of the interval. The construction is done inspired by the concept of SM in [48], relaxing the second and fourth axioms as in Definition 6 and aggregating IVREFs.

Definition 7 . *Let $\leq_L$ be an order on $L([0,1])$ and $M : [0,1]^n \to [0,1]$ be an aggregation function. A width-based interval-valued similarity measure on $IVFS(U)$ w.r.t. $\leq_L$, associated with M is a mapping $S_M : IVFS(U) \times IVFS(U) \to L([0,1])$ such that, for all $A, B, A', B' \in IVFS(U)$,*

(SM1) $S_M(A,B) = S_M(B,A)$;

(SM2) $S_M(A,A) = \Big[1 - M\big(w(A(u_1)), \ldots, w(A(u_n))\big), 1\Big]$;

(SM3) $S_M(A,B) = 0_L$ *if and only if* $\{A(u_i), B(u_i)\} = \{0_L, 1_L\}$ *for all* $i \in \{1, \ldots, n\}$;

(SM4) If $A \subseteq A' \subseteq B' \subseteq B$ w.r.t. $\leq_L$ and $w(A(u_i)) = w(A'(u_i)) = w(B'(u_i)) = w(B(u_i))$ for all $i \in \{1, \ldots, n\}$, then $S_M(A,B) \leq_L S_M(A',B')$, where, for $A, B \in IVFS(U)$, $A \subseteq B$ w.r.t. $\leq_L$ if $A(u_i) \leq_L B(u_i)$ for every $u_i \in U$.

Theorem 3 *Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be a decomposable IV aggregation function w.r.t. $\leq_L$ associated with M_L and M_U where M_L is self-dual, and let $M_{IV}(X_1, \ldots, X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Let $R_{IV} : L([0,1])^2 \to L([0,1])$ be an IV restricted equivalence function w.r.t. $\leq_L$. Then the function $S_{M_L} : IVFS(U) \times IVFS(U) \to L([0,1])$ defined by:*

$$S_{M_L}(A,B) = M_{IV}\big(R_{IV}\left(A(u_1), B(u_1)\right), \ldots, R_{IV}\left(A(u_n), B(u_n)\right)\big)$$

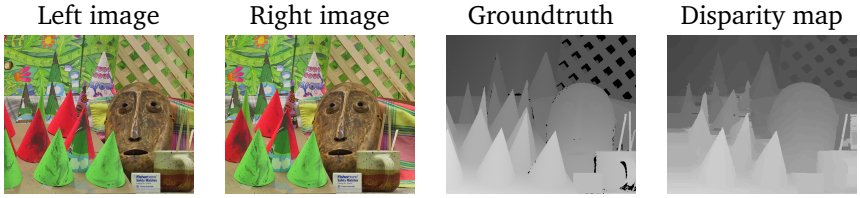


Fig. 2.3: Left and right cones image and groundtruth proposed in the dataset Middlebury along with the disparity map obtained.

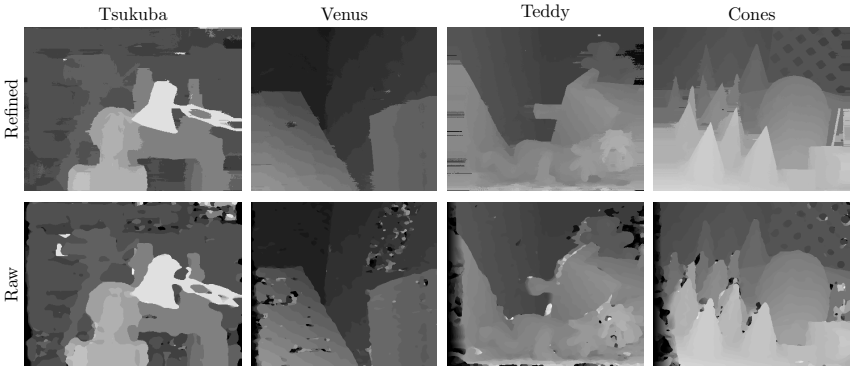


Fig. 2.4: Comparison of the raw disparity maps obtained with IV-REF and after applying refinement and outlier detection techniques

for all $A, B \in IVFS(U)$ *is a width-based IV similarity measure on* $IVFS(U)$ *w.r.t.* $\leq_L$ *associated with* M_L.

On the second part of this work we present a possible application of the new concepts defined. We use the construction process presented for IV-similarity measures preserving the interval to create an approximation of the binocular human vision. That is,we present a stereo matching application using w-preserving IV-REF. To catch the correspondent points from the two images we compare blocks of the images and build a disparity map which give the depth at which each object is located.

The correspondence procedure to build the disparity map is exposed in Algorithm 2 and the different images involved are shown in Fig. 2.3.

In step 7 each of the similarities obtained for each colour channel using a width-based IV similarity, *i.e.*$S_{M_R}(A,B), S_{M_G}(A,B), S_{M_B}(A,B)$, are aggregated using

Algorithm 2 Algorithm for constructing a disparity map using IV-similarity measures.

Input: Left anf right colour images f_l, f_r, an IV-similarity measure S_M.
Output: Disparity map $\mho_d$.

1: IV-fuzzify the images f_l, f_r, getting three IVFSs for each image, one for each colour channel.
2: **for** each pixel (x, y) of f_r **do**
3: Select a window of size $n \times m$ around the pixel;
4: **for** each possible y' until the maximal disparity (provided by the dataset) **do**
5: Select a window of size $n \times m$ around the pixel (x, y');
6: Calculate the IV-similarity between the two windows, in each of the three colour channels using the similarity S_M;
7: Aggregate the values of the IV-similarities for each colour according to Equation (2.9).
8: **end for**
9: Calculate the disparity between windows taking the pair of windows of greatest similarity according to the order relation $\preceq_{\alpha,\beta}$ with $\alpha = 1$;
10: **end for**
11: Create the disparity map with each of the disparities obtained for each position (x, y);

different alternatives as the weighted mean, the arithmetic mean, the product, the geometric mean, the harmonic mean, the median, the maximum and the minimum. Finally, as proposed in [104] the different similarities are aggregated using Eq 2.9, appearing to be the best performer.

$$S_{MT}(A, B) = 0.299 \cdot S_{M_R}(A, B) + 0.587 \cdot S_{M_G}(A, B) + 0.114 \cdot S_{M_B}(A, B) \quad (2.9)$$

We test our approach over the Middlebury dataset and compare it with classical methods from the literature as SSD [105], SAD [106], NCC [107] o ZNCC [108]. In addition to the classical proposals we test our method to the previous work from Galar *et al.*using interval-valued fuzzy sets [109]. In addition, as an improvement of our method and not focusing only on raw disparity maps we analyse the improvements obtained applying a refinement procedure, seeing the result in Fig. 2.4

Within this work we conclude that the new definition for IV-REF and IV-SM works with a total order and are able to relate the uncertainty between the extremes of the intervals. The use of these new expressions in a stereo matching application have shown that they provide a high benefit in the raw disparity map extraction with respect to the usual techniques.

This section of the thesis is associated with the following publication:

- H Bustince, C Marco-Detchart, J Fernandez, C Wagner, J. Garibaldi, and Z Takáč, "Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders", *Fuzzy Sets and Systems*, 2019

2.4 Ordered directional monotonicity in the construction of edge detectors

In the present work we propose the construction of two edge detectors by means of Ordered Directionally monotone functions. In addition we build a consensus edge detector combining both of our proposals with other methods from the literature. For the construction of the edge detector we follow the framework proposed by Bezdek *et al.* [82]:

$(S1)$ Conditioning: Smooth the image applying a Gaussian filter.

$(S2)$ Feature extraction: Obtain a feature image by means of ordered directionally monotone functions.

$(S3)$ Blending: Thin the feature image using the non-maximum suppression procedure [110].

$(S4)$ Scaling: Binarize the feature image with the hysteresis method [111].

The main proposal of this work is centred in ($(S2)$) step, using two construction methods to obtain OD-monotone functions that measure the local intensity variation from pixel to pixel within a neighbourhood.

We use the Corollary 5 (Corollary 8 in the paper) to present two alternative constructions for feature extraction.

Corollary 5 *Let $p > 0$. Let $G : [0,1]^n \to [0,1]$ be defined, for $\mathbf{x} \in [0,1]^n$ and $\sigma \in P_n$ such that $x_{\sigma(1)} \geq \ldots \geq x_{\sigma(n)}$, by*

$$G(\mathbf{x}) = \left(a + \sum_{i=1}^{n} b_i x_{\sigma(i)} \right)^{\frac{1}{p}}, \tag{2.10}$$

for some $a \in [0,1]$ and $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ such that $0 \leq a + b_1 + \cdots + b_j \leq 1$ for all $j \in \{1, \ldots, n\}$. Then G is OD $\vec{r}$-increasing for every non-null vector $\vec{r}$ such that $\vec{b} \cdot \vec{r} \geq 0$.

To extract the local features we work with an 8-neighbourhood around each position (x, y) (*i.e.* neighbours from $(x-1, y-1)$ to $(x+1, y+1)$). Then each one of the values needed in $(S2)$ is computed as:

$$x_1 = |a_{(x,y)} - a_{(x-1,y-1)}|, \ldots, x_8 = |a_{(x,y)} - a_{(x+1,y+1)}|$$

These differences are ordered in a decreasing way obtaining:

$$\vec{r} = (x_{\sigma_{(1)}}, x_{\sigma_{(2)}}, x_{\sigma_{(3)}}, x_{\sigma_{(4)}}, x_{\sigma_{(5)}}, x_{\sigma_{(6)}}, x_{\sigma_{(7)}}, x_{\sigma_{(8)}})$$

and then an ODM function is applied to each position in the image, with different a, p, and $\vec{b}$ parameters. On both of the proposals we maintain $a = 0$ and $\frac{1}{p} = 0.30$.

For the first ODM construction (*Case 1*) we use the following $\vec{b}$:

$$\vec{b} = \begin{cases} \left(\frac{x_{\sigma_{(1)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}}, \ldots, \frac{x_{\sigma_{(7)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}}, \frac{x_{\sigma_{(8)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}} \right) & \text{if } \sum_{i=1}^{8} x_{\sigma_{(i)}} \neq 0 \\ (0, \ldots, 0) & \text{otherwise.} \end{cases}$$

While for the second proposed construction (*Case 2*) we maintain the same previous parameters and change $\vec{b}$ as follows:

$$\vec{b} = \left(\frac{1}{8} \left(1 - \left| x_{\sigma_{(1)}} - \underset{i \in \{1,\ldots,8\}}{\text{median}} \{x_i\} \right| \right), \ldots, \frac{1}{8} \left(1 - \left| x_{\sigma_{(8)}} - \underset{i \in \{1,\ldots,8\}}{\text{median}} \{x_i\} \right| \right) \right);$$

Finally as a complementary experiment in order to improve the behaviour of our edge detector we propose the construction of a consensus edge detector by combining our two construction methods with well-known edge detectors in the literature like Canny [110], gravitational approach [112], [113], fuzzy morphology approach [114] and Structured Forests method [115].

From this work we can conclude that the use of ODM functions to measure the change of intensity between pixels in a neighbourhood overcomes the Canny method and most of the literature method. Only the Structured Forest surpasses our proposal, when used independently. When the consensus construction is used to extract edges then our method is the best performer.

This section of the thesis is associated with the following publication:

- C. Marco-Detchart, H. Bustince, J. Fernandez, R. Mesiar, J. Lafuente, E. Barrenechea, and J. M. Pintor, "Ordered directional monotonicity in the construction of edge detectors", *Fuzzy Sets and Systems*,

2.5 A survey on matching algorithms for boundary image comparison and evaluation

In this work we have studied different approaches for boundary quality evaluation and performed a quantitative analysis of their behaviour. Boundary image evaluation is a crucial step in order to know how a boundary detection process performs. The most extended way of evaluation is by comparing a boundary extraction method to what a human has defined as a correct solution (groundtruth). In recent years boundary evaluation has been approaches as a classification problem. Boundary matching is really a binary classification where each pixels is to be labelled as a boundary or not. Then as we are dealing with a classification problem the usual quality measures can be used: True Positives (TPs), True Negatives (TNs), etc.

We can put boundary matching and quality evaluation in mathematical terms. In this work, we consider all the images to have a dimension of $M \times N$, so that the set of positions is $\Omega = \{1, \ldots, M\} \times \{1, \ldots, N\}$. Then a binary image can be seen both as a mapping $\Omega \to \{0, 1\}$ and as a subset of Ω.

Let E_{cd} and E_{gt} represent a candidate and a ground-truth boundary image, respectively. Classification-based approaches to boundary quality evaluation are focused on generating a confusion matrix when comparing E_{cd} (usually extracted by automated method) with E_{gt} (probably defined by a human). In this context a false positive (FP) is a boundary pixel in E_{cd} with no correspondence in E_{gt}, while a false negative (FN) is a pixel in E_{gt} that is not represented in E_{cd}. With these statistical values most of the authors choose the F_α-measure to evaluate the closeness of E_{cd} to E_{gt}. The F_α-measure is given by:

$$F_\alpha = \frac{\text{PREC} \cdot \text{REC}}{\alpha\,\text{PREC} + (1-\alpha)\,\text{REC}}$$

with $\text{PREC} = \frac{\text{TP}}{\text{TP}+\text{FP}}$ and $\text{REC} = \frac{\text{TP}}{\text{TP}+\text{FN}}$, where α is a parameter modulating the relevance of PREC and REC

In order to avoid a too strict count of correct correspondence the techniques used to build confusion matrices permit to take into account a certain spatial tolerance in the boundary pixel matching.

In order to extract those measures, different approaches exist in the literature. In this work we review them and build a taxonomy based on their initial inspiration and nature. We classify them in four main strategies and experiment with one technique of each type:

- Distance-based Matching (DbM) using the Euclidean metric.

  This approach consists in validating those boundary pixels in a candidate image (E_{cd}) which are close enough (under a threshold) to the ground-truth image (E_{gt}).

- Area-based Matching (AbM) using a circular structuring element.

  This matching technique base its process in considering an area of influence around the boundary pixels, quantifying the pixels as correct boundaries when the areas of both E_{cd} and E_{gt} overlap. The area of influence is defined as a radius around boundary pixels and acts as the threshold.

- Correspondence-based Matching (CbM) using the CSA algorithm [116].

  This strategy acts as a one-to-one matching between the boundary pixels in E_{cd} to those in E_{gt} considering a given tolerance to count valid pixels.

- Pixelwise validation (Pv) using the constraints by Estrada and Jepson [72].

  This approach use the boundary pixels surroundings of both images, E_{cd} and E_{gt}, to decide the validity of the boundaries in E_{cd}.

In our study we perform three experiments using images provided by the BSDS500 set [64]:

- First experiment: We measure the Pearson correlation between the $\mathrm{F}_{0.5}$ values obtained with each one of the considered matching techniques.

- Second experiment: We measure the Equal-Sorting Ratio (ESR) between two measures, q_1, q_2, over a set of triplets of unrepeated ground truth images $(E_{\text{gt}}, E_{\text{cd1}}, E_{\text{cd2}})$. The ESR is the ratio of triplets which satisfies $q_1(E_{\text{gt}}, E_{\text{cd1}}) \geq q_1(E_{\text{gt}}, E_{\text{cd2}}))$ if and only if $q_2(E_{\text{gt}}, E_{\text{cd1}}) \geq q_2(E_{\text{gt}}, E_{\text{cd2}}))$.

- Third experiment: We repeat the first and second experiment but comparing an automatic boundary extraction (Canny method) with human-made solutions (ground-truth).

For each one of the experiments we test the considered approaches with different tolerance distances $t = \{2.5, 5, 10\}$ and analyse how this parameter affects the matching process.

After the completion of the work we can conclude that there is not a substantial difference between the different strategies and all of them perform with similar results even though their nature are substantially different. The two main events that can affect the quality of the evaluation are, on one hand, that the maximum tolerance distance to consider as correct a boundary is quite high, and on the other hand, that the boundary image to evaluate contain a higher number of spurious elements near the real boundaries.

This section of the thesis is associated with the following publication:

- C Lopez-Molina, C. Marco-Detchart, B De Baets, Member, H Bustince, and S. Member, "A survey on matching algorithms for boundary image comparison and evaluation", *IEEE Transactions on Image Processing*,

2.6 Optical images-based edge detection in Synthetic Aperture Radar images

In this work we tackle the problem edge detection for Polarimetric Synthetic Aperture Radar (PolSAR) images. We modify the gravitational edge detection proposed by Lopez-Molina *et al.* [112] using a variation in the neighbourhood selection. Concretely we use the approach proposed by Fu *et al.* [117]. Because of the lack of specifically designed techniques for SAR imagery [117] on of the possibilities to work with this type of data is to adapt techniques from classical imagery. The main problem with SAR images lies in speckle, which can be approached by adapting the properties of optical images technique to that of the SAR images or pre-process SAR images by filtering them and apply the original techniques.

In classical imagery when going through edge detection a unique band is considered, that is, we are dealing with grey-scale images. SAR imagery is composed of more than one band, recovering information from different polarizations. This variety of information permits to compensate the possible presence of speckle noise. Then, we have to tackle the fact that we need to aggregate information, but this task can be done at different moments in the process.

We consider two processing alternatives:

- DAB: edge Detection on non-binary images, Aggregation of the resulting non-binary images, Binarization.
- ADB: Aggregation of non-binary images, edge Detection on the resulting non-binary image, Binarization.

Along with those approaches we analyse in the pre-processing step the use of different filters. Among others, we experiment with the well-known Enhanced Lee filter [118] and the filter proposed by Torres *et al.* [119].

Within this work we can conclude that the use of both filtering and an adapted neighbourhood approach along with the aggregation of the different SAR image bands improve the reduction of speckle noise and hence permit the extraction of better defined edges.

This section of the thesis is associated with the following publication:

- G. P. Silva, A. C. Frery, S. Sandri, H. Bustince, E. Barrenechea, and C. Marco-Detchart, “Knowledge-Based Systems Optical images-based edge detection in Synthetic Aperture Radar images”, *Knowledge-Based Systems*, vol. 87, pp. 38–46, 2015

2.7 Conclusions

In this Ph.D. thesis we have analysed uncertainty in the context of comparison and managed to propose new ways of working with it and take it into account, analysing the effects of the new concepts in image processing. To do so we have studied different mathematical aspects of the comparison theory and presented some expressions to deal with different types of data and applications.

On one hand, we have dealt with uncertainty in the context of radial and interval-valued data and measured similarity over this specific type of data, studying the specificities of each of them. In this particular case we have studied different aspects, among others, the needed properties to take into account when dealing with radial data and the use of a total order for intervals as well as the preservation of the with of the interval considered.

On the other hand, we have studied the particular case of edge detection and performed and adaptation of the gravitational technique for edge detection to work with a specific case of Polarimetric Synthetic Aperture Radar (PolSAR) images. Also in this context we have studied the use of OD monotone functions to deal with feature extraction in images. With these experiment we show that our methods surpasses the Canny method and combining it with other edge detection approaches to build a consensus detector we obtain the best performance.

Finally, as a closing work, and to study the role of comparison in the evaluation of boundary detection we have studied the different matching methods of the literature used to compare a solution obtained for boundary detection and a real solution made by an expert. The different feature matching techniques analysed showed that they are equivalent, while the tolerance distance is not too high or there is not to much noise (spurious elements) near real boundaries.

2.8 Future research lines

From the experience gathered in the development of this dissertation, it seems clear that there is no universal, generic framework for comparison tasks. In fact, such a framework might not be reachable, considering the differences introduced in the comparison task by factors such as cardinality of data, uncertainty, or even application-related restrictions. Anyhow, we consider it would be a fair goal and, regardless of the difficulties I preview, it should be tackled.

In order to ease the task of producing this framework, an evident option is to merge the developments which we already published in a generalization effort. As an example, an intermediate framework could be considered to embrace both interval-data and radial data. Such an effort would lead to comparison tools with great application in geometrical problems, as well as a great asset for computer vision problems. As an idea, the radial data framework could be adapted to consider an interval-valued representation. This would open a great horizon of possibilities in terms of concepts and discussions on how to interpret, *e.g.* interval-valued angles or orientations, also permitting the incorporation of OD-monotone functions.

From a practical point of view, there is a list of applications which could benefit from the mathematical developments presented in this dissertation. For example, the use of radial comparison measures could be ported to multichannel images, in which gradients are multichannel. As well as, it could be also adapted to multi-scale or self-adapting approaches. In this regard, comparison measures shall be used not only for pattern matching, but also for other tasks, *e.g.* texture measurements or denoising. Such tasks should not only restrict to application-neutral computer vision, and might be also oriented towards problem-specific operators.

2.9 Conclusiones (versión en español)

En esta tesis hemos analizado la incertidumbre en el contexto de la teoría de la comparación y propuesto nuevas formas de trabajar con ella y considerarla, analizando los efectos de los nuevos conceptos desarrollados en tareas de procesamiento de imagen. Para ello hemos estudiado diferentes aspectos matemáticos de la teoría de la comparación y presentado una serie de expresiones para poder trabajar con diferentes tipos de datos y aplicaciones.

Por un lado, hemos trabajado con la incertidumbre en el contexto de datos radiales e intervalo-valorados y medido la similitud en estos tipos de datos, estudiando las especificidades de cada uno de ellos. Para ello, hemos estudiado diferentes aspectos, entre los que encontramos las propiedades específicas a tener en cuenta cuando trabajamos con datos radiales y el uso de ordenes totales para intervalos así como la conservación de la amplitud del intervalo considerado.

Por otro lado, hemos estudiado una adaptación del algoritmo gravitacional para detección de bordes para poder trabajar con el caso específico de imágenes de Radar de Apertura Polarimétrica Sintética (PolSAR en inglés), y hemos estudiado también el uso de funciones monótonas direccionalmente ordenadas para la extracción de características en imágenes. Los experimentos realizados en este contexto muestran que los métodos propuestos superan la método de Canny y a su vez, combinando nuestra propuesta con otros extractores de bordes para construir un detector basado en el consenso obtenemos los mejores resultados.

Finalmente, como trabajo de clausura, hemos estudiado los diferentes métodos de evaluación de calidad de la literatura utilizados para la comparación de una solución obtenida en extracción de bordes y su solución real dada por un experto. Las diferentes técnicas de comparación de características analizadas resultan ser equivalentes, mientras el valor de tolerancia no sea demasiado alto o no haya demasiada presencia de ruido (elementos extraños) cerca de los bordes reales.

3 Publications: published, accepted and submitted

> *The good thing about science is that it's true whether or not you believe in it.*
>
> — **Neil deGrasse Tyson**
> (Astrophysicist)

3.1 A framework for radial data comparison and its application to fingerprint analysis

- C. Marco-Detchart, J. Cerron, L. De Miguel, C. Lopez-Molina, H. Bustince, and M. Galar, "A framework for radial data comparison and its application to fingerprint analysis", *Applied Soft Computing Journal*, vol. 46, pp. 246–259, 2016
  - Journal: Applied Soft Computing
  - Status: Published
  - Impact Factor: (JCR 2016) 3,541
  - Knowledge area:
    - Computer Science, Interdisciplinary Applications: Ranking $14/105$ (Q1)
    - Computer Science, Artificial Intelligence: Ranking $21/132$ (Q1)

Applied Soft Computing 46 (2016) 246–259

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A framework for radial data comparison and its application to fingerprint analysis

C. Marco-Detchart[a,*], J. Cerron[a], L. De Miguel[a,b], C. Lopez-Molina[a,c], H. Bustince[a,b], M. Galar[a,b]

[a] Dpto. de Automatica y Computacion, Universidad Publica de Navarra, Campus Arrosadia s/n, 31006 Pamplona, Spain
[b] Institute of Smart Cities, Universidad Publica de Navarra, Campus Arrosadia s/n, 31006 Pamplona, Spain
[c] KERMIT, Dept. of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, 9000 Gent, Belgium

ARTICLE INFO

Article history:
Received 5 October 2015
Received in revised form 11 April 2016
Accepted 4 May 2016
Available online 6 May 2016

Keywords:
Radial data
Restricted Equivalence Function
Similarity Measure
Fingerprint singular point detection

ABSTRACT

This work tackles the comparison of radial data, and proposes comparison measures that are further applied to fingerprint analysis. First, we study the similarity of scalar and non-scalar radial data, elaborated on previous works in fuzzy set theory. This study leads to the concepts of restricted radial equivalence function and Radial Similarity Measure, which model the perceived similarity between scalar and vectorial pieces of radial data, respectively. Second, the utility of these functions is tested in the context of fingerprint analysis, and more specifically, in the singular point detection. With this aim, a novel Template-based Singular Point Detection method is proposed, which takes advantage of these functions. Finally, their suitability is tested in different fingerprint databases. Different Similarity Measures are considered to show the flexibility offered by these measures and the behavior of the new method is compared with well-known singular point detection methods.

1. Introduction

The ability to quantify the similarity between two objects in a given universe is a pillar in applied fields of research. Historically, this quantification has been based on metrics, which are able to capture, in a sensible (and coherent) manner, the proximity of any two objects in a measurable universe. Metrics hold very interesting properties, specifically triangular inequality, which preserves the notion that the shortest path between two objects is the straight one. However, they also impose the need for the representation of the objects in a metric space, as well as notions (e.g., transitivity), which are not natural in certain scenarios [1].

When it comes to measuring dissimilarity between multivalued data, L_p metrics often come as a straightforward option; the most relevant case is $p = 2$, which recovers the Euclidean metric. The L_p metric has been long criticized, specially regarding its low accuracy in capturing perceptual dissimilarities. For example, Attneave stated that the assumption that the *psychological space is Euclidean in its character* is *exceedingly precarious* [2]. Obviously, there exist other metrics yielding more (perceptually) accurate measurements of dissimilarity, specially when they are designed for well-defined scenarios [3,4]. The debate about the restrictivity of the requisites imposed by metrics is still open [5]. Literature contains both practical [6], and theoretical criticisms. Authors as Tversky [7] or Santini and Jain [5] criticized the necessity of imposing metric conditions to Similarity Measures, as well as the representation of objects in metric spaces, given that they are often missing in human understanding. Tversky [7,8] also revisited the necessity of symmetry and the directional nature of comparisons in certain scenarios. Finally, the low representativity of the values given by metrics for large-range comparisons has also been under debate [9,10].

Different mathematical theories have tackled the modelling of similarity with tools other than metrics, leading to what Zadeh referred to as *a vast armamentarium* of techniques for comparison [11]. In fact, even axiomatic representations of non-metric comparison frameworks have appeared in the literature (e.g., [7] for set-based similarity, or [12,13] for T-indistinguishability). In the context of fuzzy set theory, a range of authors have elaborated on the semantic interpretation of similarities and dissimilarities [5], since Zadeh introduced similarity as an extension of equivalence [11]. This is natural, considering that the concepts of proximity and similarity (as well as ordering or clustering) are strongly related to human interpretation, and hence prone to be tackled in fuzzy

* Corresponding author.
E-mail addresses: cedric.marco@unavarra.es (C. Marco-Detchart), juan.cerron@unavarra.es (J. Cerron), laura.demiguel@unavarra.es (L. De Miguel), carlos.lopez@unavarra.es (C. Lopez-Molina), bustince@unavarra.es (H. Bustince), mikel.galar@unavarra.es (M. Galar).

http://dx.doi.org/10.1016/j.asoc.2016.05.003
1568-4946/

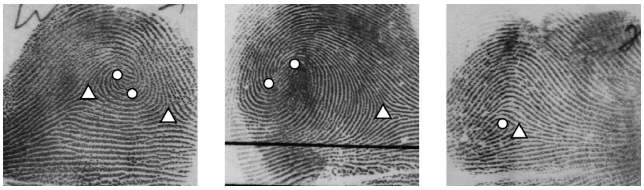


Fig. 1. Examples of singular points detected on fingerprints extracted from the NIST-4 dataset [19]. The *deltas*, represented as triangles, are triangular-like ridge confluences, while *cores*, represented as circles, take place at curly ridge structures.

terms. A large variety of proposals have appeared for modelling both similarity and dissimilarity; in this work we focus our interest on two of them: Restricted Equivalence Functions (REFs) for the comparison of membership degrees and Similarity Measures (SMs) for the comparison of fuzzy sets on discrete universes [14].

In this paper we propose a definition of the concept of REFs and SMs for radial data. This study is motivated by the increasing relevance of radial data in real applications, especially in those demanding the extraction of information by means of computer vision techniques. Very often, computer vision handles radial data in different flavours (e.g., angular, vector or tensorial data [15]) and consequently demands well-defined operators for different tasks, including data comparison. Typically, the study of radial data has been restricted to radial statistics, which mostly study the fitting and analysis of well-known distributions on radial set-ups. To the best of our knowledge, there are no studies on the quantification of the similarity of elements in a radial universe. This situation has led many researchers to use *ad-hoc* operators to deal with the special conditions of the data, instead of creating a framework in which different operators can be encompassed. For this reason in this work we develop a framework aiming at easing the comparison of radial data. More specifically, we define Restricted Radial Equivalence Functions (RREFs), as well as Restricted Similarity Measures (RSMs), which attempt to mimic the behaviour of REFs and SMs in radial universes.

As a case of study, we present an application of RREFs and RSMs to biometric identification, specifically to singular point detection in fingerprint recognition [16]. Fingerprints can be seen as a set of ridges (lines) that represent the relief of the skin in the fingertip surface. Hence, their analysis is often based on studying the line patterns in a local or semi-local basis. Within fingerprint analysis, a fundamental operation is the detection and localization of the so-called *Singular Points* (SPs), which are structural singularities in the ridges (see Fig. 1). SP detection is often related to specific occurrences in the orientation of the ridges of neighbouring regions, which are usually found using semi-local analysis [17] or complex convolution filters [18].

On this account, a simple yet effective framework for SP detection is presented in this paper by means of RREFs and RSMs, which shows the usefulness and flexibility of these new measures. Furthermore, other well-known SP detection algorithms have been used as a baseline for performance evaluation [20,21]. In this comparative analysis we have considered two different types of databases: NIST-4 database [19], the most commonly used fingerprint database and synthetic fingerprint databases generated by SFinGe.[1]

The remainder of the work is as follows. In Section 2 we review the concepts of REF and SM, as well as some standard notation on radial data. Section 3 is devoted to introduce the concepts of RREF and RSM. Both RREF and RSM are used in Section 4, in which we present our proposal for SP detection in fingerprints. Section 5 includes an experimental study in which we illustrate the performance of our SP detection method, compared to other well-known methods in the literature. Finally, Section 6 gathers some conclusions and a brief discussion on potential future evolutions of our method.

2. Preliminaries

Among the areas in which fuzzy set theory has played a relevant role, data similarity modelling is one of the most prominent. The reason is that the natural concepts of similarity, closeness or likeliness are inherently bounded to human interpretation. Hence, different proposals have appeared to effectively model the comparison of pieces of information. Among these, we find fuzzy metric spaces [6], with interesting advantages over classical metric spaces in terms of interpretability [22] or equivalence and Similarity Measures [14], which we take as inspiration to develop measures that can handle radial data. Next, we recall the concepts of REF and SM.

Definition 1. A continuous, strictly decreasing function $n:[0,1]\to[0,1]$ such that $n(0)=1$, $n(1)=0$ and $n(n(x))=x$ for all $x\in[0,1]$ (involutive property) is called strong negation.

Definition 2. [14] A mapping $r:[0,1]^2\to[0,1]$ is said to be a Restricted Equivalence Function (REF) associated with the strong negation n if it satisfies the following:

(R1) $r(x,y)=r(y,x)$ for all $x,y\in[0,1]$;
(R2) $r(x,y)=1$ if and only if $x=y$;
(R3) $r(x,y)=0$ if and only if $\{x,y\}=\{0,1\}$;
(R4) $r(x,y)=r(n(x),n(y))$ for all $x,y\in[0,1]$;
(R5) For all $x,y,z,t\in[0,1]$, such that $x\le y\le z\le t$ then $r(y,z)\ge r(x,t)$.

Note that (R5) means that, for all $x,y,z\in[0,1]$, if $x\le y\le z$ then $r(x,y)\ge r(x,z)$ and $r(y,z)\ge r(x,z)$.

REFs attempt to capture the perceived similarity between two values in $[0,1]$, which in fuzzy set theory usually represent membership degrees. It is usual to construct REFs from a pair of automorphisms of the unit interval, as proposed in [14], although alternative methods have also been studied [23].

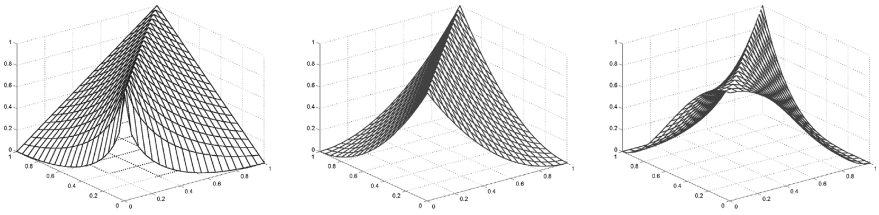
Definition 3. A continuous, strictly increasing function $\varphi:[a,b]\to[a,b]$ such that $\varphi(a)=a$ and $\varphi(b)=b$ is called automorphism of the interval $[a,b]\subset\mathbb{R}$.

Proposition 1. [14] *Let φ_1,φ_2 be two automorphisms of the interval $[0,1]$. Then*

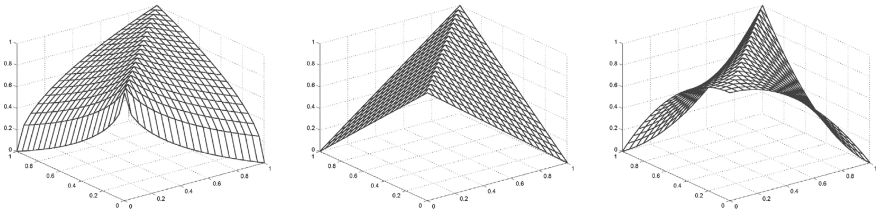
$$r(x,y)=\varphi_1^{-1}(1-|\varphi_2(x)-\varphi_2(y)|)$$

is a REF associated with the strong negation $n(x)=\varphi_2^{-1}(1-\varphi_2(x))$.

[1] Synthetic Fingerprint Generator: http://biolab.csr.unibo.it/sfinge.html.



(a) $\varphi_1(x) = \sqrt{x}, \varphi_2(x) = \sqrt{x}$ (b) $\varphi_1(x) = \sqrt{x}, \varphi_2(x) = x$ (c) $\varphi_1(x) = \sqrt{x}, \varphi_2(x) = x^2$



(d) $\varphi_1(x) = x, \varphi_2(x) = \sqrt{x}$ (e) $\varphi_1(x) = x, \varphi_2(x) = x$ (f) $\varphi_1(x) = x, \varphi_2(x) = x^2$

Fig. 2. Restricted Equivalence Functions (REFs) created from automorphisms in the unit interval as in Proposition 1.

Example. Let $\varphi_1(x) = x$ and $\varphi_2(x) = \sqrt{x}$, then

$$r(x, y) = 1 - |\sqrt{x} - \sqrt{y}| \tag{1}$$

is a REF associated with $n(x) = (1 - \sqrt{x})^2$.

Fig. 2 contains the visual representation of several REFs constructed following Proposition 1. Note that in this work we are referring to REFs in the original sense, although the concept has been exported to, e.g., interval data [24]. While REFs are useful to compare scalar data (membership degrees), Similarity Measures were developed to compare non-scalar data. Even though the measures were initially designed to compare fuzzy sets on discrete universes, they can be further applied to many other objects (e.g., vectors or matrices). Similarity Measures were originally proposed by Liu [25] and its definition is trivially applied to $[0, 1]^k$.

Definition 4. [25] A mapping $s : [0, 1]^k \times [0, 1]^k \to \mathbb{R}^+$ is called k-ary Similarity Measure (SM) associated with the strong negation n if it satisfies the following:

(S1) $s(\mathbf{x}, \mathbf{y}) = s(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in [0, 1]^k$;
(S2) $s(\mathbf{x}, n(\mathbf{x})) = 0$ if and only if $x_i \in \{0, 1\}$ for all $i \in \{1, \ldots, k\}$ and $n(\mathbf{x}) = (n(x_1), \ldots, n(x_k))$;
(S3) $s(\mathbf{z}, \mathbf{z}) = \max_{(\mathbf{x},\mathbf{y} \in [0,1]^k)} s(\mathbf{x}, \mathbf{y})$ for all $z \in [0, 1]^k$;
(S4) For all $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in [0, 1]^k$, if $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z} \leq \mathbf{t}$ then $s(\mathbf{y}, \mathbf{z}) \geq s(\mathbf{x}, \mathbf{t})$ where $\mathbf{x} \leq \mathbf{y}$ implies that $x_i \leq y_i$ for all $i \in \{1, \ldots, k\}$.

Similarly to what happens for REFs, the property (S4) is equivalent to: For all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in [0, 1]^k$, if $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z}$ then $s(\mathbf{x}, \mathbf{y}) \geq s(\mathbf{x}, \mathbf{z})$ and $s(\mathbf{y}, \mathbf{z}) \geq s(\mathbf{x}, \mathbf{z})$.

Similarity Measures (SMs) can be constructed in different ways, although the most popular method originates from the combination of REFs and aggregation functions [26].

Definition 5. A mapping $f : [0, 1]^k \to [0, 1]$ is called k-ary aggregation operator if it satisfies the following:

(AO1) If $x_i = 0$ for all $i \in \{1, \ldots, k\}$, then $f(\mathbf{x}) = 0$;
(AO2) If $x_i = 1$ for all $i \in \{1, \ldots, k\}$, then $f(\mathbf{x}) = 1$;
(AO3) f is increasing in all of its arguments.

Proposition 2. [14] *Let r be a REF and let f be a k-ary aggregation function such that $f(\mathbf{x}) = 0$ if and only if $x_i = 0$ for all $i \in \{1, \ldots, k\}$ and $f(\mathbf{x}) = 1$ if and only if $x_i = 1$ for all $i \in \{1, \ldots, k\}$. The function $s_{[f,r]} : [0, 1]^k \times [0, 1]^k \mapsto [0, 1]$, given by*

$$s_{[f,r]}(\mathbf{x}, \mathbf{y}) = f(r(x_1, y_1), \ldots, r(x_k, y_k)) \tag{2}$$

is a k-ary Similarity Measure which satisfies the following:

- $s_{[f,r]}(\mathbf{x}, \mathbf{y}) = s_{[f,r]}(\mathbf{y}, \mathbf{x})$ *for all* $\mathbf{x}, \mathbf{y} \in [0, 1]^k$;
- $s_{[f,r]}(\mathbf{x}, n(\mathbf{x})) = 0$ *if and only if* $x_i \in \{0, 1\}$ *for all* $i \in \{1, \ldots, k\}$ *and* $n(\mathbf{x}) = (n(x_1), \ldots, n(x_k))$;
- $s_{[f,r]}(\mathbf{x}, \mathbf{y}) = 1$ *if and only if* $x_i = y_i$ *for all* $i \in \{1, \ldots, k\}$;
- *For all* $\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{t} \in [0, 1]^k$, *if* $\mathbf{x} \leq \mathbf{y} \leq \mathbf{z} \leq \mathbf{t}$ *then* $s_{[f,r]}(\mathbf{y}, \mathbf{z}) \geq s_{[f,r]}(\mathbf{x}, \mathbf{t})$ *where* $\mathbf{x} \leq \mathbf{y}$ *implies that* $x_i \leq y_i$ *for all* $i \in \{1, \ldots, k\}$;
- $s_{[f,r]}(\mathbf{x}, \mathbf{y}) = s_{[f,r]}(n(\mathbf{x}), n(\mathbf{y}))$ *for all* $\mathbf{x}, \mathbf{y} \in [0, 1]^k$.

Next, we include some conventions on the use of radial data that will hold in the remainder of this work. In order to face those cases in which two angles represent the same direction (e.g., $\pi/2$ and $(5\pi)/2$), an equivalence relation R is defined.

Let $a, b \in \mathbb{R}$. We say that aRb if and only if $a = b + 2k\pi$, where $k \in \mathbb{Z}$. In this way, the equivalence class $[a] = \{b \mid bRa\}$ is the set containing all the data associated with the same direction. In the same way, any semiopen interval whose width is 2π (i.e., an interval of the form $[\omega, \omega + 2\pi[$) is the quotient set (a set which contains one element and only one of each equivalence class). In particular, the intervals $[0, 2\pi[$ and $[-\pi, \pi[$ are the most frequently used quotient sets in radial data.

In this work we consider the quotient set $\Omega = [0, 2\pi[$ and we define on Ω the classical operations sum ($\oplus$) and difference ($\ominus$), given by $a \oplus b = [a + b]$ and $a \ominus b = [a - b]$ where $[t]$ denotes the only

element $z \in \Omega$ such that zRt. Under these conditions, we refer to *mirroring* as the mapping: $m : \Omega \longrightarrow \Omega$ such that $m(a) = a \oplus \pi$.

In order to properly explain our incoming proposal on the quotient set Ω, we present the general definition of metric.

Definition 6. A function $d : U \to \mathbb{R}^+$ is called metric on U if it satisfies the following:

(D1) $d(a, b) = 0$ if and only if $a = b$;
(D2) $d(a, b) = d(b, a)$ for all $a, b \in U$;
(D3) $d(a, c) \leq d(a, b) + d(b, c)$ for all $a, b, c \in U$.

3. Comparison of radial data

As reported by Fisher [27], radial data has been a subject of analysis since mid-18th century. However, most of the literature on radial data is based on adapting the usage of distributions to the circular set-up, probably because data analysis for natural sciences was the field in which radial data was first studied [28,29].

One of the open problems in radial data is data comparison. In fact, to the best of our knowledge, no explicit mention to the quantification of similarity between two angles has been performed in the literature. There have been concepts such as the *sample median direction* [27] or the *sample modal direction* [30], and metrics such as the angular metric on $[0, 2\pi[$ ($d^*(a, b) = \min(|b - a|, 2\pi - |b - a|)$), which represents the amplitude of the shortest arc encompassing two angles. However, no development has been made on interpretable measures able to adapt to human perception or evaluation. This section is devoted to develop functions that are able to measure the perceived similarity between scalar and vector angular data. Section 3.1 covers the comparison of scalar radial data, whereas Section 3.2 covers the comparison of vector radial data.[2]

3.1. Restricted Radial Equivalence Functions

The comparison of linear data has produced a vast amount of literature, despite coming from a relatively simple concept. The concept of similarity becomes much more intricate when applied to radial data. In this section we define operators that model the comparison of elements in a radial context Ω, all inspired by the operators in Section 2.

Definition 7. A mapping $r_\theta : \Omega^2 \to [0, 1]$ is called a Restricted Radial Equivalence Function (RREF) associated with the metric d if it satisfies the following:

(RR1) $r_\theta(a, b) = r_\theta(b, a)$ for all $a, b \in \Omega$;
(RR2) $r_\theta(a, b) = 1$ if and only if $d(a, b) = 0$;
(RR3) $r_\theta(a, b) = 0$ if and only if $d(a, b)$ is maximum;
(RR4) $r_\theta(a, b) = r_\theta(m(a), m(b))$ for all $a, b \in \Omega$;
(RR5) For all $a, b, c, d \in \Omega$, if $d(b, c) \leq d(a, d)$, then $r_\theta(b, c) \geq r_\theta(a, d)$.

Definition 7 is not a direct extension of Definition 2 to radial data. The differences arise from the use of distances in (RR5) instead of orders (as in (R5)). Nevertheless, this change is due to the difficulties in the interpretation of orders in radial universes. Despite this modification, we believe that the spirit and semantics of RREFs are those of REFs.

[2] There exist in the literature certain controversy w.r.t. the most adequate name of radial data, including angular data or radial data. In this manuscript we adhere to *radial*.

In this work, we only consider RREFs associated with the angular metric $d^*(a, b) = \min(|b - a|, 2\pi - |b - a|)$ but many other metrics are also eligible, e.g.,:

$$d(a, b) = \begin{cases} 0 & \text{if } a = b, \\ \pi & \text{if } d^*(a, b) = \pi, \\ \pi/2 & \text{otherwise.} \end{cases}$$

Proposition 3. *Let r_θ be a RREF associated with the metric d^*. For all $a_1, b_1, a_2, b_2 \in \Omega$, if $d^*(a_1, b_1) = d^*(a_2, b_2)$ then $r_\theta(a_1, b_1) = r_\theta(a_2, b_2)$.*

Proof. Let $a_1, b_1, a_2, b_2 \in \Omega$ such that $d^*(a_1, b_1) = d^*(a_2, b_2)$. According to (RR5), $d^*(a_1, b_1) \leq d^*(a_2, b_2)$ implies $r_\theta(a_1, b_1) \geq r_\theta(a_2, b_2)$. Analogously, $d^*(a_2, b_2) \leq d^*(a_1, b_1)$ implies $r_\theta(a_2, b_2) \geq r_\theta(a_1, b_1)$ so the equality holds.□

Corollary 1. *Let $h : \Omega^2 \to [0, 1]$. If h satisfies (RR5) with respect to the metric d^* then it also satisfies (RR4).*

Proof. Trivial by Proposition 3 since $d^*(a, b) = d^*(m(a), m(b))$.□

Following the construction of REFs, we also define a new possible construction method for RREFs, which is also based on automorphisms.

Proposition 4. *Let φ and ψ be automorphisms of the intervals $[0, 1]$ and $[0, \pi]$, respectively. The mapping $t : \Omega^2 \to [0, 1]$ given by*

$$t(a, b) = \varphi^{-1}\left(1 - \left(\frac{1}{\pi}\psi(d^*(a, b))\right)\right) \tag{3}$$

is a RREF.

Proof. Direct by the properties of the metric d^*.□

Some examples of RREFs constructed as in Proposition 4 are included in Fig. 3.

3.2. Radial Similarity Measures

Definition 8. A mapping $s_\theta : \Omega^k \times \Omega^k \to \mathbb{R}^+$ is said to be a k-ary Radial Similarity Measure (RSM) associated with the metric d^* if it satisfies the following:

(SR1) $s_\theta(\mathbf{a}, \mathbf{b}) = s_\theta(\mathbf{b}, \mathbf{a})$ for all $\mathbf{a}, \mathbf{b} \in \Omega^k$;
(SR2) $s_\theta(\mathbf{a}, \mathbf{b}) = 0$ if and only if $d^*(a_i, b_i) = \pi$ for all $i \in \{1, \ldots, k\}$;
(SR3) $s_\theta(\mathbf{c}, \mathbf{c}) = Max_{\mathbf{a},\mathbf{b} \in \Omega^k} s_\theta(\mathbf{a}, \mathbf{b})$ for all $c \in \Omega^k$;
(SR4) For all $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \Omega^k$, if $d^*(\mathbf{a}, \mathbf{d}) \geq d^*(\mathbf{b}, \mathbf{c})$ then $s_\theta(\mathbf{a}, \mathbf{d}) \leq s_\theta(\mathbf{b}, \mathbf{c})$, where $d^*(\mathbf{a}, \mathbf{d}) \geq d^*(\mathbf{b}, \mathbf{c})$ implies that $d^*(a_i, d_i) \geq d^*(b_i, c_i)$ for all $i \in \{1, \ldots, k\}$.

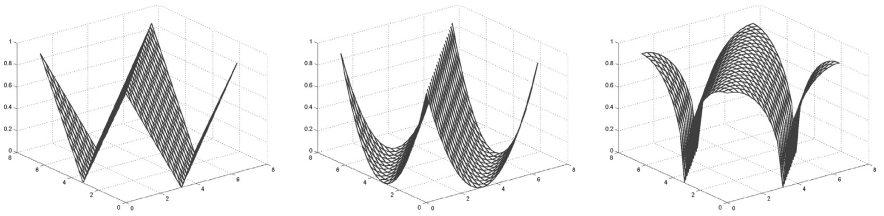
RSMs can be constructed from RREFs, aggregating their results over each element, as it is done for SMs.

Proposition 5. *Let r_θ be a RREF and let f be a k-ary aggregation function such that $f(\mathbf{x}) = 0$ if and only if $x_i = 0$ for all $i \in \{1, \ldots, k\}$ and $f(\mathbf{x}) = 1$ if and only if $x_i = 1$ for all $i \in \{1, \ldots, k\}$. The function $s_{\theta_{[f,r_\theta]}} : \Omega^k \times \Omega^k$, given by*

$$s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{b}) = f(r_\theta(a_1, b_1), \ldots, r_\theta(a_k, b_k)) \tag{4}$$

is a k-ary Radial Similarity Measure that satisfies

- $s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{b}) = s_{\theta_{[f,r_\theta]}}(\mathbf{b}, \mathbf{a})$ *for all* $\mathbf{a}, \mathbf{b} \in \Omega^k$;
- $s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{b}) = 0$ *if and only if* $d^*(a_i, b_i) = \pi$ *for all* $i \in \{1, \ldots, k\}$;
- $s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{b}) = 1$ *if and only if* $a_i = b_i$ *for all* $i \in \{1, \ldots, k\}$;



(a) $\varphi(x) = x, \psi(x) = x$ (b) $\varphi(x) = x, \psi(x) = \sqrt{\pi x}$ (c) $\varphi(x) = x^2, \psi(x) = \frac{x^2}{\pi}$

Fig. 3. Restricted radial equivalence functions created as in Proposition 4 from automorphisms in the unit interval (φ) and in $[0, \pi]$ (ψ).

- *For all* $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d} \in \Omega^k$, *if* $d^*(\mathbf{a}, \mathbf{d}) \geq d^*(\mathbf{b}, \mathbf{c})$ *then* $s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{d}) \leq s_{\theta_{[f,r_\theta]}}(\mathbf{b}, \mathbf{c})$, *where* $d^*(\mathbf{a}, \mathbf{d}) \geq d^*(\mathbf{b}, \mathbf{c})$ *implies that* $d^*(a_i, d_i) \geq d^*(b_i, c_i)$ *for all* $i \in \{1, \ldots, k\}$.
- $s_{\theta_{[f,r_\theta]}}(\mathbf{a}, \mathbf{b}) = s_{\theta_{[f,r_\theta]}}(m(\mathbf{a}), m(\mathbf{b}))$ *for all* $\mathbf{a}, \mathbf{b} \in \Omega^k$ *where* $m(\mathbf{a}) = (m(a_1), \ldots, m(a_k))$.

4. Template-based Singular Point Detection

In this section we present a SP detection method based on RSMs and RREFs. Section 4.1 introduces the problem of SP detection, while Section 4.2 presents the template based SP detection method and Section 4.3 outlines the resulting algorithm.

4.1. Fingerprint classification and singular point detection

Fingerprint-based identification is the most popular type of biometrical identity authentication systems. These systems carry out the authentication analysing the ridge patterns in the surface of the fingertips. There are two main tasks that can be performed in the context of fingerprint analysis, namely *identification* and *verification*. The former term refers to the localization of an individual in a database by the usage of an input fingerprint. The latter term refers to confirming whether an input fingerprint corresponds to a certain individual in the database. In both cases, having an accurate way to perform one-to-one fingerprint comparisons is critical [31].

The most common approach to decide whether two fingerprints are produced from the same fingertip is to compare local ridge features, usually, the so-called *minutiae*, which are local discontinuities or anomalies in the ridge pattern [32,33]. Examples of minutiae are ridge breaks and bifurcations. The whole process of deciding whether two fingerprints belong to the same individual is known as *matching* [34,35,31,36].

Fingerprint matching is not trivial, and very often demands a certain computational effort. This is not a problem in fingerprint verification (since the input fingerprint is only compared with those corresponding to the claimed identity), but it becomes critical in fingerprint identification. For this reason, several strategies have been developed to minimize the number of comparisons to be performed. Among them, the most used one is *classification* [16,37,38], which consists of classifying each fingerprint according to the general structure of its ridges. In this way, when an input fingerprint has to be matched, it only need to be compared with those belonging to the same class. Although different classification schemes for fingerprints have been proposed, most of the authors in fingerprint analysis use the five major classes in the Henry system [39], namely *arch*, *tented arch*, *left loop*, *right loop*, *whorl*. Examples of these five classes can be found in Fig. 4. As a result of this division, the number of comparisons in an identification process can be drastically reduced. Nonetheless, this task also holds great responsibility. If a fingerprint is misclassified, the system will not be able to perform a correct identification or it may lead to an increase in the computational effort due to the greater number of comparisons that must be carried out.

