Hyperspectral imaging using notions from Type-2 Fuzzy Sets

A. Lopez-Maestresalas · **L. De Miguel** ·
C. Lopez-Molina · **S. Arazuri** · **H. Bustince** ·
C. Jaren

Abstract Fuzzy set theory has developed a prolific armamentarium of mathematical tools for each of the topics that has fallen within its scope. One of such topics is data comparison, for which a range of operators has been presented in the past. These operators can be used within the fuzzy set theory, but can also be ported to other scenarios in which data is provided in various representations. In this work, we elaborate on notions for Type-2 Fuzzy Sets, specifically for the comparison of type-2 fuzzy membership degrees, to create function comparison operators. We further apply these operators to hyperspectral imaging, in which pixelwise data is provided as functions over a certain energy spectra. The performance of the functional comparison operators is put to the test in the context of in-lab hyperspectral image segmentation.

Keywords Hyperspectral Imaging · Function-Valued Arithmetics · Theory of Comparison · Type-2 Fuzzy Set

1 Introduction

Ever since fuzzy sets appeared, the community around fuzzy set theory has developed a vast literature in the topic. A large portion of the research effort paid in the field has been devoted to the generation and study of tools for dealing with logical information, provided the relationship between fuzzy sets and fuzzy logic. However, the richness of this study is such that many tools can be (and have been) ported to

A. Lopez-Maestresalas, S. Arazuri, C. Jaren
Dpto. de Proyectos e Ingenieria Rural. Univ. Publica de Navarra, Pamplona, Navarra
Tel.: +34-948-169080
E-mail: ainara.lopez@unavarra.es

L. De Miguel, C. Lopez-Molina, H. Bustince
Dpto. de Automatica y Computacion. Univ. Publica de Navarra, Pamplona, Navarra
Tel.: +34-948-169839
E-mail: laura.demiguel@unavarra.es

other fields of research, both theoretically and practically. Hence, it is not uncommon to see applications of well-known fuzzy set theory operators in contexts in which information lacks a logical interpretability.

Uses of fuzzy set theory operators can be found in a wide variety of fields. In general, most of the research fields in the context of computer sciences or applied mathematics have been somehow tackled from a fuzzy perspective. A very relevant example in which the impact of fuzzy set theory has been prolific is image processing. Image processing seems to be an appropriate application to put fuzzy logic to the test, as long as it involves two key factors. Firstly, image processing tasks are normally defined in human terms, what makes their completion and evaluation be (often) based on human, ambiguous interpretation. Secondly, image data is imprecise in very different manners, from the very capture of the information (sensor noise) to the scene-bounded factors (perspective, illumination, shading). As a consequence, a list of authors applied notions of fuzzy set theory in image processing, with or without a logical or set-based interpretation of the data. Some works elaborate on an evident, clear logical representation of concepts, e.g., image processing techniques based on fuzzy rule-based systems [40,37]. In this same direction, some authors developed methodological tools based on fuzzy notions, such as the fuzzy schema for color granulation [14] or texture labelling [13] by Chamorro-Martinez et al., or the approximate spatial relations by Vanegas [71]. Despite these works properly make a label- or granule-based interpretation of imaging data, literature is also vast in applications of fuzzy set theory operators with no clear presence of fuzzy sets or other fuzzy concepts. For example, aggregation operators have been used to develop specific low level image processing operators. In some of these cases [15,54,47] the fuzzy interpretation is missing, but authors capitalize on the properties of aggregation operators, as well as on the solidity of their theoretical basis. In fact, connections have been found between basic operators in image processing and those in fuzzy set theory, e.g., the general framework for image filtering by Wilkin and Beliakov [76]. Similar situations can be found in other subfields of image processing, e.g., in fuzzy mathematical morphology [25,7] and its relationship with grayscale morphology [29].

The transference of information between fuzzy set theory and image processing has been boosted, in our opinion, by different facts. Very importantly, by the strict nature of the work in the fuzzy set theory, which has led to very solid mathematical grounds and principles. Also, by the natural focus of fuzzy set theory on handling non-standard data. Fuzzy information has been represented in many different shapes, from scalar values (in fuzzy sets) to subsets of $[0,1]$ (in set-valued fuzzy sets) or even functions (in Type-2 Fuzzy Sets). Hence, operators in fuzzy set theory shall be prepared to deal with such variety of shapes, whether in terms of logic (as meet and join) or arithmetics (as comparison and fusion). Fuzzy set theory has produced a list of operators prepared for non-standard data, and hence provides solutions to handle mathematical objects for which other fields are not yet prepared.