Fingerprint classification is often defined as the problem of learning a classifier able to determine the class to which a (previously unseen) fingerprint belongs to. In order to do so, the classifier is usually learned from a set of labelled fingerprints. Fingerprints are classified using global features from the ridge flow, instead of local ones as it is done in fingerprint matching. Hence, fingerprint classification consists of two well-differentiated steps[3]: (a) feature extraction, where fingerprints are processed to obtain their feature vector, and (b) classification, where a classifier associates such vector to one of the classes. In this work we focus on the first one; more specifically, on the detection of the so-called singular points (SPs), which are the most commonly used feature for classification.

SPs are locations of the fingerprint in which abnormal ridge patterns occur. In a fingerprint, two types of SPs can be found: *cores* (where ridges tend to converge) and *deltas* (where the ridge flow diverges). The importance of these features for classification is clear, given that the classes in the Henry system can indeed be described in terms of SPs:

- *Arch*: There are no SPs, since the ridges flow horizontally producing a small bump in the center of the fingerprint.
- *Tented Arch*: There is one core and one delta, and the delta is under the core. The ridge flow is similar to that of the Arch type, but at least one ridge shows high curvature.
- *Left Loop*: One core and one delta, and the delta is underneath and on the right of the core. One or more ridges flow from the left side, curve back, and disappear again to the left margin of the fingertip.
- *Right Loop*: There is one core and one delta, and the delta is underneath and to the left of the core. One or more ridges flow from the right side, curve back, and disappear again on the right margin of the fingertip.
- *Whorl*: There are two cores and two deltas, and at least one of its ridges makes a full turn around the center of the fingerprint.

A proper description of SPs might be sufficient for the classification of a fingerprint, by using a fixed rule-based strategy [17,41]. Still, learning-based approaches to classification have also been proposed [42,43,21]. In general, it is accepted that accurate SP

[3] We refer to [16,37] for a detailed review on the topic.

Fig. 4. Examples of fingerprint for each of the five classes in the Henry classification system. These fingerprints have been created using the SFinGe tool [40,39].

detection is required in order to reach the highest possible accuracy in the posterior classification process. Note that, apart from their relevance in fingerprint classification, SPs are also used for some other processes on fingerprints, e.g., fingerprint alignment with respect to a reference point [44] (usually a core point).

Despite the importance of SPs, their extraction is still an open problem for which proposals are constantly being presented [16,45,46]. Most of such proposals are based on the semi-local analysis of the so-called Orientation Map (OM), which is a block-based description of the ridge flow in a fingerprint [47] (see Fig. 5(b) for an example). The most relevant proposal for SP detection is the Poincarè method [20]. In this method, each of the blocks in the Orientation Map is assigned a Poincarè index, which is computed as the total rotation of the orientations around it. This index determines the presence of a SP, as well as its type (i.e., either core or delta). A popular approach, also based on OMs, is the one proposed by Nilsson and Bigun [18]. Additionally to the usage of complex filters, an interesting novelty in [18] resides in the use of *Squared OMs* (SqOM) [48], which are obtained by multiplying by 2 the orientation at each block of an OM. This simple representation of the OM has the key advantage of producing rotation-invariant patterns at SP locations. Fig. 5 illustrates how SPs look in both conventional and Squared Orientation Maps.

In our SP detection framework we aim at exploiting the high visibility of SPs in squared OMs. More specifically, we propose to use a template-based approach to SP detection which consists of comparing templates of SPs with the actual occurrences of SqOMs. In this context, RREFs and RSMs become crucial, allowing the comparison of the directions in the SqOMs with those in the templates. Forthcoming sections provide details on our method.

4.2. Template-based Singular Point Detection

Template matching procedures are recurrent solutions in digital image processing. The reason is that the only a priori information needed for such procedures is an expression of the goal (which stems from the definition of the problem) and a comparison measure able to quantify the similarity between the input data and the template. Examples of template based methods for image processing are some low-level feature detectors [49–51], or composite object detectors (e.g., the eye detector in [52]). Moreover, despite template matching is conceptually simple, it has also evolved into rather complex theories, among which we can list, for example, mathematical morphology [53]. In this section we present a framework for SP detection based on templates, which is referred to as *Template-based SP Detection* method (TSPD method). To the best of our knowledge, no author has proposed to use templates to represent SPs, probably due to the lack of reliable comparison methods that can handle the matching score. The most similar approach is the usage of complex filters [21], which are convolved with the complex representation of the OM. Notice that we also include this method as baseline performer in the experiments. From our point of view, template-matching is a natural way to search for SPs, as long as the definition of SP is vague and based on human perception.

Any template matching-based framework is composed of three nuclear components: (a) an appropriate representation of the input data, (b) templates describing the patterns to be searched in terms of the input data, and (c) a reliable tool to quantify the similarity between both representations. Since our framework is deeply based on mimicking human perception, our aim is to maintain all three components as faithful as possible to the human

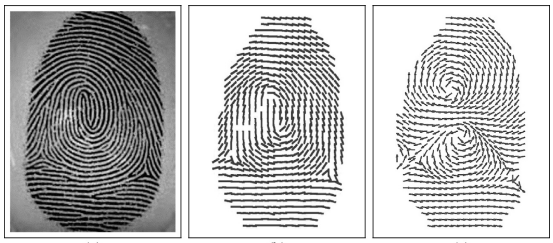
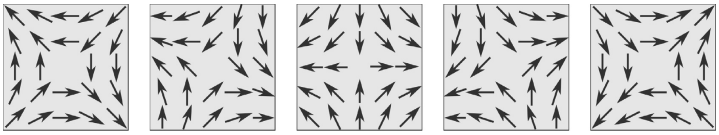


Fig. 5. Whorl image generated with SFinGe (a), together with its Orientation Map (OM) (b) and Squared Orientation Map (SqOM) (c). We have considered a size of block of 12×12 pixels to optimize the visibility of each block.



(a) $T_{\alpha=0}$ (b) $T_{\alpha=\frac{\pi}{8}}$ (c) $T_{\alpha=\frac{\pi}{4}}$ (d) $T_{\alpha=3\frac{\pi}{8}}$ (e) $T_{\alpha=\frac{\pi}{2}}$

Fig. 6. Examples of delta SP templates generated as in Eq. (6) with different values of α.

comprehension of the problem. Consequently, we elaborate on the ridge-like representation of fingerprints (a) and templates (b), while employing RSMs for (c).

(a) *Fingerprint representation using Orientation Maps.* The most obvious representation for SP detection is the fingerprint image itself, since that is all the information humans need to locate SPs. However, this representation is inconvenient, and would dramatically hinder the representation of the template. The reason is that, despite the fact that humans take as input the original image, the location of SPs is based on the analysis of the ridges. That is, the humans automatically convert the tone-based representation of the fingerprint into a ridge-based one.

The representation of the ridges in an image has been often studied, and most of the authors agree on using OMs [47]. These maps divide the fingerprint images into disjoint blocks, assigning to each of them a unique orientation given by the majority ridge orientation of its pixels. The best-known approach to OM extraction is the gradient method [45]. In this method, the orientation of the ridges is computed pixel-wise as the perpendicular to the gradient direction. In Fig. 5 we display the OM of a whorl type fingerprint. We observe how cores take oriented cup-like patterns, which are dependent on the specific orientation of each SP, and deltas produce triangular orientation patterns.

In this work we use the SqOMs, which are better fitted than OMs to our goals. This representation, as shown in Fig. 5, produces interesting changes in the representation of SPs. More specifically, it creates a rotation-invariant representation of cores, which are represented as either clockwise or anticlockwise streams. Regarding deltas, the situation is not as positive, since their appearance does not become rotation-invariant. In any case, using SqOMs simplifies the design of the templates, and is kept as standard representation in our framework.

(b) *Templates for SP representation.* The templates in our framework must be a minimal set such that it completely captures the way in which SPs appear (or are perceived) in a SqOM. Cores manifest themselves as either clockwise or anticlockwise sequences of orientations, so there is only need for two templates. Moreover, these templates can be functionally represented in a very simple manner.

Let the origin $(0, 0)$ represent the center of a template T of size $(2n+1)\times(2n+1)$. The orientation at a position $(x, y) \in [-n, n]^2$ of a core template is given by

$$T(x,y) = \begin{cases} \text{atan2}(y,x) & \text{if it is a clockwise core, and} \\ \text{atan2}(-y,-x) & \text{if it is an anticlockwise core,} \end{cases} \quad (5)$$

where $\text{atan2}(y, x)$ is the well-known sign-sensitive version of the arctangent of y/x, i.e., the angle of the vector (x, y) with respect to the positive x-axis. Note that the center of the template has no value, and hence contains no information for the matching process.

With respect to deltas, the problem becomes trickier. In a general manner, a delta is represented as a triangular pattern in the OM, and becomes a symmetric pattern with vectors opposing each other in two orthonormal directions in the SqOM (see Fig. 5). None of those representations is rotation invariant, and consequently an orientation-dependent template must be created to represent delta SPs. The orientation at a position $(x, y) \in [-n, n]^2$ of a delta SP template with orientation $\alpha \in [0, \pi]$ is given by

$$T_\alpha(x,y) = \text{atan2}(-(\cos(\alpha)y - \sin(\alpha)x), \sin(\alpha)y + \cos(\alpha)x). \quad (6)$$

Fig. 6 displays the delta SP template for different values of α. In such images we can observe how the pattern is composed of two orthonormal axis, one acting as an *attractor* to the origin, the other one being a *repeller* to it.

According to the previous template definitions, there are two decisions to be made on the set of templates. The first decision affects the number of delta SP templates to be used, i.e., the number of different values of α to produce a pattern. One can foresee that a greater number of templates will lead to more accurate detections, although it might also lead to a better fitting of abnormal ridge occurrences that do not correspond to SPs as well as a higher computational effort. The second decision relates to the size of the templates. Indeed the size of the templates must be dependent upon the size of the blocks in the SqOM, as well as upon the expected granularity of the fingerprint capturing process. These parameters are discussed in Section 5.3.

(c) *Comparison of SqOMs and templates.* The comparison of SqOMs and templates is done in the simplest possible manner. For each template we produce a similarity map with the same dimensions as the SqOM. Each position of such similarity maps corresponds to the value yielded by the RSM between the template and the neighbourhood of the block. Finally, all the similarity maps corresponding to the same type of SP are fused using the max operator. In this way, we obtain two graded representations of the presence of SPs, one for each type of SP.

4.3. Proposed algorithm

The ideas in Section 4.2 outline the complete algorithm for the detection of SPs in fingerprints.

Note that the algorithm is presented without giving the specific parameters used to make it as general as possible. Furthermore, these parameters must be chosen depending on the characteristics of the images in which it is applied. For this reason, the complete parameter specification used in this paper is shown in Section 5.3. This algorithm is composed of the following phases, which are also schematically represented in Fig. 7:

1. Dividing the image into non-overlapping blocks.
2. Segmenting the image using the previous calculated blocks.
   2.1. Normalizing the image to a desired mean and variance, 100 and 1000, respectively [33].

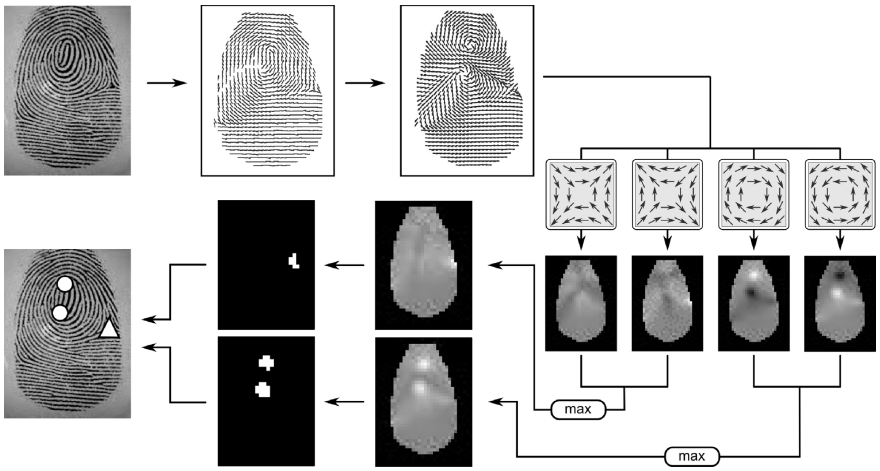


Fig. 7. Schematic representation of the proposed framework for singular point detection using orientation templates and Radial Similarity Measures, namely Template-based Singular Point Detection (TSPD).

2.2. Segmenting the fingerprint and the background, by assigning to the latter those blocks for which the variance of the pixel intensities is greater than 30 [33].
3. Calculating the Orientation Map over the segmented image.
   3.1. Computing the gradient at each pixel of the image (e.g., using Sobel masks) [45].
   3.2. Since the gradients are computed for each pixel and the result of a single pixel may not be reliable enough, the OM is smoothed to get more accurate orientations. In order to do so, the technique by Kass and Witkin [48] is used.
   3.3. Creating the OM from the regularized orientations.
4. Creating the SqOM by multiplying by two the values in the OM.
5. Detecting singular points.
   5.1. Computing the similarity map for each template. This is done by comparing the elements of the SqOM and those in each templates using the RSMs in Section 3.2.
   5.2. Fusing the similarity maps corresponding to each type of SP. This is done by obtaining, at each block, the maximum response for the cores and, in parallel, for the deltas.
   5.3. Selecting cores and deltas. This is done by taking the two points with the highest local response for core and delta similarity map in parallel. They are considered as a SP if they overall a threshold (Table 2).

Regarding computational times of the new technique, it is worth mentioning that it is comparable to the others methods we have used, Liu and Poincarè. All the methods compared share the same computational process. They go trough the fingerprint image executing a series of operations to get local maximums (SPs). One of the differences between the new method and Liu's method is that it does not use multi-scale image processing. This leads to a speed increase in execution time, but we consider it negligible. The segmentation step and the creation of the OM are common to all the methods and the complexity of the singular point extraction is equivalent to Liu's method. It may be a slight increase in complexity but it is imperceptible.

5. Experiments

The TSPD method has qualitative advantages compared to other traditional methods, e.g., the simplicity of the process and its visualization. However, it also demands a quantitative verification. In this section we check the quantitative performance of our method compared to that of the most relevant SP detection methods in the literature. In Section 5.1, we review the datasets we have used in the comparison. Section 5.2 covers the details on the quantification of the results, whereas Section 5.3 contains a detailed review of the setting and parametrization of the SP detection methods. The results, as well as a brief discussion, are included in Section 5.4.

5.1. Datasets

This experiment uses fingerprints from two different sources. The first one is National Institute of Standards and Technology Special Database 4 (NIST-4) [19]. This database contains rolled-ink fingerprints from the FBI (Federal Bureau of Investigation) and is, historically, the most used benchmark in the fingerprint classification literature. NIST-4 contains 4000 fingerprints (of 512×480 pixels) taken from 2000 fingertips. Hence, there are two captions of each fingerprint. Note that in NIST-4, as well as in most of the available databases, there is no ground truth with respect to the position and type of the SPs. On this account, we have manually labelled the first 1000 fingerprints from NIST-4 database to evaluate the proposed method. Labelling has been carried out according to the specifications given in the specialized literature and has been thoroughly revised by multiple reviewers. In order to ease the process of evaluating further proposals, our ground truth is available at [54]. For illustrative purposes, Fig. 1 shows three fingerprints from NIST-4, together with the corresponding ground truth data.

Even though NIST-4 is widely accepted, it should also be mentioned that in this dataset fingerprint classes are evenly represented, contrary to the reality in real-world databases, in which their distribution is skewed. Moreover, the representation of rolled-ink captions is limited, since most of the current applications

Table 1
Setting of SFinGe for the generation of the three datasets used in the experimental validation.

Scanner parameters
Acquisition area: 0.58″ × 0.77″ (14.6 mm × 19.6 mm)
Resolution: 500 dpi, image size: 288 × 384
Background type: optical
Background noise: default
Crop borders: 0 × 0
Generation parameters
Seed: 1
Impression per finger: 25 (only the first one is used)
Class distribution: natural
Generate pores: enabled
Save ISO templates: enabled
Output settings
Output file type: WSQ

collect fingerprint images using optical sensors. For these reasons, we also consider three additional databases generated with SFinGe synthetic fingerprint generator [40,39]. Although synthetic, SFinGe-generated fingerprints reflect in a faithful manner the difficulties in real scenarios. Moreover, the class distribution can be adjusted to that in reality, that is, 3.7%, 2.9%, 31.7%, 33.8% and 27.9% for arch, tented arch, right loop, left loop and whorl, respectively. Finally, it is relevant that SFinGe itself provides the ground truth data for the SPs, and hence the evaluation process becomes completely objective, whereas in NIST-4 there may be a certain error due to the manual labelling. The validity of the fingerprint images produced by SFinGe is, in any case, widely accepted, to the point that it has already been used in several editions of the fingerprint verification competition (FVC) [55–59] with results similar to those obtained with real fingerprint databases.

Aiming at simulating different scenarios, we have used three different quality profiles in the generation of fingerprints with SFinGe. For each quality profile, we generate 1000 fingerprint images. The following profiles are considered in our experiments (the rest of the parameters used in SFinGe tool are presented in Table 1):

- *High quality no perturbations* (HQNoPert): High quality fingerprints without any kind of perturbation;
- *Default*: Middle quality fingerprints with slight localization and rotation perturbations;
- *Varying quality and perturbations* (VQandPert): Fingerprints with different qualities are included, which are perturbed in location, rotation and geometric distortions.

The fingerprints generated for each quality profile are rather different, and also significantly different from those in NIST-4. Fig. 8 includes one fingerprint for each of the above mentioned quality profiles. In the remainder of this work, we refer to the datasets created with profiles HQNoPert, Default and VQandPert as SFinGe Dataset 1, 2 and 3, respectively.

5.2. Quantification of the results

In this experiment we have quantified the performance of each procedure in correctly and accurately detecting SPs. For each dataset we have created a confusion matrix which accounts for the success and fallout in SP detection. That is, given a dataset, a unique confusion matrix is completed from the confrontation of the SPs detected by the automatic method at each image and those in the ground truth.

After extracting the SPs for a fingerprint, we first compute the best-possible matching between the cores in the automatic solution to those in the ground truth, forcing a one-to-one correspondence. Each matched core in the automatic solution accounts for as True Positive (TP). Then, each unmatched core in the automatic solution and in the ground truth are tagged as False Positive (FP) and False Negative (FN), respectively. Finally, in case both the automatic solution and the ground truth contain less than two cores, the missing SPs are taken as correct predictions, and consequently are accounted for as True Negatives (TN). The process is analogous for the deltas, whose results are stored in a separate matrix. Note that each fingerprint can generate more than two hits in the confusion matrix, if SPs are both missed (FNs) and misdetected (FPs).

It should be considered that fingerprint analysis methods do not necessarily locate a SP at the exact location a human does. This is due to the discrete nature of data in an image and the scope of the semi-local analysis of the image needed to locate the SPs. Consequently, we consider some tolerance in the correspondence of SPs tagged by the automatic method to those in the ground truth. For the present experiment, this spatial tolerance is equivalent to 5% of the image diagonal for SFinGe databases and 10% of the image diagonal for NIST-4 database. The percentage difference between databases arises from the size of SFinGe and NIST-4 fingerprints (SFinGe ones are rectangular images, whereas NIST-4 ones are squared, but also the thickness of the ridges vary due to their different nature).

The results generated with the above-mentioned procedure lead to two confusion matrices for each dataset, one for cores and one for deltas. From such matrices, we have generated different scalar interpretations of the quality of the results. More specifically, we consider precision (Prec) and recall (Rec), given by

$$\text{Prec} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad \text{and} \quad \text{Rec} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{7}$$

(a) Profile *HQNoPert* (b) Profile *Default* (c) Profile *VQandPert*

Fig. 8. Fingerprints generated using SFinGe with different quality profiles.

respectively. Precision and recall quantify the ability of the automatic method to obtain a reliable and complete collection of SPs, respectively. They can be combined to produce a scalar representation of the overall quality of the process. In this work we adhere to the so-called F-measure, given by

$$\mathrm{F} = \frac{\mathrm{PREC} \cdot \mathrm{REC}}{0.5\mathrm{PREC} + 0.5\mathrm{REC}}. \quad (8)$$

Moreover, we also measure the percentage of fingerprints in which all the SPs (cores and deltas) have been correctly detected.

5.3. Experimental procedure

In this experiment, the results of the TSPD method have been compared with those of the Poincarè method [20], as well as to those of the one proposed by Liu [21]. The former method has been selected because it is the most used SP detection method in literature, whereas the latter method is included because it holds strong similarities to ours. Aiming at carrying out a fair comparison, the techniques used for OM computation, smoothing and segmentation are identical for each of the three SP detection methods.

Firstly, the image is divided into non-overlapping blocks of 5×5 pixels (for SFinGe databases) or 10×10 pixels (for the NIST-4 database). The different size between SFinGe and NIST-4 fingerprints makes it necessary to use different block sizes, since the ridges of the NIST-4 fingerprints are much thicker than SFinGe ones. Secondly, the image is segmented (as explained in Section 4.3) to avoid false SP detections in the ridge abnormalities occurring at the fingertip boundaries. Thirdly, to compute the gradients for the OM we use the well-known Sobel operators [60,61], which is the most common option in fingerprint analysis. The resulting matrix is the OM, which is further regularized using a flat mask of 5×5 blocks [48]. Notice that we do not use a different size of mask for NIST-4 fingerprints since the block size used to compute the OM produces blocks with equivalent information regardless of the database.

Once the OM is generated, each of the methods needs to be customized, the details being as follows:

- *TSPD* – Regarding the templates, we need to set their size and the number of delta SP templates to use. In order to preserve the fairness of the comparison, we have considered a very basic setup, which is the baseline configuration of the method. This configuration involves only 4 templates (two for each type of SP), all of them of 5×5 blocks. In the case of the delta SP templates, we take $\alpha \in \{0, 90\}$. This is, objectively, the minimum set of templates to be used.

  As for the RSMs, we consider 9 measures, in order to shed light on the impact the RSMs have on the final results. This way, we are able to show their flexibility, allowing one to define different perceived similarities. The RREFs are constructed as in Proposition 4 from pairs of automorphisms (φ, ψ) given by

$$\varphi(x) = x^{e_1} \quad \text{and} \quad \psi(x) = \frac{x^{e_2}}{\pi^{e_2 - 1}}$$

  where $e_1, e_2 \in \{0.5, 1, 2\}$. This leads to 9 different pairs of automorphisms.

  Each combination of automorphisms, together with the thresholds used for the discrimination of the SPs, is shown in Table 2. Although some authors have studied the automatic determination of thresholds [62,63], we avoid this step in order to preserve the clarity and reproducibility of the experiments. Several thresholds have been tested before selecting one that has a positive behaviour in all datasets (see Table 2).
- *Poincarè method* – This method consists of computing the difference between each orientation in a 3×3 neighbourhood and its clockwise successor. Those differences are further summed up to produce the Poincarè index in each block. This index takes value 0, 1/2 or $-(1/2)$, indicating the absence of a SP, the presence of a core or the presence of a delta, respectively. Although other authors have used other configurations of the neighbourhood [17,64], specially regarding its size, we maintain the widely accepted 3×3 size.
- *Liu's method* – In this method the SqOM is filtered with first order complex filters at different scales. More specifically, the large scale filters are used to discriminate the real SPs from spurious responses, while the fine scale ones determine their precise location. The threshold used for discrimination of SPs is set to 0.7 (this threshold is manually set to measure the performance of the method, as those in Table 2). Regarding the scales we consider, as in [21], filters of $s \times s$ blocks, with $s \in \{3, 5, 7, 9\}$.

Table 2
List of configurations of the TSPD used in the experimental validation. For each of the configurations, we list the exponents (e_1, e_2) of the automorphisms used in the construction of the RREF, as well as the threshold used for SP discrimination.

Name	e_1	e_2	Threshold
C1	0.5	0.5	0.70
C2	0.5	1.0	0.85
C3	0.5	2.0	0.95
C4	1.0	0.5	0.55
C5	1.0	1.0	0.70
C6	1.0	2.0	0.85
C7	2.0	0.5	0.35
C8	2.0	1.0	0.60
C9	2.0	2.0	0.75

5.4. Results

The results obtained for each method and dataset are listed in Tables 3–6 , including:

- The values at each position of the confusion matrix (as explained in Section 5.2), namely PREC, REC and F. This information is displayed for cores and deltas separately.
- The average distance in pixels from the position at which the matched SPs where located and their position at the ground truth. This information is also listed individually for cores and deltas.
- The arithmetic mean between the F value for cores and deltas, namely Combined F (Comb. F).
- The percentage of fingerprints for which the method achieved a perfect detection (Perfect Detection Percentage, PDP). That is, the rate of fingerprints for which each method gathered the exact number of SPs, all of them being located within the tolerance ratio of 5% and 10% of the length of the image diagonal, for SFinGe and NIST-4 fingerprints, respectively.

For each dataset, the best performer at each statistic is boldfaced.

The first fact to be noticed from the results of the experiment is the great variability of performance across datasets, especially between NIST-4 and SFinGe datasets. This is due to the low quality of NIST-4 fingerprints, which often include damaged fingertips, hand-written annotations on the fingerprint margins, etc. This does not reduce the representativity of the datasets generated with SFinGe, since modern sensors for fingerprint recording produce images that are closer to those by SFinGe than to those in NIST-4. This variable behaviour has also been shown in previous studies on the topic [37].

Table 3
Results gathered by each SP detection method on the NIST-4 dataset (1000 fingerprints).

Quant.	Template-based SP detection									Poincaré	Liu
	C1	C2	C3	C4	C5	C6	C7	C8	C9		
Cores											
TP	897	875	868	853	911	912	867	850	915	759	**922**
FP	273	147	132	161	266	241	292	138	261	**40**	329
FN	93	115	122	137	79	78	123	140	75	231	**68**
TN	776	889	905	878	777	799	756	902	782	**982**	721
Prec	.767	.856	.868	.841	.774	.791	.748	.860	.778	**.950**	.737
Rec	.906	.884	.877	.862	.920	.921	.876	.859	.924	.767	**.931**
F	.831	.870	**.872**	.851	.841	.851	.807	.859	.845	.849	.823
Avg. dist.	14.4	**13.7**	13.8	14.9	14.2	14	15.6	14.3	13.9	13.8	16.4
Deltas											
TP	871	844	839	830	874	877	814	837	**881**	728	767
FP	340	176	159	188	393	351	335	163	426	53	**29**
FN	102	129	134	143	99	96	159	136	**92**	245	206
TN	731	884	898	879	682	722	754	897	648	1006	**1011**
Prec	.719	.827	.841	.815	.690	.714	.708	.837	.674	.932	**.964**
Rec	.895	.867	.862	.853	.898	.901	.837	.860	**.905**	.748	.788
F	.798	.847	.851	.834	.780	.797	.767	.848	.773	.830	**.867**
Avg. dist.	11.2	10.4	**10.3**	11.3	11.1	10.6	12.1	11	10.7	13.5	11.5
Total											
Comb. F	.815	.859	**.861**	.843	.811	.824	.787	.853	.809	.839	.845
PDP	56.70	68.10	**69.90**	63.00	54.50	58.70	48.10	66.90	52.80	66.40	54.50

Table 4
Results gathered by each SP detection method on the Sfinge Dataset 1 (1000 fingerprints, profile HQNoPert).

Quant.	Template-based SP detection									Poincaré	Liu
	C1	C2	C3	C4	C5	C6	C7	C8	C9		
Cores											
TP	1209	1215	1228	1157	1216	1230	1151	1165	1228	1187	**1232**
FP	1	**0**	1	**0**	**0**	**0**	16	**0**	3	**0**	58
FN	33	27	14	85	26	12	91	77	14	55	**10**
TN	757	**758**	757	**758**	**758**	**758**	742	**758**	756	**758**	702
Prec	.999	**1**	.999	**1**	**1**	**1**	.986	**1**	.998	**1**	.955
Rec	.973	.978	.989	.932	.979	.990	.927	.938	.989	.956	**.992**
F	.986	.989	.994	.965	.989	**.995**	.956	.968	.993	.977	.973
Avg. dist.	7	6.7	6.4	7	6.8	6.4	7.2	6.9	6.5	**4.7**	7.2
Deltas											
TP	736	725	723	709	747	753	705	714	**754**	700	717
FP	17	1	3	13	22	17	44	8	28	**0**	12
FN	32	43	45	59	21	15	63	54	**14**	68	51
TN	1216	1231	1229	1219	1211	1215	1192	1224	1204	**1232**	1220
Prec	.977	.999	.996	.982	.971	.978	.941	.989	.964	**1**	.984
Rec	.958	.944	.941	.923	.973	.980	.918	.930	**.982**	.911	.934
F	.968	.971	.968	.952	.972	**.979**	.929	.958	.973	.954	.958
Avg. dist.	4.4	4.1	**4**	4.2	4.3	4.3	4.5	4.1	4.3	4.5	4.9
Total											
Comb. F	.977	.980	.981	.958	.980	**.987**	.943	.963	.983	.966	.966
PDP	92.40	93.50	94.30	85.20	93.60	**95.90**	80.50	87.00	94.60	88.80	89.90

Regarding the NIST-4 dataset, we find that TSPD-C3 is the best performer, obtaining the greatest PDP (69.90%). Although TSPD-C2 and TSPD-C8 stay close to this result, configurations such as TSPD-C7, TSPD-C9 and TSPD-C5 lead to the worst outcome. The relevance of the RSMs is clearly illustrated with these results. Besides, TSPD-C5 (when φ and ψ are the identity function) is not the best performer. From this fact, we infer that choosing suitable automorphisms on the construction of the RREFs significantly improves the result of the TSPD method. In Table 3 we also observe that Liu's method obtains a PDP similar to our worst configurations (54.50%). Despite being the best method detecting cores (922), Liu's method also produces 329 false cores detections, significantly more than the Poincarè method (40) and TSPD-C3 (132). The behaviour is opposed regarding deltas, where the precision of Liu's method is very high (29 FPs and 1011 TNs), at the cost of very little recall (767 TPs).[4] Otherwise, the method of Poincarè presents a high PDP, but is not the best performer because of its difficulties in detecting cores (231 FNs) and deltas (245 FNs). From these results, we understand that TSPD-C3 obtains the best results in general terms, showing the most equilibrated behaviour between successes (TPs, TNs) and failures (FPs, FNs).

[4] This behaviour of Liu's method is consistent with that observed by Galar et al. [37].

Table 5
Results gathered by each SP detection method on the Sfinge Dataset 2 (1000 fingerprints, profile Default).

Quant.	Template-based SP detection									Poincaré	Liu
	C1	C2	C3	C4	C5	C6	C7	C8	C9		
Cores											
TP	1295	1299	1312	1256	1305	**1319**	1254	1264	1315	1230	1317
FP	11	9	9	**6**	11	14	26	**6**	20	16	76
FN	37	33	20	76	27	**13**	78	68	17	102	15
TN	659	660	659	**663**	659	654	648	**663**	649	655	594
Prec	.992	.993	.993	**.995**	.992	.989	.980	**.995**	.985	.987	.945
Rec	.972	.975	.985	.943	.980	**.990**	.941	.949	.987	.923	.989
F	.982	.984	.989	.968	.986	**.990**	.960	.972	.986	.954	.967
Avg. dist.	7.3	6.9	6.7	7.4	7.1	6.7	7.7	7.2	6.7	**4.9**	7.6
Deltas											
TP	713	705	703	685	720	**727**	670	693	**727**	651	698
FP	73	31	36	44	87	80	105	38	117	29	**18**
FN	44	52	54	72	37	**30**	87	64	**30**	106	59
TN	1172	1213	1209	1202	1159	1164	1151	1208	1132	1218	**1225**
Prec	.907	.958	.951	.940	.892	.901	.865	.948	.861	.957	**.975**
Rec	.942	.931	.929	.905	.951	**.960**	.885	.915	**.960**	.860	.922
F	.924	.944	.940	.922	.921	.930	.875	.931	.908	.906	**.948**
Avg. dist.	4.8	4.6	**4.5**	4.9	4.7	4.7	5.2	4.7	4.6	4.8	5.6
Total											
Comb. F	.953	.964	**.964**	.945	.954	.960	.918	.952	.947	.930	.958
PDP	85.60	89.20	**90.00**	82.70	86.00	88.50	74.50	84.50	84.40	79.30	87.20

Table 6
Results gathered by each SP detection method on the Sfinge Dataset 3 (1000 fingerprints, profile VQandPert).

Quant.	Template-based SP detection									Poincaré	Liu
	C1	C2	C3	C4	C5	C6	C7	C8	C9		
Cores											
TP	1140	1141	1167	1080	1148	**1176**	1080	1082	1167	1019	1172
FP	27	20	21	15	30	37	30	**12**	42	26	107
FN	50	49	23	110	42	**14**	110	108	23	171	18
TN	786	792	791	798	784	776	786	**801**	772	790	705
Prec	.977	.983	.982	.986	.975	.969	.973	**.989**	.965	.975	.916
Rec	.958	.959	.981	.908	.965	**.988**	.908	.909	.981	.856	.985
F	.967	.971	**.981**	.945	.970	.979	.939	.947	.973	.912	.949
Avg. dist.	7.1	6.6	6.4	7.1	6.8	6.4	7.3	6.9	6.4	**5.1**	7.4
Deltas											
TP	656	651	649	629	663	669	620	633	**670**	612	639
FP	96	55	60	67	114	107	111	61	131	34	**13**
FN	42	47	49	69	35	29	78	65	**28**	86	59
TN	1212	1250	1245	1237	1195	1200	1197	1244	1176	1274	**1290**
Prec	.872	.922	.915	.904	.853	.862	.848	.912	.836	.947	**.980**
Rec	.940	.933	.930	.901	.950	.958	.888	.907	**.960**	.877	.915
F	.905	.927	.923	.902	.899	.908	.868	.909	.894	.911	**.947**
Avg. dist.	4.9	4.6	**4.5**	4.8	4.8	4.7	5.3	4.7	4.7	4.8	5.4
Total											
Comb. F	.936	.949	**.952**	.924	.935	.944	.904	.928	.934	.912	.948
PDP	83.10	86.60	**88.50**	77.70	83.10	86.70	72.60	79.30	83.50	74.60	84.70

For the SFinGe datasets, we observe that in Tables 4 and 5, TSPD-C3 obtains the best PDP results (95.90% and 88.50% respectively), although in Table 6 the best one is TSPD-C6 (90.00%). The general trend observed in Tables 4–6 is that the TSPD method is usually able to outperform both the Poincarè and Liu's methods, although certain configurations fail to do so. A remarkable fact is the absolute absence of core FPs when using the TSPD method, which hardly ever account for more than 30 of such mistakes over 1000 fingerprints. This leads to high Prec and, as a consequence, to high F. The situation with the delta SPs is similar as it is for core SPs, but not as positive for the TSPD method. In Dataset 1, the results are similar to those of the cores, but the situation changes in Dataset 2 and is accentuated in Dataset 3.

Summing up, from the results in the present experiment we consider the TSPD method to be competitive with the contending methods. Although the TSPD method requires setting the parameters of the RSMs and thresholds, similar situations occurs with most of the SP detection methods (including Liu's method). Interestingly, the RREF leading to the best results in the TSPD method is not that constructed with the pair of automorphisms C5, indicating that non-linear modelling of dissimilarity can play a role in real applications. Specifically, the best-performing version is that using the pair of automorphisms C3, since it generally outperforms all of the other versions of the TSPD method in terms of Combined F and PDP, the SFinGe Dataset 1 being the sole exception to this fact.

It is worth noting that the TSPD methods have advantages over its counterparts other than pure performance. For example, it holds interesting visualization properties when it comes to error correction, partly derived from the simplicity of the method. Indeed, we have not exploited the potential use of multi-scale templates yet as Liu's method does.

Attending at the results obtained by the RREFs and RSMs, we can state that this extension of the REF and SM concepts considered in Fuzzy Sets theory is appropriate to deal with radial data. Even though radial data may be different from scalar data to some extent, vagueness and imprecision are inherent to both types of data in real applications. Hence, concepts from Fuzzy Sets are also interesting to deal with radial data, as we have shown in this paper. Moreover, the parametrizable construction proposed in Section 5.3 allows us to provide a flexible and configurable model, whose results can be adapted to each application.

6. Conclusions

This work has two main contributions. First, we have adapted the concepts of Restricted Equivalence Function (REF) and Similarity Measure (SM) to radial environments. The resulting operators, namely Restricted Radial Equivalence Function (RREF) and Radial Similarity Measure (RSM), capture the expected behaviour and semantics of the original operators, but at the same time embrace the cyclic nature of radial data. In both cases, we have analysed its properties and proposed construction methods. Second, we have proved the validity of the operators in a complex scenario, such as fingerprint analysis. In order to do so, we have presented a framework for Singular Point (SP) detection based on templates, which requires the use of RSMs at the template matching stage. This framework, namely Template-based Singular Point Detection (TSPD) method, shows promising results and illustrate the usefulness of RSMs for the comparison of radial data in scenarios in which imprecision and ambiguities occur.

We expect to expand the present work in two different lines of research. As for the theoretical aspects, we aim at adapting to radial data several other operators with special relevance in fuzzy set theory, e.g., aggregation operators or dissimilarity functions. Regarding the TSPD method, we intend to improve it by incorporating notions from multi-scale image processing, as well as by designing self-adapting RREFs which are able to modify their behaviour depending on the characteristics of the fingerprint image and/or the ridge Orientation Map.

Acknowledgements

The authors gratefully acknowledge the financial support of the Spanish Ministry of Science (project TIN2013-40765-P), the Research Services of Universidad Pública de Navarra, as well as the financial support of the Research Foundation Flanders (FWO project 3G.0838.12.N).

References

[1] B. De Baets, H. De Meyer, On the cycle-transitive comparison of artificially coupled random variables, Int. J. Approx. Reason. 47 (3) (2008) 306–322.
[2] F. Attneave, Dimensions of similarity, Am. J. Psychol. (1950) 516–556.
[3] I. Bloch, On fuzzy distances and their use in image processing under imprecision, Pattern Recognit. 32 (11) (1999) 1873–1895.
[4] Y. Rubner, C. Tomasi, L.J. Guibas, The earth mover's distance as a metric for image retrieval, Int. J. Comput. Vision 40 (2) (2000) 99–121.
[5] S. Santini, R. Jain, Similarity measures, IEEE Trans. Pattern Anal. Mach. Intell. 21 (9) (1999) 871–883.
[6] I. Kramosil, J. Michálek, Fuzzy metrics and statistical metric spaces, Kybernetika 11 (5) (1975) 336–344.
[7] A. Tversky, Measures of similarity, Psychol. Rev. 84 (4) (1977) 327–352.
[8] A. Tversky, I. Gati, Studies of similarity, Cognit. Categoriz. 1 (1978) 79–98.
[9] M.A. Ruzon, C. Tomasi, Edge, junction, and corner detection using color distributions, IEEE Trans. Pattern Anal. Mach. Intell. 23 (11) (2001) 1281–1295.
[10] C. Odet, B. Belaroussi, H. Benoit-Cattin, Scalable discrepancy measures for segmentation evaluation, in: Proc. of the International Conf. on Image Processing, vol. 1, 2002, pp. 785–788.
[11] L.A. Zadeh, Similarity relations and fuzzy orderings, Inform. Sci. 3 (2) (1971) 177–200.
[12] B. De Baets, H. De Meyer, H. Naessens, A class of rational cardinality-based similarity measures, J. Comput. Appl. Math. 132 (1) (2001) 51–69.
[13] B. De Baets, H. De Meyer, Transitivity-preserving fuzzification schemes for cardinality-based similarity measures, Eur. J. Oper. Res. 160 (3) (2005) 726–740.
[14] H. Bustince, E. Barrenechea, M. Pagola, Restricted equivalence functions, Fuzzy Sets Syst. 157 (17) (2006) 2333–2346.
[15] F. Zhang, E.R. Hancock, New Riemannian techniques for directional and tensorial image data, Pattern Recognit. 43 (4) (2010) 1590–1606.
[16] M. Galar, J. Derrac, D. Peralta, I. Triguero, D. Paternain, C. Lopez-Molina, S. García, J.M. Benítez, M. Pagola, E. Barrenechea, H. Bustince, F. Herrera, A survey of fingerprint classification. Part I: Taxonomies on feature extraction methods and learning models, Knowl.-Based Syst. 81 (2015) 76–97.
[17] K. Karu, A.K. Jain, Fingerprint classification, Pattern Recognit. 29 (3) (1996) 389–404.
[18] K. Nilsson, J. Bigun, Localization of corresponding points in fingerprints by complex filtering, Pattern Recognit. Lett. 24 (13) (2003) 2135–2144.
[19] C.I. Watson, C.L. Wilson, NIST Special Database 4, Fingerprint Database, Tech. rep., U.S. National Institute of Standards and Technology, 1992.
[20] M. Kawagoe, A. Tojo, Fingerprint pattern classification, Pattern Recognit. 17 (3) (1984) 295–303.
[21] M. Liu, Fingerprint classification based on Adaboost learning from singularity features, Pattern Recognit. 43 (2010) 1062–1070.
[22] V. Gregori, S. Morillas, A. Sapena, Examples of fuzzy metrics and applications, Fuzzy Sets Syst. 170 (1) (2011) 95–111.
[23] E. Palmeira, B. Bedregal, H. Bustince, A generalization of a characterization theorem of restricted equivalence functions, in: H. Bustince, J. Fernandez, R. Mesiar, T. Calvo (Eds.), Aggregation Functions in Theory and in Practise, vol. 228 of Advances in Intelligent Systems and Computing, Springer, Berlin Heidelberg, 2013, pp. 453–464.
[24] A. Jurio, M. Pagola, D. Paternain, C. Lopez-Molina, P. Melo-Pinto, Interval-valued restricted equivalence functions applied on clustering techniques, in: Proc. of the IFSA-EUSFLAT, 2009, pp. 831–836.
[25] L. Xuecheng, Entropy, distance measure and similarity measure of fuzzy sets and their relations, Fuzzy Sets Syst. 52 (3) (1992) 305–318.
[26] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, vol. 221 of Studies in Fuzziness and Soft Computing, Springer, 2007.
[27] N.I. Fisher, Statistical Analysis of Circular Data, Cambridge University Press, 1993.
[28] E. Batschelet, Circular Statistics in Biology, Academic Press, 1981.
[29] J.C. Davis, R.J. Sampson, Statistics and Data Analysis in Geology, vol. 3, Wiley, New York, 2002.
[30] K.V. Mardia, P.E. Jupp, Directional Statistics, Wiley, 2000.
[31] D. Peralta, M. Galar, I. Triguero, D. Paternain, S. García, E. Barrenechea, J.M. Benítez, H. Bustince, F. Herrera, A survey on fingerprint minutiae-based local matching for verification and identification: taxonomy and experimental evaluation, Inform. Sci. 315 (2015) 67–87.
[32] L. Hong, Y. Wan, A. Jain, Fingerprint image enhancement: algorithm and performance evaluation, IEEE Trans. Pattern Anal. Mach. Intell. 20 (8) (1998) 777–789.
[33] D. Peralta, M. Galar, I. Triguero, O. Miguel-Hurtado, J.M. Benitez, F. Herrera, Minutiae filtering to improve both efficacy and efficiency of fingerprint matching algorithms, Eng. Appl. Artif. Intell. 32 (2014) 37–53.
[34] X. Jiang, W.-Y. Yau, Fingerprint minutiae matching based on the local and global structures, in: Proc. of the International Conf. on Pattern Recognition, vol. 2, 2000, pp. 1038–1041.
[35] X. Chen, J. Tian, X. Yang, A new algorithm for distorted fingerprints matching based on normalized fuzzy similarity measure, IEEE Trans. Image Process. 15 (3) (2006) 767–776.
[36] D. Peralta, I. Triguero, S. García, F. Herrera, J.M. Benitez, Dpd-dff: a dual phase distributed scheme with double fingerprint fusion for fast and accurate identification in large databases, Inform. Fusion 32 (Part A) (2016) 40–51.
[37] M. Galar, J. Derrac, D. Peralta, I. Triguero, D. Paternain, C. Lopez-Molina, S. García, J.M. Benítez, M. Pagola, E. Barrenechea, H. Bustince, F. Herrera, A survey of fingerprint classification. Part II: Experimental analysis and ensemble proposal, Knowl.-Based Syst. 81 (2015) 98–116.
[38] K. Cao, L. Pang, J. Liang, J. Tian, Fingerprint classification by a hierarchical classifier, Pattern Recognit. 46 (12) (2013) 3186–3197.
[39] D. Maltoni, D. Maio, A.K. Jain, S. Prabhakar, Handbook of Fingerprint Recognition, Springer-Verlag, 2009.
[40] R. Cappelli, D. Maio, D. Maltoni, SFinGe: an approach to synthetic fingerprint generation, in: International Conf. on Control, Automation, Robotics and Vision, 2004.
[41] Q. Zhang, H. Yan, Fingerprint classification based on extraction and analysis of singularities and pseudo ridges, Pattern Recognit. 37 (11) (2004) 2233–2243.

[42] J.-H. Hong, J.-K. Min, U.-K. Cho, S.-B. Cho, C.H. Leung, Fingerprint classification using one-vs-all support vector machines dynamically ordered with naive bayes classifiers, Pattern Recognit. 41 (2008) 662–671.
[43] J. Li, W.-Y. Yau, H. Wang, Combining singular points and orientation image information for fingerprint classification, Pattern Recognit. 41 (1) (2008) 353–366.
[44] A.K. Jain, S. Minut, Hierarchical kernel fitting for fingerprint classification and alignment, in: Proc. of the International Conf. on Pattern Recognition, 2002, pp. 469–473.
[45] A.M. Bazen, S.H. Gerez, Systematic methods for the computation of the directional fields and singular points of fingerprints, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 905–919.
[46] Y. Mei, R. Hou, J. Wang, An improved method for fingerprints' singular points detection based on orientation field partition, Int. J. Signal Process. Image Process. Pattern Recognit. 6 (1) (2013) 225–234.
[47] F. Turroni, D. Maltoni, R. Cappelli, D. Maio, Improving fingerprint orientation extraction, IEEE Trans. Inform. Forensics Secur. 6 (3) (2011) 1002–1013.
[48] M. Kass, A. Witkin, Analyzing oriented patterns, Comput. Vision Graph. Image Process. 37 (3) (1987) 362–385.
[49] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, IEEE Trans. Med. Imaging 8 (3) (1989) 263–269.
[50] R. Poli, G. Valli, An algorithm for real-time vessel enhancement and detection, Comput. Methods Progr. Biomed. 52 (1) (1997) 1–22.
[51] M.H. Hueckel, An operator which locates edges in digitized pictures, J. ACM 18 (1) (1971) 113–125.
[52] Y. Li, X. lin Qi, Y. jiu Wang, Eye detection by using fuzzy template matching and feature-parameter-based judgement, Pattern Recognit. Lett. 22 (10) (2001) 1111–1124.
[53] J. Serra, Introduction to mathematical morphology, Computer Vision, Graph. Image Process. 35 (3) (1986) 283–305.
[54] C. Marco-Detchart, J. Cerron, L. De Miguel, C. Lopez-Molina, H. Bustince, M. Galar, Singular Point Location for NIST-4 Database, 2015 https://giara.unavarra.es/datasets/solNIST4.zip.
[55] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain, FVC2000: fingerprint verification competition, IEEE Trans. Pattern Anal. Mach. Intell. 24 (3) (2002) 402–412.
[56] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain, FVC2002: second fingerprint verification competition, in: Proc. of the International Conf. on Pattern Recognition, 2002, pp. 811–814.
[57] D. Maio, D. Maltoni, R. Cappelli, J.L. Wayman, A.K. Jain, FVC2004: third fingerprint verification competition, in: Proc. of the International Conf. on Biometric Authentication, 2004, pp. 1–7.
[58] R. Cappelli, M. Ferrara, A. Franco, D. Maltoni, Fingerprint verification competition 2006, Biom. Technol. Today 15 (7-8) (2007) 7–9.
[59] B. Dorizzi, R. Cappelli, M. Ferrara, D. Maio, D. Maltoni, N. Houmani, S. Garcia-Salicetti, A. Mayoue, Fingerprint and on-line signature verification competitions at ICB 2009, in: M. Tistarelli, M. Nixon (Eds.), Advances in Biometrics, vol. 5558 of Lecture Notes in Computer Science, Springer, Berlin Heidelberg, 2009, pp. 725–732.
[60] I. Sobel, G. Feldman, A 3×3 Isotropic Gradient Operator for Image Processing, Presented at a Talk at the Stanford Artificial Intelligence Project, 1968.
[61] J.M.S. Prewitt, Object Enhancement and Extraction, Picture Processing and Psychopictorics, Academic Press, 1970, pp. 75–149.
[62] P.L. Rosin, Edges: saliency measures and automatic thresholding, Mach. Vision Appl. 9 (1997) 139–159.
[63] X. Liu, Y. Yu, B. Liu, Z. Li, Bowstring-based dual-threshold computation method for adaptive Canny edge detector, in: Proc. of the International Conf. of Image and Vision Computing New Zealand, 2013, pp. 13–18.
[64] H.O. Nyongesa, S. Al-Khayatt, S.M. Mohamed, M. Mahmoud, Fast robust fingerprint feature extraction and classification, J. Intell. Robot. Syst. 40 (1) (2004) 103–112.

3.2 Width-based interval-valued distances and entropies

- Z. Takáč, H. Bustince, J. M. Pintor, C. Marco-Detchart, and I. Couso, "Width-based interval-valued distances and fuzzy entropies", *IEEE Access*, vol. 7, pp. 14 044–14 057, 2019
  - Journal: IEEE Access
  - Status: Published
  - Impact Factor: (JCR 2018) 4.098
  - Knowledge area:
    * Telecommunications: Ranking $19/88$ (Q1)
    * Computer Science, Information Systems: Ranking $23/155$ (Q1)
    * Engineering, Electrical & Electronic: Ranking $52/265$ (Q1)

Z. Takáč, H. Bustince, J. M. Pintor, C. Marco-Detchart and I. Couso, "Width-Based Interval-Valued Distances and Fuzzy Entropies," in IEEE Access, vol. 7, pp. 14044-14057, 2019. DOI: 10.1109/ACCESS.2019.2893800

3.3 Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders

- H Bustince, C Marco-Detchart, J Fernandez, C Wagner, J. Garibaldi, and Z Takáč, "Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders", *Fuzzy Sets and Systems*, 2019
  - Journal: Fuzzy Sets ans Systems
  - Status: Accepted
  - Impact Factor: (JCR 2018) 2.907
  - Knowledge area:
    - Statistics & Probability: Ranking $7/123$ (Q1)
    - Computer Science Theory & Methods: Ranking $21/104$ (Q1)
    - Applied Mathematics: Ranking $16/254$ (Q1)

ARTICLE IN PRESS FSS:7635

Available online at www.sciencedirect.com

ScienceDirect

Fuzzy Sets and Systems ••• (••••) •••–•••

FUZZY sets and systems

www.elsevier.com/locate/fss

Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders

H. Bustince [a,b,c,*], C. Marco-Detchart [a,b,c], J. Fernandez [a,b,c], C. Wagner [d], J.M. Garibaldi [d], Z. Takáč [e]

[a] *Departamento de Estadistica, Informatica y Matematicas, Universidad Publica de Navarra, Campus de Arrosadia, 31006, Pamplona, Spain*
[b] *Institute of Smart Cities, Universidad Publica de Navarra, Campus de Arrosadia, 31006, Pamplona, Spain*
[c] *Laboratory Navarrabiomed, Hospital Complex of Navarre (CHN), Universidad Publica de Navarra, IdiSNA, Irunlarrea 3. 31008, Pamplona, Spain*
[d] *School of Computer Science, University of Nottingham, Nottingham, United Kingdom*
[e] *Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology in Bratislava, Radlinskeho 9, Bratislava, Slovakia*

Received 30 July 2018; received in revised form 13 February 2019; accepted 1 April 2019

Abstract

In this work we study a new class of similarity measures between interval-valued fuzzy sets. The novelty of our approach lays, firstly, on the fact that we develop all the notions with respect to total orders of intervals; and secondly, on that we consider the width of intervals so that the uncertainty of the output is strongly related to the uncertainty of the input. For constructing the new interval-valued similarity, interval valued aggregation functions and interval-valued restricted equivalence functions which take into account the width of the intervals are needed, so we firstly study these functions, both in line with the two above stated features. Finally, we provide an illustrative example which makes use of an interval-valued similarity measure in stereo image matching and we show that the results obtained with the proposed interval-valued similarity measures improve numerically (according to the most widely used measures in the literature) the results obtained with interval valued similarity measures which do not consider the width of the intervals.