This work relates to one of those fields in which non-standard data appears, as it is *hyperspectral imaging*. In hyperspectral imaging, each pixel in the image grid is associated to a functional representation of the reflected light (energy) in a certain spectral range. This functional representation is materialized as a vector in which each position holds the reflectance at equally-distant wavelengths of such range. In
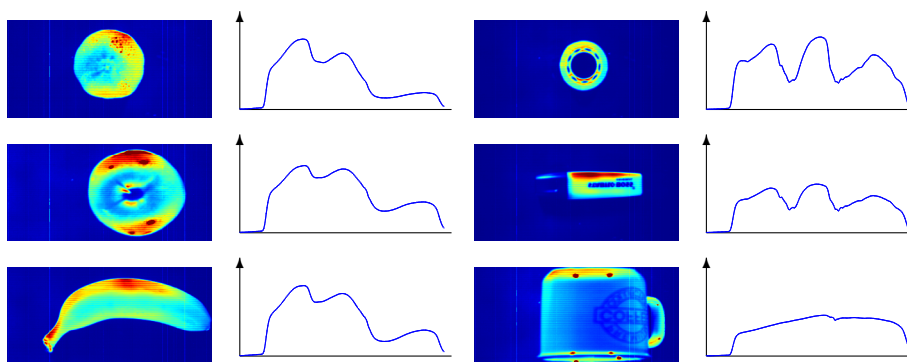
Fig. 1 Visualization of the mean spectra of different biological and artificial materials. The mean spectra is computed from the region actually occupied by the object. This region is determined using the segmentation technique in [44].

order to produce reliable image processing operators in this field one needs to define basic operations on pixelwise information, including comparison, difference, fusion, etc. As long as the information at each pixel is represented as a vector, such basic operations could be handled with standard vectorial arithmetics. However, the semantics and conditions (e.g., strong spatial correlation) of spectra make it possible to develop evolved tools which better exploit the nature of the data. These problems are, in fact, similar to those encountered in the comparison of some Type-2 Fuzzy Sets, in which the uncertainty represented by a function comes from possibility distributions. This means that the values at contiguous positions of the function are strongly correlated. In Fig. 1 we display some examples of hyperspectral profiles for in-lab imagery, featuring several living matter and human-made items. We observe, for example, that there is a strong correlation in consecutive positions at each spectra. Also, that living matter items on the left side present, in general, similar spectral patterns as a result of their similar chemical compositions.

In order to develop image processing frameworks, we first need to study operators able to work at the pixel level. In grayscale morphology, such operators are meet and join, while in image convolution they are the product, sum and difference operators. In this work, aimimg at developing a framework for hyperspectral imaging, we focus on spectral comparison, which can be seen as the problem of functional comparison. In order to do so, we study the literature and tools from type-2 fuzzy logic, which is also based on a functional representation of the information at each element. Then, we develop novel operators which can be used for the comparison of one dimensional functions defined over a partially ordered domain. As special cases, they can be used for both type-2 fuzzy logic and hyperspectral image processing. All of this is built upon the existing cumulative operators, and takes inspiration from [74].

The remainder of this work is organized as follows. First, in Section 2 we review the literature on comparison in the context of fuzzy set theory. Then, in Section 3 we propose a method for multivalued representation of functions, which leads to a comparison operator focused on the peaks and valleys in the compared functions. In Section 4 we analyze in detail the problem of spectral comparison to further put our

comparison operators to the test in Section 5. Finally, a brief discussion is included in Section 6.

2 Comparison in Fuzzy Set Theory

The boost of data-centric culture has affected most disciplines of research. As of now, scientists in different fields attempt to model their contexts and circumstances in terms of data, which needs to be further processed [68]. This has raised new concerns ranging from the ethics and accountability of the process [51] to the modelling and management of complex mathematical objects [1,6]. Data-centric science is supported by two pillars: data gathering and data processing. While the former is context-dependent, the latter allows to import experience from different fields of study.