Keywords: Interval-valued fuzzy sets; Admissible order; Total order; Interval-valued similarity measure; Equivalence and restricted equivalence functions; Interval-valued aggregation function

* Corresponding author.
E-mail addresses: bustince@unavarra.es (H. Bustince), cedric.marco@unavarra.es (C. Marco-Detchart), fcojavier.fernandez@unavarra.es (J. Fernandez), christian.wagner@nottingham.ac.uk (C. Wagner), jon.garibaldi@nottingham.ac.uk (J.M. Garibaldi), zdenko.takac@stuba.sk (Z. Takáč).

https://doi.org/10.1016/j.fss.2019.04.002
0165-0114/

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

1. Introduction

In recent times, interval-valued fuzzy sets [11] are increasingly used in the same problems as standard type-1 fuzzy sets (from now on, simply referred to as fuzzy sets). This is due to the fact that, among other factors, they provide a way to represent the uncertainty inherent to the construction of an appropriate fuzzy set to represent a given setting. In this way, they may improve the numerical results of applications, as can be seen in [1–6,16,17,27,35]. Since many applications of fuzzy sets make use of similarity measures [7,9,10,29] (see also the recent works [14,15]) in order to determine the degree of resemblance between fuzzy sets, interval-valued similarities have become also an object of interest for some authors [18,24,30,41].

However, two observations can be stated regarding the recent literature about interval-valued similarities, and, more generally, about interval-valued fuzzy sets:

1. in most of the cases, only the partial order between intervals is considered;
2. the widths of the intervals are not considered.

In our opinion these two features are obstacles to further development of the theory and applications of interval-valued fuzzy sets.

Taking these factors into account, the objective of this paper is to construct similarity measures between interval-valued fuzzy sets in such a way that:

a) a total order for intervals (not only partial) is used;
b) the widths of intervals are considered.

With respect to objective (a), while in some applications it is not necessary for intervals to be comparable, in other cases (e.g. decision making or classification) in order to get a solution, it is important that any two intervals can be compared. Furthermore, it is desirable that some well-known notion which provides good results in the fuzzy setting and which involve the ordering of elements (such as OWA operators, Choquet integrals and so on) can be generalized in a natural way.

Regarding objective (b), we assume that the width of the membership interval of an element in a given set reflects the lack of knowledge of the precise membership degree of the element to the fuzzy set. In other words, we adopt an epistemic interpretation of the membership interval capturing incomplete information in respect to the actual degree of membership. Note that this interpretation is akin to that of a *confidence interval* and is in contrast to an alternative, *ontic* representation, where the interval itself represents the actual *interval-valued* membership degree [20]. Thus, as we adopt an epistemic interpretation of the membership interval, we assume that there is one actual, real-valued membership degree of an element inside the membership interval of possible membership degrees, and consequently two elements with the same interval membership need not necessarily have the same (unknown) actual real-valued membership degree.

To achieve objectives (a) and (b), we first introduce new definitions of interval-valued aggregation functions and interval-valued restricted equivalence functions, both in line with the observations stated above. It is worth pointing out that this is the first time in the literature that interval functions are studied according to these observations, and we assume that the approach can be utilized in a wide range of problems featuring intervals in the future.

Ideally, the definition of width-preserving interval-valued restricted equivalence functions (IVREF) would have to take into account the width of the inputs in every case. In other words, the width of the output interval should always be related by some axiom to the width of the input intervals, and not only when the input intervals have the same length. In this way, the epistemic interpretation would be fully preserved. However, it is not clear which this relation exactly should be and furthermore, how such IVREF could be constructed taking into account the complexity of the analysis and the construction when general admissible orders are involved. For this reason, the proposed set of axioms provides a first step in the desired direction.

To show the validity of our approach, we present an application using an expression of the proposed interval-valued similarity measure taking into account the width of the intervals which provides better results than other methods that can be found in the literature. In particular, we describe the application of our similarity in stereo image matching and

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

show that it outperforms the classical methods that make use of interval-valued fuzzy sets but do not take into account the width of the membership intervals.

The paper is organized as follows. We start with some preliminaries, then we study the concepts of interval-valued restricted equivalence functions and interval-valued aggregation functions preserving the widths of intervals. In Section 4, we introduce the definition of width-based interval-valued similarity measures and study different construction methods. In Section 5, we present an illustrative example of application of width-based interval-valued similarity measures in stereo image matching. We finish with some conclusions and references.

2. Preliminaries

In this section, we introduce several well known notions and results which are necessary for our subsequent developments.

We start recalling the idea of aggregation function. For more details, see [31].

Definition 2.1. An aggregation function is a non-decreasing function $M : [0, 1]^n \to [0, 1]$ with $M(0, \ldots, 0) = 0$ and $M(1, \ldots, 1) = 1$.

An aggregation function $M : [0, 1]^n \to [0, 1]$ is called idempotent if $M(x, \ldots, x) = x$ for every $x \in [0, 1]$, and it is called symmetric if $M(x_1, \ldots, x_n) = M(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$ for every $x_1, \ldots, x_n \in [0, 1]$ and every permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$.

Among the most relevant classes of aggregation functions we can mention the following.

Definition 2.2. A t-norm is a symmetric aggregation function $T : [0, 1]^2 \to [0, 1]$ such that $T(x, 1) = x$ and $T(T(x, y), z) = T(x, T(y, z))$ for every $x, y, z \in [0, 1]$.

Definition 2.3. A t-conorm is a symmetric aggregation function $S : [0, 1]^2 \to [0, 1]$ such that $S(x, 0) = x$ and $S(S(x, y), z) = S(x, S(y, z))$ for every $x, y, z \in [0, 1]$.

Among the most significant t-norms we can mention the minimum or the product, whereas among the most relevant t-conorms we can cite the maximum or the probabilistic sum $S_P(x, y) = x + y - xy$ [31].

In this work we are going to deal with closed subintervals of the unit interval. We denote by $L([0, 1])$ the set of closed subintervals of the unit interval, that is:

$$L([0, 1]) = \{[\underline{X}, \overline{X}] \mid 0 \leq \underline{X} \leq \overline{X} \leq 1\}.$$

We use capital letters to denote elements in $L([0, 1])$. The width of the interval $X \in L([0, 1])$ is denoted by $w(X)$, where $w(X) = \overline{X} - \underline{X}$. An interval function $f : (L([0, 1]))^n \to L([0, 1])$ is called width-preserving (or w-preserving, for simplicity) if, for any $X_1, \ldots, X_n \in L([0, 1])$ such that $w(X_1) = \ldots = w(X_n)$, it holds that $w(f(X_1, \ldots, X_n)) = w(X_1)$.

We work on a finite universe $U = \{u_1, \ldots, u_n\}$. An interval-valued fuzzy set (IVFS) on the universe U is a mapping $A : U \to L([0, 1])$. The class of all fuzzy sets in U is denoted by $FS(U)$ and the class of all interval-valued fuzzy sets in U by $IVFS(U)$. Given $A \in IVFS(U)$, its entropy is defined as

$$\epsilon(A) = \sum_{i=1}^{n} (\overline{A}(u_i) - \underline{A}(u_i)) .$$

Note that this entropy measures how far from fuzzy sets a given IVFS is, see [8] for more details. In this sense, although we have kept the original name of entropy which appears in [8], it is a non-specificity index for interval membership grades which differs from the usual fuzzy entropy related to crispness of fuzzy sets.

Another key notion in this work is that of order relation. Recall that an order relation on $L([0, 1])$ is a binary relation $\leq$ on $L([0, 1])$ which is reflexive, symmetric and transitive. An order relation on $L([0, 1])$ is called total or linear if any two elements of $L([0, 1])$ are comparable, i.e., if for every $X, Y \in L([0, 1])$, $X \leq Y$ or $Y \leq X$. An order relation on $L([0, 1])$ is partial if it is not total.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Although many different orders can be provided in $L([0,1])$, we are interested in the lattice extension of the ordering in $[0,1]$, that we will denote by $\preceq_L$ and which is the partial order given by:

$$[\underline{X},\overline{X}] \preceq_L [\underline{Y},\overline{Y}] \quad \text{if} \quad \underline{X} \leq \underline{Y} \text{ and } \overline{X} \leq \overline{Y}. \tag{1}$$

It is worth mentioning that with this order, any two degenerate intervals (i.e., intervals of width 0) are comparable. In particular, this implies that if the interval-valued fuzzy sets we are going to deal with are in fact fuzzy sets (that is, if all the membership intervals have width 0), the order between the interval-valued fuzzy sets is the same as the order between the fuzzy sets. In other words, if we use this order (or any extension of it), we can extend algorithms in the fuzzy setting which make use of the order in a straightforward way. This is the order relation most widely used in the literature [13].

We denote by $\leq_L$ any order on $L([0,1])$ (which can be partial or total) with $0_L = [0,0]$ as its minimal element (that is, $0_L \leq_L X$ for all $X \in L([0,1])$) and $1_L = [1,1]$ as its maximal element (that is, $X \leq_L 1_L$ for all $X \in L([0,1])$). To denote a total order on $L([0,1])$ with the same minimal and maximal elements, we use the notation $\leq_{TL}$.

Example 2.4.
(i) A total order on $L([0,1])$ is, for example, the Xu and Yager's order $\leq_{XY}$ (see [40]):

$$[\underline{X},\overline{X}] \leq_{XY} [\underline{Y},\overline{Y}] \quad \text{if} \quad \begin{cases} \underline{X}+\overline{X} < \underline{Y}+\overline{Y} \text{ or} \\ \underline{X}+\overline{X} = \underline{Y}+\overline{Y} \text{ and } \overline{X}-\underline{X} \leq \overline{Y}-\underline{Y}. \end{cases} \tag{2}$$

This definition of Xu and Yager's order was originally provided for Atanassov intuitionistic fuzzy pairs [40].
(ii) Another example of total order is provided by the lexicographical order with respect to the first variable, $\leq_{lex1}$, and with respect to the second variable, $\leq_{lex2}$, which are defined, respectively, by:

$$[\underline{X},\overline{X}] \leq_{lex1} [\underline{Y},\overline{Y}] \quad \text{if} \quad \begin{cases} \underline{X} < \underline{Y} \text{ or} \\ \underline{X} = \underline{Y} \text{ and } \overline{X} \leq \overline{Y}. \end{cases}$$

$$[\underline{X},\overline{X}] \leq_{lex2} [\underline{Y},\overline{Y}] \quad \text{if} \quad \begin{cases} \overline{X} < \overline{Y} \text{ or} \\ \overline{X} = \overline{Y} \text{ and } \underline{X} \leq \underline{Y}. \end{cases}$$

Regarding total orders on $L([0,1])$, we are going to consider the so-called admissible orders, whose definition we recall now.

Definition 2.5. [12] An admissible order on $L([0,1])$ is a total order $\leq_{TL}$ on $L([0,1])$ such that it refines the partial order $\preceq_L$, that is, for every $X, Y \in L([0,1])$, if $X \preceq_L Y$ then $X \leq_{TL} Y$.

An interesting feature of admissible orders is that they can be built using aggregation functions, as stated in the following result.

Proposition 2.6. *([12]) Let $M_1, M_2 : [0,1]^2 \to [0,1]$ be two aggregation functions such that for all $X, Y \in L([0,1])$, the equalities $M_1(\underline{X},\overline{X}) = M_1(\underline{Y},\overline{Y})$ and $M_2(\underline{X},\overline{X}) = M_2(\underline{Y},\overline{Y})$ can only hold simultaneously if $X = Y$. The order $\leq_{M_1,M_2}$ on $L([0,1])$ given by*

$$X \leq_{M_1,M_2} Y \quad \text{if} \quad \begin{cases} M_1(\underline{X},\overline{X}) < M_1(\underline{Y},\overline{Y}) \text{ or} \\ M_1(\underline{X},\overline{X}) = M_1(\underline{Y},\overline{Y}) \text{ and } M_2(\underline{X},\overline{X}) \leq M_2(\underline{Y},\overline{Y}) \end{cases}$$

is an admissible order on $L([0,1])$.

Example 2.7.
(i) Xu and Yager's order is an example of admissible order with $M_1(x,y) = \frac{x+y}{2}$ and $M_2(x,y) = y$.
(ii) The lexicographical orders $\leq_{lex1}$ ($\leq_{lex2}$) are also examples of admissible orders with $M_1(x,y) = x$ ($M_1(x,y) = y$) and $M_2(x,y) = y$ ($M_2(x,y) = x$).

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

(iii) More generally, if, for $\alpha \in [0, 1]$ we define the aggregation function

$$K_\alpha(x, y) = (1 - \alpha)x + \alpha y$$

then, for $\alpha, \beta \in [0, 1]$ with $\alpha \neq \beta$, we can obtain an admissible order $\preceq_{\alpha,\beta}$ just taking $M_1(x, y) = K_\alpha(x, y)$ and $M_2(x, y) = K_\beta(x, y)$. Observe that this operator K_α corresponds to Hurwicz's criterion [26] for balancing pessimism and optimism under uncertainty. See [12] for more details.

2.1. Interval-valued aggregation functions with respect to a partial order

The definition of aggregation function has been extended to the interval-valued setting with respect to the order $\preceq_L$ in a straightforward way [28].

Definition 2.8. Let $n \geq 2$. An (n-dimensional) interval-valued (IV) aggregation function in $L([0, 1])$ with respect to $\preceq_L$ is a mapping $M_{IV} : (L([0, 1]))^n \to L([0, 1])$ which verifies:

(i) $M_{IV}(0_L, \cdots, 0_L) = 0_L$.
(ii) $M_{IV}(1_L, \cdots, 1_L) = 1_L$.
(iii) M_{IV} is a non-decreasing function with respect to $\preceq_L$.

Remark 2.9. Note that this definition does not fully recover the usefulness of the usual definition of aggregation functions in the real setting (defined with respect to a total order) since there may exist intervals which are not comparable by means of the order $\preceq_L$, so the full meaning of monotonicity is lost.

It is quite easy to get IV aggregation functions in the sense of Definition 2.8, as the following examples show.

Example 2.10. ([33]) If $A : [0, 1]^2 \to [0, 1]$ is an aggregation function, then the function $M_A : L([0, 1])^2 \to L([0, 1])$ given by

$$M_A([\underline{X}, \overline{X}], [\underline{Y}, \overline{Y}]) = [A(\underline{X}, \underline{Y}), A(\overline{X}, \overline{Y})],$$

is an IV aggregation function in $L([0, 1])$ with respect to the order $\preceq_L$.

Moreover, if $A, B : [0, 1]^2 \to [0, 1]$ are two aggregation functions such that $A(x, y) \leq B(x, y)$ for each $x, y \in [0, 1]$, then

$$M_{A,B}([\underline{X}, \overline{X}], [\underline{Y}, \overline{Y}]) = [A(\underline{X}, \underline{Y}), B(\overline{X}, \overline{Y})],$$

is an IV aggregation function in $L([0, 1])$ with respect to the order $\preceq_L$.

Example 2.11. The following functions are IV aggregation functions in $L([0, 1])$ with respect to the order $\preceq_L$.

- $M_{IV}([\underline{X}, \overline{X}], [\underline{Y}, \overline{Y}]) = [(\underline{X}\underline{Y})^2, (\overline{X}\overline{Y})^2]$,
- $M_{IV}([\underline{X}, \overline{X}], [\underline{Y}, \overline{Y}]) = [\underline{X}\underline{Y}^{1/2}, (\overline{X} + \overline{Y})/2]$.

2.2. Restricted equivalence functions

Comparison measures between fuzzy sets can be built using the notion of restricted equivalence functions. We recall now this notion. For more details, see [10].

Definition 2.12. A function $R : [0, 1]^2 \to [0, 1]$ is called a restricted equivalence function (REF) if it satisfies:

1. $R(x, y) = 0$ if and only if $\{x, y\} = \{0, 1\}$ (i.e., if and only if $|x - y| = 1$);
2. $R(x, y) = 1$ if and only if $x = y$;
3. $R(x, y) = R(y, x)$ for all $x, y \in [0, 1]$;

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

4. If $x \leq y \leq z$, then $R(x, z) \leq R(x, y)$ and $R(x, z) \leq R(y, z)$ for all $x, y, z \in [0, 1]$.

Example 2.13. For any $p \in]0, \infty[$, the function $R^p(x, y) = 1 - |x - y|^p$ is a REF.

Remark 2.14. It is worth mentioning that, for some of our developments, we can consider the weaker condition

2'. $R(x, x) = 1$ for every $x \in [0, 1]$

instead of the stronger condition 2 in Definition 2.12. With this condition, we are recovering equivalence functions, see [22].

3. Width-preserving interval valued restricted equivalence functions

In this section, we propose a new definition of restricted equivalence functions (REF) in the interval-valued setting which takes into account the width of the inputs.

Definition 3.1. Let $\leq_L$ be an order on $L([0, 1])$. An interval-valued restricted equivalence function w.r.t. the order $\leq_L$ is a function $R_{IV} : L([0, 1])^2 \to L([0, 1])$ such that:

1. $R_{IV}(X, Y) = 0_L$ if and only if $\{X, Y\} = \{0_L, 1_L\}$;
2. $R_{IV}(X, X) = [1 - w(X), 1]$ for all $X \in L([0, 1])$;
3. $R_{IV}(X, Y) = R_{IV}(Y, X)$ for all $X, Y \in L([0, 1])$;
4. If $X, Y, Z \in L([0, 1])$ are such that $X \leq_L Y \leq_L Z$ and $w(X) = w(Y) = w(Z)$, then $R_{IV}(X, Z) \leq_L R_{IV}(X, Y)$ and $R_{IV}(X, Z) \leq_L R_{IV}(Y, Z)$.

Justification of the axioms.

1. Axiom 1. recovers the property required in the definition of REF in the real-valued setting [10].
2. The main difference with respect to the definition of REFs in the fuzzy setting arises in axiom 2. Note that we consider that the width of the membership interval of an element in a given set is a measure of the lack of knowledge of the precise (real-valued) membership degree of that element, and it is assumed that the exact membership value is a value inside the membership interval. Thus, if two elements have the same interval memberships, this does not mean that their corresponding/underlying real-valued degrees of membership are the same. Hence it is natural to expect that we can not get less uncertainty when comparing them.
3. Symmetry is a natural requirement also demanded in the real-valued setting.
4. Regarding axiom 4., observe that in the real-valued case, a total order (i.e., the usual order between real numbers) is used, and hence any two valued obtained by means of a REF can be compared. If we consider that $\leq_L$ is a total order, we are also able to compare any two intervals that are obtained as the result of an interval-valued restricted equivalence function w.r.t. the order $\leq_L$. Nevertheless, this axiom is more flexible, since also partial orders can be considered. By imposing the restriction that $w(X) = w(Y) = w(Z)$, we are recovering the condition demanded in the real case, since if X, Y, Z are intervals which consist of a single point, it follows that $w(X) = w(Y) = w(Z) = 0$.

Remark 3.2. As noted in the introduction, this definition does not take into account the width of the input intervals in every case, but we consider it to be a first step in this direction.

Example 3.3.

(i) A natural example of width preserving interval-valued restricted equivalence functions is the following:

$$R_{epis}(X, Y) = \{1 - |x - y| \mid x \in X, y \in Y\},$$

which exactly reflects the epistemic nature of X and Y, and is the interval extension of $1 - |x - y|$. Note that, even if this function fulfils the four axioms in Definition 2.12, it does not take into account the width of the inputs. Note that R_{epis} can also be expressed in terms of the endpoints of the intervals:

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

$$R_{epis}(X,Y) = \left[1 - \max\left(\overline{X} - \underline{Y}, \overline{Y} - \underline{X}\right), 1 - \max\left(0, \max\left(\underline{X}, \underline{Y}\right) - \min\left(\overline{X}, \overline{Y}\right)\right)\right].$$

Furthermore, observe that in this case, if we take two intervals of the same width, as, for instance, $X = [0.5, 0.7]$ and $Y = [0.6, 0.8]$, the result $R_{epis}(X,Y) = [0.7, 1]$ does not have the same width as the inputs.

(ii) However, it is possible to provide other examples of width preserving interval-valued restricted equivalence functions such that, if the inputs have the same width, the output also has that same width. For instance, the function $R_{IV} : L([0,1])^2 \to L([0,1])$ defined as:

$$R_{IV}(X,Y) = \Big[\max\Big(0, 1 - |K_\alpha(X) - K_\alpha(Y)| - \frac{1}{2}(w(X) + w(Y))\Big), \max\Big(1 - |K_\alpha(X) - K_\alpha(Y)|, \frac{1}{2}(w(X) + w(Y))\Big)\Big]$$

is, for every $\alpha \in]0,1]$, an example of IVREF w.r.t. any admissible order.

Note that in a sense this second example can be considered an approach similar to the R_{epis}. We are taking into account specific points in the interval by means of K_α operators (without the need for bounds, as is the case for R_{epis}). In particular, note that R_{epis} can be written as:

$$R_{epis}(X,Y) = \left[\min_{\alpha,\beta\in[0,1]}\left(1 - \left|K_\alpha(X) - K_\beta(Y)\right|\right), \max_{\alpha,\beta\in[0,1]}\left(1 - \left|K_\alpha(X) - K_\beta(Y)\right|\right)\right],$$

but we can replace it by the following equivalent expression which depends only on α:

$$R_{epis}(X,Y) = \left[\min_{\alpha\in[0,1]}\left(1 - |K_\alpha(X) - K_{1-\alpha}(Y)|\right), \max_{\alpha\in[0,1]}\left(1 - |K_\alpha(X) - K_{1-\alpha}(Y)|\right)\right].$$

Observe that R_{IV}, which depends on the parameter α, is based on a similar idea as R_{epis} (in the sense that they both generalize $1 - |x - y|$ to intervals) with the distinction that it is constructed directly in line with the intuition behind the notion of $K_{\alpha,\beta}$ order, i.e., the choice of α naturally depends on the considered order. Moreover, R_{IV} also takes into account the width of the inputs intervals whereas R_{epis} does not.

At the same time, if we take $\alpha = 0.5$ and the intervals $X = [0.3, 0.7]$ and $Y = [0.5, 0.7]$, we obtain IVREF equal to $[0.6, 0.9]$, - even though it is possible that the 'real-value' of both sets is exactly the same (0.7). On the other hand, if we use R_{epis}, we see that the IVREF is equal to $[0.4, 1]$, so we would recover the value 1.

Next we give a result regarding a monotonicity with respect to the widths of the inputs.

Proposition 3.4. *Let $X, Y \in L([0,1])$. If $w(X) < w(Y)$, then, for any admissible order $\leq_{TL}$, it follows that*

$$R_{IV}(Y,Y) \leq_{TL} R_{IV}(X,X).$$

Proof. It follows straightforwardly from Definition 3.1. □

Let us consider now the construction of examples of width-preserving IVREFs with respect to an admissible order. First of all, observe that if we consider and expression of the type $R_{IV}(X,Y) = \{REF(x,y) \mid x \in X \text{ and } y \in Y\}$, as the natural interval-valued extension of a real-valued REF function, R_{IV} is a width-preserving IVREF with respect to an admissible order only if the second axiom in Definition 3.1 holds, and this happens if and only if $REF(x,y) = 1 - |x - y|$. This means that in order to get general construction methods of examples that can be useful for applications, we need to find new construction methods. Furthermore, and in order to build width-preserving similarity functions, we are interested in IVREFs which preserve the width of the input intervals if they are the same.

Our first step is the following lemma, which recovers a feature of admissible orders.

Lemma 3.5. *Let $X, Y \in L([0,1])$ be intervals such that $w(X) = w(Y)$. Then*

$$X \precsim_L Y \text{ if and only if } X \leq_{TL} Y$$

for any admissible order $\leq_{TL}$.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Proof. The proof follows from the observation that intervals with the same width are always comparable by the partial order $\precsim_L$ and admissible orders refine the partial order $\precsim_L$. □

Now we discuss a procedure to build IVREFs which preserve the width of input intervals, and which is based on the use of K_α operators. Note that these operators can be viewed as choosing one representative point inside the intervals. Thus, in order to ensure that the axioms in the definition are fulfilled, we also must take into account the widths of the input intervals.

Theorem 3.6. *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function and let $R : [0,1]^2 \to [0,1]$ be a restricted equivalence function. Then, the function $R_{IV} : L([0,1])^2 \to L([0,1])$ given by*

$$R_{IV}(X,Y) = \left[\max\left(0, R\left(K_\alpha(X), K_\alpha(Y)\right) - M(w(X), w(Y))\right), \max\left(R\left(K_\alpha(X), K_\alpha(Y)\right), M(w(X), w(Y))\right)\right] \tag{3}$$

is an IV restricted equivalence function w.r.t. any admissible order $\leq_{TL}$. Moreover, R_{IV} is w-preserving.

Proof. For simplicity we write $\mathcal{R}$ instead of $R\left(K_\alpha(X), K_\alpha(Y)\right)$, and $\mathcal{M}$ instead of $M(w(X), w(Y))$. Then (3) can be simplified:

$$R_{IV}(X,Y) = \left[\max\left(0, \mathcal{R} - \mathcal{M}\right), \max\left(\mathcal{R}, \mathcal{M}\right)\right] = \begin{cases} [\mathcal{R} - \mathcal{M}, \mathcal{R}], & \text{if } \mathcal{R} \geq \mathcal{M}, \\ [0, \mathcal{M}], & \text{otherwise.} \end{cases} \tag{4}$$

By (4) it is clear that R_{IV} is well-defined.

Observe that $R_{IV}(X,Y) = 0_L$ if and only if $\mathcal{R} = 0$ and $\mathcal{M} = 0$. The former holds if and only if $\{K_\alpha(X), K_\alpha(Y)\} = \{0,1\}$, which may happen if and only if $\{X,Y\} = \{0_L, 1_L\}$. So it follows that $w(X) = w(Y) = 0$ and we get the first condition in Definition 3.1.

The second condition in Definition 3.1 follows from the observations: $R(K_\alpha(X), K_\alpha(X)) = 1$ and $M(w(X), w(X)) = w(X)$.

Symmetry of R_{IV} directly follows from the symmetry of R and M.

The fulfilment of the fourth condition in Definition 3.1 w.r.t. any admissible order follows from the monotonicity of R, Lemma 3.5, after observing that, if $X \leq_{TL} Y \leq_{TL} Z$ and $w(X) = w(Y) = w(Z)$, then $K_\alpha(X) \leq K_\alpha(Y) \leq K_\alpha(Z)$.

Finally, the fact that R_{IV} is w-preserving directly follows from Equation (4) and idempotency of M. □

We can use any REF R and any idempotent symmetric aggregation function M in Equation (3) to build an IVREF. Equation (3) can be simplified if some additional assumptions on R and M are imposed.

Corollary 3.7. *Let $\alpha \in]0,1[$, let $M : [0,1]^2 \to [0,1]$ be an idempotent symmetric aggregation function such that $M(x,y) \leq \min\left((1-\alpha)x + \alpha y, \alpha x + (1-\alpha)y\right)$ for all $x, y \in [0,1]$, and let $R : [0,1]^2 \to [0,1]$ be a restricted equivalence function such that $R(x,y) \geq 1 - |x-y|$ for all $x, y \in [0,1]$. Then, the function $R_{IV} : L([0,1])^2 \to L([0,1])$ given by*

$$R_{IV}(X,Y) = \left[R\left(K_\alpha(X), K_\alpha(Y)\right) - M(w(X), w(Y)), R\left(K_\alpha(X), K_\alpha(Y)\right)\right] \tag{5}$$

is an IV restricted equivalence function w.r.t. any admissible order $\leq_{TL}$. Moreover, R_{IV} is w-preserving.

Proof. We only need to prove that $R\left(K_\alpha(X), K_\alpha(Y)\right) \geq M(w(X), w(Y))$ for all $X, Y \in L([0,1])$, since in that case Equation (5) is a special case of Equation (3). Due to the assumptions on M and R, it is enough to show that

$$1 - |K_\alpha(X) - K_\alpha(Y)| \geq \min\left((1-\alpha)w(X) + \alpha w(Y), \alpha w(X) + (1-\alpha)w(Y)\right). \tag{6}$$

Assume that $K_\alpha(X) \geq K_\alpha(Y)$. Then

$$1 - |K_\alpha(X) - K_\alpha(Y)| = 1 - (1-\alpha)\underline{X} - \alpha\overline{X} + (1-\alpha)\underline{Y} + \alpha\overline{Y}$$

and since

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

$$1 \geq \overline{X} - \underline{Y} = (1-\alpha)(\overline{X} - \underline{Y}) + \alpha(\overline{X} - \underline{Y}) = (1-\alpha)(\overline{X} - \underline{X} + \underline{X} - \underline{Y}) + \alpha(\overline{X} - \overline{Y} + \overline{Y} - \underline{Y}),$$

we have

$$1 - (1-\alpha)\underline{X} - \alpha\overline{X} + (1-\alpha)\underline{Y} + \alpha\overline{Y} \geq (1-\alpha)(\overline{X} - \underline{X}) + \alpha(\overline{Y} - \underline{Y}),$$

hence (6) is satisfied.

Now assume that $K_\alpha(X) < K_\alpha(Y)$. Then

$$1 - |K_\alpha(X) - K_\alpha(Y)| = 1 + (1-\alpha)\underline{X} + \alpha\overline{X} - (1-\alpha)\underline{Y} - \alpha\overline{Y}$$

and since

$$1 \geq \overline{Y} - \underline{X} = (1-\alpha)(\overline{Y} - \underline{X}) + \alpha(\overline{Y} - \underline{X}) = (1-\alpha)(\overline{Y} - \underline{Y} + \underline{Y} - \underline{X}) + \alpha(\overline{Y} - \overline{X} + \overline{X} - \underline{X}),$$

we have

$$1 + (1-\alpha)\underline{X} + \alpha\overline{X} - (1-\alpha)\underline{Y} - \alpha\overline{Y} \geq (1-\alpha)(\overline{Y} - \underline{Y}) + \alpha(\overline{X} - \underline{X}),$$

hence (6) is satisfied. □

Remark 3.8. Note that in the previous Corollary we are considering REF functions R which are greater than or equal to $1 - |x - y|$, so in general, their interval extensions

$$R_{IV}(X, Y) = \{R(x, y) \mid x \in X \text{ and } y \in Y\}$$

need not satisfy the axioms about width in Definition 3.1. However, we can make use of them to build new examples of width-preserving IVREFs, see also Example 3.11 below.

As a consequence, we show in the next Corollary that a width-preserving IVREF built as in Corollary 3.7 does not in fact decrease the width of the input intervals, even if the latter are different to each other, and the width of the resulting interval depends parametrically on α. Furthermore, if $M = \max$, then the width of the output interval equals the maximum of the widths of the input intervals. Note that in the case of R_{epis} this is not the case as, for instance, $R_{epis}([0.1, 0.9], [0.3, 0.5]) = [0.4, 1]$.

Corollary 3.9. *Consider the interval-valued restricted equivalence function R_{IV} constructed in Corollary 3.7. Then, for all $X, Y \in L([0, 1])$, it holds that*

$$\min(w(X), w(Y)) \leq w(R_{IV}(X, Y)) = M(w(X), w(Y)) \leq \min\big((1-\alpha)w(X) + \alpha w(Y), \alpha w(X) + (1-\alpha)w(Y)\big).$$

Proof. The first inequality follows from the fact that an idempotent aggregation function is always greater than or equal to the minimum and the second inequality follows from the property of M assumed in Corollary 3.7. □

We can also prove the following result.

Lemma 3.10. *If $M_1, M_2 : [0, 1]^2 \to [0, 1]$ are idempotent symmetric aggregation functions, then the function $M : [0, 1]^2 \to [0, 1]$ given by*

$$M(x, y) = \min(M_1(x, y), M_2(x, y)),$$

for all $x, y \in [0, 1]$, is an idempotent symmetric aggregation function.

Proof. The proof is straightforward. □

Example 3.11. Let us consider the construction of IVREF given by Corollary 3.7. Observe that the REF R^p defined in Example 2.13 satisfies assumption of the corollary: $R^p(x, y) \geq R^1 = 1 - |x - y|$ for all $x, y \in [0, 1]$ if and only if $p \in [1, \infty[$.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

(i) If we take $\alpha = 1/2$ and $M(x, y) = \frac{x+y}{2}$ we get a class of IVREFs w.r.t. any admissible order:

$$R_{IV}^{p}(X, Y) = \left[R^p \left(\frac{\underline{X} + \overline{X}}{2}, \frac{\underline{Y} + \overline{Y}}{2} \right) - \frac{w(X) + w(Y)}{2}, R^p \left(\frac{\underline{X} + \overline{X}}{2}, \frac{\underline{Y} + \overline{Y}}{2} \right) \right]$$

for $p \in [1, \infty[$.

(ii) If we take $M(x, y) = \min(x, y)$, a class of IVREFs w.r.t. any admissible order arises:

$$R_{IV}^{p,\alpha}(X, Y) = \left[R^p \left(K_\alpha(X), K_\alpha(Y) \right) - \min(w(X), w(Y)), R^p \left(K_\alpha(X), K_\alpha(Y) \right) \right]$$

for $p \in [1, \infty[$ and $\alpha \in]0, 1[$.

(iii) Let $\alpha \in [0, 1]$, it is easy to see that

$$\min\left((1-\beta)x + \beta y, \beta x + (1-\beta)y\right) \leq \min\left((1-\alpha)x + \alpha y, \alpha x + (1-\alpha)y\right)$$

for all $\beta \in [0, 1]$ such that $\max(\beta, 1-\beta) \geq \max(\alpha, 1-\alpha)$. Hence, we get a more general class of IVREFs w.r.t. any admissible order than that in item (ii), if we take $\alpha \in]0, 1[$ and

$$M(x, y) = \min\left((1-\beta)x + \beta y, \beta x + (1-\beta)y\right)$$

for $\beta \in [\max(\alpha, 1-\alpha), 1]$ (or equivalently for $\beta \in [0, \min(\alpha, 1-\alpha)]$):

$$R_{IV}^{p,\alpha,\beta}(X, Y) =$$
$$= \left[R^p \left(K_\alpha(X), K_\alpha(Y) \right) - \min\left((1-\beta)w(X) + \beta w(Y), \beta w(X) + (1-\beta)w(Y)\right), R^p \left(K_\alpha(X), K_\alpha(Y) \right) \right].$$

Note that for $\beta = 1$ (or equivalently for $\beta = 0$), we get the class described in item (ii).

Finally, Theorem 3.6 also allows us to provide conditions to have $R_{IV}(X, Y) = 1_L$.

Lemma 3.12. *Let $R_{IV} : L([0, 1])^2 \to L([0, 1])$ be defined as in Theorem 3.6. If $M(x, y) = 0$ if and only if $x = y = 0$, then $R_{IV}(X, Y) = 1_L$ if and only if $X = Y$ and $w(X) = 0$.*

Proof. The proof is straightforward. □

3.1. Width-preserving IV aggregation functions

In order to build width-preserving similarity measures, one possibility is to aggregate width-preserving IVREFs in a suitable way. Specifically, it is desirable that the considered aggregation function takes into account the width of the input intervals.

As a first step towards this, we recall the definition of aggregation function in the interval-valued setting.

Definition 3.13. Let $n \geq 2$. An (n-dimensional) interval-valued (IV) aggregation function in $L([0, 1])$ with respect to $\leq_L$ is a mapping $M_{IV} : (L([0, 1]))^n \to L([0, 1])$ which verifies:

(i) $M_{IV}(0_L, \cdots, 0_L) = 0_L$.
(ii) $M_{IV}(1_L, \cdots, 1_L) = 1_L$.
(iii) M_{IV} is a non-decreasing function with respect to $\leq_L$.

We say that $M_{IV} : (L([0, 1]))^n \to L([0, 1])$ is a decomposable n-dimensional IV aggregation function associated with M_L and M_U, if n-dimensional aggregation functions $M_L, M_U : [0, 1]^n \to [0, 1]$ exist such that $M_L \leq M_U$ pointwise, and where

$$M_{IV}(X_1, \ldots, X_n) = \left[M_L\left(\underline{X_1}, \ldots, \underline{X_n}\right), M_U\left(\overline{X_1}, \ldots, \overline{X_n}\right) \right] \tag{7}$$

for all $X_1, \ldots, X_n \in L([0, 1])$.

The construction of IV aggregation functions with respect to admissible orders is not a trivial task, see [1]. We devote the next results to such a construction.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Definition 3.14 *([1]).* Let $c \in [0,1]$ and $\alpha \in [0,1]$. We denote by $d_\alpha(c)$ the maximal possible width of an interval $Z \in L([0,1])$ such that $K_\alpha(Z) = c$. Moreover, for any $X \in L([0,1])$, let

$$\lambda_\alpha(X) = \frac{w(X)}{d_\alpha(K_\alpha(X))}$$

where we set $\frac{0}{0} = 1$.

Proposition 3.15 *([1]). For all $\alpha \in [0,1]$ and $X \in L([0,1])$ it holds that*

$$d_\alpha(K_\alpha(X)) = \wedge\left(\frac{K_\alpha(X)}{\alpha}, \frac{1-K_\alpha(X)}{1-\alpha}\right),$$

where we set $\frac{r}{0} = 1$ for all $r \in [0,1]$.

A construction method of IV aggregation functions w.r.t. $\leq_{\alpha,\beta}$ is proposed in the following theorem, which makes use of aggregation functions.

Theorem 3.16. *Let $\alpha, \beta \in [0,1]$, $\beta \neq \alpha$. Let $M_1, M_2 : [0,1]^n \to [0,1]$ be aggregation functions where M_1 is strictly increasing. Then $M_{IV} : (L([0,1]))^n \to L([0,1])$ defined by:*

$$M_{IV}(X_1, \ldots, X_n) = Y, \quad \text{where} \quad \begin{cases} K_\alpha(Y) = M_1(K_\alpha(X_1), \ldots, K_\alpha(X_n)), \\ \lambda_\alpha(Y) = M_2(\lambda_\alpha(X_1), \ldots, \lambda_\alpha(X_2)), \end{cases}$$

for all $X_1, \ldots, X_n \in L([0,1])$, is an IV aggregation function with respect to $\leq_{\alpha,\beta}$.

Proof. First observe that $w(Y) = \lambda_\alpha(Y) d_\alpha(K_\alpha(Y))$ and $\underline{Y} = K_\alpha(Y) - \alpha w(Y)$, $\overline{Y} = K_\alpha(Y) + (1-\alpha)w(Y)$. Clearly, M_{IV} is well defined. It suffices to show that M_{IV} is an IV aggregation function.

(i) $M_{IV}(0_L, \cdots, 0_L) = Y$ where $K_\alpha(Y) = M_1(0, \ldots, 0) = 0$. Moreover, if $\alpha \neq 0$, then $\lambda_\alpha(Y) = M_2(1, \ldots, 1) = 1$ and $w(Y) = \lambda_\alpha(Y) d_\alpha(K_\alpha(Y)) = 1 \cdot 0 = 0$. If $\alpha = 0$, then $\lambda_\alpha(Y) = M_2(0, \ldots, 0) = 0$ and $w(Y) = \lambda_\alpha(Y) d_\alpha(K_\alpha(Y)) = 0 \cdot 1 = 0$. Hence, $Y = [0,0]$.

(ii) $M_{IV}(1_L, \cdots, 1_L) = Y$ where $K_\alpha(Y) = M_1(1, \ldots, 1) = 1$. Moreover, if $\alpha \neq 1$, then $\lambda_\alpha(Y) = M_2(1, \ldots, 1) = 1$ and $w(Y) = \lambda_\alpha(Y) d_\alpha(K_\alpha(Y)) = 1 \cdot 0 = 0$. If $\alpha = 1$, then $\lambda_\alpha(Y) = M_2(0, \ldots, 0) = 0$ and $w(Y) = \lambda_\alpha(Y) d_\alpha(K_\alpha(Y)) = 0 \cdot 1 = 0$. Hence, $Y = [1,1]$.

(iii) Let $X_i \leq_{\alpha,\beta} Y_i$ for all $i = 1, \ldots, n$. Then $K_\alpha(X_i) \leq K_\alpha(Y_i)$ for all $i = 1, \ldots, n$ and there are two cases:

1. There exists $j \in \{1, \ldots, n\}$ such that $K_\alpha(X_j) < K_\alpha(Y_j)$. Since M_1 is strictly increasing, it follows that

$$M_1(K_\alpha(X_1), \ldots, K_\alpha(X_n)) < M_1(K_\alpha(Y_1), \ldots, K_\alpha(Y_n)),$$

thus $M_{IV}(X_1, \ldots, X_n) <_{\alpha,\beta} M_{IV}(Y_1, \ldots, Y_n)$.

2. $K_\alpha(X_i) = K_\alpha(Y_i)$ for all $i = 1, \ldots, n$. If $\beta > \alpha$, then $w(X_i) \leq w(Y_i)$ for all $i = 1, \ldots, n$, hence $\lambda_\alpha(X_i) \leq \lambda_\alpha(Y_i)$ for all $i = 1, \ldots, n$, thus $M_2(\lambda_\alpha(X_1), \ldots, \lambda_\alpha(X_n)) \leq M_2(\lambda_\alpha(Y_1), \ldots, \lambda_\alpha(Y_n))$, consequently $M_{IV}(X_1, \ldots, X_n) \leq_{\alpha,\beta} M_{IV}(Y_1, \ldots, Y_n)$. If $\beta < \alpha$, then $w(X_i) \geq w(Y_i)$ for all $i = 1, \ldots, n$, hence $\lambda_\alpha(X_i) \geq \lambda_\alpha(Y_i)$ for all $i = 1, \ldots, n$, thus $M_2(\lambda_\alpha(X_1), \ldots, \lambda_\alpha(X_n)) \geq M_2(\lambda_\alpha(Y_1), \ldots, \lambda_\alpha(Y_n))$. As a consequence, $M_{IV}(X_1, \ldots, X_n) \leq_{\alpha,\beta} M_{IV}(Y_1, \ldots, Y_n)$. □

The following construction methods provide IV aggregation functions w.r.t. $\leq_{\alpha,\beta}$ which preserve the width of the input intervals. First of all, given an aggregation function $M : [0,1]^n \to [0,1]$, the following two properties are considered:

(**P1**) $M(cx_1, \ldots, cx_n) \geq cM(x_1, \ldots, x_n)$ for all $c \in [0,1]$, $x_1, \ldots, x_n \in [0,1]$.
(**P2**) $M(x_1, \ldots, x_n) \leq 1 - M(1-x_1, \ldots, 1-x_n)$ for all $x_1, \ldots, x_n \in [0,1]$.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Theorem 3.17. *Let $\alpha, \beta \in [0,1]$, $\beta \neq \alpha$. Let $M_1, M_2 : [0,1]^n \to [0,1]$ be aggregation functions such that M_1 is strictly increasing, $M_1(x_1,\dots,x_n) \geq M_2(x_1,\dots,x_n)$ for all $x_1,\dots,x_n \in [0,1]$, M_1 or M_2 satisfies property (P1) and M_1 or M_2 satisfies property (P2). Then $M_{IV} : (L([0,1]))^n \to L([0,1])$ defined by:*

$$M_{IV}(X_1,\dots,X_n) = Y, \quad \text{where} \quad \begin{cases} K_\alpha(Y) = M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right), \\ w(Y) = M_2\left(w(X_1),\dots,w(X_n)\right), \end{cases}$$

for all $X_1,\dots,X_n \in L([0,1])$, is an IV aggregation function with respect to $\leq_{\alpha,\beta}$.

Moreover, if M_2 is idempotent, then M_{IV} is w-preserving.

Proof. We first show that M_{IV} is well defined. Observe that

$$Y = \left[\underline{Y}, \overline{Y}\right] = \left[K_\alpha(Y) - \alpha w(Y), K_\alpha(Y) + (1-\alpha)w(Y)\right].$$

Clearly, $\underline{Y} \leq \overline{Y}$, hence we only need to prove that

1. $\underline{Y} \geq 0$: For $\alpha = 0$ we have $\underline{Y} = M_1(\underline{X_1},\dots,\underline{X_n}) \geq 0$ and for $\alpha \in]0,1]$ we have

$$K_\alpha(Y) = M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) \geq \alpha M_2\left(\frac{K_\alpha(X_1)}{\alpha},\dots,\frac{K_\alpha(X_n)}{\alpha}\right) \geq \alpha M_2\left(w(X_1),\dots,w(X_n)\right) = \alpha w(Y)$$

   where the first inequality follows from the fact that M_2 satisfies property (P1) and the second from the observation $K_\alpha(X) = (1-\alpha)\underline{X} + \alpha\overline{X} \geq \alpha(\overline{X} - \underline{X}) = \alpha w(X)$ for all $X \in L([0,1])$.
2. $\overline{Y} \leq 1$: For $\alpha = 1$ we have $\overline{Y} = M_1(\overline{X_1},\dots,\overline{X_n}) \leq 1$ and for $\alpha \in [0,1[$ we have

$$\begin{aligned} &K_\alpha(Y) + (1-\alpha)w(Y) = M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) + (1-\alpha)M_2\left(w(X_1),\dots,w(X_n)\right) \leq \\ &\leq M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) + (1-\alpha)M_2\left(\frac{1-K_\alpha(X_1)}{1-\alpha},\dots,\frac{1-K_\alpha(X_n)}{1-\alpha}\right) \leq \\ &\leq M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) + M_2\left(1-K_\alpha(X_1),\dots,1-K_\alpha(X_n)\right) \leq \\ &\leq M_1\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) + 1 - M_2\left(K_\alpha(X_1),\dots,K_\alpha(X_n)\right) = 1 \end{aligned}$$

   where the first inequality follows from the observation $1 - K_\alpha(X) = 1 - (1-\alpha)\underline{X} - \alpha\overline{X} \geq (1-\alpha)(\overline{X} - \underline{X}) = (1-\alpha)w(X)$ for all $X \in L([0,1])$, and the second and third ones from the assumptions of the theorem.

Now we prove that M_{IV} is an IV aggregation function. (i) $M_{IV}(0_L,\cdots,0_L) = Y$ where $K_\alpha(Y) = M_1(0,\dots,0) = 0$ and $w(Y) = M_2(0,\dots,0) = 0$, hence $Y = 0_L$. (ii) $M_{IV}(1_L,\cdots,1_L) = Y$ where $K_\alpha(Y) = M_1(1,\dots,1) = 1$ and $w(Y) = M_2(0,\dots,0) = 0$, hence $Y = 1_L$. (iii) Let $X_i \leq_{\alpha,\beta} Y_i$ for all $i = 1,\dots,n$. Then $K_\alpha(X_i) \leq K_\alpha(Y_i)$ for all $i = 1,\dots,n$ and there are two cases:

1. There exists $j \in \{1,\dots,n\}$ such that $K_\alpha(X_j) < K_\alpha(Y_j)$. Then

$$M_1(K_\alpha(X_1),\dots,K_\alpha(X_n)) < M_1(K_\alpha(Y_1),\dots,K_\alpha(Y_n)),$$

   since M_1 is strictly increasing, thus $M_{IV}(X_1,\dots,X_n) <_{\alpha,\beta} M_{IV}(Y_1,\dots,Y_n)$.
2. $K_\alpha(X_i) = K_\alpha(Y_i)$ for all $i = 1,\dots,n$. If $\beta > \alpha$, then $w(X_i) \leq w(Y_i)$ for all $i = 1,\dots,n$, hence $M_2(w(X_1),\dots,w(X_n)) \leq M_2(w(Y_1),\dots,w(Y_n))$, consequently $M_{IV}(X_1,\dots,X_n) \leq_{\alpha,\beta} M_{IV}(Y_1,\dots,Y_n)$. If $\beta < \alpha$, then $w(X_i) \geq w(Y_i)$ for all $i = 1,\dots,n$, thus $M_2(w(X_1),\dots,w(X_n)) \geq M_2(w(Y_1),\dots,w(Y_n))$, consequently $M_{IV}(X_1,\dots,X_n) \leq_{\alpha,\beta} M_{IV}(Y_1,\dots,Y_n)$.

   Finally, it is easy to check that M_{IV} is w-preserving from the idempotency of M_2. □

To clarify our results, we consider now the case where the aggregation functions that we use to build the width-preserving IV aggregation function are given by a weighted arithmetic mean. This result will be relevant in the next section and for the application.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Proposition 3.18. *Let $(v_1, \ldots, v_n) \in]0,1]^n$ be a weighting vector with $v_1 + \ldots + v_n = 1$. Under the assumptions of Theorem 3.17, if $M_1(x_1, \ldots, x_n) = M_2(x_1, \ldots, x_n) = v_1x_1 + \ldots + v_nx_n$ for all $x_1, \ldots, x_n \in [0,1]$, then M_{IV} is the decomposable IV aggregation function associated with M_L and M_U where $M_L = M_U = M_1$.*

Proof. Let $X_1, \ldots, X_n \in L([0,1])$ and $M_{IV}(X_1, \ldots, X_n) = Y$. According to Theorem 3.17 we have

$$w(Y) = \sum_{i=1}^{n} v_i w(X_i) \qquad \text{and} \qquad K_\alpha(Y) = \sum_{i=1}^{n} v_i K_\alpha(X_i).$$

Taking $M_L = M_U = M_1$ we obtain

$$M_U\left(\overline{X_1}, \ldots, \overline{X_n}\right) - M_L\left(\underline{X_1}, \ldots, \underline{X_n}\right) = \sum_{i=1}^{n} v_i \overline{X_i} - \sum_{i=1}^{n} v_i \underline{X_i} = \sum_{i=1}^{n} v_i w(X_i) = w(Y)$$

and

$$(1-\alpha)M_L\left(\underline{X_1}, \ldots, \underline{X_n}\right) + \alpha M_U\left(\overline{X_1}, \ldots, \overline{X_n}\right) = (1-\alpha)\sum_{i=1}^{n} v_i \underline{X_i} + \alpha \sum_{i=1}^{n} v_i \overline{X_i} = \sum_{i=1}^{n} v_i K_\alpha(X_i) = K_\alpha(Y).$$

Hence, according to Equation (7), M_{IV} is decomposable and associated with M_L and M_U. □

Finally, we present some properties that hold for the functions $M_{IV} : (L([0,1]))^n \to L([0,1])$ defined as in Theorem 3.17.

Lemma 3.19. *Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be defined as in Theorem 3.17.*

(i) *If*
- *$M_1(x_1, \ldots, x_n) = 0$ if and only if $x_1 = \ldots = x_n = 0$ and*
- *$M_2(x_1, \ldots, x_n) = 0$ if and only if $x_1 = \ldots = x_n = 0$,*

then $M_{IV}(X_1, \ldots, X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Moreover, if $\alpha \neq 0$, then the restriction on M_2 can be skipped.

(ii) *If*
- *$M_1(x_1, \ldots, x_n) = 1$ if and only if $x_1 = \ldots = x_n = 1$ and*
- *$M_2(x_1, \ldots, x_n) = 0$ if and only if $x_1 = \ldots = x_n = 0$,*

then $M_{IV}(X_1, \ldots, X_n) = 1_L$ if and only if $X_1 = \ldots = X_n = 1_L$. Moreover, if $\alpha \neq 1$, then the restriction on M_2 can be skipped.

(iii) *M_{IV} is idempotent if and only if M_1 and M_2 are idempotent.*

Proof. The proof is straightforward. □

Example 3.20. A function $M_{IV} : (L([0,1]))^n \to L([0,1])$ defined as in Theorem 3.17, is a w-preserving IV aggregation function with respect to $\leq_{\alpha,\beta}$, if, for instance:

1. (i) $M_1(x_1, \ldots, x_n) = M_2(x_1, \ldots, x_n) = \frac{x_1+\ldots+x_n}{n}$ for all $x_1, \ldots, x_n \in [0,1]$, or
2. (ii) $M_1(x_1, \ldots, x_n) = \frac{x_1+\ldots+x_n}{n}$, $M_2(x_1, \ldots, x_n) = \min\{x_1, \ldots, x_n\}$ for all $x_1, \ldots, x_n \in [0,1]$.

4. Width-based interval valued similarity measures

The same arguments that we have used to justify the introduction of IV restricted equivalence functions which take into account the width of the intervals are valid for the case of IV similarity measures. For this reason we propose the following definition.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Definition 4.1. Let $\leq_L$ be an order on $L([0,1])$ and $M:[0,1]^n \to [0,1]$ be an aggregation function. A width-based interval-valued similarity measure on $IVFS(U)$ w.r.t. $\leq_L$, associated with M is a mapping $S_M : IVFS(U) \times IVFS(U) \to L([0,1])$ such that, for all $A, B, A', B' \in IVFS(U)$,

(SM1) $S_M(A,B) = S_M(B,A)$;
(SM2) $S_M(A,A) = \left[1 - M\left(w(A(u_1)), \ldots, w(A(u_n))\right), 1\right]$;
(SM3) $S_M(A,B) = 0_L$ if and only if $\{A(u_i), B(u_i)\} = \{0_L, 1_L\}$ for all $i \in \{1,\ldots,n\}$;
(SM4) If $A \subseteq A' \subseteq B' \subseteq B$ w.r.t. $\leq_L$ and $w(A(u_i)) = w(A'(u_i)) = w(B'(u_i)) = w(B(u_i))$ for all $i \in \{1,\ldots,n\}$, then $S_M(A,B) \leq_L S_M(A',B')$, where, for $A, B \in IVFS(U)$, $A \subseteq B$ w.r.t. $\leq_L$ if $A(u_i) \leq_L B(u_i)$ for every $u_i \in U$.

The definition is motivated by that given in [10]. However, the second axiom is changed in line with Definition 3.1 and the fourth axiom is relaxed in a similar way as the fourth axiom in Definition 3.1.

Now, a construction method of IV similarity measure by aggregation of IVREFs is given.

Recall that an aggregation function $M:[0,1]^n \to [0,1]$ is called self-dual with respect to the standard negation if

$$M(x_1,\ldots,x_n) = 1 - M(1-x_1,\ldots,1-x_n)$$

for all $x_1,\ldots,x_n \in [0,1]$.

Theorem 4.2. *Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be a decomposable IV aggregation function w.r.t. $\leq_L$ associated with M_L and M_U where M_L is self-dual, and let $M_{IV}(X_1,\ldots,X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Let $R_{IV} : L([0,1])^2 \to L([0,1])$ be an IV restricted equivalence function w.r.t. $\leq_L$. Then the function $S_{M_L} : IVFS(U) \times IVFS(U) \to L([0,1])$ defined by:*

$$S_{M_L}(A,B) = M_{IV}\left(R_{IV}\left(A(u_1), B(u_1)\right), \ldots, R_{IV}\left(A(u_n), B(u_n)\right)\right)$$

for all $A, B \in IVFS(U)$ is a width-based IV similarity measure on $IVFS(U)$ w.r.t. $\leq_L$ associated with M_L.

Proof. (SM1) Directly follows from Axiom 3 of Definition 3.1.

(SM2) Observe that

$$\begin{aligned} S_{M_L}(A,A) &= M_{IV}\left(R_{IV}\left(A(u_1), A(u_1)\right), \ldots, R_{IV}\left(A(u_n), A(u_n)\right)\right) = \\ &= M_{IV}\left([1 - w(A(u_1)), 1], \ldots, [1 - w(A(u_n)), 1]\right) = \left[M_L\left(1 - w(A(u_1)), \ldots, 1 - w(A(u_n))\right), 1\right] = \\ &= \left[1 - M_L\left(w(A(u_1)), \ldots, w(A(u_n))\right), 1\right], \end{aligned}$$

where the second equality follows from Axiom 2 of Definition 3.1, the third from the fact that M_{IV} is decomposable associated with M_L, M_U and the last from the self-duality of M_L.

(SM3) Since $S_{M_L}(A,B) = 0_L$ if and only if $R_{IV}\left(A(u_1), B(u_1)\right) = 0_L$ for all $i \in \{1,\ldots,n\}$, which only holds if $\{A(u_i), B(u_i)\} = \{0_L, 1_L\}$ for all $i \in \{1,\ldots,n\}$, (SM3) is satisfied.

(SM4) Follows from Axiom 4 of Definition 3.1 and the monotonicity of M_{IV}. □

Example 4.3. Let us consider the IVREF R_{epis} given in Example 3.3 (i), $M_L(x_1,\ldots,x_n) = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $M_U(x_1,\ldots,x_n) = \max(x_1,\ldots,x_n)$. Then the assumptions of Theorem 4.2 are satisfied and we obtain the following width-based IV similarity measure on $IVFS(U)$:

$$S(A,B) = \left[\frac{1}{n}\sum_{i=1}^{n}\left(1 - \max\left(\overline{A(u_i)} - \underline{B(u_i)}, \overline{B(u_i)} - \underline{A(u_i)}\right)\right),\right.$$
$$\left.\max\left(1 - \max\left(0, \max\left(\underline{A(u_i)}, \underline{B(u_i)}\right) - \min\left(\overline{A(u_i)}, \overline{B(u_i)}\right)\right)\right)\right].$$

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

We now study the conditions under which the IVREF R_{IV} given by Theorem 3.6 and the IV aggregation function M_{IV} given by Theorem 3.17 can be applied in the previous theorem to obtain an IV similarity measure that preserves the width of intervals.

Corollary 4.4. *Let $\alpha, \beta \in]0,1[$ where $\beta \neq \alpha$. Let $(v_1, \ldots, v_n) \in]0,1]^n$ be a weighting vector such that $v_1 + \ldots + v_n = 1$ and let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be the IV aggregation function w.r.t. $\leq_{\alpha,\beta}$ defined as in Theorem 3.17 where $M_1(x_1, \ldots, x_n) = M_2(x_1, \ldots, x_n) = v_1x_1 + \ldots + v_nx_n$ for all $x_1, \ldots, x_n \in [0,1]$. Let $R_{IV} : (L([0,1]))^2 \to L([0,1])$ be an IVREF defined as in Theorem 3.6. Then the function $S_M : IVFS(U) \times IVFS(U) \to L([0,1])$ defined by:*

$$S_M(A,B) = M_{IV}\big(R_{IV}(A(u_1), B(u_1)), \ldots, R_{IV}(A(u_n), B(u_n))\big),$$

for all $A, B \in IVFS(U)$, is a width-based IV similarity measure on $IVFS(U)$ w.r.t. $\leq_L$ associated with M_1. Moreover, S_M satisfies the following for all $A, B \in IVFS(U)$:

$$w(S_M(A,B)) = w(A(u_1)) \quad \text{whenever} \quad w(A(u_1)) = w(B(u_1)) = \ldots = w(A(u_n)) = w(B(u_n)). \tag{8}$$

Proof. Observe that, by Proposition 3.18, M_{IV} is the decomposable IV aggregation function associated with M_L, M_U where $M_L = M_U = M_1$. Moreover, by Lemma 3.19 we have that $M_{IV}(X_1, \ldots, X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Since a weighted arithmetic mean is self-dual and idempotent, from Theorem 4.2 it follows that S_M is a width-based IV similarity measure associated with M_1. Finally, since M_{IV} and R_{IV} are w-preserving, we have (8). □

Example 4.5. Let $(v_1, \ldots, v_n) \in]0,1]^n$ be a weighting vector with $v_1 + \ldots + v_n = 1$. Consider the IVREF given by Example 3.11 (iii) for $p = 1$, $\alpha = 0.5$, $\beta = 1$; and the IV aggregation function defined as in Theorem 3.17 for $\alpha = 0.5$ (and $\beta = 1$, for instance) and $M_1(x_1, \ldots, x_n) = M_2(x_1, \ldots, x_n) = v_1x_1 + \ldots + v_nx_n$. Then, applying Corollary 4.4, we obtain a width-based IV similarity measure S_{M_1} on $IVFS(U)$ w.r.t. $\leq_L$ associated with M_1; moreover, S_{M_1} satisfies (8).

As a special case, which is later used in the stereo matching application in the following section, we give an expression of S_{M_1} for weighting vector $\left(\frac{1}{n}, \ldots, \frac{1}{n}\right)$. Note that, since the considered IV aggregation function is decomposable, the expression can be simplified significantly (see Lemma 3.18):

$$S_{M_1}(A,B) = \left[1 - \frac{\sum_{i=1}^{n} \left|\underline{A(u_i)} + \overline{A(u_i)} - \underline{B(u_i)} - \overline{B(u_i)}\right|}{2n} - \frac{\sum_{i=1}^{n} \min(w(A(u_i)), w(B(u_i)))}{n}, \right.$$
$$\left. 1 - \frac{\sum_{i=1}^{n} \left|\underline{A(u_i)} + \overline{A(u_i)} - \underline{B(u_i)} - \overline{B(u_i)}\right|}{2n}\right]. \tag{9}$$

Theorem 4.6. *Let $M_{IV} : (L([0,1]))^n \to L([0,1])$ be an IV aggregation function w.r.t. $\leq_L$ satisfying $M_{IV}(X_1, \ldots, X_n) = 1_L$ if and only if $X_1 = \ldots = X_n = 1_L$ and $M_{IV}(X_1, \ldots, X_n) = 0_L$ if and only if $X_1 = \ldots = X_n = 0_L$. Let $R_{IV} : (L([0,1]))^2 \to L([0,1])$ be a function satisfying axioms 1, 3, 4 from Definition 3.1 and $R_{IV}(X,Y) = 1_L$ if and only if $X = Y$ and $w(X) = 0$ for all $X, Y \in L([0,1])$. Then the function $S_M : IVFS(U) \times IVFS(U) \to L([0,1])$ defined by:*

$$S_M(A,B) = M_{IV}\big(R_{IV}(A(u_1), B(u_1)), \ldots, R_{IV}(A(u_n), B(u_n))\big)$$

for all $A, B \in IVFS(U)$ satisfies axioms (SM1), (SM3), (SM4) and

(SM2′) $S(A,B) = 1_L$ if and only if $A = B$ and $A, B \in FS(U)$.

Proof. The proof is obtained just by a straight calculation. □

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

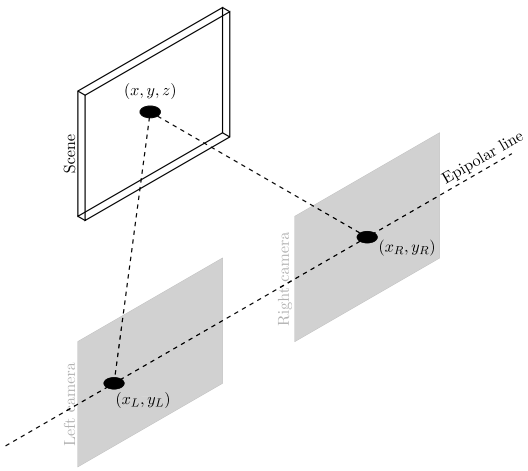


Fig. 1. Schema of epipolar geometry for stereo vision.

Remark 4.7. Note that replacing axiom (SM2) in Definition 4.1 by Axiom (SM2′), we get an alternative approach to IV similarity measures which takes the width of intervals into account and is in the line with the ideas stated after Definition 3.1.

In the next Section we apply our developments to a stereo matching problem.

5. Stereo matching with w-preserving interval-valued restricted equivalence functions

Stereo vision arises as a model to capture the information around us in the same way as human vision does. Human visual perception is based on the formation of 3D images of the environment. Each eye captures a different scene, and from both, our brain builds a 3D image of the world around us. In stereo vision, human vision is mimicked by means of two cameras pointing at the same scene, such that each camera acts effectively as one 'eye'. In order to obtain a perception of depth from these two images, the so-called correspondence between stereo images is used. Correspondence between stereo images is one of the main problems in computer vision [36,44] and it is very relevant in applications such as 3D scene reconstruction, autonomous movement or robotics.

The camera system used to catch stereo images is arranged according to an epipolar geometry, in such a way that all the points which correspond to the same camera lay on a plane, as shown in Fig. 1. Points captured from the scene by each of the cameras corresponding to the same point in the space are called corresponding points. Finding these corresponding points in the captured images is not a trivial task. Corresponding points may be affected by noise, occlusions or distortion during the capturing procedure. This task becomes easier if it is required that points in the captured images are subject to the epipolar restriction, i.e., each point should be in its corresponding epipolar line. This requirement implies that the search for the best correspondence should only be done along one of the dimensions of the image (horizontal axis).

Once the corresponding points have been located, a disparity map can be calculated. This map represents the depth of the objects in the image. Each disparity value is obtained by calculating the difference between the positions of the corresponding points along the horizontal axis. In the next subsection, we discuss a specific way for calculating the disparity map. Finally, we will apply the proposed IV similarity measures, taking into the width of the intervals into account.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Algorithm 1 Algorithm for constructing a disparity map using IV-similarity measures.

Input: Left and right colour images f_l, f_r, an IV-similarity measure S_M.
Output: Disparity map $\mathbb{f}_d$.
1: IV-fuzzify the images f_l, f_r, getting three IVFSs for each image, one for each colour channel.
2: **for** each pixel (x, y) of f_r **do**
3: Select a window of size $n \times m$ around the pixel;
4: **for** each possible y' until the maximal disparity (provided by the dataset) **do**
5: Select a window of size $n \times m$ around the pixel (x, y');
6: Calculate the IV-similarity between the two windows, in each of the three colour channels using the similarity S_M;
7: Aggregate the values of the IV-similarities for each colour according to Equation (10).
8: **end for**
9: Calculate the disparity between windows taking the pair of windows of greatest similarity according to the order relation $\preceq_{\alpha,\beta}$ with $\alpha = 1$;
10: **end for**
11: Create the disparity map with each of the disparities obtained for each position (x, y);

Fig. 2. Choice of windows for calculating the IV similarity and the disparity.

5.1. Methods for calculating the disparity map

For us, an image is a function $f : X \times Y = \{1, ..., r\} \times \{1, ..., c\} \mapsto L$, where r represents the number of rows in the image, that is, its height; c is the number of columns in the image, that is, its width; and L is a finite lattice whose elements are used to capture the intensities of the pixels in the image. Different choices of L allow the representation of different types of images. In particular, if we take $L = \{0, 1\}$, we represent black and white images, whereas for greyscale images we take $L = \{0, ..., 255\}$, and for colour images in the RGB colour space we take $L = \{0, ..., 255\}^3$.

In order to do the correspondence procedure and to build the disparity map, we use the block-correspondence method. This technique consists of selecting a window of size $n \times m$ in the right image and calculating the similarity of this window with each of the windows of the same size, centred at pixels in the same epipolar line in the left image. To calculate this similarity, classical measures have been usually considered. In stereo vision, the most common ones are SSD [39], SAD [43], NCC [21] o ZNCC [37]. Furthermore, some authors have used fuzzy measures [38]. Among the latter, we find some studies which consider extensions of fuzzy sets, in particular interval-valued fuzzy sets [23] or Atanassov intuitionistic fuzzy sets [32].

In Algorithm 1 we present a method for constructing a disparity map in an RGB image a using IV-similarity measures.

For steps 3 and 5 we proceed as indicated in Fig. 2, comparing the window in the right image to the different windows in the left image. To calculate the similarity between intervals, in step 6 we use the expression of width-based IV similarity proposed in Eq. (9). Besides, as we take into account the colour information in the images, in step 7 we aggregate the information as suggested in [19], by means of the expression:

$$S_{MT}(A, B) = 0.299 \cdot S_{M_R}(A, B) + 0.587 \cdot S_{M_G}(A, B) + 0.114 \cdot S_{M_B}(A, B) \tag{10}$$

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

where $S_{M_R}(A,B)$, $S_{M_G}(A,B)$, $S_{M_B}(A,B)$ represent the similarity values for each of the colour channels calculated using a width-based IV similarity measure S_M. Beyond the weighted mean (**wMean**) considered in Eq. (10), in the experiments, a number of other aggregation functions are also used to merge the similarity values, including the arithmetic mean (**mean**), the product (**prod**), the geometric mean (**gmean**), the harmonic mean (**hmean**), the median (**median**), the maximum (**max**) and the minimum (**min**). Finally, as in [23], for all the images we consider windows of size 7×11.

5.2. IV-fuzzification

A digital image is the result of a discretization of a real world scene. In order to represent such digital images, fuzzy sets have been used in the literature. However, due to the uncertainty which is inherent to the process of discretization, it is very hard to provide an accurate (real-valued) membership value for each pixel. In this sense, the use of interval-valued fuzzy sets allow us to consider such uncertainty by means of the width of the membership intervals associated to each pixel, as the precise (real-valued) fuzzy membership value is an unknown value considered to be inside the provided membership interval. For these reasons, and as it was done in [23], we represent the images by means of interval-valued fuzzy sets, assigning an interval-valued membership to each pixel. In particular, we are going to build the IVFSs representing a given image from different membership functions, as was done in [23]. It is worth mentioning that this IV-fuzzification method has been successfully used in image segmentation problems, obtaining results better than those methods which only take into account one single membership function.

Given k restricted equivalence functions R_i, the procedure to get an IV-fuzzified image is the following:

- Consider an image f with L intensity values.
  - For each level of intensity t, $(t = 0, t = 1, \ldots, L-1)$:
    - Build k fuzzy sets $Q_t^1 \ldots Q_t^k$ where for each $i = 1 \ldots k$
      $Q_t^i = \left\{(q, \mu_{Q_t^i}(q)) | q \in \{0, 1, \ldots L-1\})\right\}$, with

$$\mu_{Q_t^i}(q) = \begin{cases} R_i\left(\frac{q}{L-1}, \frac{m_b(t)}{L-1}\right), & \text{si } q \leq t, \\ R_i\left(\frac{q}{L-1}, \frac{m_o(t)}{L-1}\right), & \text{si } q > t, \end{cases} \tag{11}$$

      where $m_o(t)$ and $m_b(t)$ are the mean intensities of the object and the background, given by
      $m_o(t) = \frac{\sum_{q=0}^{t} q \cdot h(q)}{\sum_{q=0}^{t} h(q)}$
      $m_b(t) = \frac{\sum_{q=t+1}^{L-1} q \cdot h(q)}{\sum_{q=t+1}^{L-1} h(q)}$
      where $h(q)$ denotes the number of pixels with intensity q.
    - For each level of intensity t, build an IVFS $\tilde{Q}_t$ from the fuzzy sets $Q_t^1 \ldots Q_t^k$, taking, for each $q \in \{0, \ldots, L-1\}$,
      $\mu_{\tilde{Q}_t}(q) = \left[T\left(\mu_{Q_t^1}(q), \ldots, \mu_{Q_t^k}(q)\right), S\left(\mu_{Q_t^1}(q), \ldots, \mu_{Q_t^k}(q)\right)\right]$ with T a t-norm and S a t-conorm.
    - Calculate the entropy ϵ_F of each of the L interval valued fuzzy sets $\tilde{Q}_t$ using the expression:
      $\epsilon_F(\tilde{A}) = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{\mu}_{\tilde{A}}(x_i) - \underline{\mu}_{\tilde{A}}(x_i)\right)$, proposed in [34]
      For representing the image f, select the IVFS $\tilde{Q}_t$ of smallest entropy.

Note that we take as best interval valued fuzzy set for representing the image the one of smallest entropy ϵ_F. This is so because this interval-valued fuzzy set displays the smallest amount of uncertainty in order to build the precise value of membership for the pixels.

In the experimentation, we use the configuration proposed in the first experiment in [23], i.e., we take $k = 2$ and the restricted equivalence functions:

$$\begin{cases} R_1(x,y) = 1 - |x - y| \\ R_2(x,y) = 1 - |\sqrt[10]{x} - \sqrt[10]{y}| \end{cases} \tag{12}$$

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

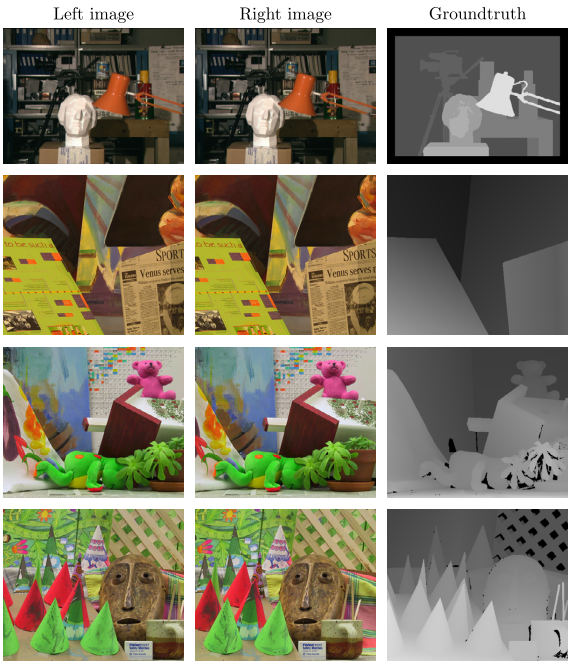


Fig. 3. Left and right images and groundtruth image proposed in the dataset Middlebury.

5.3. *w*-preserving IVREF Stereo Matching

To test the behaviour of the proposed width-depending IV similarity measure, in this work we use the Middlebury dataset [36], composed of the images in Fig. 3. One of the advantages of this dataset is that each pair of images is associated with a model of the disparity map (groundtruth) provided by an expert.

The evaluation of the error in a stereo matching procedure is done by calculating the absolute error percentage between the obtained disparity maps and the groundtruth provided by the dataset. To measure the error percentage, three types of regions, defined by the dataset, are considered:

- nonocc: considers only those pixels that are not occluded (occluded pixels are those which only appear in one of the images).
- all: all the pixels of the image are considered.
- disc: pixels near the discontinuous regions are considered.

In Fig. 4, an example of the evaluation regions is shown. For each one of the regions only pixels labelled in white are considered as part of the error computation.

5.4. Comparison of the proposed interval methods

We compare Algorithm 1 using our width-based IV similarity measure with the one proposed in [23], which also makes use of interval-valued fuzzy sets but without taking into account the width of the intervals.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

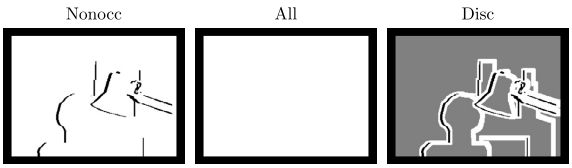


Fig. 4. Regions considered for the evaluation of a disparity map.

Table 1
Comparison of the two methods, IVFS and IVREF, with different aggregation functions to merge colour information. The first column represents the technique and aggregation used. The remaining columns represent for each image the percentage of incorrect disparities obtained for each evaluated region. Finally, the last column represents the mean error.

Algorithm	Tsukuba			Venus			Teddy			Cones			%Et
	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	
IVREF wMean	**5.70**	**7.56**	19.20	9.00	10.51	35.61	**14.90**	**23.61**	35.02	**7.46**	**17.44**	20.69	**17.23**
IVFS wMean	8.02	9.94	**17.12**	14.98	16.38	**32.12**	16.60	25.16	33.44	7.79	17.97	**19.34**	18.24
IVREF greyscale	6.29	8.22	20.46	**7.48**	**9.01**	34.87	15.85	24.43	33.67	8.16	18.02	22.16	17.38
IVFS greyscale	8.87	10.81	18.84	11.71	13.19	32.13	18.91	27.21	**33.24**	9.54	19.54	22.76	18.90
IVREF mean	6.06	7.86	19.37	11.16	12.62	35.53	15.75	24.37	35.62	7.80	17.75	21.34	17.94
IVFS mean	8.44	10.30	17.29	17.92	19.28	32.73	18.61	26.96	34.66	8.34	18.48	20.02	19.42
IVREF prod	6.07	7.86	19.37	11.18	12.64	35.54	15.79	24.41	35.67	7.80	17.75	21.33	17.95
IVFS prod	8.44	10.31	17.31	17.94	19.30	32.74	18.69	27.03	34.76	8.34	18.49	20.02	19.45
IVREF gmean	6.07	7.86	19.37	11.18	12.64	35.54	15.79	24.41	35.67	7.80	17.75	21.33	17.95
IVFS gmean	8.44	10.31	17.31	17.94	19.30	32.74	18.69	27.03	34.76	8.34	18.49	20.02	19.45
IVREF hmean	6.07	7.86	19.39	11.20	12.66	35.51	15.84	24.46	35.73	7.81	17.75	21.35	17.97
IVFS hmean	8.45	10.31	17.32	17.96	19.32	32.74	18.76	27.09	34.82	8.35	18.50	20.04	19.47
IVREF max	6.25	7.98	20.42	10.61	12.10	35.28	16.42	24.92	35.79	9.72	19.53	24.03	18.59
IVFS max	9.13	10.90	19.06	18.68	20.05	34.57	21.29	29.35	35.63	11.83	21.63	25.50	21.47
IVREF median	7.07	8.89	21.03	14.09	15.50	35.99	17.49	25.98	35.98	9.24	19.07	23.89	19.52
IVFS median	9.72	11.60	18.93	20.87	22.20	33.81	20.98	29.11	34.53	11.04	20.91	23.71	21.45
IVREF min	7.63	9.45	20.30	22.49	23.76	37.06	29.21	36.37	43.77	12.51	21.96	26.79	24.27
IVFS min	11.05	12.91	18.63	30.60	31.74	36.52	35.41	41.97	45.50	17.01	26.19	26.83	27.86

Table 1 shows the results obtained in the experiments, highlighting the best result in bold. In this table, IVREF denotes the method using Eq. (9) and the proposed definitions of w-preserving IV restricted equivalence function and similarity, and IVFS denotes the method in [23]. Furthermore, the aggregation function applied to merge the colour channels is specified after the terms IVREF or IVFS, respectively. In addition, we also consider greyscale images.

We can observe that the new similarity method outperforms – in respect to the total error, the one based on fuzzy sets presented in [23], even with different aggregation functions for colour merging. The only exceptions to this are when the median or min are used. It is worth mentioning that although the best performance is achieved when the weighted mean is used in respect to the overall result, in the case of the *Venus*, the image the performance obtained with greyscale images is considerably better for non occluded regions and for all regions.

5.5. Comparison of the proposed method with other methods which do not use intervals

When we compare the best results (using the weighted mean) with the classical methods in the literature, as **SSD**, **SAD**, **NCC**, **ZNCC** and fuzzy measures (**FUZZY**), we clearly see in Table 2 that the new proposal gets globally better results.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

Table 2
Comparison of our similarity measure using the weighted mean for colour aggregation with the classical approaches SSD, SAD, ZNCC and Fuzzy similarity. The first column represents the technique and aggregation operator used. The remaining columns represent for each image the percentage of incorrect disparities obtained for each evaluated region. Finally, the last column represents the mean error.

Algorithm	Tsukuba			Venus			Teddy			Cones			%Et
	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	
IVREF wMean	**5.70**	**7.56**	19.20	9.00	10.51	35.61	14.90	23.61	35.02	**7.46**	**17.44**	**20.69**	**17.23**
IVREF greysacle	6.29	8.22	20.46	7.48	9.01	34.87	15.85	24.43	33.67	8.16	18.02	22.16	17.38
FUZZY wMean	7.75	9.63	**16.17**	13.67	15.09	31.51	18.56	26.90	**29.01**	15.00	24.33	23.24	19.24
FUZZY greyscale	8.35	10.24	16.84	8.41	9.93	**31.04**	20.26	28.39	29.38	16.83	25.89	25.55	19.26
ZNCC wMean	10.53	12.42	29.44	13.01	14.48	42.08	17.00	25.57	39.09	10.68	20.60	29.91	22.07
ZNCC greyscale	10.80	12.67	29.75	**6.05**	**7.64**	42.36	**14.41**	**23.26**	39.04	10.65	20.59	30.92	20.68
SAD wMean	8.74	10.75	22.48	13.57	15.01	38.88	20.00	28.23	35.20	14.84	24.39	28.40	21.71
SAD greyscale	9.26	11.21	22.86	8.52	10.07	38.60	21.55	29.61	34.94	16.43	25.70	29.42	21.51
NCC wMean	10.37	12.30	30.38	17.16	18.54	43.67	21.47	29.58	42.07	13.26	23.03	35.89	24.81
NCC greyscale	10.60	12.52	30.04	6.10	7.68	43.73	14.55	23.38	40.22	12.10	21.96	35.38	21.52
SSD wMean	9.74	11.79	28.75	13.09	14.55	43.02	19.62	27.91	39.62	13.88	23.61	33.47	23.25
SSD greyscale	10.14	12.12	29.11	7.92	9.48	43.58	21.23	29.34	39.73	14.29	23.89	33.72	22.88

Table 3
Comparison of our similarity measure without post-processing and our same method combined with exponential cost aggregation and cost guided filtering in addition to LRC check.

Algorithm	Tsukuba			Venus			Teddy			Cones			%Et
	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	%noocc	%all	%disc	
IVREF wMean	5.70	7.56	19.20	9.00	10.51	35.61	14.90	23.61	35.02	7.46	17.44	20.69	17.23
IVREF modified	**3.10**	**3.76**	**13.42**	**2.43**	**2.99**	**19.79**	**11.49**	**17.14**	**26.09**	**4.72**	**10.74**	**12.34**	**10.67**

The proposed method obtains the best results in the images *Tsukuba* and *Cones*, whereas for the other two ones, (*Venus* and *Teddy*) it remains below the ZNCC method using greyscale images. As we can see, our method outperforms the other ones in non occluded regions, in images where there is a wide variety of colour tones (*Tsukuba* and *Cones*). In regions near discontinuities, the best approach is the FUZZY similarity approach, with greyscale images. In Fig. 5, we show the approaches that obtain the best results in each of the images and regions. Note also that our method leads to disparity maps with less incorrect disparities, mainly in the case of *Tsukuba* and *Cones*. Nevertheless, the disparity values obtained tend to fail at getting the right value near discontinuous pixels. In these regions of discontinuity, better disparity values are obtained with the FUZZY approach, except in the case of *Cones*, where these regions get worse values.

Note that the inclusion of colour information in the process is useful in some cases where the correct value is not clear – here, it offsets some ambiguities. Despite being the best performer, our method still exhibits incorrect disparity values. This is because we have applied a basic methodology without refinement. The results shown are those from the raw disparity maps.

5.6. Illustrative example with refinement steps

As an illustrative example and to show the possibility of improvement of the proposed method, we discuss here an example using some refinement steps.

For the sake of this experiment we have chosen steps 3 and 5 of Algorithm 1, a window of size 3×3, coupled with an Exponential step aggregation function [42] and a Cost Guided filter [25]. Finally we use an LRC [36] to check the raw disparity map for possible outliers.

We can observe that most of the outliers are removed, obtaining a more homogeneous disparity map, although, some objects in the images lose their shape. For example, in the case of the image *Tsukuba* in Fig. 6 the electrical wire

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

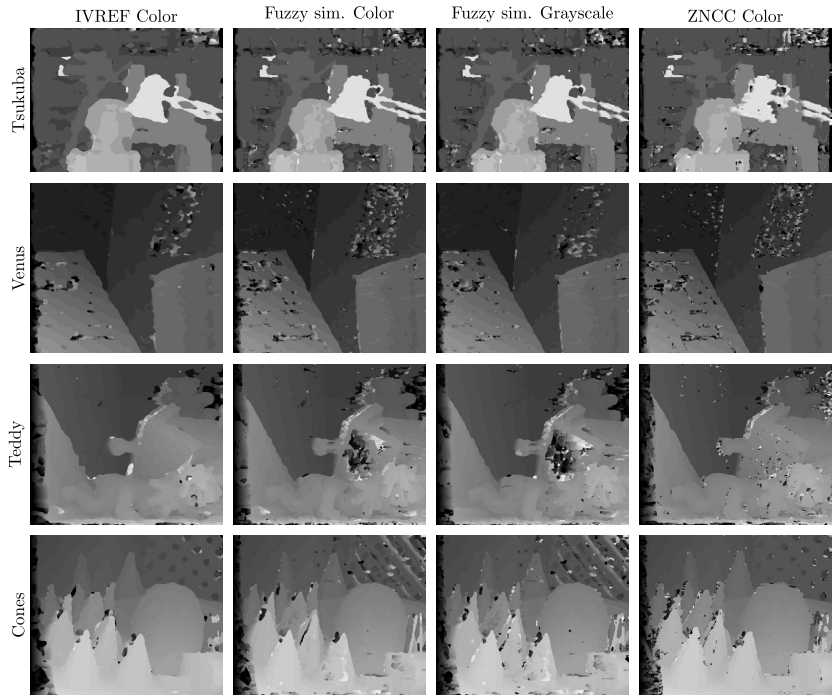


Fig. 5. Disparity maps with the best obtained results. That is, IVREF aggregating colour with the weighted mean, FUZZY with colour aggregation and greyscale and ZNCC with greyscale images.

in the lamp disappears. Also, the use of a filter in the smoothing process makes the shapes of the objects better defined just prior to obtaining the disparity value. In the case of the *Tsukuba* and *Cones* images, the lamp and the cones have almost recovered their shape.

The quantitative results exposed in Table 3 confirm the benefits of using refinement processes during the similarity computation and in the raw disparity map processing. We see that the global error goes down to 10.67% and that in all the analysed images and in all the regions used for error measuring, the refinement improves the results. The most significant decrease occurs in the *Venus* image, where the disparity values homogeneize the image and outliers almost disappear.

6. Conclusions and future research

We have proposed a new definition of interval-valued restricted equivalence functions with respect to total order considering the width of the intervals and described a construction method for restricted equivalence functions which preserves the widths of intervals. In a similar way, we have presented a construction method for w-preserving interval-valued aggregation functions with respect to total orders. Consequently, we have introduced a new definition of a width-based interval-valued similarity measure with respect to total order and proposed a construction method by aggregating restricted equivalence functions. Finally, we have discussed an illustrative example in stereo image matching where the width-based interval-valued similarity measures were used.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

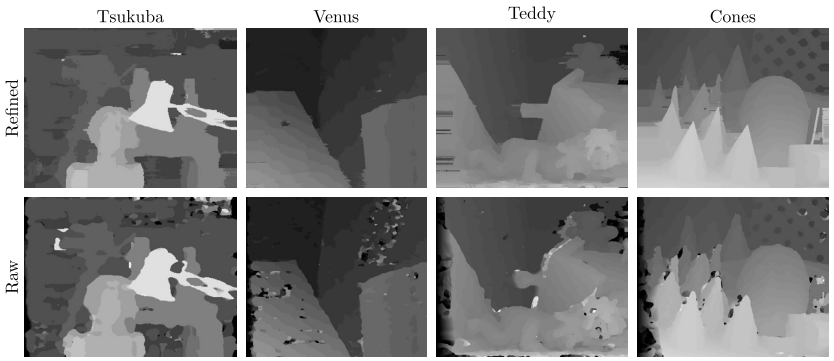


Fig. 6. Comparison of the raw disparity maps obtained with raw IVREF and after applying refinement and outlier detection techniques.

The presented experimental study allows us to conclude that considering the width of the intervals in order to measure similarity between intervals is highly beneficial. The proposed method obtains better results, both compared to other interval-valued methods which do not take into account the width of the intervals, and compared to classical methods such as SSD, SAD, NCC, ZNCC or FUZZY. We note that the time complexity of our proposal is the same as the fuzzy method not using intervals ($\mathcal{O}(N^2)$) and slightly higher than the classical ones due to the fuzzyfication process.

It is worth mentioning that the use of colour in the extraction of the disparity values is a key point that leads to an improvement in the results. Finally, the presented method can be improved combining filtering steps during the similarity computation and searching for inconsistent disparities in the raw disparity map.

In our future work we are going to study the notions and constructions of interval-valued dissimilarity and interval-valued entropy in line with the presented approach, i.e., with respect to total order and considering the widths of intervals. We also intend to analyze the use of refinement steps to further improve our method.

Acknowledgements

H. Bustince, C. Marco-Detchart and J. Fernandez were partially supported by Spanish research project TIN2016-77356-P (AEI/FEDER, UE). Z. Takáč has been supported by Project VEGA 1/0614/18. This contribution is also the partial result of the Research & Development Operational Programme for the project University Scientific Park STU in Bratislava, ITMS 26240220084. C. Wagner and J. M. Garibaldi were partially supported by the UK EPSRC grant EP/P011918/1.

References

[1] M.J. Asiáin, H. Bustince, R. Mesiar, A. Kolesárová, Z. Takáč, Negations with respect to admissible orders in the interval-valued fuzzy set theory, IEEE Trans. Fuzzy Syst. 26 (2018) 556–568.

[2] E. Barrenechea, H. Bustince, B. De Baets, C. Lopez-Molina, Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images, IEEE Trans. Fuzzy Syst. 19 (5) (2011) 819–830.

[3] E. Barrenechea, J. Fernandez, M. Pagola, F. Chiclana, H. Bustince, Construction of interval-valued fuzzy preference relations from ignorance functions and fuzzy preference relations. Application to decision making, Knowl.-Based Syst. 58 (2014) 33–44.

[4] U. Bentkowska, H. Bustince, A. Jurio, M. Pagola, B. Pekala, Decision making with an interval-valued fuzzy preference relation and admissible orders, Appl. Soft Comput. 35 (2015) 792–801.

[5] P. Burillo, H. Bustince, Construction theorems for intuitionistic fuzzy sets, Fuzzy Sets Syst. 84 (1996) 271–281.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

[6] H. Bustince, Indicator of inclusion grade for interval-valued fuzzy sets. Application to approximate reasoning based on interval-valued fuzzy sets, Int. J. Approx. Reason. 23 (3) (2000) 137–209.
[7] H. Bustince, E. Barrenechea, M. Pagola, Relationship between restricted dissimilarity functions, restricted equivalence functions and normal E_N-functions: image thresholding invariant, Pattern Recognit. Lett. 29 (4) (2008) 525–536.
[8] P. Burillo, H. Bustince, Entropy on intuitionistic fuzzy sets and on interval-valued fuzzy sets, Fuzzy Sets Syst. 78 (1996) 305–316.
[9] H. Bustince, E. Barrenechea, M. Pagola, Image thresholding using restricted equivalence functions and maximizing the measure of similarity, Fuzzy Sets Syst. 128 (5) (2007) 496–516.
[10] H. Bustince, E. Barrenechea, M. Pagola, Restricted equivalence functions, Fuzzy Sets Syst. 157 (17) (2006) 2333–2346.
[11] H. Bustince, E. Barrenechea, M. Pagola, J. Fernández, Z. Xu, B. Bedregal, J. Montero, H. Hagras, F. Herrera, B. De Baets, A historical account of types of fuzzy sets and their relationship, IEEE Trans. Fuzzy Syst. 24 (1) (2016) 179–194.
[12] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar, Generation of linear orders for intervals by means of aggregation functions, Fuzzy Sets Syst. 220 (2013) 69–77.
[13] C. Cornelis, G. Deschrijver, E.E. Kerre, Implication in intuitionistic fuzzy and interval-valued fuzzy set theory: construction, classification, application, Int. J. Approx. Reason. 35 (1) (2004) 55–95.
[14] I. Couso, H. Bustince, From fuzzy sets to interval-valued and Atanassov intuitionistic fuzzy sets: a unified view of different axiomatic measures, IEEE Trans. Fuzzy Syst. 27 (2) (2019) 362–371.
[15] I. Couso, L. Sánchez, Additive similarity and dissimilarity measures, Fuzzy Sets Syst. 322 (2017) 35–53.
[16] H.M. Choi, G.S. Mun, J.Y. Ahn, A medical diagnosis based on interval-valued fuzzy sets, Biomed. Eng. Appl. Basis Commun. 24 (4) (2012) 349–354.
[17] P. Couto, A. Jurio, A. Varejao, M. Pagola, H. Bustince, P. Melo-Pinto, An IVFS-based image segmentation methodology for rat gait analysis, Soft Comput. 15 (10) (2011) 1937–1944.
[18] G. Deng, L. Song, Y. Jiang, J. Fu, Monotonic similarity measures of interval-valued fuzzy sets and their applications, Int. J. Uncertain. Fuzziness Knowl.-Based Syst. 25 (4) (2017) 515–544.
[19] E. Deza, M.-M. Dezam, Image and audio distances, in: Dictionary Distances, Elsevier, 2006, pp. 262–278.
[20] D. Dubois, H. Prade, Gradualness, uncertainty and bipolarity: making sense of fuzzy sets, Fuzzy Sets Syst. 192 (2012) 3–24.
[21] O. Faugeras, T. Viéville, E. Theron, J. Vuillemin, B. Hotz, Z. Zhang, L. Moll, P. Bertin, H. Mathieu, P. Fua, G. Berry, C. Proy, Real-Time Correlation-Based Stereo: Algorithm, Implementations and Applications, Research Report RR-2013, INRIA, 1993.
[22] J. Fodor, M. Roubens, Fuzzy preference modelling and multicriteria decision support, in: Theory and Decision Library, Kluwer Academic Publishers, Dordrecht, 1994.
[23] M. Galar, J. Fernandez, G. Beliakov, H. Bustince, Interval-valued fuzzy sets applied to stereo matching of color images, IEEE Trans. Image Process. 20 (7) (2011) 1949–1961.
[24] A. Heidarzade, A new similarity measure for interval type-2 fuzzy sets: application in fuzzy risk analysis, Int. J. Appl. Decis. Sci. 9 (4) (2016) 400–412.
[25] A. Hosni, C. Rhemann, M. Bleyer, C. Rother, M. Gelautz, Fast cost-volume filtering for visual correspondence and beyond, IEEE Trans. Pattern Anal. Mach. Intell. 35 (2) (Feb 2013) 504–511.
[26] L. Hurwicz, The Generalised Bayes Minimax Principle. A Criterion for Decision Making Under Uncertainty, Cowles Comission Discussion Paper 355, 1951.
[27] A. Jurio, M. Pagola, R. Mesiar, G. Beliakov, H. Bustince, Image magnification using interval information, IEEE Trans. Image Process. 20 (11) (2011) 3112–3123.
[28] M. Komorníková, R. Mesiar, Aggregation functions on bounded partially ordered sets and their classification, Fuzzy Sets Syst. 175 (1) (2011) 48–56.
[29] X. Liu, Entropy, distance measure and similarity measure of fuzzy sets and their relations, Fuzzy Sets Syst. 52 (1992) 305–318.
[30] Z. Lu, J. Ye, Logarithmic similarity measure between interval-valued fuzzy sets and its fault diagnosis method, Information (Switzerland) 9 (2018) 36.
[31] R. Mesiar, M. Komorníková, Aggregation functions on bounded posets, in: C. Cornelis, et al. (Eds.), 35 Years of Fuzzy Sets Theory, in: Studies in Fuzziness and Soft Computing, vol. 261, Springer, Berlin, 2010, pp. 3–17.
[32] G. Nieradka, B.S. Butkiewicz, Features stereo matching based on fuzzy logic, in: Proceedings IFSA/EUSFLAT Conference, 2009, pp. 1188–1193.
[33] A. Pradera, G. Beliakov, H. Bustince, B. De Baets, A review of the relationship between implication, negation and aggregation functions from the point of view of material implication, Inf. Sci. 329 (2016) 357–380.
[34] R. Sambuc, Function phi-flous application a l'aide au diagnostic en pathologie thyroidienne, PhD thesis, University of Marseille, 1975.
[35] J.A. Sanz, A. Fernandez, H. Bustince, F. Herrera, IVTURS: a linguistic fuzzy rule-based classification system based on a new interval-valued fuzzy reasoning method with tuning and rule selection, IEEE Trans. Fuzzy Syst. 21 (3) (2013) 399–411.
[36] D. Scharstein, R. Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, Int. J. Comput. Vis. 47 (1/3) (2002) 7–42.
[37] C. Sun, Fast stereo matching using rectangular subregioning and 3D maximum-surface techniques, Int. J. Comput. Vis. 47 (1–3) (2002) 99–117.
[38] G. Tolt, I. Kalaykov, Measures based on fuzzy similarity for stereo matching of color images, Soft Comput. 10 (12) (Oct 2006) 1117–1126.
[39] E. Trucco, V. Roberto, S. Tinonin, M. Corbatto, SSD disparity estimation for dynamic stereo, in: BMVC, 1996, pp. 1–10.
[40] Z.S. Xu, R.R. Yager, Some geometric aggregation operators based on intuitionistic fuzzy sets, Int. J. Gen. Syst. 35 (2006) 417–433.
[41] J. Ye, Multicriteria decision-making method based on cosine similarity measures between interval-valued fuzzy sets with risk preference, Econ. Comput. Econ. Cybern. Stud. Res. 50 (4) (2016) 205–215.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

[42] W. Yu, T. Chen, J.C. Hoe, Real time stereo vision using exponential step cost aggregation on GPU, in: Proc. - Int. Conf. Image Process. ICIP, 2009, pp. 4281–4284.
[43] R. Zabih, J. Woodfill, Non-parametric local transforms for computing visual correspondence, in: Jan-Olof Eklundh (Ed.), Comput. Vis. — ECCV '94, Springer, Berlin, Heidelberg, 1994, pp. 151–158.
[44] M.Z. Brown, D. Burschka, G.D. Hager, Advances in computational stereo, IEEE Trans. Pattern Anal. Mach. Intell. 25 (8) (Aug 2003) 993–1008.

Please cite this article in press as: H. Bustince et al., Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders, Fuzzy Sets Syst. (2019), https://doi.org/10.1016/j.fss.2019.04.002

3.4 Ordered directional monotonicity in the construction of edge detectors

- C. Marco-Detchart, H. Bustince, J. Fernandez, R. Mesiar, J. Lafuente, E. Barrenechea, and J. M. Pintor, "Ordered directional monotonicity in the construction of edge detectors", *Fuzzy Sets and Systems*,
  - Journal: Fuzzy Sets and Systems
  - Status: Submitted
  - Impact Factor: (JCR 2018) 2.907
  - Knowledge area:
    - Statistics & Probability: Ranking $7/123$ (Q1)
    - Computer Science Theory & Methods: Ranking $21/104$ (Q1)
    - Applied Mathematics: Ranking $16/254$ (Q1)

Elsevier Editorial System(tm) for Fuzzy Sets and Systems

Manuscript Draft

Manuscript Number: FSS-D-19-00196

Title: Ordered directional monotonicity in the construction of edge detectors

Article Type: Full Length Article (FLA)

Keywords: Ordered directionally monotone function; Directional monotonocity; Edge detection;
Consensus image

Corresponding Author: Dr. Humberto Bustince, Dr

Corresponding Author's Institution: Public University of Navarra

First Author: Cedric Marco-Detchart

Order of Authors: Cedric Marco-Detchart; Humberto Bustince, Dr; Javier Fernandez; Radko Mesiar; Julio Lafuente; Edurne Barrenechea; Jesus Pintor

Abstract

In this paper we provide a specific contruction method of ordered directionally monotone functions. We show that the functions obtained with this construction method can be used to build edge detectors for grayscale images. We compare the results of these detector to those obtained with some other ones that are widely used in the literature. Finally, we show how a consensus edge detector can be built that improves the results that are obtained both by our proposal and by those in the literature individually.

Ordered directional monotonicity in the construction of edge detectors

C. Marco-Detchart[a,b,c], H. Bustince[a,b,c], J. Fernandez[a,b,c], R. Mesiar[e], J. lafuente[a], E. Barrenechea[a,b,c], J. M. Pintor[d]

[a] *Departamento de Estadistica, Informatica y Matematicas, Universidad Publica de Navarra, Campus de Arrosadia, 31006, Pamplona, Spain*
[b] *Institute of Smart Cities, Universidad Publica de Navarra, Campus de Arrosadia, 31006, Pamplona, Spain*
[c] *Laboratory Navarrabiomed, Hospital Complex of Navarre (CHN), Universidad Publica de Navarra, IdiSNA, Irunlarrea 3. 31008, Pamplona, Spain*
[d] *Departamento de Ingeniería, Universidad Publica de Navarra, Campus de Arrosadia, 31006, Pamplona, Spain*
[e] *Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology in Bratislava, Radlinskeho 9, Bratislava, Slovakia*

Abstract

In this paper we provide a specific contruction method of ordered directionally monotone functions. We show that the functions obtained with this construction method can be used to build edge detectors for grayscale images. We compare the results of these detector to those obtained with some other ones that are widely used in the literature. Finally, we show how a consensus edge detector can be built that improves the results that are obtained both by our proposal and by those in the literature individually.

Keywords: Ordered directionally monotone function; Directional monotonocity; Edge detection; Consensus image.

1. Introduction

In recent years, the analysis of functions which fulfill monotonicity conditions weaker than those required to aggregations is attracting a growing interest. In particular, this study has led to the development of new concepts as, for instance, weak monotonicity (where monotonicity is required only along the ray defined by the vector $(1, \ldots, 1)$, see [42]) or directional monotonicity (where monotonicity is required along some ray defined by a vector in the first quadrant, see [11]). These extensions have shown themselves very useful in different application fields, specially in classification [31].

A common feature shared by all this extensions is that the direction along which monotonicity is considered is the same for every considered input. But, in some applied problems, this can be a too hard restriction.

Let us consider, for instance, the problem of automatically identify objects on an image [3, 17, 24, 25, 32] Many times, an important step for this identification consists in calculating the edges of the objects in the image [10]. Recall that the contours of the visible objects are denoted as edges if there exists a big enough jump between the intensity of a pixel and those of its neighbours. Clearly, this is an imprecise definition, but this is due to the uncertainty inherent to the concept of border itself [28].

Many edge detection methods use, among other things, the *gradient* vector to represent intensity jumps between pixels [37, 41], as for example the well-known approaches by Sobel and Feldman [40], Prewitt [38] and Canny [15], being the latter still considered as one of the most important references for comparison in order to determine the quality of a given method. In recent years, different methods have appeared

Email addresses: cedric.marco@unavarra.es (C. Marco-Detchart), bustince@unavarra.es (H. Bustince), fcojavier.fernandez@unavarra.es (J. Fernandez), mesiar@math.sk (R. Mesiar), lafuente@unavarra.es (J. lafuente), edurne.barrenechea@unavarra.es (E. Barrenechea), txma@unavarra.es (J. M. Pintor)

which also use Machine Learning techniques, as the ones based on Random Forest [16]; other methods using Newton's Gravitational Law, as for instance,[28], etc. Furthermore, deep learning methods (see [43, 44, 45], for instance) have become the most widely used ones due to their high performance, which improves those of the other mentioned algorithms. Nevertheless, deep learning techniques have in general a high computational cost and require of a previous training, so in some cases it is enough to use easier, less costly methods.

In order to focus our problem, let us consider Bezdek standardization of the edge detection problem. Bezdek *et al.* [4] introduced a framework encompassing many of the existing methods in the literature in that moment, which is a breakdown structure formed by four phases:

($S1$) Smoothing the image. To apply a Gaussian filter (with $\sigma_G = 2$) to the image I obtaining a new smoother image, I_G.

($S2$) Obtaining the feature image, i.e., a new image I_M obtained from I_G, where each pixel in I_M represents information about the intensity jump between the corresponding pixel in I_G and its neighbours in I_G.

($S3$) Thinning the feature image. To use the non-maximum suppression procedure [15], computing prior orientations by Kovesis' function [27], obtaining a thinned image, I_H.

($S4$) Binarizing the thinned image. To apply a hysteresis method [36] to obtain the binary edge image I_B. Such representations are commonly demanded to every edge detection approach fulfilling the restrictions imposed by Canny [14, 15].

If we center in the *feature extraction* phase ($S2$), we observe that we must consider the information provided by a pixel and its neighbours simultaneously. But, if we only consider the difference between the maximal and the minimal intensities [6] it is clear that we are losing information is lost as we do not take into account all the differences of intensity between the central pixel and its neighbours. Furthermore, the relevant information must take into account all the intensity jumps in the neighbourhood and how they are related to each other according to their relative size. In particular, it is natural to order the intensity jumps in a decreasing order, as the bigger an intensity jump is, the larger its influence is in the possible existence of an edge. In this sense, it is natural to consider the directions defined by the increasingness or decreasingness of the intensity, but such directions may vary from one pixel to another one. For this reason, we are led to consider that the use or ordered directionally monotone functions may be natural for this problem.