When it comes to data processing, knowledge is normally built upon simplistic mathematical tools, in the same manner as complex mathematics are built upon basic operations. Apart from object equality, the most basic operation on data is object comparison. Historically, metrics (a.k.a. distance functions) have been a mainstream tool for complex object comparison. This is mostly due to the interesting properties they satisfy, and to the fact that metrics abilitate the use of certain key techniques in, e.g., optimization [4]. Besides, the use of metrics preserves an Euclidean-like interpretation of the data universe, which certainly facilitates the understanding of subsequent computational goals. The study of metrics and metrical spaces has been a key topic of research in mathematics, leading to different sub-concepts and related studies (e.g. multidistances or multi-argument distances [50,57]). Although akin to human interpretation, the properties imposed by metrics shall not always be natural for some types of data [19]. In the past century, different psychologies have also questioned whether humans truly perform object interpretation and comparison based on metrical spaces. Such questioning has been done from a psychological perspective [3], but also involves interdisciplinar approaches [81]. Criticisms can be found to be practical [39], and also theoretical [3]. The inconsistencies between human judgements and triangular inequality have led authors as Tversky [69] or Santini and Jain [64] to disregard the use of metrics. Other authors [69,70] even question the role of the symmetry, mainly because it negates a directional interpretation of human comparison. Apart from purely theoretical analysis, applied research has sometimes found the results by metrics to be meaningless in large-range comparisons [63,58]. Overall, a quorum seems to exist that humans do not always behave according to metrical properties, specially when it comes to triangular inequality, but it is also acknowledged that its properties greatly ease the use of certain data.

Considering the vagueness of the data comparison problem (we are still uncertain of the properties satisfied by general human comparison) and the potential complexity of the mathematical objects involved in it, soft computing appears as a straighforward tool to analyze comparison. In fact, over the years a *vast armamentarium* [83] of techniques for comparison has been presented. Some of these techniques embrace the metrical approach [27,26], while some others developed parallel theories (as the T-indistinguishability by De Baets [20,18]). Still, all of them are born from Zadeh's introduction of similarity as an extension of equivalence [83]. Note that binary equiv-

alence shall not be totally disregarded from the debate in a fuzzy approach to object comparison. In fact, the very idea of granulation and *wide fuzzy logic*, as presented by Zadeh [84], can be understood under the prism of classes of equivalence for input data. Also, there has been relevant research in the generation of graded equivalence relations from both crisp and graded relational data [72].

In a historical account of comparison operators for fuzzy set theory we find that each generalization of fuzzy sets has come coupled to new definitions of comparison operators. For example, scalar membership functions in fuzzy sets can be compared using equivalence functions or restricted equivalence functions [8]. Similar developments were created for comparing interval-valued fuzzy sets [36,61], hesitant fuzzy sets [85] or *L*-fuzzy sets [60]. In this work, we focus on the tools and means for the comparison membership degrees in Type-2 Fuzzy Sets. Type-2 Fuzzy Sets (T2FSs, for short) were introduced in 1975 by Zadeh [82]. Soon after its definition, some theoretical works about the algebra of fuzzy truth values of T2FSs were introduced [55, 56]. However, it is not until the recent literature that T2FSs have grown in importance. This lack of early literature is, at least partially due to the fact that the uncertainty modelled by T2FSs is more intricate and its treatment more complicated than in other generalizations of fuzzy sets. In order to solve this problem, Mendel et al. [42,53] introduced the notion of interval T2FSs. However, the notation used by Mendel et al. is different from the notation by authors as Walker et al. [30,31,73,74] and Hernández et al. [32,33]. The existence of different notations has, in fact, slowed down the development of T2FSs.

In [9] it is shown that the pointwise interpretation of functions, which is the most usual one in literature [21], is not appropriate for introducing operations in T2FSs. The significance of the function, as mathematical object, cannot be observed individually at each point of its domain x; it can only be observed, hence, in its indivisible behaviour over the domain. For that reason, the definition of operations in T2FSs (see, for example [32,38,74]) lies on Zadeh's extension principle [82]. This problem also affects, among others, most of the comparison operators for T2FSs. Most of the operations defined in the literature are only applicable to interval T2FSs [78–80]. McCulloch et al. have recently tried to generalize the concept of similarity measure to deal with any T2FS [52], but their definition heavily depends on the use of (some) z-Slides which are not congruent with the semantics of the present work.