Taking all the previous considerations into account, we have considered the following goals in this work:

1. To develop a new method to build mathematically ordered directionally monotone (ODM).
2. To show that the ODM functions built using this new method can be useful to develop a new method to build a feature image for phase ($S2$) in the edge detection procedure

Besides, as we have already commented previously, there exists many different methods to get the feature image in phase ($S2$). For this reason, we consider the following goal to complete our experimental study:

3. To build, using penalty functions [13, 8, 9], a consensus feature image from those obtained with different algorithms in phase ($S2$).

The structure of this work is the following. In Section 2 some of the basic notation and concepts needed for our proposal are exposed. Section 3 is devoted to expose a specific case of OD monotone function and Section 4 develops the previous section ideas in the context of image feature extraction along with two alternative constructions (Section 5). Section 6 presents a consensus edge detector build upon the combination of the different considered methods. The experimental framework is exposed in Section 7. Then, Section 8 shows the quantitative evaluation of the proposed methods compared with different alternatives of the literature. Finally in Section 9 some conclusions and future works are presented.

2. Preliminaries

2.1. Basic notations and concepts

In this subsection we fix some notations and concepts which will be useful for the remainder of the work.

Let $n > 1$. We use bold letters to denote points in the hypercube $[0,1]^n$, i.e., $\mathbf{x} = (x_1, \ldots, x_n) \in [0,1]^n$. In particular, we write $\mathbf{0} = (0, \ldots, 0)$ and $\mathbf{1} = (1, \ldots, 1)$. Given $\mathbf{x}, \mathbf{y} \in [0,1]^n$ we write $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for every $i \in \{1, \ldots, n\}$. Note that this relation is a partial order which extends the usual linear order between real numbers.

For $n > 1$, we denote by P_n the set of permutations of $\{1, \ldots, n\}$. That is,

$$P_n = \{\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\} \mid \sigma \text{ is bijective}\}.$$

Given $\sigma \in P_n$, $\mathbf{x} \in [0,1]^n$ and $\vec{r} \in \mathbb{R}^n$, we define:

$$\mathbf{x}_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

and

$$\vec{r}_\sigma = (r_{\sigma(1)}, \ldots, r_{\sigma(n)}).$$

We recall now that, for the sake of commodity, we call fusion function (of dimension n) to any function $F : [0,1]^n \to [0,1]$.

A distinguished class of fusion functions, specially from the point of view of applications, is that of aggregation functions [20, 2].

Definition 1. *[1, 12] A mapping $M : [0,1]^n \to [0,1]$ is an aggregation function if it is monotone non-decreasing in each of its components and satisfies $M(\mathbf{0}) = M(0, 0, \ldots, 0) = 0$ and $M(\mathbf{1}) = M(1, 1, \ldots, 1) = 1$.*

2.2. Ordered directionally monotonicity

Imposing monotonicity might be too restrictive for some specific applications (e.g., the mode is not increasing with respect to all its arguments but it is a valid function for certain applications). This consideration led Wilkin and Beliakov [42] to introduce the notion of weak monotonicity.

This concept of weak monotonicity can be further extended by the notion of directional monotonicity, that we define now.

Definition 2. *[11] Let $\vec{r} = (r_1, \ldots, r_n)$ be a real n-dimensional vector, $\vec{r} \neq \vec{0}$. A fusion function $F : [0,1]^n \to [0,1]$ is $\vec{r}$-increasing if for all points $(x_1, \ldots, x_n) \in [0,1]^n$ and for all $c > 0$ such that $(x_1 + cr_1, \ldots, x_n + cr_n) \in [0,1]^n$ it holds*

$$F(x_1 + cr_1, \ldots, x_n + cr_n) \geq F(x_1, \ldots, x_n)\ .$$

That is, an $\vec{r}$-increasing function is a function which is increasing along the ray (direction) determined by the vector $\vec{r}$. For this reason, we say that F is directionally monotone, or, more specifically, directionally $\vec{r}$-increasing. For an in-depth study of the concept of directional monotonicity see [11]. Nevertheless, it is worth to mention that directional monotonicity combined with appropriate boundary conditions leads to the notion of pre-aggregation function, see [31].

In [7] the notion of ordered directionally monotone function (ODM) is presented. To motivate its introduction, note that by means of directional monotonicity, usual monotonicity may be relaxed, just requiring increasingness along some fixed ray. However, the direction along which monotonicity is demanded is the same for every point in the domain $[0,1]^n$, and it is independent of the particular point which is being considered.

For ODM functions, on the contrary, the direction along which monotonicity is required varies depending on the relative size of the coordinates of the considered input. The formal definition reads as follows.

Definition 3. *[7] Let $F : [0,1]^n \to [0,1]$ be a fusion function and let $\vec{r} \neq \vec{0}$ be a n-dimensional vector. F is said to be ordered directionally (OD) $\vec{r}$-increasing if for any $\mathbf{x} \in [0,1]^n$, for any $c > 0$ and for any permutation $\sigma \in P_n$ with $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(n)}$ and such that*

$$1 \geq x_{\sigma(1)} + cr_1 \geq \cdots \geq x_{\sigma(n)} + cr_n \geq 0,$$

it holds that

$$F(\mathbf{x} + c\vec{r}_{\sigma^{-1}}) \geq F(\mathbf{x}),$$

where $\vec{r}_{\sigma^{-1}} = (r_{\sigma^{-1}(1)}, \ldots, r_{\sigma^{-1}(n)})$.

Analogously, F is said to be OD $\vec{r}$-decreasing if for any $\mathbf{x} \in [0,1]^n$, for any $c > 0$ and for any permutation $\sigma \in P_n$ with $x_{\sigma(1)} \geq \cdots \geq x_{\sigma(n)}$ and such that

$$1 \geq x_{\sigma(1)} + cr_1 \geq \cdots \geq x_{\sigma(n)} + cr_n \geq 0,$$

it holds that

$$F(\mathbf{x} + c\vec{r}_{\sigma^{-1}}) \leq F(\mathbf{x}),$$

where $\vec{r}_{\sigma^{-1}} = (r_{\sigma^{-1}(1)}, \ldots, r_{\sigma^{-1}(n)})$.

By an OD $\vec{r}$-monotone function we mean a function which is OD $\vec{r}$-increasing or ordered directionally $\vec{r}$-decreasing.

Example 4. *The weighted Lehmer mean*

$$L_\lambda(x,y) = \frac{\lambda x^2 + (1-\lambda)y^2}{\lambda x + (1-\lambda)y},$$

with the convention $0/0 = 0$, is $(1-\lambda, \lambda)$-increasing. It follows that the function

$$G_\lambda(x,y) = \frac{\lambda(\vee(x,y))^2 + (1-\lambda)(\wedge(x,y))^2}{\lambda \vee (x,y) + (1-\lambda) \wedge (x,y)}$$

is OD $(\lambda, 1-\lambda)$-increasing. Note that, if $\lambda \in]0,1[$, then L_λ is a pre-aggregation function which is not an aggregation function.

3. A particular case of OD monotone function

In this section we discuss an affine construction method for ODM functions.

Theorem 5. *Let $G : [0,1]^n \to [0,1]$ be defined, for $\mathbf{x} \in [0,1]^n$ and $\sigma \in P_n$ such that $x_{\sigma(1)} \geq \ldots \geq x_{\sigma(n)}$, by*

$$G(\mathbf{x}) = a + \sum_{i=1}^{n} b_i x_{\sigma(i)},$$

for some $a \in [0,1]$ and $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ such that $0 \leq a + b_1 + \cdots + b_j \leq 1$ for all $j \in \{1, \ldots, n\}$. Then G is OD $\vec{r}$-increasing for every non-null vector $\vec{r}$ such that $\vec{b} \cdot \vec{r} \geq 0$. In particular, for every non-null vector $\vec{r}$ which is orthogonal to $\vec{b}$.

Proof.

Take $\sigma \in P_n$ such that $x_{\sigma(1)} \geq x_{\sigma(2)} \geq \cdots \geq x_{\sigma(n)}$. Take also $c > 0$ such that $(x_{\sigma(1)} + cr_1, \ldots, x_{\sigma(n)} + cr_n) \in [0,1]^n$. Then we have that

$$\begin{aligned} G(\mathbf{x} + c\vec{r}_{\sigma^{-1}}) &= a + \sum_{i=1}^{n} b_i (x_{\sigma(i)} + cr_i) \\ &= a + \sum_{i=1}^{n} b_i x_{\sigma(i)} + c \sum_{i=1}^{n} b_i r_i \\ &\geq G(\mathbf{x}), \end{aligned}$$

as we wanted to show. ∎

Theorem 5 can be generalized taking into account the following lemma.

Lemma 6. *Let $\varphi : [0,1] \to [0,1]$ be an automorphism (i.e., an increasing bijection). Then, if $G : [0,1]^n \to [0,1]$ is an ordered directionally increasing function, the function $\varphi \circ G$ is also an ordered directionally increasing function.*

Proof.

It follows straightforwardly from the definition of ordered directionally increasing functions and the fact that φ is increasing. ∎

From Lemma 6, we have the following corollary of Theorem 5, which is very relevant for our edge detectors, since the value of p allows us to darken or lighten the considered image.

Corollary 7. *Let $p > 0$. Let $G : [0,1]^n \to [0,1]$ be defined, for $\mathbf{x} \in [0,1]^n$ and $\sigma \in P_n$ such that $x_{\sigma(1)} \geq \ldots \geq x_{\sigma(n)}$, by*

$$G(\mathbf{x}) = \left(a + \sum_{i=1}^{n} b_i x_{\sigma(i)} \right)^{\frac{1}{p}}, \tag{1}$$

for some $a \in [0,1]$ and $\vec{b} = (b_1, \ldots, b_n) \in \mathbb{R}^n$ such that $0 \leq a + b_1 + \cdots + b_j \leq 1$ for all $j \in \{1, \ldots, n\}$. Then G is OD $\vec{r}$-increasing for every non-null vector $\vec{r}$ such that $\vec{b} \cdot \vec{r} \geq 0$.

Proof.

It follows from Lemma 6 taking into account that the function $\varphi(x) = x^{\frac{1}{p}}$ is an automorphism. ∎

The following result is very relevant from the point of view of applications in image processing, since it can be straightforwardly applied to the cases in which an edge exists or not.

Corollary 8. *Let $p > 0$ and let $G : [0,1]^n \to [0,1]$ be defined as in Corollary 7. Then*

(i) $G(\mathbf{0}) = 0$ if and only if $a = 0$.

(ii) Assume that $a = 0$. Then, $G(\mathbf{1}) = 1$ if and only if $b_1 + \cdots + b_n = 1$.

Proof. To see (i), observe that $G(\mathbf{0}) = a^{\frac{1}{p}}$, so the result is straightforward. Regarding (ii), $G(\mathbf{1}) = (\sum_{i=1}^{n} b_i \cdot 1)^{\frac{1}{p}}$, and the result follows. ∎

4. A new algorithm to construct a feature image using Ordered Directionally Monotone functions

In Algorithm 1 we present the process to obtain a feature image by means of ODM functions, phase ($S2$), which by definition consider both the intensity jumps between a pixel and its neighbours and the direction along which such jumps vary, in a decreasing way.

Algorithm 1 Algorithm to construct a feature image using ODM functions

Input: A normalized greyscale image I_G and a parameter $p > 0$ to build an ODM function G as in Corollary 7.

Output: A feature image I_M.

1: **for** each pixel (x, y) of I_G **do**
2: Calculate the 8 values obtained by applying the absolute value of the difference between $\mathbb{I}_g(x, y)$ and its 8-neighbourhood;
3: Order the eight values of step 2 in a decreasing way;
4: Calculate the parameters a, $\vec{r}$ y $\vec{b}$ according to the vector obtained in step 3.
5: Build the ODM function G as in Corollary 7 with the parameters obtained in step 4.
6: Apply the ODM function G to the values obtained in step 3;
7: Assign as intensity of the pixel (x, y) of I_M the value obtained in step 6.
8: **end for**

Next, let us expose how to carry out the execution of the Algorithm 1. Firstly, let us consider that the pixel (x, y) of I_G is the pixel a_{22} of Fig. 1 and then calculate the 8 values indicated in step 2; the outcome of these calculations are given by:

$$x_1 = |a_{22} - a_{11}|, \; x_2 = |a_{22} - a_{12}|,$$
$$x_3 = |a_{22} - a_{13}|, \; x_4 = |a_{22} - a_{23}|,$$
$$x_5 = |a_{22} - a_{33}|, \; x_6 = |a_{22} - a_{32}|,$$
$$x_7 = |a_{22} - a_{31}|, \; x_8 = |a_{22} - a_{21}|.$$

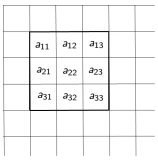


Figure 1: Pixel a_{22} and its 8-neighbourhood.

In step 3, these differences are ordered in a decreasing way; that is,

$$x_{\sigma(1)} \geq x_{\sigma(2)} \geq \ldots \geq x_{\sigma(7)} \geq x_{\sigma(8)}.$$

Note that, as we have already said, there exists an edge if there is a big enough intensity jump between a pixel and its neighbours. So the greatest intensity differences are the most relevant ones in order to determine if there is an edge or not.

In step 4 we calculate the parameters a, $\vec{r}$ y $\vec{b}$ necessary to get in step 5 an ODM function G as in Corollary 7. Finally, in step 6, the ODM function G is applied to get the feature image.

5. Two specific cases of constructions of the ODM function for step 5 in Algorithm 1

In this section we discuss two expressions for ODM functions for Algorithm 1. These expressions are obtained from Corollary 7 using Eq. (1) and giving specific values to the parameters a, p, vector $\vec{r}$ and $\vec{b}$. It is important to remark that these expressions are a first approach and they have not been optimized. On

the contrary, we discuss them due to their simplicity. In a future work, we intend to optimize the value of the different parameters depending on the specific type of images that we consider.

If we use the expression given in Eq. (1) for step 6 in Algorithm 1, then the parameter p allows us to darken or lighten the resulting feature image. It is enough to observe that if $p > 1$, then we get a lighter feature image, and if $0 < p < 1$, then we get a darker one (see [18]).

5.1. Case 1

We consider:

$$\vec{r} = (x_{\sigma_{(1)}}, x_{\sigma_{(2)}}, x_{\sigma_{(3)}}, x_{\sigma_{(4)}}, x_{\sigma_{(5)}}, x_{\sigma_{(6)}}, x_{\sigma_{(7)}}, x_{\sigma_{(8)}});$$

$$\vec{b} = \begin{cases} \left(\frac{x_{\sigma_{(1)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}}, \ldots, \frac{x_{\sigma_{(7)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}}, \frac{x_{\sigma_{(8)}}}{\sum_{i=1}^{8} x_{\sigma_{(i)}}} \right) & \text{if } \sum_{i=1}^{8} x_{\sigma_{(i)}} \neq 0 \\ (0, \ldots, 0) & \text{otherwise.} \end{cases}$$

$a = 0$ and $\frac{1}{p} = 0.30$.

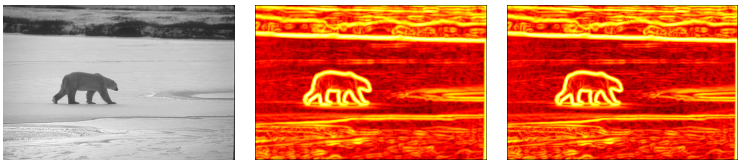
Regarding $\vec{r}$, the highest value , $x_{\sigma(1)}$, is the most relevant for the possible edge, since it corresponds to the biggest intensity jump between the central pixel and its neighbours. With respect to the value of a, we take $a = 0$ because if all the values to be aggregated are null, i.e., if $x_{\sigma(1)} = \cdots = x_{\sigma(8)} = 0$, this means that in the considered 8-neighbourhood all the pixels have the same intensity and hence there is no edge. So, in this case, we should have $G(0, \ldots, 0) = 0$, and from Corollary 8, this is so if and only if $a = 0$. Besides, if $x_{\sigma(1)} = \cdots = x_{\sigma(8)} = 1$, the intensity jump between the pixels in the 8-neighbourhood and the central pixel is as large as possible. So $G(1, \ldots, 1)$ must be equal to 1 and, again from Corollary 8, since $a = 0$, we should require that $b_1 + \cdots + b_n = 1$.

Regarding the calculation of the parameter $1/p$ for constructing the ODM functions, according to Corollary 7, we do as follows: We consider the points $1/p_1^1 = 0.1$, $1/p_1^2 = 0.2$, $1/p_1^3 = 0.3$,..., $1/p_1^9 = 0.9$ (i.e., a uniform partition of the interval $[0, 1]$). We apply Algorithm 1 with each of these nine values and we evaluate the quality of the resulting edge images (in terms of the average of the values of $F_{0.5}$ on the considered training images). Let i_1 be the index such that, if we apply Algorithm 1 with $1/p = 1/p_1^{i_1}$, we get the best result among the nine considered values. Denote $1/p_1^{fix} = p_1^{i_1}$. Next, consider the interval $[1/p_1^{fix} - 0.05, 1/p_1^{fix} + 0.05]$. We take again a uniform partition of this interval with nine points, $1/p_2^1 = 1/p_1^{fix} - 0.04$, $1/p_2^2 = 1/p_1^{fix} - 0.03$, ... ,$1/p_2^9 = 1/p_1^{fix} + 0.04$ and we repeat again the procedure of running and evaluating the results of Algorithm 1 for each of these 9 values of the parameter $1/p$ to get a new point $1/p_2^{fix}$. We repeat the procedure n times, where the new interval around the point $1/p_i^{fix}$ is $[1/p_i^{fix} - 5 \times 10^{-i}, 1/p_i^{fix} + 5 \times 10^{-i}]$ $(i = 2, \ldots, n)$ and we finish when the variation of the average value of the measure $F_{0.5}$ between the images obtained after applying Algorithm 1 with $1/p = 1/p_n^{fix}$ and those obtained with $1/p = 1/p_{n+1}^{fix}$ is smaller than a given tolerance, in our case 0.001.

As a result of following these procedure with the training images in the considered dataset, we get the cited value $1/p = 0.30$. Finally, observe that the expression that we are actually recovering in this situation is $G(\mathbf{x}) = \sum_{i=1}^{8} \frac{x_i^2}{\sum_{i=1}^{8} x_i}$. However, this simple expression is justified by the previous considerations.

5.2. Case 2

In this case we consider:



(a) *Original* (b) *Case 1* (c) *Case 2*

Figure 2: Original image from BSDS [33] (100007) along with feature images obtained after applying Algorithm 1 with Case 1 and case 2.

$$\vec{r} = (x_{\sigma_{(1)}}, x_{\sigma_{(2)}}, x_{\sigma_{(3)}}, x_{\sigma_{(4)}}, x_{\sigma_{(5)}}, x_{\sigma_{(6)}}, x_{\sigma_{(7)}}, x_{\sigma_{(8)}});$$
$$\vec{b} = \left(\frac{1}{8}\left(1 - \left|x_{\sigma_{(1)}} - \underset{i\in\{1,\dots,8\}}{\text{median}}\{x_i\}\right|\right), \dots \right.$$
$$\left. \dots, \frac{1}{8}\left(1 - \left|x_{\sigma_{(8)}} - \underset{i\in\{1,\dots,8\}}{\text{median}}\{x_i\}\right|\right)\right);$$

$a = 0$ and $\frac{1}{p} = 0.30$. The justification for the choice of these parameters is analogous to that in Case 1.

Note that in this case necessarily $a = 0$. Indeed, it suffices to consider the case $x_1 = \dots = x_8$.

In Fig. 2 we show the results obtained by applying Algorithm 1 with the two ODM functions, Case 1 and Case 2, to an original image, Fig. 2a.

6. Consensus feature images

We have seen in the previous sections that, from a given image, we can build different feature images in phase (S2). In this section we discuss a method to build a consensus feature image from the different feature images. This consensus image can be built using aggregation function pixel by pixel, and there are several ways of doing it. Given a set of aggregation functions $M_1, \dots, M_s$, we may:

1. fix one of this aggregation functions, M_i, and apply it to the intensities of the first pixel in every feature image to get the intensity of the first pixel in the consensus image. Then we can apply it again to the intensities of the second pixel in every feature image to get the intensity of the second pixel in the feature image, and we can repeat the process, always with the same M_i, for the intensities of all the other pixels; or we may
2. use the notion of penalty function [5], which allows us to choose the best aggregation (among the s considered ones) for each pixel. In this case, it may happen that, in order to build the intensity of the first pixel in the consensus image, we use an aggregation function M_j, whereas to build the intensity of the second pixel we use an aggregation function M_k different from M_j, and so on.

The advantage of using penalty functions is that the aggregation used to build the intensity of each pixel in the consensus image is that one which provides the least dissimilar result from the intensities of the corresponding pixels in the feature images. So it seems logical to use penalty functions to build the consensus feature image. But the main problem with this method is the choice of the best penalty function. In our experimentation we will use the following expression (see [5]):

$$P_\nabla(\mathbf{X}, \mathbf{y}) = \sum_{q=1}^{m}\sum_{p=1}^{n} |x_p^q - y_q|^2 \tag{2}$$

where we have n feature images to aggregate, each of them with m pixels. In this way, x_p^q denotes the intensity of the pixel q in image p and y_q is the result of aggregating $x_1^q, \ldots, x_n^q$ by means of some of the considered aggregation functions. Finally, we denote $\mathbf{X} = ((x_1^1, x_2^1, \ldots, x_n^1), \ldots, (x_1^m, x_2^m, \ldots, x_n^m))$ and $\mathbf{y} = (y_1, \ldots, y_m)$.

We use Eq. 2 since it is among the most used ones in image comparison, as it is based on the mean squared error; and, moreover, due to the following property.

If the arithmetic mean is among the considered aggregation functions $M_1, \ldots, M_s$ set then the least dissimilar result is obtained applying in every case the arithmetic mean.

To analyze, this property, let us recall the concept of pnalty function in a Cartesian product of lattices, that was deeply studied in [5]. Consider the following goal: *given a set of n numerical values $x_1, \ldots, x_n$ and q averaging aggregation (i.e., between the minimum and the maximum, see [2]) functions $M_1, \ldots, M_q$ penalty functions (see [13, 8, 9]) allow us to select, between the q functions, the one that provides the output least dissimilar to all the inputs. That is, we choose the aggregation functions using a consensus procedure based on testing several functions until we find the one providing the least dissimilar result with respect to the values of the inputs.*

The definition of penalty function in a Cartesian product of lattices reads as follows [9]:

Definition 9. *A function $P_\nabla : ([0,1]^n)^m \times [0,1]^m \to [0, \infty[$ is a penalty function if, for every $\mathbf{X} = (\mathbf{x^1}, \ldots, \mathbf{x^m}) \in ([0,1]^n)^m$ (with $\mathbf{x^i} = (x_1^i, \ldots, x_n^i)$ for every $i \in \{1, \ldots, m\}$) and for every $\mathbf{y} = (y_1, \ldots, y_m) \in [0,1]^m$, it satisfies that:*

1. *$P_\nabla(\mathbf{X}, \mathbf{y}) \geq 0$;*
2. *$P_\nabla(\mathbf{X}, \mathbf{y}) = 0$ if and only if $x_1^i = \cdots = x_n^i = y^i$ for every $i \in \{1, \ldots, m\}$;*
3. *P_∇ is convex in y_i or every $i \in \{1, \ldots, m\}$.*

To understand how the penalty function works, let us consider the following example:

Let us assume that we have three feature images A, B, C, obtained with three different methods in phase (S2).

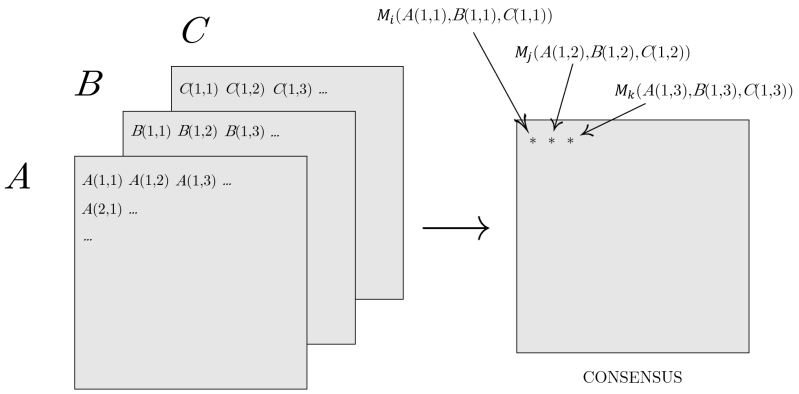


Figure 3: Consensus image.

We want to choose, among a set of aggregation functions $M_1, \ldots, M_s$, the combinations of three aggregation functions M_i, M_j y M_k which provide the smallest value of the considered penalty function.

Looking at Fig. 3, where M_i, M_j and M_k are three aggregation functions, it is clear that inputs can be seen as vectors, so we should use penalty functions defined over a Cartesian product of lattices. That is, functions defined as in Fig. 4, where it is clear that each combination of three aggregation functions M_i, M_j y M_k gives us a value P_{ijk}. so the idea is to determine which combination of three aggregation functions provides the smallest value of P_{ijk} for the considered pixels.

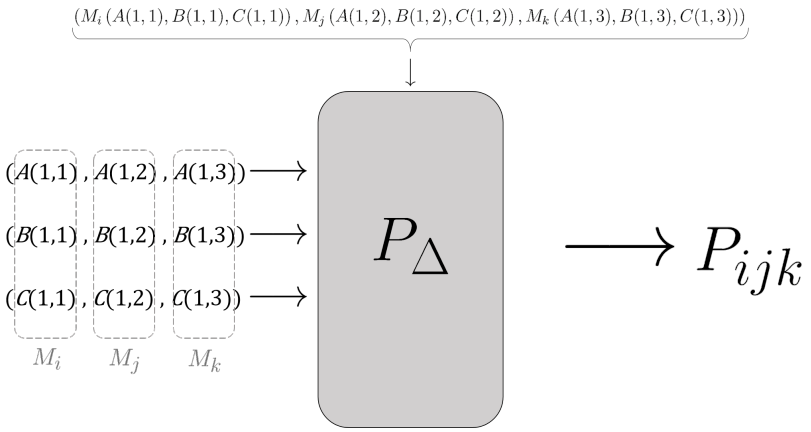


Figure 4: Penalty function

As we have said, for the experimentation we use the following specific expression:

$$P_\nabla(\mathbf{X},\mathbf{y}) = \sum_{q=1}^{m}\sum_{p=1}^{n}|x_p^q - y_q|^2. \tag{3}$$

In the literature, it is used the notation of penalty based function [13] or P-function [8] (for short) to refer to the function which selects the value y which minimizes the value of the penalty function for any $(x_1,\ldots,x_n)$, i.e.,

$$f(x_1,\ldots,x_n) = \underset{y}{\operatorname{argmin}}\ P(x_1,\ldots,x_n,y)$$

if y is the unique minimizer and $y = \frac{(a+b)}{2}$ if the set of minimizers is the interval (a,b) (open or closed). It is also known that in some cases the penalty based function can be expressed analytically, while in other cases it is not possible. For the proposed penalty function of the experimentation given in Eq. (2), we have the following result.

Theorem 10. *Let P be the penalty function given in Eq. (2). The penalty based function of P can be expressed analytically and its expression is the arithmetic mean, i.e., the value of y which minimizes the penalty function is*

$$f(x_1,\ldots,x_n) = P(x_1,\ldots,x_n,\frac{x_1+\ldots+x_n}{n})$$

Proof. It is straight taking into account that the arithmetic mean is the P-function of a given $P(x_1,\ldots,x_n,y) =$

$\sum_{i=1}^{n}(x_i - y)^2$ (see for example [8]). ∎

So, with the considered expression, the least dissimilar result is obtained using in every case the arithmetic mean, i.e., for $\mathbf{y} = (\frac{1}{n}\sum_{i=1}^{n} x_i^1, \ldots, \frac{1}{n}\sum_{i=1}^{n} x_i^m)$.

Taking into account this property and the results in the Annex, in this work we use the arithmetic mean to build the consensus feature image.

We leave for future works the analysis of other possible expressions for the penalty functions, and hence, of different aggregation functions.

7. Experimental framework

In this section, we present the set-up of the experimental framework used to develop the empirical comparison in this work. Given a grayscale image I, we consider it as a matrix of elements (pixels) arranged in rows and columns, where each pixel takes an intensity value in $\{0, 1, \ldots, L-1\}$. A prior step is to normalize the intensities to values belonging to $[0, 1]$. As we have said, our proposal to perform the experiments for a given grayscale image I, considering Bezdek *et al.* [4] includes phases $(S1)$-$(S4)$ described in the introduction. Moreover, we also need to evaluate the quality of the edges that we obtain, so we also consider the following fifth step to compare to ground truth images, i.e,.

$(S5)$ Comparing with ground truth images. To compare the binary edge image I_B with hand-labeled segmentations, getting measurements in terms of *Precision* ($PREC$), *Recall* (REC) and the measure $F_{0.5}$ [29].

7.0.1. The dataset

For our experiments we have used the images of Berkeley Segmentation Dataset (BSDS500) [33], specifically 200 test natural images in grayscale with dimensions of 321×481 or 481×321. Associated to each original image there exist several hand-labeled segmentations (done by humans), which, as we have already said, are denoted as *ground truth images*; usually there are 5 ground truthimages associated to each original image, but this number can vary between 4 and 9 (see Fig. 5). In this sense, the ground truth images are considered as ideal images and they serve to test if the results obtained by an edge detection method are similar, or not, to the ones segmented by a human.

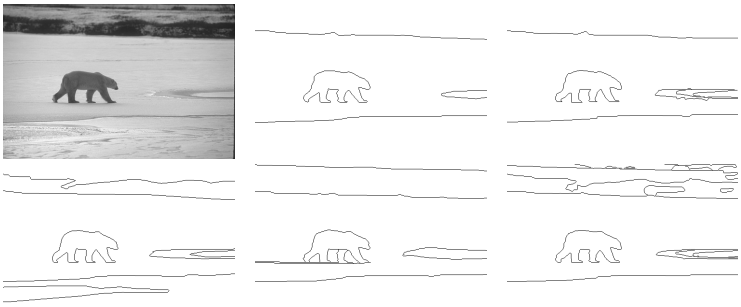


Figure 5: The original image and its five ground truth images (BSDS500).

7.0.2. Edge detection approaches and parameter configuration

Firstly, in step ($S1$) all the approaches perform a Gaussian filter with standard deviation $\sigma_G = 2$. Secondly, to obtain a feature image, ($S2$), we consider the following edge detection approaches: Canny's [15], gravitational approach [28, 30], fuzzy morphology approach [21], Structured Forests method [16] and the one based on ODM functions (Algorithm 1); all these methods have different configurations to be executed, which are listed below:

- Canny [15]. We obtain the magnitude of the gradient estimation with the convolution operator proposed by Canny, with $\sigma_C = 2.25$ which is a common value considering the size of the images [23, 30, 35]. This σ_C is different from that required for calculating the Gaussian filter, i.e., σ_G. Note that the unsupervised calculation of an optimal value of σ_C has not been solved so far. We denote this approach by C.

- Gravitational [28, 30]. This approach is based on gravitational forces using relief functions where, as particular cases, t-norms and t-conorms are used. In our experiments, we denote by G_{S_P} and G_{S_M}, when the probabilistic sum ($S_P(x, y) = x + y - xy$) and the maximum ($S_M(x, y) = \vee(x, y)$) are considered as t-conorms, respectively.

- Fuzzy Morphology [21, 22]. The authors proposed a generalization of the morphological operators based on considering general t-conorms and t-norms in erosion and dilation definitions. We have carried out in our experiments the best configurations claimed in [21, 22]: for erosion the Kleene-Dienes implication operator (Eq. (4)) and for dilation, the nilpotent minimum as t-norm (Eq. (5)).

$$I_{KD}(x, y) = \vee(1 - x, y); \tag{4}$$

$$T_{nM}(x, y) = \begin{cases} 0, & \text{if } x + y \leq 1, \\ \wedge(x, y) & \text{otherwise.} \end{cases} \tag{5}$$

  Considering the pair (I_{KD}, T_{nM}) we use the best two parameterizations given in [21, 22]:

  - FM_{SS}. The Schweizer-Sklar t-norm (T_λ^{SS}) [39] is applied to Eq. (4) to get the fuzzy erosion and for fuzzy dilation the Schweizer-Sklar t-conorm (defined as the dual of Schweizer-Sklar t-norm) is applied to Eq. (5).

$$T_\lambda^{SS}(x, y) = \begin{cases} \wedge(x, y), & \text{if } \lambda = -\infty, \\ xy & \text{if } \lambda = 0, \\ T_D(x, y) & \text{if } \lambda = +\infty, \\ (\vee(x^\lambda + y^\lambda - 1, 0))^{\frac{1}{\lambda}} & \text{otherwise.} \end{cases}$$

    where $T_D(x, y) = \begin{cases} 0, & \text{if } x, y \in [0, 1), \\ \wedge(x, y) & \text{otherwise.} \end{cases}$.

    We have taken the value of $\lambda = -5$ given in [22] as best result.

- Structured Forests [16]. In this approach the authors propose an edge detector learned from information of local image patches, using Random Decision Forests. Originally the method works with *RGB* color images, so we have trained the model for computing edges in grayscale with the configuration parameters exposed in [16]. We denote this approach by *SF*.

- The two edge detector introduced in the next Section, *Case 1* (*C1*) and *Case 2* (*C2*), built from ODM functions.

Anew for all methods, to thin the feature image, ($S3$), we compute prior orientations and subsequently NMS process, both by Kovesis' functions [27]. We finish binarizing the thinned image, ($S4$), [15, 36].

7.0.3. Comparison method

It is well-known that the procedure to evaluate the performance of an edge detector is an open problem [29]. In this paper, to make the comparison, ($S5$), we have applied the approach given by Martin *et al.* [34]. This approximation considers that we are dealing with a binary classification problem (each pixel is classified as an element of the edge or is not classified as an element of the edge) and compares the output image given by an edge detector method with the ones generated by humans (ground truth images). To do so, a confusion matrix is performed.

In this work we use $F_{0.5}$.

Thereupon, bearing in mind all previous indications, for each image in BSDS500, the measures $PREC$, REC and $F_{0.5}$ of a given edge detection method are those produced in the comparison with the ground truth images for which $F_{0.5}$ is maximal. As we have previously stated, there exist a number of ground truth images for each original image, so in our experiments (see also [30]) we compare the solution given by any edge detection method with all of the ground truth images. Then, the triplet $(PREC, REC, F_{0.5})$ having the greatest $F_{0.5}$ is considered as the evaluation of the detection method for that image, i.e., for us, the highest $F_{0.5}$ value means that the solution obtained by the edge detection method is the closest to a human, so, in this sense, is the best solution.

It is necessary to remark that we have also used a one-to-one pixel matching algorithm to map the edge pixels in the output edge image (generated by an edge detection method) and the ground truth ones. This matching allows some spatial tolerance (in our case, as much as 2.5% of the diagonal of the image), so that an edge pixel can be slightly displaced from its true position, yet being considered as correctly classified.

In order to do the pixel-to-pixel matching, we use the method presented by Estrada and Jepson [19] (the implementation can be found in [26]).

8. Experimental results

In this section we provide the results obtained by our new proposal, as an individual feature extractor and as a consensus feature image, combined with other feature extraction techniques.

8.1. Case 1 and Case 2 vs. the other methods

In Table 1 the results of each edge detection method are indicated displaying the average of $PREC$, REC and $F_{0.5}$ for the 200 test set images. In terms of REC we can deduce that we have obtained better results than the Canny method with all of our two methods (*Case 1*, *Case 2*), i.e., not including a lot of false positives. On the other hand we may observe that FM_{SS} combines a medium precision with a very high recall, therefore the majority of edges are detected at the cost of including a high number of false positives. In general, the best results are when $PREC$ and REC have similar values, that is, considering $F_{0.5}$ as an overall quality measure. In this case, the results achieved with *Case 1* and *Case 2* are competitive with the ones obtained with Canny and gravitational forces. Particularly, we obtain better results than the ones with Canny and similar to those using the gravitational edge detector. If we focus on the results obtained by the Structured Forest method, it obtains the highest values in terms of $F_{0.5}$, this is due to the high value of $PREC$ as it detects more edges than the rest of the methods although having a lowest REC. So, the results obtained with Case 1 and with Case 2 are better than those obtained with Canny, Fuzzy Morphology, Gravitational with the maximum and Gravitational with the probabilistic sum, but worse than those obtained with SF.

Next, in Table 2 we show the number of images in the dataset for which it is the best or worst performer (in terms of $F_{0.5}$).

We observe that *C1* and *C2* have the lowest values in terms of worst images. Moreover, the number of images where we are the best result outperforms the ones obtained by all the other methods except Structured Forest, which remains the best performer.

Method	$PREC$	REC	$F_{0.5}$
$C1$	0.579	**0.794**	0.653
$C2$	0.602	0.765	0.654
FM_{SS}	0.572	0.719	0.615
C	0.687	0.618	0.631
G_{S_P}	0.649	0.649	0.650
G_{S_M}	0.661	0.665	0.641
SF	**0.753**	0.645	**0.682**

Table 1: Comparison of ODM functions approach with other edge detection methods as Gravitational, Fuzzy Morphology, Structured Forest and Canny in terms of $PREC$, REC and $F_{0.5}$.

	Method											
	$*$		FM_{SS}		C		G_{S_P}		G_{S_M}		SF	
	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗
$C1$	44	17	14	86	6	39	29	**16**	21	19	**86**	23
$C2$	50	**9**	16	89	6	39	27	20	16	16	**86**	27

Table 2: Comparison of best (✓) and worst (✗) approaches for 200 images of (BSDS500) in terms of $F_{0.5}$.

8.2. Consensus edge detector vs. the other methods

In this subsection we are going to build combinations without repetition of the feature images obtained with the seven edge detection methods that we have considered in this work (Case 1, Case 2, Canny, Gravitational with the probabilistic sum, Gravitational with the maximum, Fuzzy Morphology and SF). Specifically, we are going to consider combinations of two (21 combinations), three (35 combinations), four (35 combinations) and five (21 combinations) feature images. For each of the considered combinations, we calculate the value of the measures $PREC$, REC and $F_{0.5}$ (Tables 3- 6).

8.2.1. Combinations of two feature images

First, in Table 3 we consider combinations of two feature images obtained with two different edge detectors. Since in this work we consider 7 edge detectors, we have 21 possible combinations, which appear in the first column of the Table. For each of these combinations, we calculate the values of the PREC, REC and $F_{0.5}$ measures. we observe that the combination of Case 1 with SF provides the best result (in terms of the measure $F_{0.5}$). Moreover, the $F_{0.5}$ value that we get for this combination of Case 1 and SF is better than the $F_{0.5}$ value obtained for SF individually (see Table 1).

8.2.2. Combinations of three feature images

Next, we consider combinations of three feature images obtained with three different edge detectors. The 35 possible combinations appear in the first column of Table 4 and for each of them we make the same study for the PREC, REC and $F_{0.5}$ values. In this case, we observe that the best result in terms of the measure $F_{0.5}$ is the combination of Case 1 with SF and the detector based on fuzzy morphology. But the $F_{0.5}$ value that we obtain, although better that the corresponding one for SF in Table 1, is worse than the one for the combination of Case 1 and SF considered in Table 3.

8.2.3. Combinations of more than three feature images

Finally we study combinations of four (Table 5) and five (Table 6) feature images. In both cases, we see that the best result according to $F_{0.5}$ is obtained from the combination of our cases, SF and some of the other methods, but the resulting $F_{0.5}$ value, although bigger than the one of SF in Table 1, does not overcome neither the combination of Case 1 with SF nor the combination of Case 1 with SF and the detector

based on fuzzy morphology. Even more, the best combination in Table 6 gets a worst result (according to $F_{0.5}$) than the best combination in Table 5.

8.3. Experimental results in terms of the best and worst performance

In Tables VII-X we compare the different combinations of methods with the individual methods. We study for how many images from the set of 200 each method provides the best and the worst results. The first column contains the different combinations of methods. In the row corresponding to each of these methods the 200 studied images are considered. In the column marked with * it is displayed in how many of the 200 images the method to which the row corresponds provides the best (tick) and the worst (x) results. In the other columns it is displayed the number of images among the 200 in which the method at the heading of the column provides the best and the worst results.

8.3.1. Combinations of two feature images

In Table VII we observe that the combination of Case 1 with SF and the combination of Case 2 with SF provide extremely competitive results, much better than those obtained with each method individually, including SF.

8.3.2. Combinations of three feature images

Regarding the number of images in which the best and the worst results are obtained by combinations of three methods, the result is even more relevant in Table VIII than in VII, since the combination of Cases 1 and 2 with SF provides the best result in 94 images and the worst in just 1.

8.3.3. Combinations of more than three feature images

In Tables IX and X we see that the best results are obtained by combination in which Cases 1 and 2 always appear, and moreover, such results are always very competeitive. Furthermore, in all the cases the combination of methods provides the worst result in a very small number of cases.

8.4. Note about the time complexity

In Algorithm 1, regardless we use $C1$ or $C2$ we have a complexity of $O(N)$, with N the number of pixels in an image. In Algorithm 2 we only use the arithmetic mean, so we also have lineal complexity $O(N)$.

If we consider the complete process until calculating the $PREC$, REC and $F_{0.5}$ values, this is equivalent for all the compared methods, so the time complexities are equivalent. But it is worth to mention that the highest cost in our proposal corresponds to the calculation of the optimal value for $1/p$. To do so, we must run Algorithm 1 for each considered value of $1/p$. So, if we consider N values of $1/p$, we have a time complexity of $O(N^2)$. But this calculation is done only once, in the training phase.

8.5. Conclusions from the experimental study

From the experimental study that we have done, we arrive at the following conclusions:

- The combination of methods improves the results obtained by each method separately.
- As more methods are combined, the results for the best combination (according to the $F_{0.5}$ measure) worsen.
- The best results correspond always to combinations in which Case1 or Case 2 (or both) and SF appear.

These results justify the introduction and the study of Case 1 and Case 2. So we propose to use always combinations of two methods. We leave for a future work the analysis of the causes of this improvement in the performance when different methods are combined.

Method	$PREC$	REC	$F_{0.5}$
C1-C2	0.586	**0.785**	0.654
C1-Canny	0.661	0.682	0.652
C1-FM_{SS}	0.582	0.784	0.651
C1-G_{S_P}	0.620	0.749	0.660
C1-G_{S_M}	0.625	0.731	0.654
C1-SF	0.715	0.724	**0.705**
C2-Canny	0.668	0.669	0.650
C2-FM_{SS}	0.602	0.754	0.649
C2-G_{S_P}	0.628	0.731	0.657
C2-G_{S_M}	0.635	0.715	0.653
C2-SF	0.720	0.710	0.701
Canny-FM_{SS}	0.675	0.650	0.644
Canny-G_{S_P}	0.677	0.666	0.651
Canny-G_{S_M}	0.674	0.648	0.641
Canny-SF	**0.728**	0.671	0.683
FM_{SS}*-*G_{S_P}	0.624	0.734	0.655
FM_{SS}*-*G_{S_M}	0.631	0.711	0.649
FM_{SS}*-SF*	0.722	0.708	0.701
G_{S_P}*-*G_{S_M}	0.654	0.687	0.650
G_{S_P}*-SF*	0.722	0.688	0.687
G_{S_M}*-SF*	0.725	0.673	0.681

Table 3: Comparison of penalty functions combining two feature images in terms of $PREC$, REC and $F_{0.5}$.

9. Conclusions and Future Research

As a conclusion for this paper, we can make the following remarks. The goal of evaluating the validity of ODM functions, used to measure the changes of intensity between a pixel and its neighbours taking into account the direction defined by the vector obtained by ordering in a decreasing way such intensity changes, has been achieved, so such functions can be used to determine the existence of an edge. Furthermore, using different ODM functions in phase (S2) and bearing in mind $F_{0.5}$ measure,we can conclude that, among the considered cases, Case 1 and Case 2 individually are only overcome by SF. This justifies the consideration of the new method to build ODM functions and the introduction of both Case 1 and Case 2.

Furthermore, if we consider combinations of methods, the combinations of Case 1, Case 2 or both with SF overcome the results obtained by any of the methods individually. But the more combined methods, the worse the results are, so we consider that the best option is to combine Case 1 or Case 2 with SF.

With respect to future research lines, we consider that it is necessary to make a study for optimizing the parameters in Eq. 1, both in a general way and with specific types of images.

As we have already mentioned in the introduction, our proposal is outperformed by deep learning methods. But we would like to insist once more on the simplicity and interpretability of our method, in particular when the feature image is built using the expression in Eq. (1). Note that, contrary to the case of deep neural networks, no training is need in our case, and furthermore, the computational cost of our algorithms is very low compared to that of deep learning procedures. Nevertheless, in future works we intend to study the possible combination of our algorithms and those based in deep learning techniques.

[1] G. Beliakov, A. Pradera and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, ser. Studies In Fuzziness and Soft Computing, Berlin, Germany: Springer-Verlag, 2007.
[2] G. Beliakov, H. Bustince, T. Calvo, *A practical Guide to Averaging Functions*, Springer, Heidelberg, 2016.
[3] H. Benninghoff, H. Garcke, "Image Segmentation Using Parametric Contours With Free Endpoints", *IEEE Transactions on Image Processing*, vol. 25, no. 4, pp. 1639–1648, 2016.

Method	$PREC$	REC	$F_{0.5}$
$C1$-$C2$-$Canny$	0.648	0.704	0.656
$C1$-$C2$-FM_{SS}	0.591	**0.779**	0.654
$C1$-$C2$-G_{S_P}	0.614	0.758	0.661
$C1$-$C2$-G_{S_M}	0.620	0.741	0.656
$C1$-$C2$-SF	0.699	0.738	0.702
$C1$-$Canny$-FM_{SS}	0.654	0.693	0.655
$C1$-$Canny$-G_{S_P}	0.656	0.701	0.659
$C1$-$Canny$-G_{S_M}	0.656	0.686	0.652
$C1$-$Canny$-SF	0.710	0.700	0.689
$C1$-FM_{SS}-G_{S_P}	0.615	0.757	0.660
$C1$-FM_{SS}-G_{S_M}	0.620	0.737	0.654
$C1$-FM_{SS}-SF	0.701	0.737	**0.704**
$C1$-G_{S_P}-G_{S_M}	0.638	0.723	0.658
$C1$-G_{S_P}-SF	0.695	0.725	0.693
$C1$-G_{S_M}-SF	0.701	0.706	0.687
$C2$-$Canny$-FM_{SS}	0.660	0.683	0.653
$C2$-$Canny$-G_{S_P}	0.660	0.692	0.656
$C2$-$Canny$-G_{S_M}	0.664	0.674	0.649
$C2$-$Canny$-SF	0.713	0.692	0.687
$C2$-FM_{SS}-G_{S_P}	0.621	0.746	0.659
$C2$-FM_{SS}-G_{S_M}	0.626	0.728	0.653
$C2$-FM_{SS}-SF	0.704	0.726	0.700
$C2$-G_{S_P}-G_{S_M}	0.643	0.713	0.656
$C2$-G_{S_P}-SF	0.700	0.717	0.691
$C2$-G_{S_M}-SF	0.707	0.700	0.686
$Canny$-FM_{SS}-G_{S_P}	0.664	0.689	0.657
$Canny$-FM_{SS}-G_{S_M}	0.665	0.670	0.648
$Canny$-FM_{SS}-SF	**0.719**	0.691	0.689
$Canny$-G_{S_P}-G_{S_M}	0.673	0.668	0.650
$Canny$-G_{S_P}-SF	0.714	0.680	0.679
$Canny$-G_{S_M}-SF	**0.719**	0.663	0.672
FM_{SS}-G_{S_P}-G_{S_M}	0.640	0.714	0.655
FM_{SS}-G_{S_P}-SF	0.704	0.716	0.693
FM_{SS}-G_{S_M}-SF	0.712	0.696	0.688
G_{S_P}-G_{S_M}-SF	0.711	0.682	0.678

Table 4: Comparison of penalty functions combining three feature images in terms of $PREC$, REC and $F_{0.5}$.

Method	$PREC$	REC	$F_{0.5}$
$C1$-$C2$-$Canny$-FM_{SS}	0.644	0.711	0.657
$C1$-$C2$-$Canny$-G_{S_P}	0.647	0.719	0.662
$C1$-$C2$-$Canny$-G_{S_M}	0.650	0.699	0.654
$C1$-$C2$-$Canny$-SF	0.702	0.709	0.690
$C1$-$C2$-FM_{SS}-G_{S_P}	0.611	**0.763**	0.661
$C1$-$C2$-FM_{SS}-G_{S_M}	0.617	0.744	0.655
$C1$-$C2$-FM_{SS}-SF	0.694	0.739	**0.700**
$C1$-$C2$-G_{S_P}-G_{S_M}	0.629	0.738	0.660
$C1$-$C2$-G_{S_P}-SF	0.684	0.738	0.693
$C1$-$C2$-G_{S_M}-SF	0.685	0.724	0.688
$C1$-$Canny$-FM_{SS}-G_{S_P}	0.649	0.714	0.661
$C1$-$Canny$-FM_{SS}-G_{S_M}	0.651	0.698	0.654
$C1$-$Canny$-FM_{SS}-SF	0.705	0.709	0.691
$C1$-$Canny$-G_{S_P}-G_{S_M}	0.656	0.698	0.657
$C1$-$Canny$-G_{S_P}-SF	0.697	0.707	0.685
$C1$-$Canny$-G_{S_M}-SF	0.704	0.687	0.678
$C1$-FM_{SS}-G_{S_P}-G_{S_M}	0.629	0.737	0.659
$C1$-FM_{SS}-G_{S_P}-SF	0.683	0.738	0.694
$C1$-FM_{SS}-G_{S_M}-SF	0.688	0.721	0.687
$C1$-G_{S_P}-G_{S_M}-SF	0.688	0.715	0.683
$C2$-$Canny$-FM_{SS}-G_{S_P}	0.654	0.706	0.660
$C2$-$Canny$-FM_{SS}-G_{S_M}	0.655	0.688	0.652
$C2$-$Canny$-FM_{SS}-SF	0.707	0.703	0.689
$C2$-$Canny$-G_{S_P}-G_{S_M}	0.662	0.689	0.655
$C2$-$Canny$-G_{S_P}-SF	0.703	0.698	0.683
$C2$-$Canny$-G_{S_M}-SF	**0.709**	0.680	0.676
$C2$-FM_{SS}-G_{S_P}-G_{S_M}	0.631	0.731	0.658
$C2$-FM_{SS}-G_{S_P}-SF	0.690	0.729	0.693
$C2$-FM_{SS}-G_{S_M}-SF	0.697	0.711	0.687
$C2$-G_{S_P}-G_{S_M}-SF	0.693	0.708	0.682
$Canny$-FM_{SS}-G_{S_P}-G_{S_M}	0.662	0.687	0.654
$Canny$-FM_{SS}-G_{S_P}-SF	0.707	0.696	0.685
$Canny$-FM_{SS}-G_{S_M}-SF	**0.709**	0.678	0.676
$Canny$-G_{S_P}-G_{S_M}-SF	0.705	0.678	0.673
FM_{SS}-G_{S_P}-G_{S_M}-SF	0.696	0.707	0.683

Table 5: Comparison of penalty functions combining four feature images in terms of $PREC$, REC and $F_{0.5}$.

Method	$PREC$	REC	$F_{0.5}$
$C1$-$C2$-$Canny$-FM_{SS}-G_{S_P}	0.642	0.726	0.663
$C1$-$C2$-$Canny$-FM_{SS}-G_{S_M}	0.645	0.706	0.655
$C1$-$C2$-$Canny$-FM_{SS}-SF	0.695	0.715	0.689
$C1$-$C2$-$Canny$-G_{S_P}-G_{S_M}	0.651	0.707	0.658
$C1$-$C2$-$Canny$-G_{S_P}-SF	0.688	0.718	0.685
$C1$-$C2$-$Canny$-G_{S_M}-SF	0.690	0.702	0.678
$C1$-$C2$-FM_{SS}-G_{S_P}-G_{S_M}	0.625	0.742	0.660
$C1$-$C2$-FM_{SS}-G_{S_P}-SF	0.675	**0.744**	**0.692**
$C1$-$C2$-FM_{SS}-G_{S_M}-SF	0.678	0.733	0.687
$C1$-$C2$-G_{S_P}-G_{S_M}-SF	0.676	0.724	0.682
$C1$-$Canny$-FM_{SS}-G_{S_P}-G_{S_M}	0.650	0.709	0.659
$C1$-$Canny$-FM_{SS}-G_{S_P}-SF	0.689	0.718	0.686
$C1$-$Canny$-FM_{SS}-G_{S_M}-SF	0.694	0.699	0.679
$C1$-$Canny$-G_{S_P}-G_{S_M}-SF	0.692	0.698	0.676
$C1$-FM_{SS}-G_{S_P}-G_{S_M}-SF	0.679	0.726	0.685
$C2$-$Canny$-FM_{SS}-G_{S_P}-G_{S_M}	0.654	0.701	0.658
$C2$-$Canny$-FM_{SS}-G_{S_P}-SF	0.694	0.710	0.685
$C2$-$Canny$-FM_{SS}-G_{S_M}-SF	0.699	0.692	0.678
$C2$-$Canny$-G_{S_P}-G_{S_M}-SF	0.697	0.691	0.675
$C2$-FM_{SS}-G_{S_P}-G_{S_M}-SF	0.685	0.717	0.684
$Canny$-FM_{SS}-G_{S_P}-G_{S_M}-SF	**0.700**	0.691	0.677

Table 6: Comparison of penalty functions combining five feature images in terms of $PREC$, REC and $F_{0.5}$.

[4] J. Bezdek, R. Chandrasekhar, Y. Attikouzel, "A geometric approach to edge detection," *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 52–75, 1998.

[5] H. Bustince, E. Barrenechea, T. Calvo, S. James, G. Beliakov, "Consensus in multi-expert decision making problems using penalty functions defined over a Cartesian product of lattices," *Information Fusion*, vol. 17, pp. 56–64, 2014.

[6] H. Bustince, E. Barrenechea, M. Pagola, J. Fernandez , "Interval-valued fuzzy sets constructed from matrices: Application to edge detection," *Fuzzy Sets and Systems* , vol. 160, no. 13, pp. 1819–1840, 2009.

[7] H. Bustince, E. Barrenechea, M. Sesma-Sara, J. Lafuente, R. Mesiar, A. Kolesárová, "Ordered directionally monotone functions. Justification and application," *IEEE Transactions on Fuzzy Systems*, vol. 26, no. 4, pp. 2237–2250, 2018.

[8] H. Bustince, G. Beliakov, G. Pereira Dimuro, B. Bedregal, R. Mesiar , "On the definition of penalty functions in data aggregation," *Fuzzy Sets and Systems*, vol. 323, no. 15, pp. 1–18, 2017.

[9] H. Bustince, E. Barrenechea, T. Calvo, S. James, G. Beliakov , "Consensus in multi-expert decision making problems using penalty functions defined over a Cartesian product of lattices," *Information Fusion*, vol. 17, pp. 56–64, 2014.

[10] H. Bustince, M. Pagola, E. Barrenechea, J. Fernandez, P. Melo-Pinto, P. Couto, H. Tizhoosh, and J. Montero, "Ignorance functions. an application to the calculation of the threshold in prostate ultrasound images," *Fuzzy Sets and Systems*, vol. 161, no. 1, pp. 20 – 36, 2010

[11] H. Bustince, J. Fernandez, A. Kolesárová, R. Mesiar , "Directional monotonicity of fusion functions," *European Journal of Operational Research*, vol. 244, no. 1, pp. 300–308, 2015.

[12] T. Calvo, A. Kolesárová, M. Komorníková and R. Mesiar, "Aggregation operators: properties, classes and construction methods," in *Aggregation Operators New Trends and Applications*, T. Calvo, G. Mayor and R. Mesiar, Eds. Physica-Verlag, Heidelberg, 2002, pp. 3-104.

[13] T. Calvo, G. Beliakov, "Aggregation functions based on penalties," *Fuzzy Sets and Systems*, vol. 161, no. 10, pp. 1420–1436, 2010.

[14] J. Canny, (1983). "Finding edges and lines in images," Tech. Rep., Massachusetts Institute of Technology, Cambridge, MA, USA, 1983.

[15] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, 1986.

[16] P. Dollar and C. L. Zitnick, "Fast edge detection using structured forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 8, pp. 1558–1570, Aug 2015.

[17] A. Ferraz, C. Mallet, and N. Chehata, "Large-scale road detection in forested mountainous areas using airborne topographic lidar data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 112, pp. 23 – 36, 2016.

[18] M.G. Forero, "Fuzzy thresholding and histogram analysis," in *Fuzzy Filters for Image Processing*, M. Nachtegael, D. Van

	Method											
	$*$		FM_{SS}		C		G_{S_P}		G_{S_M}		SF	
	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗
C1-C2	42	**11**	16	88	6	39	31	20	18	16	**87**	26
C1-Canny	18	**2**	26	94	6	39	41	21	18	16	**91**	28
C1-FM_{SS}	43	**15**	11	87	6	39	34	19	20	16	**86**	24
C1-G_{S_P}	43	**6**	22	93	7	39	18	18	19	16	**91**	28
C1-G_{S_M}	30	**4**	22	93	6	39	33	21	17	15	**92**	28
C1-SF	**97**	**1**	20	96	4	39	26	21	15	16	38	27
C2-Canny	15	**1**	28	95	6	39	42	21	18	16	**91**	28
C2-FM_{SS}	38	**11**	14	88	6	39	37	19	17	15	**88**	28
C2-G_{S_P}	38	**3**	25	95	7	39	22	19	18	16	**90**	28
C2-G_{S_M}	26	**2**	24	94	6	39	36	21	15	16	**93**	28
C2-SF	**96**	**1**	20	96	4	39	29	21	15	16	36	27
Canny-FM_{SS}	16	**1**	29	96	3	38	42	21	20	16	**90**	28
Canny-G_{S_P}	15	**3**	28	96	7	38	37	20	19	15	**94**	28
Canny-G_{S_M}	9	**5**	28	96	7	37	43	21	19	13	**94**	28
Canny-SF	52	**2**	27	95	5	39	36	20	16	16	**64**	28
FM_{SS}*-*G_{S_P}	42	**9**	23	91	6	38	19	18	19	16	**91**	28
FM_{SS}*-*G_{S_M}	24	**4**	22	93	6	39	39	21	17	15	**92**	28
FM_{SS}*-SF*	**98**	**1**	20	96	4	39	28	21	16	16	34	27
G_{S_P}*-*G_{S_M}	18	**6**	26	95	7	38	37	19	17	14	**95**	28
G_{S_P}*-SF*	61	**6**	28	95	5	38	28	18	16	16	**62**	27
G_{S_M}*-SF*	48	**3**	29	95	5	39	36	20	15	16	**67**	27

Table 7: Comparison of best (✓) and worst (✗) approaches in terms of $F_{0.5}$ considering penalty approaches with two feature images.

	Method											
	*		FM_{SS}		C		G_{S_P}		G_{S_M}		SF	
	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗
C1-C2-Canny	30	**1**	21	95	6	39	38	21	16	16	**89**	28
C1-C2-FM_{SS}	41	**10**	14	90	6	39	33	19	18	16	**88**	26
C1-C2-G_{S_P}	44	**2**	22	95	6	39	22	20	17	16	**89**	28
C1-C2-G_{S_M}	34	**4**	22	92	6	39	34	21	17	16	**87**	28
C1-C2-SF	**93**	**2**	19	96	5	39	25	21	11	16	47	26
C1-Canny-FM_{SS}	22	**1**	24	95	6	39	40	21	17	16	**91**	28
C1-Canny-G_{S_P}	28	**3**	27	95	6	38	30	20	17	16	**92**	28
C1-Canny-G_{S_M}	22	**1**	26	96	6	39	37	21	17	15	**92**	28
C1-Canny-SF	**63**	**1**	26	95	5	39	30	21	13	16	63	28
C1-FM_{SS}*-*G_{S_P}	46	**11**	22	90	7	38	19	18	20	16	**86**	27
C1-FM_{SS}*-*G_{S_M}	33	**6**	20	92	6	39	32	21	18	14	**91**	28
C1-FM_{SS}*-SF*	**95**	**1**	19	96	3	39	23	21	13	16	47	27
C1-G_{S_P}*-*G_{S_M}	34	**3**	25	95	7	38	25	21	17	15	**92**	28
C1-G_{S_P}*-SF*	**85**	**0**	24	96	5	39	13	21	15	16	58	28
C1-G_{S_M}*-SF*	**67**	**1**	22	95	6	39	29	21	11	16	65	28
C2-Canny-FM_{SS}	20	**1**	25	95	5	39	41	21	18	16	**91**	28
C2-Canny-G_{S_P}	21	**2**	27	95	6	38	31	21	20	16	**95**	28
C2-Canny-G_{S_M}	15	**1**	28	95	7	39	41	21	18	16	**91**	28
C2-Canny-SF	52	**3**	27	94	6	39	36	20	16	16	**63**	28
C2-FM_{SS}*-*G_{S_P}	39	**6**	23	93	7	38	22	19	19	16	**90**	28
C2-FM_{SS}*-*G_{S_M}	25	**4**	21	94	6	39	38	21	18	14	**92**	28
C2-FM_{SS}*-SF*	**88**	**1**	16	96	3	39	29	21	15	16	49	27
C2-G_{S_P}*-*G_{S_M}	30	**3**	24	96	7	38	30	21	17	14	**92**	28
C2-G_{S_P}*-SF*	**81**	**0**	24	96	5	39	20	21	15	16	55	28
C2-G_{S_M}*-SF*	62	**1**	24	95	6	39	31	21	12	16	**65**	28
Canny-FM_{SS}*-*G_{S_P}	25	**2**	27	95	6	39	29	20	21	16	**92**	28
Canny-FM_{SS}*-*G_{S_M}	18	**3**	28	94	6	39	40	21	17	15	**91**	28
Canny-FM_{SS}*-SF*	**60**	**2**	26	95	5	39	33	21	16	16	60	27
Canny-G_{S_P}*-*G_{S_M}	16	**5**	27	94	7	38	39	20	17	15	**94**	28
Canny-G_{S_P}*-SF*	44	**3**	28	95	4	38	34	20	14	16	**76**	28
Canny-G_{S_M}*-SF*	37	**5**	28	93	4	39	39	20	12	16	**80**	27
FM_{SS}*-*G_{S_P}*-*G_{S_M}	28	**5**	25	95	7	38	30	21	16	13	**94**	28
FM_{SS}*-*G_{S_P}*-SF*	**85**	**2**	25	95	5	39	17	20	15	16	53	28
FM_{SS}*-*G_{S_M}*-SF*	62	**1**	23	95	5	39	34	21	13	16	**63**	28
G_{S_P}*-*G_{S_M}*-SF*	43	**2**	29	95	6	38	31	21	15	16	**76**	28

Table 8: Comparison of best (✓) and worst (✗) approaches in terms of $F_{0.5}$ considering penalty approaches with three feature images.

	Method											
	*		FM_{SS}		C		G_{S_P}		G_{S_M}		SF	
	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗
C1-C2-Canny-FM_{SS}	29	**0**	22	96	6	39	36	21	18	16	**89**	28
C1-C2-Canny-G_{S_P}	34	**1**	24	96	6	38	31	21	16	16	**89**	28
C1-C2-Canny-G_{S_M}	26	**2**	25	94	6	39	35	21	16	16	**92**	28
C1-C2-Canny-SF	**71**	**1**	23	95	4	39	28	21	13	16	61	28
C1-C2-FM_{SS}-G_{S_P}	46	**6**	19	93	7	38	22	19	18	16	**88**	28
C1-C2-FM_{SS}-G_{S_M}	36	**3**	18	94	6	39	34	21	17	15	**89**	28
C1-C2-FM_{SS}-SF	**97**	**3**	14	95	3	39	22	21	12	16	52	26
C1-C2-G_{S_P}-G_{S_M}	33	**2**	25	95	7	38	26	21	18	16	**91**	28
C1-C2-G_{S_P}-SF	**88**	**1**	20	96	5	39	11	21	16	16	60	27
C1-C2-G_{S_M}-SF	68	**0**	18	96	6	39	27	21	11	16	**70**	28
C1-Canny-FM_{SS}-G_{S_P}	35	**3**	23	95	7	38	27	20	17	16	**91**	28
C1-Canny-FM_{SS}-G_{S_M}	27	**1**	23	96	7	39	37	21	16	15	**90**	28
C1-Canny-FM_{SS}-SF	**70**	**1**	22	95	5	39	27	21	16	16	60	28
C1-Canny-G_{S_P}-G_{S_M}	24	**3**	25	95	7	38	32	20	19	16	**93**	28
C1-Canny-G_{S_P}-SF	65	**1**	25	95	5	39	22	21	15	16	**68**	28
C1-Canny-G_{S_M}-SF	47	**1**	25	95	6	39	35	21	14	16	**73**	28
C1-FM_{SS}-G_{S_P}-G_{S_M}	37	**2**	24	95	7	38	26	21	18	16	**88**	28
C1-FM_{SS}-G_{S_P}-SF	**86**	**1**	18	96	5	39	14	21	15	16	62	27
C1-FM_{SS}-G_{S_M}-SF	**70**	**1**	17	95	6	39	29	21	10	16	68	28
C1-G_{S_P}-G_{S_M}-SF	62	**0**	25	96	6	39	21	21	13	16	**73**	28
C2-Canny-FM_{SS}-G_{S_P}	24	**2**	25	95	6	38	33	21	19	16	**93**	28
C2-Canny-FM_{SS}-G_{S_M}	22	**2**	26	95	6	39	39	21	16	15	**91**	28
C2-Canny-FM_{SS}-SF	**68**	**2**	22	94	5	39	30	21	16	16	59	28
C2-Canny-G_{S_P}-G_{S_M}	22	**3**	25	95	7	38	35	20	17	16	**94**	28
C2-Canny-G_{S_P}-SF	63	**2**	27	95	5	38	20	21	14	16	**71**	28
C2-Canny-G_{S_M}-SF	40	**1**	26	95	5	39	34	21	14	16	**81**	28
C2-FM_{SS}-G_{S_P}-G_{S_M}	31	**3**	23	95	7	38	30	20	17	16	**92**	28
C2-FM_{SS}-G_{S_P}-SF	**87**	**1**	22	95	5	39	14	21	13	16	59	28
C2-FM_{SS}-G_{S_M}-SF	67	**1**	19	95	6	39	29	21	11	16	**68**	28
C2-G_{S_P}-G_{S_M}-SF	62	**1**	25	96	6	38	23	21	13	16	**71**	28
Canny-FM_{SS}-G_{S_P}-G_{S_M}	22	**3**	26	94	7	38	33	21	18	16	**94**	28
Canny-FM_{SS}-G_{S_P}-SF	65	**1**	25	96	4	38	24	21	13	16	**69**	28
Canny-FM_{SS}-G_{S_M}-SF	41	**3**	28	94	5	39	34	21	14	16	**78**	27
Canny-G_{S_P}-G_{S_M}-SF	37	**3**	29	95	4	38	34	20	16	16	**80**	28
FM_{SS}-G_{S_P}-G_{S_M}-SF	65	**2**	24	95	5	38	23	21	13	16	**70**	28

Table 9: Comparison of best (✓) and worst (✗) approaches in terms of $F_{0.5}$ considering penalty approaches with four feature images.

	Method											
	*		FM_{SS}		C		G_{S_P}		G_{S_M}		SF	
	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗	✓	✗
C1-C2-Canny-FM_{SS}-G_{S_P}	34	**2**	22	95	7	38	30	21	17	16	**90**	28
C1-C2-Canny-FM_{SS}-G_{S_M}	26	**0**	24	96	6	39	37	21	16	16	**91**	28
C1-C2-Canny-FM_{SS}-SF	**69**	**1**	21	95	5	39	29	21	13	16	63	28
C1-C2-Canny-G_{S_P}-G_{S_M}	28	**3**	25	94	7	38	31	21	17	16	**92**	28
C1-C2-Canny-G_{S_P}-SF	69	**1**	24	95	5	39	19	21	13	16	**70**	28
C1-C2-Canny-G_{S_M}-SF	48	**1**	24	95	5	39	32	21	15	16	**76**	28
C1-C2-FM_{SS}-G_{S_P}-G_{S_M}	36	**3**	22	94	7	38	28	21	17	16	**90**	28
C1-C2-FM_{SS}-G_{S_P}-SF	**86**	**1**	17	95	6	39	14	21	15	16	62	28
C1-C2-FM_{SS}-G_{S_M}-SF	**72**	**1**	16	95	6	39	26	21	11	16	69	28
C1-C2-G_{S_P}-G_{S_M}-SF	66	**1**	23	95	6	39	19	21	13	16	**73**	28
C1-Canny-FM_{SS}-G_{S_P}-G_{S_M}	25	**3**	25	94	7	38	31	21	19	16	**93**	28
C1-Canny-FM_{SS}-G_{S_P}-SF	**70**	**1**	23	95	5	39	22	21	15	16	65	28
C1-Canny-FM_{SS}-G_{S_M}-SF	53	**1**	24	95	6	39	32	21	12	16	**73**	28
C1-Canny-G_{S_P}-G_{S_M}-SF	52	**1**	24	95	5	39	27	21	12	16	**80**	28
C1-FM_{SS}-G_{S_P}-G_{S_M}-SF	66	**1**	20	95	6	39	22	21	14	16	**72**	28
C2-Canny-FM_{SS}-G_{S_P}-G_{S_M}	27	**4**	25	94	6	38	33	20	16	16	**93**	28
C2-Canny-FM_{SS}-G_{S_P}-SF	64	**2**	25	95	5	38	24	21	15	16	**67**	28
C2-Canny-FM_{SS}-G_{S_M}-SF	48	**1**	26	95	5	39	33	21	12	16	**76**	28
C2-Canny-G_{S_P}-G_{S_M}-SF	46	**1**	27	95	5	39	30	21	15	16	**77**	28
C2-FM_{SS}-G_{S_P}-G_{S_M}-SF	63	**2**	24	95	6	38	23	21	14	16	**70**	28
Canny-FM_{SS}-G_{S_P}-G_{S_M}-SF	48	**2**	26	95	5	38	28	21	15	16	**78**	28

Table 10: Comparison of best (✓) and worst (✗) approaches in terms of $F_{0.5}$ considering penalty approaches with five feature images.

der Weken, D. Van de Ville and E.E. Kerre, Eds. Springer, 2003, pp. 129–152.
[19] F. J. Estrada, A. D. Jepson, "Benchmarking Image Segmentation Algorithms," *International Journal of Computer Vision*, vol. 85, no. 2, pp. 167–181, 2009.
[20] M. Grabisch, J.-L. Marichal, R. Mesiar, E. Pap, *Aggregation Functions*, Cambridge University Press, Cambridge, 2009.
[21] M. Gonzalez-Hidalgo, S. Massanet, A. Mir, D. Ruiz-Aguilera, "On the choice of the pair conjunction-implication into the fuzzy morphological edge detector," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 4, pp. 872–884, 2015.
[22] M. Gonzalez-Hidalgo, A. Mir-Torres, D. Ruiz-Aguilera, J. Torrens, "On the generalization of the fuzzy morphological operators for edge detection," in *Proc. of the IFSA-EUSFLAT Conference*, 2015, pp. 1082–1089.
[23] M. Heath, S. Sarkar, T. Sanocki, K. Bowyer, "A robust visual method for assessing the relative performance of edge-detection algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 12, pp. 1338-1359, 1997.
[24] J. Juszczyk, E. Pietka, and B. Pycinski, "Granular computing in model based abdominal organs detection," *Computerized Medical Imaging and Graphics*, vol. 46, pp. 121 – 130, 2015, information Technologies in Biomedicine.
[25] A. Karaali, C. R. Jung, "Edge-Based Defocus Blur Estimation With Adaptive Scale Selection", *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1126–1137, 2017
[26] Kermit Research Unit (Ghent University), The Kermit Image Toolkit (KITT) [Online]. Available: www.kermitimagetoolkit.com
[27] P. D. Kovesi, (2012). "MATLAB and octave functions for computer vision and image processing," Centre for Exploration Targeting, School of Earth and Environment, Univ. Western Australia, Perth, Australia. Available: http://www.peterkovesi.com/matlabfns/.
[28] C. Lopez-Molina, H.Bustince, J. Fernandez, P. Couto, B. De Baets, "A gravitational approach to edge detection based on triangular norms," *Pattern Recognition*, vol. 43, no. 11, pp. 3730–3741, 2010.
[29] C. Lopez-Molina, B. De Baets, H. Bustince, "Quantitative error measures for edge detection," *Pattern Recognition*, vol. 46, no. 4, pp. 1125-1139, 2013.
[30] C. Lopez-Molina, B. De Baets, H. Bustince, "A frame work for edge detection based on relief functions," *Information Sciences*, vol. 278, no. 10, pp. 127–140, 2014.
[31] G. Lucca, J. Sanz, G. Pereira Dimuro, B. Bedregal, R. Mesiar, A. Kolesárová, H. Bustince, "Pre-aggregation functions: construction and an application," *IEEE Transactions on Fuzzy Systems*, vol. 23, no. 2, pp. 260–272, 2016.
[32] J. M. Malof, K. Bradbury, L. M. Collins, and R. G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied Energy*, vol. 183, pp. 229 – 240, 2016.
[33] D. Martin, C. Fowlkes, D. Tal, J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proc. of the 8th International Conference on Computer Vision*, vol. 2, 2001, pp. 416–423.
[34] D. Martin, C. Fowlkes, J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530-549, 2004.
[35] R. Medina-Carnicer, F. Madrid-Cuevas, A. Carmona-Poyato, R. Muñoz-Salinas, "On candidates selection for hysteresis thresholds in edge detection," *Pattern Recognition*, vol. 42, no. 7, pp. 1284–1296, 2009.
[36] R. Medina-Carnicer, R. Muñoz-Salinas, E. Yeguas-Bolivar, L. Diaz-Mas, "A novel method to look for the hysteresis thresholds for the Canny edge detector," *Pattern Recognition*, vol. 44, no. 6, pp. 1201-1211, 2011.
[37] S. Ono, "L_0 gradient projection", *IEEE Transactions on Image Processing*, vol. 26, no. 4, pp. 1554–1564, 2017.
[38] J.M.S. Prewitt, "Object Enhancement and Extraction," *Picture Processing and Psychopictorics*, Academic Press, pp. 75-149, 1970.
[39] B. Schweizer, A. Sklar, "Associative functions and statistical triangle inequalities," *Pub. Math. Debrecen*, vol. 8, pp. 169–186, 1961.
[40] I. Sobel, G. Feldman, (1968). "A 3×3 isotropic gradient operator for image processing. Presented at a talk at the Stanford Artificial Project, 1968.
[41] V. Torre, T. Poggio, "On edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, pp. 147–163, 1984.
[42] T. Wilkin, G. Beliakov, "Weakly Monotonic Averaging Functions," *International Journal of Intelligent Systems*, vol. 30, no. 2, pp. 144–169, 2015.
[43] X. Wang, H. Ma, X. Chen, S. You, "Edge Preserving and Multi-Scale Contextual Neural Network for Salient Object Detection", *IEEE Transactions on Image Processing*, vol. 27, no.1, pp. 121–134, 2018.
[44] S. Xie, Z. Tu, "Holistically nested edge detection", *International Journal of Computer Vision*, vol. 125, no. 1-3, pp. 3–18, 2017.
[45] Q. Zou, Z. Zhang, Q. Li, X. Qi, Q. Wang, S. Wang, "DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection", *IEEE Transactions on Image Processing*, vol. 28, no.3, pp. 1498–1512, 2018.

Cover Letter

Dear Editor
This manuscript is the authors' original work and has not been published nor has it been submitted simultaneously elsewhere.
All authors have checked the manuscript and have agreed to the submission.

3.5 A survey on matching algorithms for boundary image comparison and evaluation

- C Lopez-Molina, C. Marco-Detchart, B De Baets, Member, H Bustince, and S. Member, "A survey on matching algorithms for boundary image comparison and evaluation", *IEEE Transactions on Image Processing*,
  - Journal: IEEE Transactions On Image Processing
  - Status: Submitted
  - Impact Factor: (JCR 2018) 6.790
  - Knowledge area:
    - Computer Science, Artificial Intelligence: Ranking $9/133$ (Q1)
    - Engineering, Electrical & Electronic: Ranking $18/265$ (Q1)

IEEE Transactions on Image Processing

A survey on matching algorithms for boundary image comparison and evaluation

Journal:	*Transactions on Image Processing*
Manuscript ID	TIP-20774-2019
Manuscript Type:	Regular Paper
Date Submitted by the Author:	04-Jun-2019
Complete List of Authors:	Lopez-Molina, Carlos; Universidad Publica de Navarra, Estadistica, Informatica y Matematicas Marco-Detchart, Cedric; Estadistica, Informatica y Matematicas Bustince, Humberto; Universidad Publica de Navarra, Automatica y Computacion; De Baets, Bernard; Ghent University, Applied Mathematics, Biometrics and Process Control
EDICS:	31. ARS-RBS Region, Boundary, and Shape Analysis < Image and Video Analysis, Synthesis and Retrieval

A survey on matching algorithms for boundary image comparison and evaluation

C. Lopez-Molina, C. Marco-Detchart, H. Bustince, *Senior Member, IEEE*, and B. De Baets

***Abstract*—Among the different alternatives to evaluate boundary images, the ones gaining more interest are those based on comparison to ground truth solutions. More specifically, recent years have seen a dominance of techniques based on the use of a confusion matrix. These techniques require a correspondence between the boundary pixels in the candidate boundary image and those in the ground-truth; that correspondence is further used to create the confusion matrix, from which a statistic is computed. This correspondence faces different challenges, mainly related to the matching of potentially displaced boundaries. In this work, we survey all existing strategies for establishing such correspondence, we propose a taxonomy to embrace them all, and perform a usability-driven quantitative analysis of their behaviour.**

***Index Terms*—Boundary image, Displaced boundary, Linear feature matching, Image comparison**

I. Introduction

In the context of boundary detection, quality evaluation has long been studied, the first references dating back to the 1970s [15], [21]. The reasons are manifold, including the need to rank different boundary detection methods or the development of training-based methods, which demand reliable objective functions. It is generally accepted that the best way to evaluate boundary detection methods is by comparing their results to those by humans. However, there is no agreement on which is the most reliable way to (quantitatively) perform that comparison, either it is carried out in terms of similarity (how close both images are) or dissimilarity (how different they are). As of today, several different techniques and configurations are used; this has led to a rather disorganized situation in which very few standards are kept.

In the past, most of the measures for boundary image comparison relied on distance transformations [1], [4], which are convenient to overcome counting dilemmas due to variable boundary position or boundary cardinality [33]. However, in recent years, a new family of proposals approached the problem with a classification-based inspiration. The reason is simple: at a very basic level, boundary detection is nothing else but binary classification, since every pixel in the image must be labelled as either boundary or not. The comparison between two images is then phrased in the usual terms of binary classification: the matching boundary pixels becoming True Positives (TPs), etc. Once a confusion matrix has been created, well-known statistics and quantifiers yield final evaluations [24], [73].

C. Lopez-Molina, C. Marco-Detchart and H. Bustince are with the Dept. Estadistica, Informatica y Matematicas, Universidad Publica de Navarra, Spain.

C. Lopez-Molina and B. De Baets are with KERMIT, Dept. of Data Analysis and Mathematical Modelling, Ghent University, Belgium.

Manuscript received -, 2019; revised -, 2019.

The classification-based inspiration is very convenient and attractive, since it relates to machine learning and classification, two fields in which quality evaluation is deeply studied [14], [57], [67]. However, it also comes coupled with a critical handicap that is unresolved in the literature: if the boundary of an object appears at slightly displaced (non-overlapping) positions in two images, the comparison method should be able to recognize the circumstance and count its pixels as correctly classified. That is, the counting of the correctly/wrongly classified pixels cannot be done based on mere a pixel-to-pixel comparison. A more elaborated *matching* is needed to map the boundaries in one image to those in the other, up to some spatial tolerance. This matching would ideally be able to tolerate small spatial deviations, yet not pairing the boundaries of objects to those due to different objects, texture or noise.

The boundary matching problem is, at a broad level, that of linear feature matching, which has been regularly addressed in computer vision literature. An evident solution is to treat the problem as bipartite graph matching, *i.e.* one-to-one mapping of the boundary pixels in one image to those in the other, typically minimizing the distance between paired pixels. However, this solution is not perfect. Firstly, the computational cost of deterministic optimal algorithms (such as the Hungarian/Munkres algorithm [25]) is exorbitant. Second, there is also a theoretical problem, since bipartite graph matching requires one-to-one correspondence in the matching. This is a problem in the context of boundary matching strategies because non-overlapping boundaries might be composed of a different (yet similar) number of pixels. Several authors have proposed alternatives to this algorithm, either focusing on the computational cost (as Martin [42] using the CSA algorithm by Goldberg and Kennedy [16]) or presenting alternatives able to cope with the one-to-many correspondences (as Estrada and Jepson [13]). Moreover, alternatives based on area overlapping or mathematical morphology have also been employed. Currently, there is no clear way of knowing which is the best alternative for boundary matching. The most accepted one is the CSA algorithm, but this might due to the fact that it is used in the most popular boundary detection benchmark (the BSDS [2]).

In this work we review the most relevant boundary matching techniques with application to boundary quality evaluation. Moreover, we investigate and compare their performance. Since there is no clear way of knowing *which one works better*, we pose a different question: do different alternatives

have a real impact on the resulting confusion matrices? That is, should we expect significantly different results when using different matching strategies? Although this does not answer the question on the best possible option, it clearly reduces the controversy about it. Section II describes such techniques, while Section III includes some experimental comparisons between them and Section IV recaps a brief discussion.

II. Boundary matching techniques

While understanding the human visual system remains a challenge, it seems clear that humans make intensive use of features to compose objects and scenes [62], [59]. So is the underlying idea of Marr's Primal Sketch [39], arguably the best computational model for human vision. While there exists a diversity of features, the ones represented as lines or curves play a key role in the human visual system. In fact, some famous experiments in the context of cognitive sciences are based on the recognition of contours and shape-based objects, such as the Shepard-Metzler study [63] and subsequent evolutions into the Mental Rotation Test [46], [65].

Not surprisingly, linear features[1] became one of the most relevant low-level features in computer vision [40], either representing visually salient linear structures or any other artefact. In this work we concentrate on boundary images, which use linear features to represent the silhouette of relevant objects in a scene. Nevertheless, examples of use of linear features can also be found in biometrics [75], computational biology [7], [66] or photogrammetry [5], [64].

The ubiquity of linear features in computer vision makes it necessary to design comparison methods, either at individual or whole-scene level. The generic problem of linear feature comparison and matching has been recurrently materialized in specific challenges. The tools and techniques used to resolve it are, nevertheless, heavily dependent upon contextual matters. Such matters include, e.g., the characteristics of the linear features, the semantics of the problem to be solved, or the expected output of the matching process. As a result we find a diversity of linear feature matching techniques, most of them being applicable exclusively to the context for which they were designed.

In Section II-A we review different tasks that are partially or totally based on linear feature matching. This analysis is driven by their interestingness for our final goal: boundary comparison and matching. Then, a mathematical framework for such goal is provided in Section II-B. Finally, Section II-C introduces a novel taxonomy for boundary matching techniques.

A. Linear feature matching for image processing

Linear feature matching is the process of comparing, in quantitative terms, the linear features in two scenes. There are two factors with major influence in linear feature matching: (a) the nature and constraints of the potential displacements of the same feature in different images and (b) the expected output of the matching. The first factor relates to the nature of the linear features and the circumstances under which they were gathered. The second factor is bounded to the utility of the matching process, and the application in which it is intended to be useful. The study of linear feature matching techniques in the literature must therefore be performed under a dual perspective.

The most evident application of linear feature matching algorithms is line pattern recognition, with applications ranging from biometrics to aerial imagery registration. The complexity of linear patterns has often led to the use of derived information in the matching process, instead of the linear features themselves [23]. However, some authors propose an explicit matching of linear features (e.g. for palmprint matching [11] or for fingerprint matching [38]).

A task slightly different from line pattern matching is that of silhouette (also, shape) matching, often with the goal of object recognition or classification. Generally, silhouettes contain most of the information for object recognition, while avoiding problems related to image texture, shading or colour. In some scenarios, *texture or colour cannot be used as a cue for recognition* [47], while in others silhouettes contain enough information to complete the recognition task [22], [69]. Literature contains biometric systems based on silhouette matching and recognition, e.g. for person/gait recognition [32] or hand shape-based identification [26]. Also within the context of silhouette recognition we can list linear feature matching tasks, in applications such as stereo matching or image correspondence [5]. Although stereo matching is often carried out using disparity maps [60], early proposals were based on silhouette recognition and matching [41], [45], [51]. In recent works, line matching is still used for other delicate calibration tasks, e.g. pose estimation in bifocal camera setups [71].

Recognition tasks, either based on linear patterns or shapes, can be roughly divided according to their output: verification (matching of two elements or not), identification (which element, out of a pool of candidates, is a good matching) [37] or offset computation (e.g. modelling the physical 3D setting of two or more scenes)[31]. None of them fits the semantics of boundary image matching, mostly because recognition is grounded on the idea that a perfect (or fair enough) version of the instance is stored in the system. This often turns these tasks into the computation of an optimal deformation model.

Silhouette matching and recognition is so relevant that it led to prominent mathematical developments. A good example is the Curvature Scale-Space (CSS), which has as final goal the comparison and matching of curves and shapes. Introduced by Mokhtarian and Mackworth [48], the CSS uses a parametric representation of non-overlapping curves based on paired functions. Using this representation, the similarity between two curves is quantified as the comparison of such functions, more specifically of the zero-crossings in their second derivatives when projected into the Gaussian Scale-Space (GSS). The CSS is aimed at scale- and rotation-invariant silhouette recognition. These aims are possibly critical to explain human shape recognition (whose behaviour in this regard is still under analysis, see recent works by Chen et al. [10] or Han et al. [19]), but play no positive role in boundary comparison,

[1] In this work, we refer as *linear features* to any visible artefact whose representation is a line. This holds regardless of their origin (edges, countours, ridges) or characteristics (lines, curves, shapes).

in which rotation or scale invariance is undesired. Subsequent evolutions of the CSS (see, e.g. [49]) extended the original proposal, but do not substantially enhanced its utility for boundary matching.

An interesting variation of the problem of shape matching is that of shape-based object tracking, examples being [12], [44], [74]. The difference between shape matching and shape-based object tracking is twofold. First, tracking is performed across several images, not only two of them. This enables, among other possibilities, simultaneous multi-image evaluation. Second, there is contextual information that takes relevance in modelling the movement of the objects. That information involves, for example, the appearance of other (potentially occluding, or occluded) objects in the scene [72], as well as inertia influencing the object movement. Moreover, restrictions in the movement of objects can apply, e.g. in [61], [18], rotating motion matching relies on an expected geometrical transformation. Silhouette-based object tracking is usually solved by applying non-rigid models based on active contours or B-snakes [17], [68]. None of these solutions are available for boundary image matching, where the variation in the displacement of the boundaries can greatly differ for different objects in the same image, or even different segments of the same object boundary.

Object tracking is the most prominent example involving multi-image analysis, but not the only one. Another very relevant example is multiscale image analysis. Within multiscale image analysis we often find inter-scale linear feature matching, especially when it involves tracking. As proposed by Bergholm [6], evolving the ideas by Witkin [70], tracking-based linear feature detection methods intend to discriminate the relevant boundaries in an image at a coarse scale; then, those boundaries are tracked down to the position they occupy at a finer scale. In this way, a feature detector can combine the good discrimination properties of large scales with the accuracy of the fine ones. The tracking process starts out by matching the boundaries at the coarsest scale to those in the immediately finer scale, then repeats the process until the image corresponding to the finest one. Normally, cross-scale feature displacement is limited by a maximum scalewise displacement, which dramatically eases the problem. Although tracking has received considerably less attention than other components of multiscale image analysis (e.g. the theoretical properties of the scale-spaces), some alternatives have appeared in the literature [53], [33]. Note that multiscale edge and ridge analysis often demand linear feature matching, even if they do not involve tracking. For example, the practical implementation of Lindeberg's ideas in the GSS [28], which involve the location of the optimal scale for each boundary, demand the explicit construction of the so-called *edge surfaces in scale-space*. Such surfaces are typically given by the correspondence of the positions of each boundary at each scale.

Although linear feature tracking (or edge surface construction) and boundary matching for quality evaluation seem to be close tasks, relevant differences arise in the detailed analysis. Linear feature tracking involves a sense of inertia and, despite being carried out as consecutive matchings between pairs of images, it aims at the construction of multi-image, cross-scale structure. Also, the scale-spaces under which the image is projected can incorporate relevant constraints to the matching problem. For example, the causality principle in the GSS imposes that any feature at a coarse scale corresponds to a *not-necessarily unique* [55] feature at a lower scale. This allows, and in a sense encourages, that a single feature at a given scale is matched to spatially diverging features at a finer scale. While some similarities between boundary matching and edge tracking are evident, in particular the fact that no explicit matching is required, solutions from linear feature tracking can hardly be ported to boundary matching for comparison.

A completely different approach to line-based object recognition is that mimicking human recognition abilities through the implementation of CNNs. This research, which roots back to early Mental Rotation Tests [65], attempt to model human abilities in early stages of the Human Visual System (HVS), specially regarding rotation, scale and eccentricity invariance [56], [10]. Since some of these experiments are based on line based draws (e.g., in [19], Korean characters), it is to be expected that eventually CNNs might mimic human line pattern recognition abilities. Such a solution would, however, raise evident questioning when used for boundary matching in the context of quantitative evaluation. First, some of the features (as scale invariance) are hardly desirable when applied to comparison, even if it seems to be part of the HVS [56], [19]. Second, if many differently trained CNNs yield similar results, how to produce a canonical comparer out of them? Alternatively, if using *any* of such CNNs for the task, how to audit their results [27]? Third, as reported by Nguyen, Yosinski and Clune [52], well-performing CNNs can lead to aberrant results which are not only misleading, but also yield in almost-full ($> 99\%$) confidence. As the authors state, their findings raise *questions about the true generalization capabilities of DNNs* [52], which are of paramount importance for the present task.

As a recap, we find very different tasks for which linear feature comparison or matching becomes relevant, but exporting their solutions to the context of boundary quality comparison and matching is troublesome. Recognition or classification systems usually look for a closest-possible match, assuming that the same object appears in two different images. Tracking systems extend that assumption to a large number of images, even if the object(s) to be tracked is (are) potentially occluded at some of them. Moreover, the semantics of the output of such tasks do not properly match that of boundary image evaluation.

B. A model for boundary evaluation and matching

Boundary matching for quality evaluation is oriented to the recognition (and quantification) of the amount of coincidental information in two boundary images. Although some controversies hold on the interpretation of boundaries [54], most authors accept that boundaries are the silhouette of the relevant objects in an image. Often, the position of such silhouettes is hard to determine even for humans, either due to limited resolution in the image or to the very configuration of the scene in terms of lightning, shading, occlusions, etc. Matching for boundary quality evaluation needs to tolerate a certain

displacement, which can take place, *a priori*, in any possible direction. However, it is desirable to have some control on the displacement of nearby boundary pixels. For example, it does not seem natural that pixels that are contiguous in one image get matched to pixels that are very distant in the other one. With respect to the expected output, boundary matching for quality evaluation imposes very few restrictions. In fact, the matching does not need to be in terms of an explicit item-by-item matching, but in terms of the amount of common information instead. That is, since the final question to be cleared out is how similar the images are, the correspondence of the boundary pixels at each image does not need to be explicitly enunciated. Any other output is acceptable, as long as it serves as support for computing a subsequent confusion matrix.

The problem of boundary matching, as well as that of quality evaluation, can be put to mathematical terms. In this work, we consider all the images to have generic dimensions $M \times N$, so that the set of positions is $\Omega = \{1, \ldots, M\} \times \{1, \ldots, N\}$. A binary image can be seen both as a mapping $\Omega \to \{0, 1\}$ and as a subset of Ω.

Let E_{cd} and E_{gt} represent a candidate and a ground-truth boundary image, respectively. Classification-based approaches to boundary quality evaluation aim at generating a confusion matrix from the comparison of E_{cd} (probably due to an automated method) with E_{gt} (probably due to a human). In this context, a False Positive (FP) is a boundary pixel in E_{cd} with no correspondence in E_{gt}, a false negative (FN) is a pixel in E_{gt} that is not represented in E_{cd}, etc. Although some authors have proposed to use confusion matrices to compute χ^2 or ROC-derived measures [8], [73], most authors choose the F_α-measure to evaluate the closeness of E_{cd} to E_{gt}. The F_α-measure is given by:

$$F_\alpha = \frac{\text{PREC} \cdot \text{REC}}{\alpha\,\text{PREC} + (1-\alpha)\,\text{REC}}$$

with $\text{PREC} = \frac{\text{TP}}{\text{TP}+\text{FP}}$ and $\text{REC} = \frac{\text{TP}}{\text{TP}+\text{FN}}$, where α is a parameter modulating the relevance of PREC and REC (typically set to 0.5 [33], [43]), and TP, FP and FN are the usual quantities in a confusion matrix.

The F_α-measure is preferred over other quantities for a list of reasons, including the fact that it avoids using the true negatives (TN), which is typically (much) greater than the other quantities in a confusion matrix for boundary quality evaluation.

Given E_{cd} and E_{gt}, the generation of a confusion matrix is far from trivial. The reason is that the same boundary might appear at nearby, yet not overlapping, positions at each image. When comparing the images in a pixel-to-pixel manner, the non-overlapping boundary pixels at each image would be accounted for as both a FP (those in E_{cd}) and a FN (those in E_{gt}). This is undesirable, since the information contained in both images is similar and, as long as the distance between both appearances of the boundary is not excessive, the boundary should be taken as a correct detection. While some previous studies [24], [35], [36] analyze different quantitative measures to be extracted from a confusion matrix, this work focuses on the strategies for the generation of such matrix; that is, for the correspondence between E_{cd} and E_{gt}.

C. A taxonomy for boundary matching techniques

We consider that boundary matching techniques should be classified according to the inspiration they take, not according to the mathematical tools they make use of. Consequently, next, we present a fourfold taxonomy in which the categories are not completely disjoint. At the end of this section, we review such overlappings.

Note that this taxonomy can be seen as a specialization of a part of the taxonomy for error measures in [35], [34], since it appears as a subdivision of the statistical error measures. Also, it relates to the so-called *Confusion matrix-based error assessments* by Adbulrahman et al. [**?**], and to the statistics derived from confusion matrices in the survey by Magnier [36].

There exist four strategies for boundary pixel matching in the context of boundary evaluation:

a) *Distance-based Matching (DbM).*- This strategy roots in validating (matching) boundary pixels in E_{cd} as long as they are closer to a boundary pixel in E_{gt} than a given threshold [8]. Such pixels become true positive detections, while the remaining pixels in E_{cd} are taken as false positives. In the context of classification, the confusion matrix is given by

$$\begin{aligned} \text{TP} &= |\{p \in E_{\text{cd}} \mid d(p, E_{\text{gt}}) \leq t\}|\,, \\ \text{FP} &= |\{p \in E_{\text{cd}} \mid d(p, E_{\text{gt}}) > t\}|\,, \text{ and} \\ \text{FN} &= |\{p \in E_{\text{gt}} \mid d(p, E_{\text{cd}}) > t\}|\,, \end{aligned} \tag{1}$$

where $d(p, E)$ represents the distance from a pixel p to the closest boundary pixel in E, $|\cdot|$ represents the cardinality of a set, and $t \in \mathbb{R}^+$ is the maximum allowed distance between matched pixels[2].
This strategy can also be put in terms of mathematical morphology, as done by Arbelaez in [3] (Ch. 7). In this case, a circular structuring element [20] represents the potential displacement of a boundary pixel, so that a pixel in E_{cd} is validated if it falls within the dilation scope of E_{gt} (and *vice versa*):

$$\begin{aligned} \text{TP} &= |E_{\text{cd}} \cap \text{dil}_S(E_{\text{gt}})|\,, \\ \text{FP} &= |E_{\text{cd}} \cap \neg\,\text{dil}_S(E_{\text{gt}})|\,, \text{ and} \\ \text{FN} &= |E_{\text{gt}} \cap \neg\,\text{dil}_S(E_{\text{cd}})|\,, \end{aligned} \tag{2}$$

where dil_S is the dilation operation with structuring element S, which here is a circular structuring element of radius t. The results using the formulation in Eq. (1) and that in Eq. (2) are equivalent.
The DbM strategy is extremely simple and computationally cheap, especially if using implementations based on either distance transformation or mathematical morphology. However, it lacks refinement in the discrimination of boundaries and spurious responses. As an example, spurious responses due to texture or noise might be accounted for as true positives when appearing relatively

[2]Note that $d(p, E)$ can also be seen as the value of the pixel p in the distance transform of E by means of d.

close to the actual boundaries. Hence, although simple and understandable, its use is often disregarded.

b) *Area-based Matching (AbM).*- The AbM grounds on the idea that boundaries have a certain area of influence, which is simply defined as the area surrounding it. Then, it quantifies the TP as the overlapping of the areas of influence of the boundaries at each image.
This strategy renders in a formulation similar to that of the DbM, but allows for a more delicate setting of the allowed displacement. In AbM, the tolerated displacement is modelled by a structuring element [20], leading to

$$\begin{aligned} \mathrm{TP} &= |\,\mathrm{dil}_S(E_{\mathrm{gt}}) \cap \mathrm{dil}_S(E_{\mathrm{cd}})\,|\,, \\ \mathrm{FP} &= |\,\mathrm{dil}_S(E_{\mathrm{gt}}) \cap \neg\,\mathrm{dil}_S(E_{\mathrm{cd}})\,|\,, \text{ and} \\ \mathrm{FN} &= |\neg\,\mathrm{dil}_S(E_{\mathrm{gt}}) \cap \mathrm{dil}_S(E_{\mathrm{cd}})\,|\,, \end{aligned} \tag{3}$$

where S represents the structuring element and dil_S stands, again, for morphological dilation.
Although the formulation in Eq. (3) resembles that in Eq. (2), it yields significantly different results. Firstly, the interpretation of the quantities in the confusion matrices is completely different, since those for DbM represent the number of pixels in the boundaries, but those for AbM represent area sizes. Secondly, displaced boundaries can yield perfect matching in terms of DbM, as long as the displacement is lower than t. However, in AbM, any boundary in E_{cd} slightly displaced from its position in E_{gt} will produce a certain number of FPs and FNs, since $\mathrm{dil}_S(E_{\mathrm{cd}})$ and $\mathrm{dil}_S(E_{\mathrm{gt}})$ will not overlap completely.
Note that both AbM and DbM do not perform explicit matchings between the images. Instead, they count the pixels that are matchable to the counterpart image, which leads to the generation of the confusion matrix.

c) *Correspondence-based matching (CbM).*- This strategy attempts to create an explicit one-to-one matching of the pixels in E_{cd} to those in E_{gt}. Such matching can lead to conclusions on the amount of information in E_{cd} also present in E_{gt} (and vice versa), as well as to the average displacements of the matched pixels [29].
Put to mathematical terms, CbM is presented as the problem of finding a minimal-cost assignment of the pixels in E_{cd} to those in E_{gt}. That is, finding a (largest possible) subset $Q = \{(p_1, q_1), \ldots, (p_n, q_n)\}$ so that $p_i \in E_{\mathrm{cd}}$, $q_i \in E_{\mathrm{gt}}$ and $\sum_{i\in\{1,\ldots,n\}} d(p_i, q_i)$ is minimal. Note that some constraints apply to the set Q. Firstly, each boundary pixel in E_{cd} (resp. E_{gt}) can only be matched to one boundary pixel in E_{gt} (resp. E_{cd}). Secondly, since the number of boundary pixels in each image might be different, the set Q is likely not to cover any of them completely. Thirdly, pairs of pixels $(p_i, q_j) \in \Omega \times \Omega$ are eligible to belong to Q iff $d(p_i, q_j) < t$, where t represents a certain threshold in terms of a metric d. This approach to boundary matching has evident links to both the assignment problem and the transportation problem [25], [50].
Liu and Haralick [29] studied the problem deeply, covering such constraints and proposing strategies to overcome them. Finally, the authors convert the problem into an unconstrained assignment problem by adding *ghost* pixels (to equalize the cardinality of both sets) and setting artificially long distances to the *unmatchable* pairs of pixels. Another detailed analysis of this problem, leading to a different proposal, can be found in [42] (Ch. 3).
Regardless of the algorithm used in the process, CbM leads to the creation of Q, containing the matched pairs of boundary pixels. Once Q is created, the confusion matrix is constructed as:

$$\begin{aligned} \mathrm{TP} &= |\{p \in E_{\mathrm{cd}} \mid (p, y) \in Q \text{ for some } y\}|\,, \\ \mathrm{FP} &= |\{p \in E_{\mathrm{cd}} \mid (p, y) \notin Q \text{ for any } y\}|\,, \text{ and} \\ \mathrm{FN} &= |\{p \in E_{\mathrm{gt}} \mid (x, p) \notin Q \text{ for any } x\}|\,. \end{aligned}$$

CbM has interesting properties and is mathematically sound, but also requires a careful revision. Some of its features are controversial, especially when it comes to the restriction of the one-to-one assignment. On the one hand, this helps avoiding situations in which one single pixel in E_{gt} is matched to (or validates) multiple boundary pixels in E_{cd}, potentially far from each other. Also, the fact that CbM creates an explicit matching allows for the direct computation of related, contextual information, e.g. the average distance between matched boundaries. On the other hand, the fact that each pixel can be matched only once in CbM produces problems in matching slightly displaced boundaries, which are usually composed of a similar, yet different, number of pixels. Although unmatched pixels can be seen as an implicit penalization for the displacement, such penalization is hard to interpret or predict.
Apart from the one-to-one restriction, finding the optimal set Q in CbM is computationally prohibitive. Some alternatives have been presented based on the Hungarian/Munkres algorithm [50], even including pre- and post-processing of E_{cd} and E_{gt} to reduce the computational load. However, most authors use pseudo-optimal algorithms to produce Q, more specifically the implementation of the CSA algorithm [16] distributed within the BSDS300 and the BSDS500 [42], [2].

d) *Pixelwise validation (Pv).*- This strategy takes individual decisions on the validity of each of the boundary pixels in E_{cd}, using information from the surrounding region in both images. The validation process can be expressed as a mapping $\psi : \mathcal{P}(\Omega)\times\mathcal{P}(\Omega) \mapsto \mathcal{P}(\Omega)$ so that $\psi(A, B) \subseteq A$ is the subset of boundary pixels in A that are validated w.r.t. B. Depending on the rules used for validation, ψ might not be symmetric (*i.e.* it might happen that $\psi(A, B) \neq \psi(B, A)$). Given ψ, a confusion matrix can be constructed as

$$\begin{aligned} \mathrm{TP} &= |\{p \in E_{\mathrm{cd}} \mid p \in \psi(E_{\mathrm{cd}}, E_{\mathrm{gt}})\}|\,, \\ \mathrm{FP} &= |\{p \in E_{\mathrm{cd}} \mid p \notin \psi(E_{\mathrm{cd}}, E_{\mathrm{gt}})\}|\,, \text{ and} \\ \mathrm{FN} &= |\{p \in E_{\mathrm{gt}} \mid p \notin \psi(E_{\mathrm{gt}}, E_{\mathrm{cd}})\}|\,. \end{aligned}$$

However, as proposed in [13], it is more natural to skip the confusion matrix and to compute PREC/REC as

$$\mathrm{PREC} = \frac{|\psi(E_{\mathrm{cd}}, E_{\mathrm{gt}})|}{|E_{\mathrm{cd}}|} \quad \text{and} \quad \mathrm{REC} = \frac{|\psi(E_{\mathrm{gt}}, E_{\mathrm{cd}})|}{|E_{\mathrm{gt}}|}\,.$$

The Pv strategy has properties of great interest. First, it is able to produce a deterministic one-to-many matching of the boundary pixels in each image at an acceptable cost (among the previous strategies, only certain cases of CbM are able to do so). Secondly, it allows for the application of rather complex, yet meaningful, validation rules, including boundary orientation or interference of different boundaries. However, the fact that it is based on a local analysis makes it computationally inefficient compared to other strategies, especially to those based on mathematical morphology or distance transformations.