In this work we present a novel strategy for functional comparison. This strategy is not focused on the pointwise analysis of functions, but on the idea of indivisible function, instead. In this sense, our approach is heavily inspired by the ideas of Zadeh [82], as well as by the mathematical developments by Walker, Walker and Harding [31,74].

3 A novel strategy for function comparison

In this section we present our proposal for functional comparison, which will later be applied to hyperspectral imaging. Our idea is based on two principles. Firstly, a function represents the information as a piece and it should have the consideration of an indivisible object. Secondly, most of the information in functions is in their

maxima and minima, either local or global. From these two principles we deduce that a pointwise treatment is not suitable for the purpose of this work. Also, that the measurement of the local and semi-local variations of a function are a key for the interpretation and comparison of functions. With these two ideas we present in this section a novel similarity measure for one-dimensional functions, which will be further applied to hyperspectral image processing.

3.1 Preliminary definitions

In order to better understand our proposal, it is necessary to recall the expression of union and intersection of two membership functions of T2FSs [55,74].

Definition 1 Let $f, g : [0,1] \to [0,1]$ be two functions which represent the membership degrees of two T2FSs. The union and intersection are defined as follows:

$$(f \sqcup g)(x) = \bigvee_{y \vee z = x} f(y) \wedge g(z), \quad \text{for any } x \in [0,1];$$

and

$$(f \sqcap g)(x) = \bigvee_{y \wedge z = x} f(y) \wedge g(z), \quad \text{for any } x \in [0,1].$$

In [74], Walker and Walker proved that the union and intersection can be equivalently formulated in terms of the so-called cumulative functions.

Definition 2 Let $f : [0,1] \to [0,1]$ be a function which represents the membership degree of a T2FS.

– The left-cumulative function f^L of f is given by

$$f^L(x) = \bigvee_{y \leq x} f(y), \quad \text{for any } x \in [0,1]. \tag{1}$$

– The right-cumulative function f^R of f is given by

$$f^R(x) = \bigvee_{y \geq x} f(y), \quad \text{for any } x \in [0,1]. \tag{2}$$

Let $f, g : [0,1] \mapsto [0,1]$. For any $x \in [0,1]$, it holds that

$$\begin{aligned}(f \sqcup g)(x) &= (f(x) \wedge g^L(x)) \vee (f^L(x) \wedge g(x)) = (f(x) \vee g(x)) \wedge (f^L(x) \wedge g^L(x)), \\ (f \sqcap g)(x) &= (f(x) \wedge g^R(x)) \vee (f^R(x) \wedge g(x)) = (f(x) \vee g(x)) \wedge (f^R(x) \wedge g^R(x)).\end{aligned} \tag{3}$$

Cumulative functions are an initial step to treat functions as indivisible objects, and capture most of the information needed to interpret convex functions. However, they do not hold enough information to represent general functions, since they are oblivious of local minima (valleys) and, potentially, of the local maxima (peaks) in the function f. Hence, despite valid to model union and intersection operations in type-2 fuzzy logic, we need to complete the representation of a general function f with additional information.

3.2 A four-fold representation of one-dimensional functions

The aim in this section is to produce a representation of functions which is able to capture, in an evident manner, the presence of maxima and minima in a one-dimensional function. Our proposal intends to build upon cumulative functions, so that functions can be treated globally, and not pointwise. Note that the cumulative functions themselves do not completely fulfill our goals, since left (resp. right) cumulative function is plain when the function is decreasing (resp. increasing). However, we make use of cumulative functions to produce a richer interpretation of the original function.

Definition 3 Let $f : X \mapsto [0,1]$, where X is an ordered, discrete set. We consider a four-fold representation of the function f, namely $\hat{f} = (f_1, f_2, f_3, f_4)$, given by:

- $f_1(x) = f^L(x) = \bigvee_{y \leq x} f(y)$;
- $f_2(x) = f^R(x) = \bigvee_{y \geq x} f(y)$;
- $f_3(x) = f^L(x) - f(x) = f_1(x) - f(x)$;
- $f_4(x) = f^R(x) - f(x) = f_2(x) - f(x)$.