The four categories in this taxonomy are not completely disjoint. For example, DbM can be seen as a specialization of Pv in which no information other than the distance between pixels is used. Also, DbM with $d = 1$ would produce results equal to those by AbM with a (rather useless) circular structuring element with radius 1. Still, we consider each category to have different semantics, even if leading to equivalent instantiations.

The theoretical comparison of the four strategies can be performed from different perspectives, none of them providing clear advantage to any of the strategies. Pv matching is the strategy able to use most of the available local information, while all of the other ones only consider the area/length of a potential displacement of each boundary pixel. However, the fact that CbM considers a global solution (instead of pixelwise validation) seems more adequate for the task. With respect to the robustness against small variations in the input, we find AbM to be the most adequate strategy, since it is robust against boundary rugging and progressively penalizes displaced edges; also, this displacement-derived penalty is easily interpretable. Interestingly, DbM, CbM and Pv might judge a boundary pixel to be either completely matchable or completely unmatchable on the basis of a 1-position displacement, (the displacement making the boundary pixel inside or outside the allowed displacement scope). Alternatively, if robustness refers to that against the inclusion of noise in one of the images, CbM appears as the best option, since each noisy pixel contributes with at least one unit to TN or FN (maybe both).

On top of the previous considerations, we find that some desirable properties are not provided by any of the strategies. For example, one might consider that neighbouring pixels in a boundary image should be matched to nearby pixels in the other image with some spatial coherence. That is, there should be some spatial coherence between the neighbouring *source* pixels matched in one image and their *destination* pixels in the other. No such proposal have appeared in the literature, and only Pv, due to its flexibility in neighbourhood analysis, seems to allow for similar goals.

III. Experimental results

Ideally, we could investigate which is the *best* matching strategy in boundary image comparison, or at least the one that fits better some specific goals. However, there is no clear way of doing so, and, moreover, there is no evident way of generating ground-truth for the boundary matching. As an alternative, the present work questions whether there are practical differences in using different boundary matching techniques. That is, whether there exist significant differences in the results (quality evaluations) obtained when the matching is performed based on different techniques.

The matching methods considered are the following:

- Distance-based Matching (DbM) using the Euclidean metric;
- Area-based Matching (AbM) using a circular structuring element;
- Correspondence-based Matching (CbM) using the CSA algorithm [16];
- Pixelwise validation (Pv) using the constraints by Estrada and Jepson [13].

All the matching methods above have a parameter representing the maximum allowed displacement of a boundary pixel, *i.e.* the maximum distance between two matched pixels. This parameter is embodied differently in each method: the maximum distance in the DbM, the radius of a circular structuring element in the AbM, and the maximum matching distance in both CbM and Pv. Apart from that common parameter, which we refer to as t, the only parameter required in the experiment is the maximum angular distance used by Estrada and Jepson, which we set to $\frac{\pi}{2}$ (see [13] for more details).

A. First experiment

Our first inquiry relates to how similar it is, the evaluation of a candidate image E_{cd} w.r.t. E_{gt} using different matching strategies. Hence, we measure the Pearson correlation of the $\text{F}_{0.5}$ evaluations after performing the matching with each of the considered techniques. The set of images used in the experiment are the ground truth (human-made) images in the BSDS500 set (500 original images, 2696 ground truth images [2]). Note that this dataset contains several human-made solutions for each of the original images.

Table I displays the Pearson correlation between the values of $\text{F}_{0.5}$ in the one-to-one comparison of the images. The results are displayed separately for the intra-class comparisons (comparisons of non-identical images that are ground-truth to the same image), and the the inter-class ones. The results are discriminated in this way for two reasons. Firstly, because the matching problem is completely different when a large number of coincidences exists, as is the case of the intra-class comparisons, from when very few elements are to be matched. Secondly, since the number of inter-class comparisons is about 300 times greater than the number of intra-class ones, a joint display would obfuscate the visibility of the latter[3]. We use $t = 2.5$, $t = 5$ and $t = 10$ pixels, what corresponds to around 0.5%, 1%, 2% of the length of the image diagonal, respectively.

In Table I we observe that there exists a high correlation between the values yielded by any pair of matching strategies. This seems to be true for both intra- and inter-class comparisons. Also, there is a slight decrease in the correlation coefficients as t increases. This is natural, since the opportunities for disparities in the matching of the boundaries

[3]Ignoring the comparison of images with themselves, there is around $3 \cdot 10^4$ intra-class and 10^7 inter-class pairs of images to be compared.

Matching strs.	Maximum matching distance $t = 2.5$	$t = 5$	$t = 10$
DbM - AbM	.994 / .993	.985 / .984	.974 / .966
DbM - CbM	.999 / .996	.993 / .986	.974 / .966
DbM - Pv	.999 / .997	.996 / .988	.987 / .974
AbM - CbM	.992 / .994	.979 / .988	.955 / .971
AbM - Pv	.994 / .995	.983 / .990	.972 / .982
CbM - Pv	.999 / .999	.998 / .995	.991 / .989

TABLE I: Pearson correlation coefficient between the $F_{0.5}$ values obtained in the comparison of the ground truth images in the BSDS500 using four different strategies for boundary matching. For each pair of strategies and maximum distance between matched pixels, we list the correlation coefficients in the intra-class (left) and inter-class (right) comparisons.

Matching strs.	Maximum matching distance $t = 2.5$	$t = 5$	$t = 10$
DbM - AbM	.955 / .960	.935 / .948	.922 / .926
DbM - CbM	.981 / .973	.954 / .949	.908 / .914
DbM - Pv	.990 / .977	.967 / .951	.932 / .926
AbM - CbM	.952 / .961	.924 / .949	.890 / .918
AbM - Pv	.953 / .964	.930 / .949	.915 / .932
CbM - Pv	.986 / .983	.967 / .966	.934 / .949

TABLE II: Equal-Sorting Ratio (ESR) obtained in the comparison of the ground truth images in the BSDS500 using four different strategies for boundary matching. For each pair of strategies and maximum distance between matched pixels, we list the ratios for intra-class (left) and inter-class (right) triplets.

increase as t increases. We also observe that the pairs AbM-CbM and AbM-Pv produce slightly lower correlations than the other pairs. Still, the correlation is very high for any possible combination of matching strategies and maximum matching distance.

B. Second experiment

The results in Section III-A indicate that all of the matching techniques lead to highly correlated results. The linear correlation does not necessarily mean that the $F_{0.5}$ values are similar, but we can expect the results with a given matching algorithm to be a scaled version of the others. However, quality evaluation is often not about scoring, but about ranking and/or picking the best contender. Although the results by different matching algorithms are similar, would they lead to similar rankings of boundary image? With this second experiment we intend to shed light on this fact.

Let $\mathbf{R}$ be the set of triplets of unrepeated ground truth images (A, B, C) in the BSDS500 Test Set. That is, every triplet (A, B, C) so that $A \neq B \neq C$. Let q_1 and q_2 be any two comparison measures.

The Equal-Sorting Ratio (ESR) between two measures q_1 and q_2[4] in a dataset $\mathbf{R}$ is

$$\mathrm{ESR}^{\mathbf{R}}_{q_1,q_2} = \frac{|\mathbf{G}|}{|\mathbf{T}|} , \tag{4}$$

with $\mathbf{G} \subseteq \mathbf{R}$ defined as

$$\mathbf{G} = \{r \in \mathbf{R} \mid q_1(A,B) \geq q_1(A,C) \text{ iff } q_2(A,B) \geq q_2(A,C)\} . \tag{5}$$

The ESR is hence the proportion of triplets $r = (A, B, C)$ in $\mathbf{R}$ for which the measures q_1 and q_2 agree on whether B or C is closer to A. Hence, it aims at measuring how consistent would rankings be if created with different boundary matching techniques.

[4]In fact, in this experiment, we have the same measure ($F_{0.5}$) embodied with different matching strategies.

Table II displays the ESR for the intra- and inter-class triplets in $\mathbf{R}$ separately. Note that in the BSDS500 Dataset there are over $2 \cdot 10^5$ and 10^{10} of intra- and inter-class triplets, respectively.

Table II displays the ESR for different pairs of matching strategies, and exposes facts similar to those in Table I. The ESR is very high for any possible combination of matching strategies and maximum matching distance. As in the first experiment, the increase of t comes coupled to a greater divergence between the matching strategies (in this case, lower ESR). It is also noticeable that the pairs AbM-CbM and AbM-Pv again produce lower ESR values than any other pair. It is worth mentioning that, in the case of inter-class comparison, the pair DbM-AbM shows a significantly lower ESR, closer to the two previous mentioned pairs of strategies. From the results in Table II, we conclude that the rankings obtained by different matching techniques are very similar for any two matching strategies.

C. Third experiment

From the first and second experiments, we can infer that the matching technique used to create the confusion matrix is nearly irrelevant, since any choice leads to similar conclusions in terms of amount of matched information (first experiment), and also fine grained ranking (second experiment). However, there is still a question to be posed, related to the validity of the dataset used in the experiments, *i.e.* to the fact that the comparisons are always carried out between human-made images. Truly, the goal of the community is to produce methods whose results look like human-made images. Nevertheless, this is not always true, and current boundary image evaluations might involve images whose characteristics are different from those in the ground truth of the BSDS500. The final goal in quality evaluation is not to measure the quality of human-made images, but that of computer-generated ones.

The first and second experiments have been repeated using a set of boundary images generated with the Canny method [9]. That is, comparing the human-made images in the BSDS Test Set with those generated automatically. The Canny method has been selected because it is a good representative of boundary detection methods based on gradient magnitude, a historically

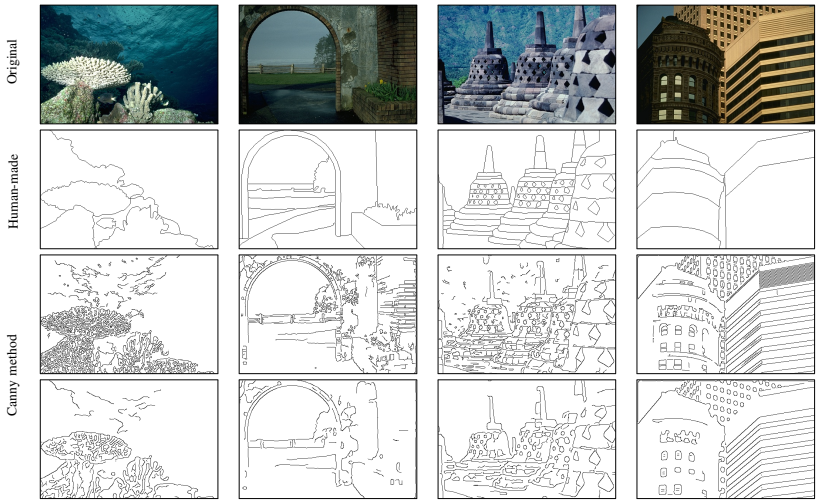


Fig. 1: Comparison of boundary image when produced by hand by human labellers or automatic methods. The upmost row contains the original images in the BSDS. The second contains human-made ground-truth images. The two lowest rows contain the result by the Canny method on two different configurations: $\sigma_1 = 1, \sigma_2 = 1$ and $\sigma_1 = 2, \sigma_2 = 2$, where $\sigma = 1$ and σ_2 refer to the standard diviation for the smoothing filter and the differentiation filter, respectively.

significant class of methods. Even if it is not a state-of-the art competitor in the BSDS, this fits our intention of using not-very-human boundary images in the comparison. Using more advanced boundary detection methods, whose results look more like those by human labellers, would have led to the replication of the results in the first and second experiments. The Canny method is hence both significant for its position in the literature and appropriate for the characteristics of its results.

In this experiment, intra-class comparisons are those between images generated (either by humans or by the Canny method) from the same original image. In the Canny method we take $\sigma_1 \in \{1, 2\}$ for the Gaussian smoothing and $\sigma_2 \in \{1, 2, 3\}$ for the differentiation kernels. The binarization is performed using NMS, hysteresis and the double-threshold determination technique by Liu [30]. In this manner, the (six) boundary images in each class are similar, while differing in the position (and potentially, in the selection) of some boundaries due to the variation in the filter sizes[5]. Moreover, the intra-class divergences are due to the parameter setting, not to human interpretation, what leads to a more realistic scenario. In fact, the Canny method is preferred over other alternatives for its simplicity, and the clear way in which it manifests the most common errors in edge detection methods: texture false detection, edge displacement, edge breaking, etc.

[5]This set of boundary images is, as well as the code of the comparisons, available at [58].

Some examples of the differences between the human-made images and those obtained with the Canny method are shown in Figure 1. We can observe that, in textured areas, typically tag no boundaries (first, second and third column from the left), while the Canny method might do it. Also, salient, but semantically unimportant structures, might also be tagged by automatic, yet not by humans. Examples of such are the windows on the building facades in the rightmost column of Figure 1.

In the replication of the second experiment, triplets are created so that (A, B, C), with A a human-made ground truth and B, C automatically-generated images by the Canny method. In this way we simulate a realistic setup in which two computer-generated images need to be ranked taken a human-made one as reference.

The results gathered in the comparisons are listed in Table III. These results illustrate a rather acute decrease of correlation and, very specially, of ESR. This indicates that, even if results are highly correlated, the small perturbations in the $F_{0.5}$ values might lead to different rankings. A direct hypothesis which would explain all results is that such missortings are decided by very small margins. Hence, we attempt to verify it.

Let $r = (A, B, C)$ be any triplet of images. According to the definition of ESR, r is missorted by two comparison measures q_1 and q_2 iff

$$(q_1(A, B) - q_1(A, C)) \cdot (q_2(A, B) - q_2(A, C)) < 0.$$

Matching strs.	Maximum matching distance $t = 2.5$	$t = 5$	$t = 10$
DbM - AbM	.992 / .997	.982 / .990	.960 / .973
DbM - CbM	.987 / .977	.956 / .900	.883 / .785
DbM - Pv	.999 / .999	.991 / .983	.968 / .947
AbM - CbM	.982 / .979	.933 / .921	.832 / .831
AbM - Pv	.992 / .997	.967 / .986	.922 / .957
CbM - Pv	.988 / .981	.979 / .955	.960 / .928

(a) Correlation coefficients, as in Table I

Matching strs.	Maximum matching distance $t = 2.5$	$t = 5$	$t = 10$
DbM - AbM	.914 / .958	.870 / .932	.808 / .882
DbM - CbM	.875 / .909	.784 / .770	.699 / .646
DbM - Pv	.979 / .981	.873 / .909	.768 / .825
AbM - CbM	.829 / .907	.732 / .784	.669 / .669
AbM - Pv	.917 / .962	.810 / .909	.720 / .830
CbM - Pv	.874 / .917	.895 / .847	.910 / .805

(b) Equal-Sorting Ratio (ESR), as in Table II

TABLE III: Repetition of the experiments in Tables I and II, using different images. In this case the comparisons are run between hand-made images in the BSDS500 Test Set and automatically-generated boundary images. The second set is created running the Canny method with 6 different parameter settings on each of the (grayscale) images in the BSDS500 Test Set. For each pair of strategies and maximum distance between matched pixels, we list the ratios for intra-class (left) and inter-class (right) triplets.

Capitalizing on this idea, given two error measures q_1 and q_2, we define the Sorting Margin (SM) of a triplet $t = (A, B, C)$ as:

$$\mathrm{SM}_{q_1,q_2}(r) = \mathrm{sign}(\alpha) \cdot \sqrt{|\alpha|}\ , \tag{6}$$

with

$$\alpha = (q_1(A,B) - q_1(A,C)) \cdot (q_2(A,B) - q_2(A,C)). \tag{7}$$

The SM is positive for well-sorted triplets, and becomes negative in case of missorted ones. At the same time, SMs of small absolute value indicate marginal missortings, while SMs of large absolute value are due to severe differences between q_1 and q_2. For every pair of comparison measure and matching distance t we have computed the distribution of SM over the triplets in $\mathbf{R}$. Figure 2 displays the distribution of triplets with negative SM for each combination of q_1, q_2 and matching distance t. In this figure we can see how most of the missorted triplets are so by a margin which rarely exceeds 0.03. Even in the cases in which the ESR is relatively low, most of the SMs are near zero. For example, in the case of CbM-DbM, with $t = 5$, the 2.5 percentile stays in -0.049, being the lowest 0.025 percentile of the distribution represented in the figure. Overall, it can be concluded that the SM is always relatively low.

Three factors can, still, raise the SM. First, increasing the matching distance t increases the SM, as seen in Fig. 2. Second, interclass triplets (which are less representative for real comparison) normally yield lower SM, meaning that more realistic (intraclass) comparisons are missorted by smaller margins. Third, some combinations of distances clearly produce greater ESR and SM than other. A paradigmatic case is that of CbM-DbM, mostly due to the fact that CbM enforces a 1-to-1 correspondence of matched pixels, while DbM allows any number of nearby pixels to be matched. This, in images containing strong textures near actual edges, can lead to significant variability in the matching and, hence, in the measured F. Such cases are, however, rare over a standard dataset as the BSDS, and mostly produced by high-frequency gradient characterization filters (in this experiment, the Canny method with $\sigma_1 = \sigma_2 = 1$).

Figure 3 contains one of the triplets having a greatest SM in the overall comparison. The triplet composed by those three images produces, for example, a SM of -0.094 when using the strategies DbM-CbM with matching distance $t = 0.05$. It can be observed that the image recaps the characteristics that can create a perfect storm in terms of discrepancies between different matching strategies. First, there is a large number of objects which human ignore due to contextual facts. For example, the public leaning on the fence, which is relatively salient, yet contextually unimportant. Also, it combines two facts that create divergent results by different matching strategies: (a) strong textures being detected near edge structures, so that strategies as DbM or AbM will count as TPs boundary pixels which will become FPs in CbM; (b) multiple lines that are uniquely tagged by users as a single line, yet being multiply labelled by automatic methods. Examples of the former are the people leaning on fence, or the elements in the cyclist bodies, while the latter manifests at the limits of the cycling track.

Figure 4 contains a visual representation of the matching by each strategy in the triplet featured in Fig. 3, setting the matching distance to $t = 5$. Specifically, for each image and matching method, two areas are highlighted for better illustration of the relevant characteristics of the images and the output of the matching strategies. In the two upper rows we observe how strong textures near human-labelled boundaries produce a variable response depending on the matching strategies. This is mostly due to the fact that CbM restricts the matching to be a 1-to-1 correspondence, which in this case results in judicious responses. In the two lower rows we observe the behaviour of the matching strategies when multiple boundaries appear. While DbM or AbM validate all responses, CbM restricts the matching (again) to a 1-to-1 correspondence. The behaviour by DbM or AbM seems, in this case, more appropriate. Anyhow, the combination of both types of situations lead to divergent interpretations by different matching strategies.

In general, even if some triplets can show the problematics seen in the triplet in Fig. 3, such cases are exceptional. In order for a triplet to generate a SM under -0.03, a list of circumstances must co-occur, including poor performance by the automatic method, and adversary situations in the image (as those seen in Fig. 4) leading to variable interpretation by the matching algorithms. In general, most of the missorted triplets produce SMs in the range [-0.03,0], what proves that such co-ocurrences are rather unusual. This, additionally, explains the results in Table III.

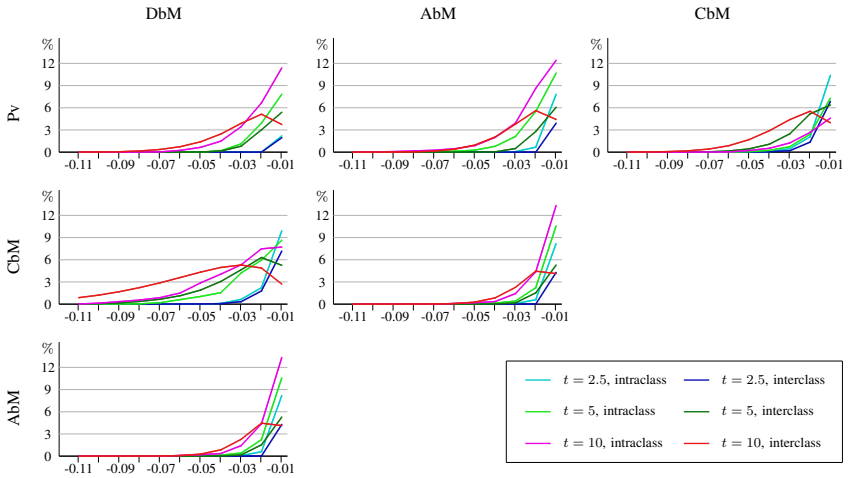
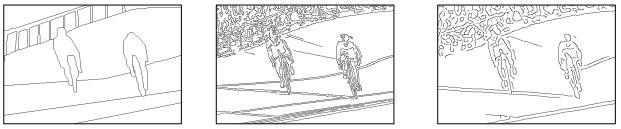


Fig. 2: Percentual distribution of triplets with negative ESM when combining different matching strategies and matching distances. The images used for the experiment are those in the BSDS500 Test Set, compared to automated Canny method-generated solutions, as in Table III. The distribution is binned with 0.01 granularity. It is clearly seen that, even in case of a large ESR (as is the case for CbM-DbM and AbM-CbM), the ESM is mostly in the range $[0, 0.03]$. Triplets for which the ESM is greater than 0.05 are rather unusual.



Human-labelled ground truth | Canny method, $\sigma_1 = 1$, $\sigma_2 = 1$ | Canny method, $\sigma_1 = 2$, $\sigma_2 = 3$

Fig. 3: Ground truth image 226022_05 extracted from the BSDS500 Test Set, together with two automatically-generated images using the Canny method. The parameter setting of the Canny method are specified for each image, with σ_1 representing the standard deviation of the regularization filter and σ_2 represneting the standard deviation of the differentiation filter.

IV. Discussion

In this work, we have reviewed the problem of boundary matching for boundary image quality evaluation. This problem can be seen as an instantiation of a frequent problem in image processing, namely that of linear feature matching. Also, it can be seen as a practical realization of some of the most studied problems in pattern matching, *i.e.* as embodiments of the assignment and the transportation problems. However, most of practical solutions found in literature with similar goals are unapplicable to the present task. We have proposed a novel taxonomy for the different strategies present in the literature. Finally, we have questioned whether the different strategies might lead to significantly different (quantitative) results in boundary quality evaluation. We have found that most of the strategies lead to similar results in terms of evaluation, despite their fundamental differences in inspiration and realization. We report very high similarity in the results by different matching strategies, with two factors negatively influencing such similarity: (a) increasing maximum matching distance t and (b) the boundary images to be rated/ranked contain a large number of spurious responses near the boundaries. Overall, for any fixed matching distance t and dataset, almost all strategies hold very high correlation in terms of the $F_{0.5}$ measure. Also, they all produce very similar rankings, and most of the discrepancies are due to small margins which, given the imperfect reliability of all matching strategies, should not be taken into account.

With this analysis, we can state that none of the strategies for boundary matching leads to significant differences for

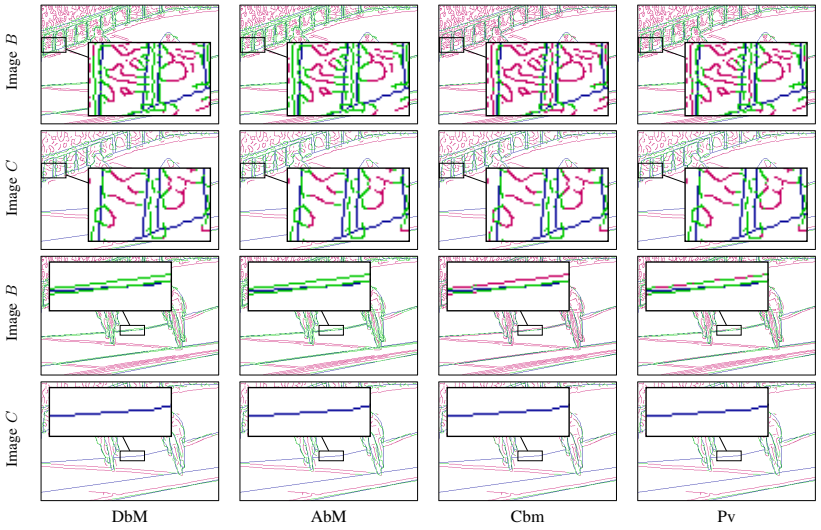


Fig. 4: Visual comparison of the matching made on the triplet of images in Fig. 3. Each image displays the result of the matching, so that (a) green pixels are validated boundaries, (b) pink pixels are unmatched ones and (c) black pixels are the boundaries in the ground truth..

boundary quality evaluation. Regarding the potential variation in rankings when using one or the other, we consider that the strategies might actually alter rankings. However, authors shall understand that neither the matching strategies analyzed in this work, nor any other boundary quality evaluation method, are precise, reliable or exhaustive enough to produce trustworthy, detailed rankings.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the Spanish Ministry of Science (project TIN2016-77356-P, AEI/FEDER, UE), ass well as that of the Research Foundation Flanders (FWO project 3G.0838.12.N).

REFERENCES

[1] I. Abdou and W. Pratt, "Quantitative design and evaluation of enhancement/thresholding edge detectors," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 753–763, 1979.

[2] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, pp. 898–916, 2011.

[3] P. A. Arbelaez, "Une approche métrique pour la segmentation d'images," Ph.D. dissertation, Université Paris-Dauphine, 2005.

[4] A. J. Baddeley, "An error metric for binary images," in *Robust Computer Vision: Quality of Vision Algorithms*, W. Förstner and S. Ruwiedel, Eds. Karlsruhe: Wichmann Verlag, 1992, pp. 59–78.

[5] C. Baillard, C. Schmid, A. Zisserman, and A. Fitzgibbon, "Automatic line matching and 3D reconstruction of buildings from multiple views," in *ISPRS Conf. on Automatic Extraction of GIS Objects from Digital Imagery*, vol. 32, 1999, pp. 69–80.

[6] F. Bergholm, "Edge focusing," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 6, pp. 726–741, 1987.

[7] G. P. Boswell and F. A. Davidson, "Modelling hyphal networks," *Fungal Biology Reviews*, vol. 26, no. 1, pp. 30–38, 2012.

[8] K. Bowyer, C. Kranenburg, and S. Dougherty, "Edge detector evaluation using empirical ROC curves," *Computer Vision and Image Understanding*, vol. 84, no. 1, pp. 77–103, 2001.

[9] J. Canny, "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[10] F. X. Chen, G. Roig, L. Isik, X. Boix, and T. Poggio, "Eccentricity dependent deep neural networks: Modeling invariance in human vision," in *AAAI Spring Symposium Series*, 2017.

[11] J. Dai, J. Feng, and J. Zhou, "Robust and efficient ridge-based palmprint matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1618–1632, 2012.

[12] R. Deriche and O. Faugeras, "Tracking line segments," in *Proc. of the European Conference on Computer Vision*, 1990, pp. 259–268.

[13] F. J. Estrada and A. D. Jepson, "Benchmarking image segmentation algorithms," *International Journal of Computer Vision*, vol. 85, no. 2, pp. 167–181, 2009.

[14] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[15] J. Fram and E. S. Deutsch, "Quantitative evaluation of edge detection algorithms and their comparison with human performance," *IEEE Trans. on Computers*, vol. 24, no. 6, pp. 616–628, 1975.

[16] A. V. Goldberg and R. Kennedy, "An efficient cost scaling algorithm for the assignment problem," *Mathematical Programming*, vol. 71, pp. 153–177, 1995.

[17] J. Ha, "Real-time visual tracking using image processing and filtering methods," Ph.D. dissertation, Georgia Institute of Technology, 2008.

[18] J. H. Han and J. S. Park, "Contour matching using epipolar geometry," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 4, pp. 358–370, 2000.

[19] Y. Han, G. Roig, G. Geiger, and T. Poggio, "Is the human visual system

invariant to translation and scale?" in *AAAI Spring Symposium Series*, 2017.

[20] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 9, no. 4, pp. 532–550, 1987.

[21] A. Herskovits and T. O. Binford, "On boundary detection," Massachusetts Institute of Technology, Tech. Rep., 1970, project MAC Memo no. 183.

[22] S. Jaggi, W. C. Karl, S. G. Mallat, and A. S. Willsky, "Silhouette recognition using high-resolution pursuit," *Pattern Recognition*, vol. 32, no. 5, pp. 753–771, 1999.

[23] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition*, vol. 42, no. 7, pp. 1408–1418, 2009.

[24] R. Koren and Y. Yitzhaky, "Automatic selection of edge detector parameters based on spatial and statistical measures," *Computer Vision and Image Understanding*, vol. 102, no. 2, pp. 204–213, 2006.

[25] H. W. Kuhn, "The Hungarian method for the assignment problem," *Naval Research Logistics Quarterly*, vol. 2, no. 1-2, pp. 83–97, 1955.

[26] A. Kumar and D. Zhang, "Personal recognition using hand shape and texture," *IEEE Trans. on Image Processing*, vol. 15, no. 8, pp. 2454–2461, 2006.

[27] H. Kuwajima, M. Tanaka, and M. Okutomi, "Improving transparency of deep neural inference process," *Progress in Artificial Intelligence*, vol. 8, no. 2, pp. 273–285, 2019.

[28] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, pp. 117–156, 1998.

[29] G. Liu and R. M. Haralick, "Optimal matching problem in detection and recognition performance evaluation," *Pattern Recognition*, vol. 35, no. 10, pp. 2125–2139, 2002.

[30] X. Liu, Y. Yu, B. Liu, and Z. Li, "Bowstring-based dual-threshold computation method for adaptive Canny edge detector," in *Proc. of the International Conf. of Image and Vision Computing New Zealand*, 2013, pp. 13–18.

[31] Y. Liu, T. S. Huang, and O. D. Faugeras, "Determination of camera location from 2-D to 3-D line and point correspondences," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 28–37, 1990.

[32] Z. Liu and S. Sarkar, "Effect of silhouette quality on hard problems in gait recognition," *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 2, pp. 170–183, 2005.

[33] C. Lopez-Molina, B. De Baets, and H. Bustince, "Quantitative error measures for edge detection," *Pattern Recognition*, vol. 46, no. 4, pp. 1125–1139, 2013.

[34] ——, "Twofold consensus for boundary detection ground truth," *Knowledge-Based Systems*, vol. 98, pp. 162–171, 2016.

[35] C. Lopez-Molina, B. De Baets, H. Bustince, J. Sanz, and E. Barrenechea, "Multiscale edge detection based on Gaussian smoothing and edge tracking," *Knowledge-Based Systems*, vol. 44, pp. 101–111, 2013.

[36] B. Magnier, "Edge detection: a review of dissimilarity evaluations and a proposed normalized measure," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9489–9533, 2018.

[37] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, *Handbook of fingerprint recognition*, Springer, Ed. Springer-Verlag, 2009.

[38] A. N. Marana and A. K. Jain, "Ridge-based fingerprint matching using hough transform," in *Proc. of the Brazilian Symposium on Computer Graphics and Image Processing*, 2005, pp. 112–119.

[39] D. Marr, *Vision*. MIT Press, 1982.

[40] D. Marr and E. Hildreth, "Theory of edge detection," *Proceedings of the Royal Society of London*, vol. 207, no. 1167, pp. 187–217, 1980.

[41] D. Marr and T. Poggio, "A theory of human stereo vision." Massachusetts Institute of Technology, Tech. Rep. No. AI-M-451, 1977.

[42] D. R. Martin, "An empirical approach to grouping and segmentation," Ph.D. dissertation, University of California, Berkeley, 2003.

[43] D. Martin, C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 530–549, 2004.

[44] J. C. McEachen, J. S. Duncan *et al.*, "Shape-based tracking of left ventricular wall motion," *IEEE Trans. on Medical Imaging*, vol. 16, no. 3, pp. 270–283, 1997.

[45] G. Medioni and R. Nevatia, "Segment-based stereo matching," *Computer Vision, Graphics, and Image Processing*, vol. 31, no. 1, pp. 2–18, 1985.

[46] J. Metzler and R. N. Shepard, *Theories in cognitive psychology: The Loyola Symposium*. Lawrence Erlbaum, 1974, ch. Transformational studies of the internal representation of three-dimensional objects.

[47] K. Mikolajczyk, A. Zisserman, and C. Schmid, "Shape recognition with edge-based features," in *Proc. of the British Machine Vision Conference*, vol. 2, 2003, pp. 779–788.

[48] F. Mokhtarian and A. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, no. 1, pp. 34–43, 1986.

[49] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, no. 8, pp. 789–805, 1992.

[50] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the Society for Industrial and Applied Mathematics*, vol. 5, no. 1, pp. 32–38, 1957.

[51] N. M. Nasrabadi, "A stereo vision technique using curve-segments and relaxation matching," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, no. 5, pp. 566–572, 1992.

[52] A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2015, pp. 427–436.

[53] G. Papari, P. Campisi, N. Petkov, and A. Neri, "A biologically motivated multiresolution approach to contour detection," *EURASIP Journal on Advances in Signal Processing*, vol. 2007, 2007, article ID 71828.

[54] G. Papari and N. Petkov, "Edge and line oriented contour detection: State of the art," *Image and Vision Computing*, vol. 29, no. 2-3, pp. 79–103, 2011.

[55] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[56] T. Poggio, J. Mutch, and L. Isik, "Computational role of eccentricity dependent cortical magnification," Center for Brains, Minds and Machines, Massachusets Inst. of Technology, Tech. Rep., 2014.

[57] F. J. Provost and T. Fawcett, "Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions." in *Proc. of the International Conf. on Knowledge Discovery & Data Mining*, vol. 97, 1997, pp. 43–48.

[58] K. Research Unit (Ghent University), "The Kermit Image Toolkit (KITT)." [Online]. Available: www.kermitimagetoolkit.com

[59] M. Riesenhuber and T. Poggio, "Hierarchical models of object recognition in cortex," *Nature Neuroscience*, vol. 2, no. 11, p. 1019, 1999.

[60] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 7–42, 2002.

[61] C. Schmid and A. Zisserman, "Automatic line matching across views," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 1997, pp. 666–671.

[62] T. Serre, L. Wolf, and T. Poggio, "Object recognition with features inspired by visual cortex," Dept. of Brain and Cognitive Sciences, Massachusets Inst. of Technology, Tech. Rep., 2006.

[63] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects," *Science*, vol. 171, no. 3972, pp. 701–703, 1971.

[64] F. Tupin, H. Maitre, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: Application to road network extraction," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 36, no. 2, pp. 434–453, 1998.

[65] S. G. Vandenberg and A. R. Kuse, "Mental rotations, a group test of three-dimensional spatial visualization," *Perceptual and Motor Skills*, vol. 47, no. 2, pp. 599–604, 1978.

[66] G. Vidal-Diez de Ulzurrun, J. Baetens, J. Van den Bulcke, and B. De Baets, "Modelling three-dimensional fungal growth in response to environmental stimuli," *Journal of Theoretical Biology*, vol. 414, pp. 35–49, 2017.

[67] W. Waegeman, B. De Baets, and L. Boullart, "ROC analysis in ordinal regression learning," *Pattern Recognition Letters*, vol. 29, no. 1, pp. 1–9, 2008.

[68] Y. Wang, E. K. Teoh, and D. Shen, "Lane detection and tracking using B-Snake," *Image and Vision Computing*, vol. 22, no. 4, pp. 269–280, 2004.

[69] Z. Wang, Z. Chi, and D. Feng, "Shape based leaf image retrieval," in *IEE Proc.- Vision, Image and Signal Processing*, vol. 150, no. 1. IET, 2003, pp. 34–43.

[70] A. P. Witkin, "Scale-space filtering," in *Proc. of the International Joint Conf. on Artificial Intelligence*, vol. 2, 1983, pp. 1019–1022.

[71] C. Xu, L. Zhang, L. Cheng, and R. Koch, "Pose estimation from line correspondences: A complete analysis and a series of solutions," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1209–1222, 2017.

[72] A. Yilmaz, X. Li, and M. Shah, "Contour-based object tracking with occlusion handling in video acquired using mobile cameras," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 11, pp. 1531–1536, 2004.

[73] Y. Yitzhaky and E. Peli, "A method for objective edge detection evaluation and detector parameter selection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 1027–1033, 2003.

[74] M. Yokoyama and T. Poggio, "A contour-based moving object detection and tracking," in *Proc. of the International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, 2005, pp. 271–276.

[75] F. Zana and J.-C. Klein, "A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform," *IEEE Trans. on Medical Imaging*, vol. 18, no. 5, pp. 419–428, 1999.

3.6 Optical images-based edge detection in Synthetic Aperture Radar images

- G. P. Silva, A. C. Frery, S. Sandri, H. Bustince, E. Barrenechea, and C. Marco-Detchart, "Knowledge-Based Systems Optical images-based edge detection in Synthetic Aperture Radar images", *Knowledge-Based Systems*, vol. 87, pp. 38–46, 2015

  - Journal: Knowledge-Based Systems
  - Status: Published
  - Impact Factor: (JCR 2015) 3,433
  - Knowledge area:
    * Computer Science, Artificial Intelligence: Ranking $17/130$ (Q1)

Knowledge-Based Systems 87 (2015) 38–46

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Optical images-based edge detection in Synthetic Aperture Radar images

Gilberto P. Silva Junior [a], Alejandro C. Frery [b], Sandra Sandri [a,*], Humberto Bustince [c], Edurne Barrenechea [c], Cédric Marco-Detchart [c]

[a] Laboratório Associado de Computação e Matemática Aplicada, Instituto Nacional de Pesquisas Espaciais (LAC/INPE), Av. dos Astronautas, 1758, 12227-010 São José dos Campos, SP, Brazil
[b] Laboratório de Computação Científica e Análise Numérica (LaCCAN/UFAL), Universidade Federal de Alagoas, Av. Lourival Melo Mota, s/n, 57072-970 Maceió, AL, Brazil
[c] Departamento de Automática y Computación, Universidad Pública de Navarra, Campus Arrosadía, 31006 Pamplona, Spain

ARTICLE INFO

Article history:
Received 2 November 2014
Received in revised form 26 June 2015
Accepted 24 July 2015
Available online 1 August 2015

Keywords:
Edge detection
SAR images
Computational Intelligence
Gravitational method

ABSTRACT

We address the issue of adapting optical images-based edge detection techniques for use in Polarimetric Synthetic Aperture Radar (PolSAR) imagery. We modify the gravitational edge detection technique (inspired by the Law of Universal Gravity) proposed by Lopez-Molina et al., using the non-standard neighbourhood configuration proposed by Fu et al., to reduce the speckle noise in polarimetric SAR imagery. We compare the modified and unmodified versions of the gravitational edge detection technique with the well-established one proposed by Canny, as well as with a recent multiscale fuzzy-based technique proposed by Lopez-Molina et al. We also address the issues of aggregation of gray level images before and after edge detection and of filtering. All techniques addressed here are applied to a mosaic built using class distributions obtained from a real scene, as well as to the true PolSAR image; the mosaic results are assessed using Baddeley's Delta Metric. Our experiments show that modifying the gravitational edge detection technique with a non-standard neighbourhood configuration produces better results than the original technique, as well as the other techniques used for comparison. The experiments show that adapting edge detection methods from Computational Intelligence for use in PolSAR imagery is a new field worthy of exploration.

1. Introduction

Edge detection seeks to identify sharp differences automatically in the information associated with adjacent pixels in an image [1]. Edge detection for optical images is nowadays quite an established field. It is traditionally carried out using gradient-based techniques, such as the well-known Canny algorithm [2]. Techniques based on Computational Intelligence have also been proposed in the recent literature. Sun et al. [3] proposed the gravitational edge detection method, inspired by Newton's Universal Law of Gravity. Lopez-Molina et al. [4] proposed a fuzzy extension for this technique, allowing the use of T-norms, a large class of fuzzy operators; they also proposed small modifications in the basic formalism (see Section 3). Danková et al [5] proposed the use of a fuzzy-based function, the F-transform; the original universe of functions is transformed into a universe of their skeleton models (vectors of F-transform components), making further computations easier to perform. Barrenechea et al. [6] proposed the use of interval-valued fuzzy relations for edge detection, using a T-norm and a T-conorm to produce a fuzzy edge image, that is then binarized. This approach was extended by Chang and Chang [7]. First of all, two new images are created—one rather dark and the other rather bright—by applying two different parameters on the linear combinations of the images obtained using min and max operators, respectively. Then, the fuzzy edge image is created by the difference between these two new images. Another recent approach from Computational Intelligence is the multiscale edge detection method proposed by Lopez-Molina et al. [8], using Sobel operators for edge extraction and the concept of Gaussian scale-space.

SAR sensors are not as adversely affected by atmospheric conditions and the presence of clouds as optical sensors. Moreover, unlike the optical counterparts, SAR sensors can be used at any time of day or night. For these reasons, remote sensing applications using SAR imagery have been growing over the years [9]. SAR images, however, contain a great amount of noise, known as *speckle*, that degrades the visual quality of the images. Caused by inherent characteristics of radar technology, this multiplicative

* Corresponding author.
E-mail address: sandra.sandri@inpe.br (S. Sandri).

http://dx.doi.org/10.1016/j.knosys.2015.07.030
0950-7051/

non-Gaussian noise is proportional to the intensity of the received signal.

Contrary to what happens with optical images, there are still few algorithms specifically dedicated to SAR images [10]. One interesting means to create edge detection algorithms for SAR images is to modify those created for optical images. However, the use of these methods on SAR images is not straightforward, due to speckle. One can either adapt optical image techniques to meet SAR data properties, or first preprocess the images using filters and then apply the original optical techniques.

The main purpose of our study is to investigate the application of the gravitational edge detection, Here we modify the original 3×3 window: the value in each cell in the window is no longer the original one, but the aggregation of a set of neighbouring pixels, according to the larger 9×9 neighbourhood configuration proposed by Fu et al. [10]. We propose a typology of experiments to study the behaviour of the modified edge detection method, considering polarization, image aggregation, and image binarization. We focus on the use of the following processes: DAB (edge Detection on non-binary images, Aggregation of the resulting non-binary images, Binarization) and ADB (Aggregation of non-binary images, edge Detection on the resulting non-binary image, Binarization).

We also investigate the use of noise-reduction filters in preprocessing the images, by making use of the well-known Enhanced Lee filter [11] and a filter recently proposed by Torres et al. [12].

da Silva et al. [13] describe a classification experiment, based on a full polarimetric image from an agricultural area in the Amazon region in Brazil. In that study, the authors estimated the parameters for probability distributions associated to each of the classes of interest, such as water and different types of vegetation and their phenology. They assessed their results in an image formed by a mosaic of the classes, with pixel values generated using the parameters found for each class. We apply all techniques addressed in this study on twenty simulated mosaics, using the parameters estimated in [13], considering amplitude images derived from different polarizations. We assess the quality of the results, according to Baddeley's Delta Metric (BDM) [14].

We also apply the methods on the real images, but assessment is only visual. We compare our results with those produced by the use of Canny's algorithm [2] and the recently proposed multiscale method by Lopez-Molina et al. [8].

The present study is an extended version of [15], in which some of the main ideas of this paper were first delineated. However, the present study and [15] differ in the scope of the proposed approach as well as in the reliability of the results. Indeed, in [15], only one simulated image was used in the experiments and only Canny's technique was compared to its results. Moreover, in the previous paper we only addressed the edge detection of the image resulting from the aggregation of the three simulated polarization images. In our first paper only ADB was addressed; edge detection on the individual polarization images as well as DAB strategy were not considered.

The results from our current study show that adapting edge detection methods from Computational Intelligence to use in radar imagery is a new field worthy of exploration. In particular, our experiments show that modifying the gravitational method with Fu's 9×9 neighbourhood produces better results than the unmodified method. They also show the importance of filtering when adapting edge detection techniques from optical to radar images.

2. Basic concepts on SAR images

Optical and SAR sensors measure the amount of energy reflected by a target in various bands of the electromagnetic spectrum. The bands employed in most imaging radars use frequencies in the 2 MHz to 12.5 GHz range, with wavelengths ranging from 2.4 cm to 1 m. In this study, we used only the L-band with wavelengths of [30 cm, 1 m] and frequencies of [1 MHz, 2 GHz].

SAR systems generate the image of a target area by moving along a usually linear trajectory, and transmitting pulses in lateral looks towards the ground, in either horizontal (H) or vertical (V) polarizations [16], respectively denoted as H and V (see Fig. 1). In the past, the reception of the transmitted energy was made solely on the same polarization of the transmission, generating images in the HH and VV polarizations. Currently, with the advent of polarized and fully polarimetric radars (PolSAR – *Polarimetric Synthetic Aperture Radar*), information about intensity and phase of the cross signals are also obtained, generating images relating to HV and VH polarizations. Usually, applications only consider the HH, VV, and HV polarizations.

The imaging can be obtained by gathering all the intensity and phase information data from the electromagnetic signal after it has been backscattered by the target in a given polarization [18]. Each polarization in a given a scene generates a complex image, which can be thought of as two images, containing the real and imaginary values for the pixels, respectively.

We denote the complex images from HH, VV, and HV polarizations as S_{HH}, S_{HV}, and S_{VV}. Multiplying the vector $[S_{HH}\ S_{HV}\ S_{VV}]$ by its transposed conjugated vector $[S^*_{HH}\ S^*_{HV}\ S^*_{VV}]^t$, we obtain a 3×3 covariance matrix. The main diagonal contains intensity values; taking their square root, we obtain amplitude values. We denote the intensity images by I_{HH}, I_{HV}, and I_{VV} and their corresponding amplitude counterparts by A_{HH}, A_{HV}, and A_{VV}. In this paper, we only considered the amplitude images, such as those depicted in Fig. 2.

Speckle noise is multiplicative, non-Gaussian, and is proportional to the intensity of the received signal. Speckle degrades the visual quality of the displayed image by sudden variations in image intensity with a salt and pepper pattern, as can be seen in Fig. 2. It can be reduced with multiple looks in the generation of the complex images, causing degradation in spatial resolution. Another way to reduce noise is to employ filters, as will be discussed in the next section.

In SAR image classification, one often uses samples from the classes in order to estimate the parameters of the distribution believed to underlie each class. Synthetic images can then be created using Monte Carlo simulation by taking the realization of the random variable associated to the class of each classified pixel. This artifice is useful to choose the most apt classifier for a given application: instead of relying solely on the original image, one takes the classifier that obtains the best average accuracy on the set of synthetic images. This methodology can also be used in other tasks, such as edge detection.

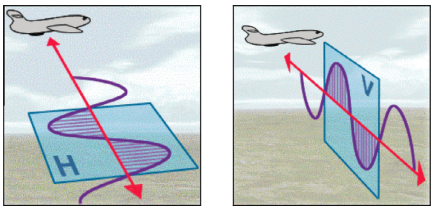


Fig. 1. Horizontal and vertical signal polarizations transmitted by an antenna. *Source:* [17].

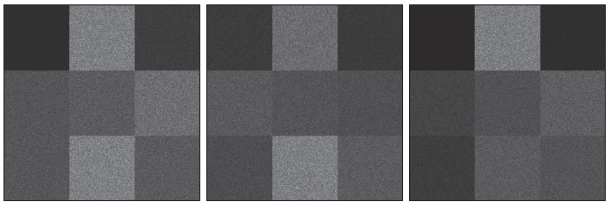


Fig. 2. Amplitude images for polarizations HH, VV, and HV from the same scene.

3. Related work

One of the most successful edge detection algorithms for optical images was proposed by Canny [2], based on the following guidelines: (i) the algorithm should mark as many real edges in the image as possible; (ii) the marked edges should be as close as possible to the edge in the real image; (iii) a given edge in the image should only be marked once; and (iv) image noise should not create false edges. It makes use of numerical optimization to derive optimal operators for ridge and roof edges. The usual implementation of this method uses a 3×3 neighbourhood.

A more recent multi-scale edge detection method was proposed by Lopez-Molina et al. [8], using Sobel operators for edge extraction and the concept of Gaussian scale-space. More specifically, the Sobel edge detection method is applied on increasingly smoother versions of the image. Then, the edges which appear on different scales are combined by performing coarse-to-fine edge tracking.

The gravitational edge detection approach was first proposed by Sun et al. [3] and applied to optical images. It is based on Newton's Universal Law of Gravity, described by Eq. (1):

$$f_{1,2} = G \times \frac{m_1 \times m_2}{\|\vec{r}\|^2} \times \frac{\vec{r}}{\|\vec{r}\|}, \tag{1}$$

where m_1 and m_2 are the masses of two bodies; $\vec{r}$ is the vector connecting them; $\vec{f}_{1,2}$ is the gravitational force between them; $\|\cdot\|$ denotes the magnitude of a vector; and G is the gravitational constant. In the analogy proposed by Sun et al. [3]; the bodies are the gray level values of pixels in a grid; G is a function of the values of the pixels in a given window; the distance between any two adjacent pixels is equal to 1; and, when computing the resulting force of the pixel in the center of a window; the pixels outside that window are considered negligible. Lopez-Molina et al. [4] extended this technique, proposing the use of a Triangular Norm [19] in place of the product between the two masses,[1] by first normalizing the gray level values to $[0, 1]$. The authors treat edges as fuzzy sets for which membership degrees are extracted from the resulting gravitational force on each pixel. They take G as a normalization constant, calculated so as to guarantee that the resulting forces lie in $[0, 1]$. Also, in the normalization of gray level values into $[0, 1]$, a small value δq is added beforehand to both the numerator and denominator so as avoid pixels with value 0, which would have too strong an effect on neighbouring pixels. The authors used 3×3 and 5×5 windows as well as several prototypical triangular norms.

The so-called Lee (or sigma) filter introduced in 1983 [20], is still in use today due to its simplicity, its effectiveness in speckle reduction, and its computational efficiency. It is based on the fact that, under Gaussian distribution, approximately 95.5% of the probability is concentrated within two standard deviations from the mean. The filter estimates the mean and the standard deviation of samples around each pixel, and only those values within this interval are used to compute the local mean. Lopes et al. [11] proposed an adaptive version for this filter, here referred to as "Enhanced Lee".

Torres et al. [12] recently proposed a nonlocal means approach for PolSAR image speckle reduction based on stochastic distances; the method can be tailored to any distribution, both univariate of multi-variate. It consists of comparing the distributions which describe the central observation for each pixel, and each of the observations which comprise a search region. The comparison is made through a goodness-of-fit test, and the p-value of the test statistic is used to define the convolution matrix which will define the filter: the higher the p-value the larger the confidence and, thus, the importance, each observation will have in the convolution. In Torres et al.'s proposal, the tests are derived from h–ϕ divergences between multi-look scaled complex Wishart distributions for fully PolSAR data [21]. Their results are competitive with classical and advanced polarimetric filters, with respect to usual quantitative measures of quality.

Fu et al. [10] proposed a statistical edge detector suitable for SAR images which uses the squared successive difference of averages to estimate the edge strength from the sliding window. An interesting feature of this paper is the proposal of a specific type of 9×9 neighbourhood, shown in Fig. 3. In a previous paper [15], we proposed a modification of the gravitational approach using Fu et al.'s neighbourhood: given a central pixel in a 3×3 window in an image, the values considered for the surrounding pixels in the window are no longer the ones in the original image, but the mean values in this new configuration.

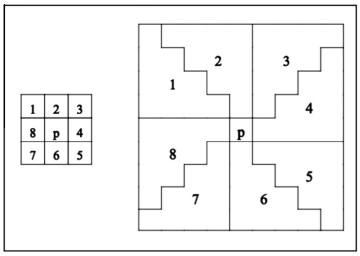


Fig. 3. Standard 3×3 and Fu's neighbourhood [10].

[1] Triangular norm operators are mappings from $[0, 1]^2$ to $[0, 1]$, that are commutative, associative, monotonic, and have 1 as neutral element.

4. Materials and methods

We compare the edge detection methods proposed by Canny [2] and by Lopez-Molina et al. [8] to the modified gravitational approach using the product T-norm, followed by thresholding. The effect of preprocessing the images through filtering is also studied, using the filter described by Torres et al. [12] and the Enhanced Lee filter [11]. We study the behaviour of Lopez-Molina's method with the usual 3×3 window as well as a modified version of this approach, proposed in a previous paper [15], involving the neighbourhood proposed by Fu et al. [10].

The input for Canny's and Lopez-Molina's edge detector are images in, respectively, $\{0, \ldots, 255\}$ and $[0, 1]$. Image values are, thus, mapped into these sets prior to edge detection. For the Lopez-Molina methods (the original and modified versions), we normalize further to $[0, 1]$, using $\delta q = 1$ and making $q' = (q + 1)(255 + 1)^{-1}$, where q and q' are the old and new value of a given pixel, respectively.

4.1. Working image

We apply the methods on data derived from a fully polarimetric image, presented by da Silva et al. [13], from an agricultural area in the Amazon region in Brazil (see Fig. 4). The authors describe a classification experiment using classes of interest from that area, such as water and different types of crops and natural vegetation, at different stages of growth. Samples from the classes from band L are used to estimate the parameters of the complex Wishart distribution associated to each class. The results are assessed using a mosaic with the classes that was created using the derived Wishart distributions. Fig. 5 illustrates the approach. For our study we apply the edge detection methods on twenty independently simulated mosaics amplitude images, using the parameters estimated in [13] to assess the quality of the methods.

4.2. Quality assessment

The quality of the results is assessed by the Baddeley's Delta Metric (BDM) [14], by comparison with what would be the perfect result, discarding those pixels close to the outer frame.

Let $\mathbf{x}$ and $\mathbf{y}$ be two binary images, seen as mappings from Λ to $[0, 1]$, where Λ is a set of sites arranged in a grid (positions). Let ρ be a metric on Λ, such as the Euclidean distance, and $d(i, A)$ be the distance between a site i and a set $A \subseteq \Lambda$, defined as

$$d(i, A) = \min_{j \in A} \rho(i, j).$$

Let $b(\mathbf{x}) = \{i \in \Lambda | x_i = 1\}$ denote the set of foreground sites in $\mathbf{x}$. BDM between $\mathbf{x}$ and $\mathbf{y}$, denoted as $\Delta_{p,w}(\cdot, \cdot)$, is then defined as

$$\Delta_{p,w}(\mathbf{x}, \mathbf{y}) = \left(\frac{1}{|\Lambda|} \sum_{i \in \Lambda} |w(d(i, b(\mathbf{x})) - w(d(i, b(\mathbf{y}))|^p \right)^{\frac{1}{p}}, 1 \leqslant p \leqslant \infty \tag{2}$$

where w is a strictly increasing concave function satisfying $w(0) = 0$. Here we use $w(t) = t$ and $p = 2$, as in [4].

Throughout the text, we display BDM results in $[0, 100]$ instead of $[0, 1]$, for the sake of readability.

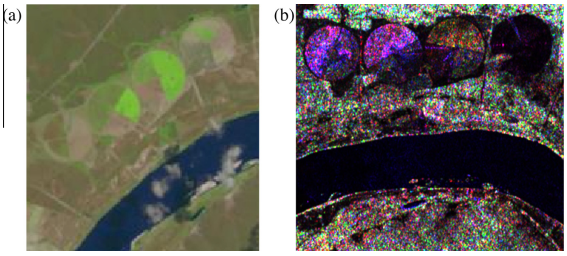


Fig. 4. Images derived from a scene in Bebedouro in Brazil (not registered): (a) Landsat RGB composition and (b) SAR L-band RGB composition. *Source:* [13].

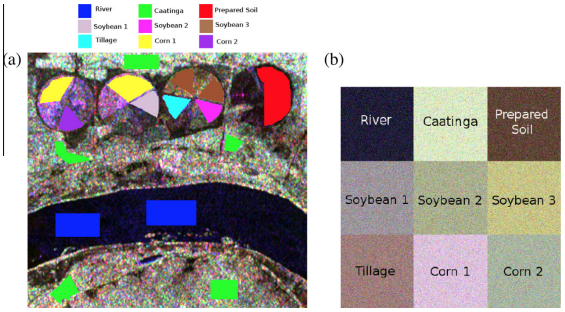


Fig. 5. Images derived from a scene in Bebedouro in Brazil: (a) training samples used to generate Wishart distributions and (b) synthetic mosaic images generated using the Wishart distributions estimated in [13] from image samples. *Source:* [13].

5. Proposed methodologies

Edge detectors use a window around a center pixel to verify whether that pixel belongs to an edge or not. When adapting optical image edge detectors to radar imagery, we have to find the means to deal with speckle. The main contribution of this study is to modify the original 3×3 window used by the edge detection method proposed in [4] for use in radar imagery such that the value in each cell in the window is no longer the original one but the aggregation of set of neighbouring pixels, according to a larger 9×9 non-standard neighbourhood proposed by Fu et al. [10]. We here investigate this particular combination of method and neighbourhood, but the same procedure can be applied using other edge detection methods and/or non-standard filters.

Frequently, a single band is used in edge-detection, resulting in a gray level-image that is then binarized at some point (the usual implementation of some methods, like Canny's, already involve a binarization step). In radar imagery, very often one deals with more than one band at the same time (e.g. intensity images coming from different polarizations, or complex images in the fully polarimetric case), aiming at using the richness of information to compensate for the speckle noise. Therefore, the question of when to aggregate results has to be addressed. One may, for instance, first aggregate the bands and then apply the edge detector on the aggregated image, or else apply the edge detector on the individual bands and then aggregate the edge images. These two methods usually yield different results.

Here, we propose a typology for experiments using radar imagery, considering different orderings of three steps: edge detection on gray level images, binarization of gray level images, and aggregation of results. In the aggregation step, the input may be either gray level or binary images, depending on whether the binarization is made immediately after edge detection or not. Three strategies can then be envisaged to perform edge detection experiments with radar images:

- DAB (edge Detection on non-binary images, Aggregation of the resulting non-binary images, Binarization).
- ADB (Aggregation of non-binary images, edge Detection on the resulting non-binary image, Binarization).
- DBA (edge Detection on non-binary images, Binarization, Aggregation of the resulting binary images).

Options ABD, BAD and BDA are not considered, since that would mean applying edge detectors on the binary images.

In this work, we focus on the DAB and ADB strategies. For both of them, we use the arithmetic mean to aggregate gray level images. Strategy DBA, involving the aggregation of binary images, is left for future study.

When no aggregation is considered, the strategies are reduced to only edge detection and binarization. For example, for the HH, HV and VV polarizations, we obtain strategies DB-HH, DB-HV and DB-HH. Note that some methods already incorporate the binarization step in the edge detector. That is for instance the case of all methods discussed previously. However, to be consistent with the notation, we will denote by ADB the strategy in a method that includes binarization, such as Canny's and the multi-scale method, when it is applied to the gray level image resulting from the aggregation of the images from the HH, HV and VV polarizations.

6. Experimental results

The output of the Lopez-Molina gravitational method is an image with values in $[0, 1]$. In order to obtain binary indicators of edges, the authors use a hysteresis transformation. Here, we use a simple threshold and search for the value in the $[0.05, 0.15]$ interval which produces the best BDM. For Canny's method, we search for the best value for the noise standard deviation parameter σ in the interval $[0.3, 1.5]$. The intervals above for both Canny and Lopez-Molina are the ones that presented the best results by trial-and-error. The following parameters were used for the Lopez-Molina multi-scale method, as suggested in [8]: $\delta_\sigma = 0.25$, $\sigma \in \{0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00, 2.25, 2.50, 2.75, 3.00, 3.25, 3.50, 3.75, 4.00\}$.

We applied two filters Torres et al. [12] and Enhanced Lee [11] on intensity values, which were then transformed in amplitude before further processing.

Tables 1–4 show the results for BDM mean and standard deviation after applying four methods to twenty simulated mosaic images: Canny's method, Lopez-Molina et al.'s multi-scale method, Lopez-Molina et al.'s original gravitational method, and Lopez-Molina et al.'s method modified using Fu's 9×9 neighbourhood.

We see that the best BDM average values were obtained with the use of Lopez-Molina et al.'s gravitational method modified by Fu's neighbourhood, using the ADB and DAB strategies, both preprocessed with the Enhanced Lee filter. Both are significantly higher than the other procedures.

Table 1
Average BDM results for Canny's method, with standard deviation inside parentheses.

Strategy	No filter	Torres filter	Enh. Lee filter
DB-HH	26.72 (1.22)	23.40 (1.92)	28.85 (1.86)
DB-HV	30.70 (1.39)	26.74 (1.33)	66.40 (0.42)
DB-VV	29.81 (2.17)	26.28 (1.09)	24.83 (1.98)
ADB	27.87 (1.16)	48.16 (0.003)	30.24 (1.20)

Table 2
Average BDM results for the multi-scale method, with standard deviation inside parentheses.

Strategy	No filter	Torres filter	Enh. Lee filter
DB-HH	28.23 (0.89)	28.00 (0.93)	25.36 (2.11)
DB-HV	28.34 (1.13)	24.55 (2.97)	19.56 (2.36)
DB-VV	25.62 (1.10)	28.42 (0.56)	28.89 (2.25)
ADB	25.15 (3.23)	24.37 (1.52)	20.71 (3.97)

Table 3
Average BDM results for the gravitational method; with standard deviation inside parentheses.

Strategy	No filter	Torres filter	Enh. Lee filter
DB-HH	33.89 (1.98)	26.61 (2.00)	38.97 (0.79)
DB-HV	31.95 (0.46)	27.14 (1.22)	32.26 (2.78)
DB-VV	32.35 (1.47)	28.95 (0.95)	43.65 (1.13)
DAB	29.26 (1.62)	25.91 (1.55)	27.71 (2.62)
ADB	31.50 (0.82)	26.63 (1.27)	18.24 (3.41)

Table 4
Average BDM results for the gravitational method modified by Fu's neighbourhood, with standard deviation inside parentheses.

Strategy	No filter	Torres filter	Enh. Lee filter
DB-HH	25.27 (0.76)	22.18 (0.48)	17.79 (3.05)
DB-HV	26.48 (1.00)	24.21 (0.65)	18.40 (5.75)
DB-VV	21.41 (1.97)	18.14 (0.77)	17.83 (2.54)
DAB	22.67 (2.29)	18.97 (1.62)	5.43 (1.68)
ADB	23.80 (2.23)	22.74 (0.50)	5.16 (0.36)

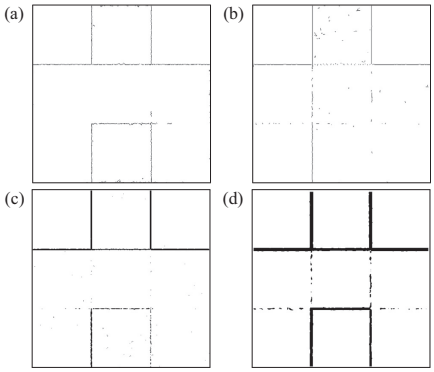


Fig. 6. Best BDM results obtained from the best methods (average): (a) Canny (DB-HH, with Torres filtering, BDM = 18.51), (b) multi-scale (DB-HV, with Enh. Lee filtering, BDM = 14.94), (c) gravitational (ADB with Enh. Lee filtering, BDM = 10.96) and (d) gravitational and Fu (ADB with Enh. Lee filtering, BDM = 3.05).

In Table 1, we see that filtering did not have a significative impact on Canny's detector. The same is true, to a lesser degree, for most results of the multi-scale and (unmodified) gravitational methods, as can be seen in Table 2 and 3. In these methods, there is a slight advantage in preprocessing the images using the Enhanced Lee filter. However, filtering has an impressive effect on the Lopez-Molina gravitational method modified with Fu's neighbourhood. In particular, the best results are obtained for strategies DAB and ADB with preprocessing with the Enhanced Lee filter.

Fig. 6 shows the negative images corresponding to the best results, according to BDM, obtained by the edge detection methods and the filtering strategies with the best average values; note that the image boundaries are depicted only for illustrative purposes. We see that according to BDM the best binary image (depicted in Fig. 6b) presents little noise and most of the regions are separated, even though the lines are rather thick. We also see that BDM was able to distinguish the best image from the others.

Fig. 6b shows the best results from the methods come from filtered images, which raises the question of how important preprocessing by filtering is. In what follows we discuss the details of the gravitational method using the original 3×3 and Fu's 9×9 neighbourhood in relation to filtering. We take the simulation that obtained the best BDM results for each type of neighbourhood. We see in these examples, that filtering does, indeed, ameliorate the results for all methods.

Fig. 7a–c respectively show that: using the 3×3 neighbourhood for the original gravitational method, the unfiltered image is very noisy; the Torres filter reduced the noise and separated the regions; and the Lee filter detected false edges. In Fig. 8, we see that Fu's 9×9 neighbourhood detected almost all the edges, especially when using Lee's filter. Filtering for the modified method presented a larger trade-off between detection of edges and reduction of noise (some edges were detected using Torres filter with an increase of noise when compared to the unfiltered image).

When we compare the results in Figs. 7 and 8 we see that the gravitational method modified with Fu's 3×3 neighbourhood clearly produced better results than the method with the original 3×3 window, which agrees with the BDM evaluation.

Figs. 9–11 show the best edge detection results obtained for the Bebedouro SAR image in terms of visual analysis, with the application of the Enhanced Lee filter, for the parameterizations used here. The figures present the results for HV polarization; in general, HH (respec. VV) polarization produced images with more noise (respec. less information) than HV. Figs. 10 and 11, respectively, show the results of the application of ADB and DAB aggregation

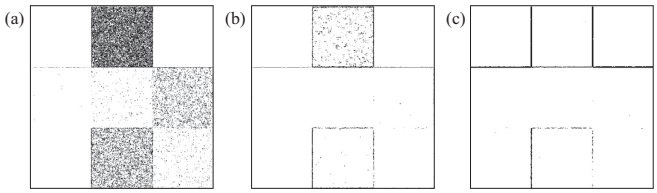


Fig. 7. Results for the gravitational method with original 3×3 window, on a single simulated image from ADB: (a) no filtering (BDM = 31.90), (b) Torres (BDM = 27.73) and (c) Enh. Lee (BDM = 10.96).

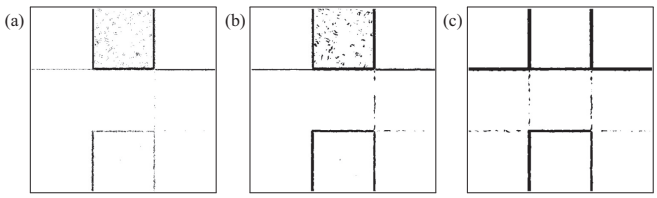


Fig. 8. Results for the gravitational method modified with Fu's neighbourhood, on a single simulated image from ADB: (a) no filtering (BDM = 23.80), (b) Torres (BDM = 23.76) and (c) Enh. Lee (BDM = 3.05).

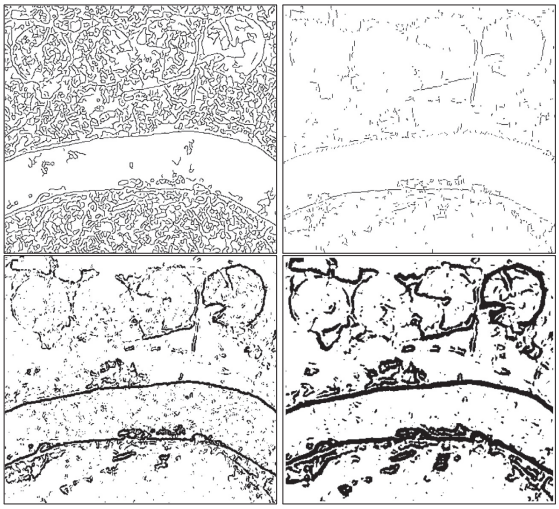


Fig. 9. HV Bebedouro binary images, with Enh. Lee filter: the first row depicts results obtained using the Canny and the Multi-scale methods; and the second row depicts results obtained using the original Gravitational method and its modification with Fu's neighbourhood, (the latter two methods use binarization threshold = .2).

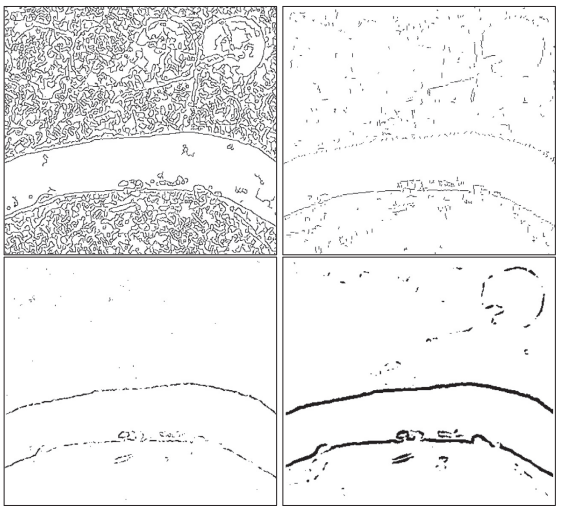


Fig. 10. ADB Bebedouro binary images, with Enh. Lee filter: the first row depicts results obtained using the Canny and the Multi-scale methods; and the second row depicts results obtained using the original Gravitational method and its modification with Fu's neighbourhood, (the latter two methods use binarization threshold = .2).

strategies. ADB in general produced binary images with very little information for all methods; the decrease of noise in relation to the individual polarizations does not compensate the lack of information. In general, the DBA aggregation method produced results with less noise for the gravitational method, with and without modification, than the results obtained with the individual

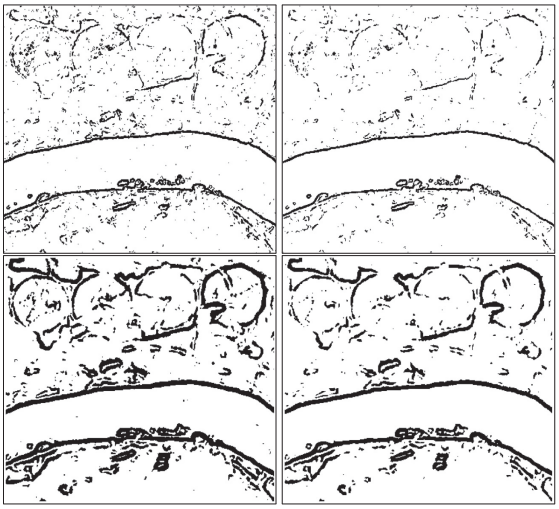


Fig. 11. DAB Bebedouro binary images, with Enh. Lee filter: the first and second rows respectively depict results using the original Gravitational and Gravitational modified with Fu's neighbourhood, the first and second columns respectively depict results obtained with binarization thresholds .1 and .2.

polarizations. The best results for the gravitational method, both with and without modification with Fu's neighbourhood, were obtained with thresholds around the same interval that produced the best results using the mosaics.

7. Conclusions

Contrary to what happens with optical imagery, few algorithms are specifically dedicated to PolSAR image edge detection [10]. One interesting means to create edge detection algorithms for SAR images is to modify those created for optical images, in such a way as to reduce the non-Gaussian noise. Here we have investigated the modification of a method issued from Computational Intelligence for optical imagery, the gravitational edge detection method extension proposed in [4] (see also [3]), to Synthetic Aperture Radar imagery. In order to deal with speckle, we modified the gravitational method with a non-standard 9×9 neighbourhood configuration proposed by Fu et al. [10]: considering a 3×3 window centered around a given pixel, the value of any pixel in the window becomes the average value of the region associated to that pixel in the non-standard neighbourhood configuration.

Considering that SAR imagery has different polarizations, and that their joint use may compensate for the presence of speckle, we also proposed a typology of experiments regarding aggregation of these images. In particular, we addressed two procedures: DAB (edge Detection on non-binary images, Aggregation of the resulting non-binary images, Binarization) and ADB (Aggregation of non-binary images, edge Detection on the resulting non-binary image, Binarization).

For means of comparison, we also addressed the use of two other edge detector methods stemming from the realm of optical images: the traditional method proposed by Canny [2] and a recent multi-scale one coming from Computational Intelligence, based on Sobel operators for edge extraction and the concept of Gaussian scale-space [8].

We studied the effect of filtering the images prior to edge detection by two procedures: Enhanced Lee [11] and Torres et al. [12] filters. The methods were applied on twenty samples of a scene, which were simulated using Wishart distributions derived from a fully polarimetric image [13]. Using both visual inspection and the Baddeley Delta metric [14] we verified that the combination with the Lopez-Molina technique with the 9×9 neighbourhood proposed by Fu et al. [10] and preprocessing with the Enhanced Lee filter produced the best results.

This paper is an extended version of [15]; together, these studies represent a first step towards investigating the use of edge detection methods derived from Computational Intelligence techniques for use in SAR images. The main implication of our results is that the joint use filtering and neighbourhood modification on those methods, as well as the use of aggregation of the individual polarization images, are able to deal with speckle, which is crucial when detecting edges in radar imagery.

Future work includes modifying the Lopez-Molina method with other types of neighbourhoods, such as Nagao–Matsuyama [22]; to verify the performance of other T-norms than the product to calculate the gravitational forces; and to perform preprocessing with other filters. We also intend to investigate the use of the proposed procedure with other edge detection methods, such as the one described in [23], involving fuzzy sets.

We would like to better address the issue of aggregation. Here we have dealt exclusively with the aggregation of non-binary images, using the arithmetic means in strategies DAB and ADB. In the future, we intend to explore aggregation of the images considering families of operators in general, such as weighted means, ordered weighted means (OWA), T-norms, and T-conorms [19]. Also, we intend to study other operators than the average to perform aggregation of pixel values in regions of a non-standard neighbourhood.

Moreover, we intend to assess the results using other methods than BDM, such as the one proposed recently by Frery et al. [24]. We would also like to draw comparisons with other edge detection algorithms, such as the one proposed by Fu in 2012 [10].

Last but not least, we intend to verify the use of the approach considering fully polarimetric images (PolSAR), instead of just intensity images. In this case, Torres filter, designed specifically for PolSAR images, can be more adequately employed.

Acknowledgements

The authors are grateful for Wagner Barreto Silva, Leonardo Torres and Corina Freitas for help in the preparation of this manuscript. They are also thankful for the editor and reviewers for comments and suggestions. The Brazilian authors acknowledge support from CNPq (Projeto Universal 487032/2012-8). The Spanish authors have been supported by project TIN2013-40765-P of the Spanish Government.

References

[1] R.C. Gonzalez, R.E. Woods, Digital Image Processing, 3rd ed., Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.
[2] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 8 (6) (1986) 679–698, http://dx.doi.org/10.1109/TPAMI.1986.4767851.
[3] G. Sun, Q. Liu, C. Ji, X. Li, A novel approach for edge detection based on the theory of universal gravity theory of universal gravity, Pattern Recogn. 40 (10) (2007) 2766–2775.
[4] C. Lopez-Molina, H. Bustince, J. Fernandez, P. Couto, B.D. Baets, A gravitational approach to edge detection based on triangular norms, Pattern Recogn. 43 (11) (2010) 3730–3741, http://dx.doi.org/10.1016/j.patcog.2010.05.035.
[5] M. Danková, P. Hodáková, I. Perfilieva, M. Vajgl, Edge detection using F-transform, in: 2011 11th International Conference on Intelligent Systems Design and Applications (ISDA), IEEE, 2011, pp. 672–677.
[6] E. Barrenechea, H. Bustince, B.D. Baets, C. Lopez-Molina, Construction of interval-valued fuzzy relations with application to the generation of fuzzy edge images, IEEE Trans. Fuzzy Syst. 19 (5) (2011) 819–830.
[7] J.-Y. Chang, Y.-H. Chang, Applying weighted mean aggregation to edge detection of images, in: 2013 International Conference on System Science and Engineering (ICSSE), IEEE, 2013, pp. 153–157.
[8] C. Lopez-Molina, B.D. Baets, H. Bustince, J. Sanz, E. Barrenechea, Multiscale edge detection based on Gaussian smoothing and edge tracking, Knowl.-Based Syst. 44 (5) (2013) 010–111.
[9] H. Mott, Remote Sensing with Polarimetric Radar, John Wiley & Sons, 2006.
[10] X. Fu, H. You, K. Fu, A statistical approach to detect edges in SAR images based on square successive difference of averages, IEEE Geosci. Remote Sens. Lett. 9 (6) (2012) 1094–1098, http://dx.doi.org/10.1109/LGRS.2012.2190378.
[11] A. Lopes, R. Touzi, E. Nezry, Adaptive speckle filters and scene heterogeneity, IEEE Trans. Geosci. Remote Sens. 28 (6) (1990) 992–1000.
[12] L. Torres, S.J.S. Sant'Anna, C.C. Freitas, A.C. Frery, Speckle reduction in polarimetric SAR imagery with stochastic distances and nonlocal means, Pattern Recogn. 47 (2014) 141–157, http://dx.doi.org/10.1016/j.patcog.2013.04.001.
[13] W.B. da Silva, C. da Costa Freitas, S.J.S. Sant'Anna, A.C. Frery, Classification of segments in PolSAR imagery by minimum stochastic distances between Wishart distributions, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. 6 (3) (2013) 1263–1273.
[14] A.J. Baddeley, An error metric for binary images, in: W. Förstner, H. Ruwiedel (Eds.), Robust Computer Vision: Quality of Vision Algorithms, Wichmann, Karlsruhe, 1992, pp. 59–78.
[15] G.P. Silva Junior, A. Frery, S. Sandri, Synthetic aperture radar edge detection with canny's procedure and a gravitational approach, in: 11th International FLINS Conference on Decision Making and Soft Computing (FLINS), 2014, pp. 149–154. doi:http://dx.doi.org/10.1142/9789814619998_0027.
[16] J. Richards, Remote Sensing with Imaging Radar, Signals and Communication Technology, Springer, 2009.
[17] P.R. Meneses, T. d. Almeida, Intodução ao Processamento de Imagens de Sensoriamento Remoto, Universidade de Brasília – CNPq, Brasilia, Br, 2012.
[18] J. Lee, E. Pottier, Polarimetric Radar Imaging: From Basics to Applications, Optical Science and Engineering, Taylor & Francis, 2009. <http://books.google.com.br/books?id=1nAvp2HW_gwC>.
[19] D. Dubois, H. Prade, Possibility Theory: An Approach to Computerized Processing of Uncertainty, Plenum Press, New York, USA, 1988.
[20] J.-S. Lee, A simple speckle smoothing algorithm for synthetic aperture radar images, IEEE Trans. Syst. Man Cybernet. 13 (1) (1983) 85–89, http://dx.doi.org/10.1109/TSMC.1983.6313036.
[21] A.C. Frery, A.D.C. Nascimento, R.J. Cintra, Analytic expressions for stochastic distances between relaxed complex Wishart distributions, IEEE Trans. Geosci. Remote Sens. 52 (2) (2014) 1213–1226, http://dx.doi.org/10.1109/TGRS.2013.2248737.
[22] M. Nagao, T. Matsuyama, Edge preserving smoothing, Comput. Graph. Image Process. 9 (4) (1979) 394–407.
[23] C. Lopez-Molina, B.D. Baets, H. Bustince, Generating fuzzy edge images from gradient magnitudes, Comput. Vis. Image Understand. 115 (11) (2011) 1571–1580, http://dx.doi.org/10.1016/j.cviu.2011.07.003.
[24] M.E. Buemi, A.C. Frery, H.S. Ramos, Speckle reduction with adaptive stack filters, Pattern Recogn. Lett. 36 (2014) 281–287, http://dx.doi.org/10.1016/j.patrec.2013.06.005.

Bibliography

[1] J. J. Gibson, *The ecological approach to visual perception: classic edition*. Psychology Press, 2014 (cit. on pp. 1, 4).

[2] A. S. Reber, "Implicit learning and tacit knowledge", *Journal of experimental psychology: General*, vol. 118, no. 3, p. 219, 1989 (cit. on p. 1).

[3] D. Gentner and J. Medina, "Similarity and the development of rules", *Cognition*, vol. 65, no. 2-3, pp. 263–297, 1998 (cit. on p. 1).

[4] A. Bandura and R. H. Walters, *Social learning theory*. Prentice-hall Englewood Cliffs, NJ, 1977, vol. 1 (cit. on p. 1).

[5] R. N. Shepard and J. Metzler, "Mental rotation of three-dimensional objects", *Science*, vol. 171, no. 3972, pp. 701–703, 1971 (cit. on p. 2).

[6] S. G. Vandenberg and A. R. Kuse, "Mental rotations, a group test of three-dimensional spatial visualization", *Perceptual and motor skills*, vol. 47, no. 2, pp. 599–604, 1978 (cit. on p. 2).

[7] T. Poggio, J. Mutch, and L. Isik, "Computational role of eccentricity dependent cortical magnification", *arXiv preprint arXiv:1406.1770*, 2014 (cit. on p. 2).

[8] F. X. Chen, G. Roig, L. Isik, X. Boix, and T. Poggio, "Eccentricity dependent deep neural networks: Modeling invariance in human vision", in *2017 AAAI Spring Symposium Series*, 2017 (cit. on p. 2).

[9] Y. Han, G. Roig, G. Geiger, and T. Poggio, "Is the human visual system invariant to translation and scale?", in *2017 AAAI Spring Symposium Series*, 2017 (cit. on p. 2).

[10] A. Nguyen, J. Yosinski, and J. Clune, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images", in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 427–436 (cit. on p. 2).

[11] L. A. Zadeh, "Fuzzy logic, neural networks, and soft computing", *Communications of the ACM*, vol. 37, no. 3, pp. 77–85, 1994 (cit. on p. 3).

[12]M. Wertheimer, “Untersuchungen zur lehre von der gestalt”, *Psychological Research*, vol. 1, no. 1, pp. 47–58, 1922 (cit. on p. 3).

[13]G Kanizsa, “Grammatica del vedere/la grammaire du voir”, *Il Mulino, Bologna/Éditions Diderot, arts et sciences*, vol. 1997, 1980 (cit. on pp. 3, 4).

[14]D. Marr, “Early processing of visual information”, *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, vol. 275, no. 942, pp. 483–519, 1976 (cit. on p. 4).

[15]F. Attneave, “Dimensions of similarity”, *The American journal of psychology*, vol. 63, no. 4, pp. 516–556, 1950 (cit. on pp. 6, 7).

[16]R. N. Shepard, “The analysis of proximities: Multidimensional scaling with an unknown distance function. i.”, *Psychometrika*, vol. 27, no. 2, pp. 125–140, 1962 (cit. on pp. 6, 7, 13).

[17]A. J. Baddeley, “An error metric for binary images”, *Robust computer vision*, vol. 5978, 1992 (cit. on p. 7).

[18]S. Santini and R. Jain, “Similarity measures”, *IEEE Transactions on pattern analysis and machine Intelligence*, vol. 21, no. 9, pp. 871–883, 1999 (cit. on pp. 8, 13).

[19]A. Tversky and I. Gati, “Similarity, separability, and the triangle inequality.”, *Psychological review*, vol. 89, no. 2, p. 123, 1982 (cit. on p. 8).

[20]A. Tversky, “Features of similarity.”, *Psychological review*, vol. 84, no. 4, p. 327, 1977 (cit. on pp. 8, 10).

[21]C. L. Krumhansl, “Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density.”, 1978 (cit. on p. 8).

[22]D. L. Davies and D. W. Bouldin, “A cluster separation measure”, *IEEE transactions on pattern analysis and machine intelligence*, no. 2, pp. 224–227, 1979 (cit. on p. 9).

[23]P. Resnik, “Semantic similarity in a taxonomy: An information-based measure and its application to problems of ambiguity in natural language”, *Journal of artificial intelligence research*, vol. 11, pp. 95–130, 1999 (cit. on p. 9).

[24]D. Comaniciu, V. Ramesh, and P. Meer, “Kernel-based object tracking”, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 564–575, 2003 (cit. on p. 9).

[25]C. E. Shannon, “A mathematical theory of communication”, *Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948 (cit. on p. 9).

[26]A. De Luca and S. Termini, “A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory”, *Information and control*, vol. 20, no. 4, pp. 301–312, 1972 (cit. on p. 9).

[27]J. Martín and G. Mayor, "Dispersion measures and multidistances on $\mathbb{R}^k$", in *International Conference on Soft Methods in Probability and Statistics*, Springer, 2016, pp. 347–354 (cit. on p. 9).

[28]——, "Multi-argument distances", *Fuzzy Sets and Systems*, vol. 167, no. 1, pp. 92–100, 2011 (cit. on p. 9).

[29]V. Barnett, "The ordering of multivariate data", *Journal of the Royal Statistical Society: Series A (General)*, vol. 139, no. 3, pp. 318–344, 1976 (cit. on p. 9).

[30]D. Eckert and C. Klamler, "Distance-based aggregation theory", in *Consensual processes*, Springer, 2011, pp. 3–22 (cit. on p. 10).

[31]R. N. Shepard, "Representation of structure in similarity data: Problems and prospects", *Psychometrika*, vol. 39, no. 4, pp. 373–421, 1974 (cit. on p. 10).

[32]A. M. Leslie, "Pretense and representation: The origins of "theory of mind".", *Psychological review*, vol. 94, no. 4, p. 412, 1987 (cit. on p. 10).

[33]M. Iacoboni, R. P. Woods, M. Brass, H. Bekkering, J. C. Mazziotta, and G. Rizzolatti, "Cortical mechanisms of human imitation", *science*, vol. 286, no. 5449, pp. 2526–2528, 1999 (cit. on p. 10).

[34]P. N. Johnson-Laird, R. M. Byrne, and W. Schaeken, "Propositional reasoning by model.", *Psychological review*, vol. 99, no. 3, p. 418, 1992 (cit. on p. 10).

[35]P. N. Johnson-Laird, "Mental models, deductive reasoning, and the brain", *The cognitive neurosciences*, vol. 65, pp. 999–1008, 1995 (cit. on p. 10).

[36]M.-M. Mesulam, "From sensation to cognition.", *Brain: a journal of neurology*, vol. 121, no. 6, pp. 1013–1052, 1998 (cit. on p. 10).

[37]J. Łukasiewicz, *Selected works, l. borkowski*, 1970 (cit. on p. 11).

[38]S. Blamey, "Partial logic", in *Handbook of Philosophical Logic*, D. M. Gabbay and F. Guenthner, Eds. Dordrecht: Springer Netherlands, 2002, pp. 261–353 (cit. on p. 11).

[39]L. A. Zadeh, "Fuzzy sets", *Information and control*, vol. 8, no. 3, pp. 338–353, 1965 (cit. on p. 12).

[40]L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoningi", *Information sciences*, vol. 8, no. 3, pp. 199–249, 1975 (cit. on p. 12).

[41]K. T. Atanassov, "Intuitionistic fuzzy sets", in *Intuitionistic fuzzy sets*, Springer, 1999, pp. 1–137 (cit. on p. 12).

[42]L. A. Zadeh, "Fuzzy logic = computing with words", *IEEE Transactions on Fuzzy Systems*, vol. 4, no. 2, pp. 103–111, May 1996 (cit. on p. 13).

[43]E. Beckenbach and R. Bellman, *An Introduction to Inequalities*, ser. An Introduction to Inequalities. Random House, New York, 1961 (cit. on p. 13).

[44]J. Kacprzyk, “Fuzzy-set-theoretic approach to the optimal assignment of work places”, *IFAC Proceedings Volumes*, vol. 9, no. 3, pp. 123 –132, 1976, IFAC Symposium on Large Scale Systems Theory and Applications, Milano, Italy, 16-20 June (cit. on p. 13).

[45]A Kaufmann and M. Gupta, “Introduction to fuzzy arithmetic. 1985”, *Reinhold, New York*, 1985 (cit. on p. 13).

[46]R. Zwick, E. Carlstein, and D. V. Budescu, “Measures of similarity among fuzzy concepts: A comparative analysis”, *International journal of approximate reasoning*, vol. 1, no. 2, pp. 221–242, 1987 (cit. on p. 13).

[47]C. P. Pappis and N. I. Karacapilidis, “A comparative assessment of measures of similarity of fuzzy values”, *Fuzzy sets and systems*, vol. 56, no. 2, pp. 171–174, 1993 (cit. on p. 13).

[48]H. Bustince, E. Barrenechea, and M. Pagola, “Restricted equivalence functions”, *Fuzzy Sets and Systems*, vol. 157, no. 17, pp. 2333–2346, 2006 (cit. on pp. 13, 16, 22, 37).

[49]A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: A review”, *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999 (cit. on p. 13).

[50]M. R. Boutell, J. Luo, X. Shen, and C. M. Brown, “Learning multi-label scene classification”, *Pattern recognition*, vol. 37, no. 9, pp. 1757–1771, 2004 (cit. on p. 13).

[51]D. Milner and M. Goodale, *The visual brain in action*. Oxford University Press, 2006 (cit. on p. 13).

[52]R. T. Collins, A. J. Lipton, H. Fujiyoshi, and T. Kanade, “Algorithms for cooperative multisensor surveillance”, *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456–1477, 2001 (cit. on p. 14).

[53]M. Greiffenhagen, D. Comaniciu, H. Niemann, and V. Ramesh, “Design, analysis, and engineering of video monitoring systems: An approach and a case study”, *Proceedings of the IEEE*, vol. 89, no. 10, pp. 1498–1517, 2001 (cit. on p. 14).

[54]T. McInerney and D. Terzopoulos, “Deformable models in medical image analysis: A survey”, *Medical image analysis*, vol. 1, no. 2, pp. 91–108, 1996 (cit. on p. 14).

[55]R. J. Gillies, P. E. Kinahan, and H. Hricak, “Radiomics: Images are more than pictures, they are data”, *Radiology*, vol. 278, no. 2, pp. 563–577, 2015 (cit. on p. 14).

[56]T Rumpf, A.-K. Mahlein, U Steiner, E.-C. Oerke, H.-W. Dehne, and L Plümer, “Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance”, *Computers and electronics in Agriculture*, vol. 74, no. 1, pp. 91–99, 2010 (cit. on p. 14).

[57]I. B. Strachan, E. Pattey, and J. B. Boisvert, “Impact of nitrogen and environmental conditions on corn as detected by hyperspectral reflectance”, *Remote sensing of Environment*, vol. 80, no. 2, pp. 213–224, 2002 (cit. on p. 14).

[58]F. Turroni, D. Maltoni, R. Cappelli, and D. Maio, “Improving fingerprint orientation extraction”, *IEEE Trans. on Information Forensics and Security*, vol. 6, no. 3, pp. 1002–1013, 2011 (cit. on p. 14).

[59]J. Canny, “A computational approach to edge detection”, in *Readings in computer vision*, Elsevier, 1987, pp. 184–203 (cit. on p. 14).

[60]D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis”, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 5, pp. 603–619, 2002 (cit. on p. 14).

[61]D. G. Lowe, “Distinctive image features from scale-invariant keypoints”, *International journal of computer vision*, vol. 60, no. 2, pp. 91–110, 2004 (cit. on p. 14).

[62]S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories”, in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’06)*, IEEE, vol. 2, 2006, pp. 2169–2178 (cit. on p. 14).

[63]R. M. Haralick, K. Shanmugam, *et al.*, “Textural features for image classification”, *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610–621, 1973 (cit. on p. 14).

[64]P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation”, *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 5, pp. 898–916, 2011 (cit. on pp. 14, 46).

[65]R. Wilson and G. H. Granlund, “The uncertainty principle in image processing”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 6, pp. 758–767, 1984 (cit. on p. 15).

[66]J. M. Keller and C. L. Carpenter, “Image segmentation in the presence of uncertainty”, *International Journal of Intelligent Systems*, vol. 5, no. 2, pp. 193–208, 1990 (cit. on p. 15).

[67]E. E. Kerre and M. Nachtegael, *Fuzzy techniques in image processing*, ser. Studies in Fuzziness and Soft Computing. Physica-Verlag, 2000, vol. 52 (cit. on p. 15).

[68]D. Van De Ville, M. Nachtegael, D. Van der Weken, E. E. Kerre, W. Philips, and I. Lemahieu, “Noise reduction by fuzzy image filtering”, *IEEE transactions on fuzzy systems*, vol. 11, no. 4, pp. 429–436, 2003 (cit. on p. 15).

[69]N. R. Pal and S. K. Pal, “A review on image segmentation techniques”, *Pattern recognition*, vol. 26, no. 9, pp. 1277–1294, 1993 (cit. on p. 15).

[70]B. Moghaddam, T. Jebara, and A. Pentland, “Bayesian face recognition”, *Pattern Recognition*, vol. 33, no. 11, pp. 1771–1782, 2000 (cit. on p. 15).

[71]S. Lee and M. M. Crawford, “Unsupervised multistage image classification using hierarchical clustering with a bayesian similarity measure”, *IEEE Transactions on Image Processing*, vol. 14, no. 3, pp. 312–320, 2005 (cit. on p. 15).

[72]F. J. Estrada and A. D. Jepson, “Benchmarking image segmentation algorithms”, *International Journal of Computer Vision*, vol. 85, no. 2, pp. 167–181, 2009 (cit. on pp. 15, 46).

[73]S. Krinidis and V. Chatzis, “A robust fuzzy local information c-means clustering algorithm”, *IEEE Transactions on Image Processing*, vol. 19, no. 5, pp. 1328–1337, 2010 (cit. on p. 15).

[74]T. Uricchio, L. Ballan, L. Seidenari, and A. Del Bimbo, “Automatic image annotation via label transfer in the semantic space”, *Pattern Recognition*, vol. 71, pp. 144–157, 2017 (cit. on p. 15).

[75]L. Xuecheng, “Entropy, distance measure and similarity measure of fuzzy sets and their relations”, *Fuzzy Sets and Systems*, vol. 52, no. 3, pp. 305–318, 1992 (cit. on pp. 16, 23).

[76]G. Beliakov, A. Pradera, and T. Calvo, *Aggregation Functions: A Guide for Practitioners*, ser. Studies in Fuzziness and Soft Computing. Springer, 2007, vol. 221 (cit. on pp. 16, 23).

[77]E. Batschelet, *Circular statistics in biology*. Academic Press, 1981 (cit. on p. 16).

[78]J. C. Davis and R. J. Sampson, *Statistics and data analysis in geology*. Wiley New York, 2002, vol. 3 (cit. on p. 16).

[79]N. I. Fisher, *Statistical Analysis of Circular Data*, C. U. Press, Ed. Cambridge University Press, 1993 (cit. on p. 16).

[80]K. V. Mardia and P. E. Jupp, *Directional statistics*, J. Wiley and sons, Eds. Wiley, 2000 (cit. on p. 16).

[81]D. R. Martin, “An empirical approach to grouping and segmentation”, PhD thesis, EECS Department, University of California, Berkeley, Aug. 2003 (cit. on p. 17).

[82]J. Bezdek, R. Chandrasekhar, and Y. Attikouzel, “A geometric approach to edge detection”, *IEEE Transactions on Fuzzy Systems*, vol. 6, no. 1, pp. 52–75, 1998 (cit. on pp. 17, 41).

[83]C. Marco-Detchart, J. Cerron, L. De Miguel, C. Lopez-Molina, H. Bustince, and M. Galar, “A framework for radial data comparison and its application to fingerprint analysis”, *Applied Soft Computing Journal*, vol. 46, pp. 246–259, 2016 (cit. on pp. 19, 27, 55).

[84]Z. Takáč, H. Bustince, J. M. Pintor, C. Marco-Detchart, and I. Couso, “Width-based interval-valued distances and fuzzy entropies”, *IEEE Access*, vol. 7, pp. 14 044–14 057, 2019 (cit. on pp. 19, 33, 70).

[85]H Bustince, C Marco-Detchart, J Fernandez, C Wagner, J. Garibaldi, and Z Takáč, “Similarity between interval-valued fuzzy sets taking into account the width of the intervals and admissible orders”, *Fuzzy Sets and Systems*, 2019 (cit. on pp. 19, 40, 85).

[86]C. Marco-Detchart, H. Bustince, J. Fernandez, R. Mesiar, J. Lafuente, E. Barrenechea, and J. M. Pintor, "Ordered directional monotonicity in the construction of edge detectors", *Fuzzy Sets and Systems*, (cit. on pp. 19, 43, 111).

[87]C Lopez-Molina, C. Marco-Detchart, B De Baets, Member, H Bustince, and S. Member, "A survey on matching algorithms for boundary image comparison and evaluation", *IEEE Transactions on Image Processing*, (cit. on pp. 19, 47, 139).

[88]G. P. Silva, A. C. Frery, S. Sandri, H. Bustince, E. Barrenechea, and C. Marco-Detchart, "Knowledge-Based Systems Optical images-based edge detection in Synthetic Aperture Radar images", *Knowledge-Based Systems*, vol. 87, pp. 38–46, 2015 (cit. on pp. 19, 50, 154).

[89]F. Zhang and E. R. Hancock, "New Riemannian techniques for directional and tensorial image data", *Pattern Recognition*, vol. 43, no. 4, pp. 1590–1606, 2010 (cit. on p. 21).

[90]L. Hong, Y. Wan, and A. Jain, "Fingerprint image enhancement: Algorithm and performance evaluation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 8, pp. 777–789, 1998 (cit. on p. 25).

[91]D. Peralta, M. Galar, I. Triguero, O. Miguel-Hurtado, J. M. Benitez, and F. Herrera, "Minutiae filtering to improve both efficacy and efficiency of fingerprint matching algorithms", *Engineering Applications of Artificial Intelligence*, vol. 32, pp. 37–53, 2014 (cit. on pp. 25, 26).

[92]X. Jiang and W.-Y. Yau, "Fingerprint minutiae matching based on the local and global structures", in *Proc. of the International Conf. on Pattern Recognition*, vol. 2, 2000, pp. 1038–1041 (cit. on p. 25).

[93]A. M. Bazen and S. H. Gerez, "Systematic methods for the computation of the directional fields and singular points of fingerprints", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 905–919, 2002 (cit. on p. 26).

[94]M. Kass and A. Witkin, "Analyzing oriented patterns", *Computer Vision Graphics and Image Processing*, vol. 37, no. 3, pp. 362–385, 1987 (cit. on p. 26).

[95]K.-S. Fu and J. Mui, "A survey on image segmentation", *Pattern recognition*, vol. 13, no. 1, pp. 3–16, 1981 (cit. on p. 32).

[96]M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation", *Journal of Electronic imaging*, vol. 13, no. 1, pp. 146–166, 2004 (cit. on p. 32).

[97]L.-K. Huang and M.-J. J. Wang, "Image thresholding by minimizing the measures of fuzziness", *Pattern Recognition*, vol. 28, no. 1, pp. 41 –51, 1995 (cit. on p. 32).

[98]R Sambuc, *Function Φ-flous, application a laide au diagnostic en pathologie thyroidienne [ph. d. thesis]*, 1975 (cit. on p. 33).

[99]E. Szmidt and J. Kacprzyk, "Entropy for intuitionistic fuzzy sets", *Fuzzy sets and systems*, vol. 118, no. 3, pp. 467–477, 2001 (cit. on p. 33).

[100]I. K. Vlachos and G. D. Sergiadis, “Intuitionistic fuzzy information–applications to pattern recognition”, *Pattern Recognition Letters*, vol. 28, no. 2, pp. 197–206, 2007 (cit. on p. 33).

[101]N. Otsu, “A threshold selection method from gray-level histograms”, *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979 (cit. on p. 33).

[102]H. Bustince, E. Barrenechea, and M. Pagola, “Image thresholding using restricted equivalence functions and maximizing the measures of similarity”, *Fuzzy Sets and Systems*, vol. 158, no. 5, pp. 496–516, 2007 (cit. on p. 33).

[103]H. R. Tizhoosh, “Image thresholding using type ii fuzzy sets”, *Pattern recognition*, vol. 38, no. 12, pp. 2363–2372, 2005 (cit. on p. 33).

[104]E. Deza and M.-M. Deza, “Chapter 21 - Image and Audio Distances”, in *Dict. Distances*, Elsevier, 2006, pp. 262–278 (cit. on p. 39).

[105]E. Trucco, V. Roberto, S Tinonin, and M Corbatto, “Ssd disparity estimation for dynamic stereo.”, in *BMVC*, 1996, pp. 1–10 (cit. on p. 39).

[106]R. Zabih and J. Woodfill, “Non-parametric local transforms for computing visual correspondence”, in *European conference on computer vision*, Springer, 1994, pp. 151–158 (cit. on p. 39).

[107]O. Faugeras, B. Hotz, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, *et al.*, “Real time correlation-based stereo: Algorithm, implementations and applications”, Inria, Tech. Rep., 1993 (cit. on p. 39).

[108]C. Sun, “Fast stereo matching using rectangular subregioning and 3d maximum-surface techniques”, *International Journal of Computer Vision*, vol. 47, no. 1-3, pp. 99–117, 2002 (cit. on p. 39).

[109]M. Galar, J. Fernandez, G. Beliakov, and H. Bustince, “Interval-valued fuzzy sets applied to stereo matching of color images.”, *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 1949–61, 2011 (cit. on p. 39).

[110]J. F. Canny, “A Computational Approach to Edge Detection”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, 1986 (cit. on pp. 41, 42).

[111]R. Medina-Carnicer, R. Muñoz-Salinas, E. Yeguas-Bolivar, and L. Diaz-Mas, “A novel method to look for the hysteresis thresholds for the Canny edge detector”, *Pattern Recognit.*, vol. 44, no. 6, pp. 1201–1211, 2011 (cit. on p. 41).

[112]C. Lopez-Molina, H. Bustince, J. Fernandez, P. Couto, and B. De Baets, “A gravitational approach to edge detection based on triangular norms”, *Pattern Recognit.*, vol. 43, no. 11, pp. 3730–3741, 2010 (cit. on pp. 42, 49).

[113]C. Lopez-Molina, B. De Baets, and H. Bustince, “A framework for edge detection based on relief functions”, *Inf. Sci. (Ny).*, vol. 278, pp. 127–140, 2014 (cit. on p. 42).

[114]M. Gonzalez-Hidalgo, S. Massanet, A. Mir, and D. Ruiz-Aguilera, “On the choice of the pair conjunction-implication into the fuzzy morphological edge detector”, *IEEE Trans. Fuzzy Syst.*, vol. 23, no. 4, pp. 872–884, 2015 (cit. on p. 42).

[115]P. Dollár and C. L. Zitnick, “Fast Edge Detection Using Structured Forests”, *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 8, pp. 1558–1570, 2014 (cit. on p. 42).

[116]A. V. Goldberg and R. Kennedy, “An efficient cost scaling algorithm for the assignment problem”, *Mathematical Programming*, vol. 71, pp. 153–177, 2 1995 (cit. on p. 46).

[117]X. Fu, H. You, and K. Fu, “A statistical approach to detect edges in sar images based on square successive difference of averages”, *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 6, pp. 1094–1098, 2012 (cit. on p. 49).

[118]A. Lopes, R. Touzi, and E Nezry, “Adaptive speckle filters and scene heterogeneity”, *IEEE transactions on Geoscience and Remote Sensing*, vol. 28, no. 6, pp. 992–1000, 1990 (cit. on p. 49).

[119]L. Torres, S. J. Sant’Anna, C. da Costa Freitas, and A. C. Frery, “Speckle reduction in polarimetric sar imagery with stochastic distances and nonlocal means”, *Pattern Recognition*, vol. 47, no. 1, pp. 141–157, 2014 (cit. on p. 49).

List of Figures

1.1 Figures from the Shepard and Metzler's experiment on mental rotation. The experiment shows two equal figures rotated and mirrored (A and B) and a third one which is impossible to obtain by either rotation or mirroring (C). 2
1.2 Illustration showing some laws from the Gestalt Theory [13]. 4
1.3 General scheme of the thesis, representing the three main pillars analysed, Comparison, Logic and Image, which serves to build the idea of Perception. 5
1.4 Illustration used by Tversky and Gati [19] to show that in terms of human perception the triangle inequality does not always hold. **A** and **B** are totally different, but the transition from **A** to **B** and from **B** to **C**, is perceived as shorter. 8
1.5 Fuzzy sets representing the transitions between human concepts of temperature sensations. 12

2.1 Schematic representation of the proposed framework for singular point detection using orientation templates and radial similarity measures, namely Template-based singular point detection (TSPD). 26
2.2 Comparison of the different approaches for thresholding an image compared with the groundtruth provided by the dataset. 32
2.3 Left and right cones image and groundtruth proposed in the dataset Middlebury along with the disparity map obtained. 38
2.4 Comparison of the raw disparity maps obtained with IV-REF and after applying refinement and outlier detection techniques 38