Note that f_1 and f_2 are exactly the cumulative functions as described in Eqs. (1) and (2). The function f_3 is called *subtracted left cumulative function* while the function f_4 is called *subtracted right cumulative function*.
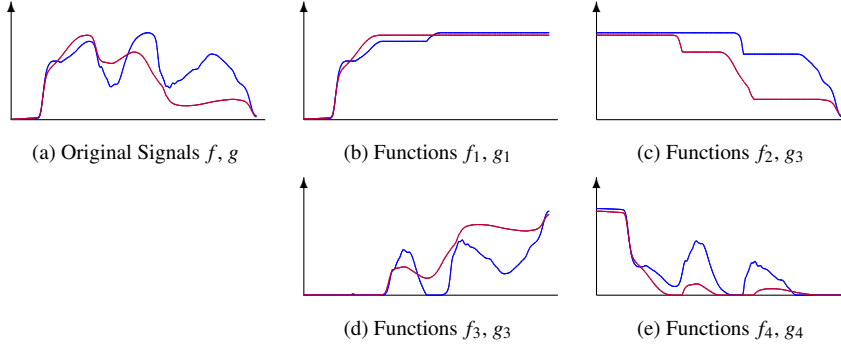
A visual example of the representation proposed in Definition 3 is given in Fig. 2. Note that the functions used in this example are taken from the spectral profiles from Fig. 1. In Fig. 2 we observe how the differences between the two functions f and g are modelled, from different perspectives, in the functions f_i and g_i, for $i \in \{1,2,3,4\}$. Since both functions have a similar increase on the left, f_1 is rather similar to g_1. However, the increase on the right side of the function is quite different, and so the differences between f_2 and g_2 are notorious. Although both functions feature two valleys in its range, their visual interpretation is rather different in terms of depth and sharpness. Accordingly, functions f_i and g_i, for $i \in \{3,4\}$, show differences which allow us to model such dissimilarity.

In the remainder of this section we study some mathematical properties of the functions of the four-fold representation. Some of the properties, specifically those only referring to the left and right cumulative functions f_1 and f_2, can be found in [74].

Let $\underline{X}$ and $\overline{X}$ respectively denote the smallest and greatest elements of the discrete, ordered domain X. We can then consider the domain X as the set containing all the elements in the interval $[\underline{X}, \overline{X}]$, i.e., we can write $X = [\underline{X}, \overline{X}]$

Proposition 1 *Let* $f : [\underline{X}, \overline{X}] \mapsto [0,1]$ *be a function and let* $\hat{f} = (f_1, f_2, f_3, f_4)$ *be its four-fold representation as introduced in Definition 3. The following holds:*

- $0 \leq f_i(x) \leq 1$ *for any* $x \in [\underline{X}, \overline{X}]$ *and for* $i = \{1,2,3,4\}$.
- f_1 *is an increasing function.*
- f_2 *is a decreasing function.*

(a) Original Signals f, g (b) Functions f_1, g_1 (c) Functions f_2, g_3

(d) Functions f_3, g_3 (e) Functions f_4, g_4

Fig. 2 Visual example of the proposed four-fold representation of one dimensional features. We display the original functions f and g, together with the representations included in $\hat{f}$ and $\hat{g}$, as in Definition 3.

- Let $x' = \min(\operatorname{argmax}_{x\in[\underline{X},\overline{X}]} f(x))$. It holds that $f_1(x) = M$ for all $x \geq x'$, where $M = f(x')$. Note that, since the X is discrete, x' always exists.
- Let $x'' = \max(\operatorname{argmax}_{x\in[\underline{X},\overline{X}]} f(x))$. It holds that $f_2(x) = M$ for all $x \leq x''$, where $M = f(x'')$.
- Let $x' \in [\underline{X},\overline{X}]$ satisfy that $f_1(x') = f_2(x') = f(x')$. Then, $x' \in \operatorname{argmax}_{x\in[\underline{X},\overline{X}]} f(x)$.
- Let x^* be a strict local maximum of the function, i.e., let x^* satisfy that $f(x^*) > f(x)$, for all x in a neighbourhood of x^*. It holds that
  - $f_3(x^*) \leq f_3(x)$, for all x in a neighbourhood of x^*;
  - $f_4(x^*) \leq f_4(x)$, for all x in a neighbourhood of x^*.
  Otherwise said, a local maximum generates a local minimum in both f_3 and f_4 functions.
- Let x^* be a strict local minimum of the function, i.e., let x^* satisfy that $f(x^*) < f(x)$, for all x in a neighbourhood of x^*. It holds that
  - $f_3(x^*) \geq f_3(x)$, for all x in a neighbourhood of x^*;
  - $f_4(x^*) \geq f_4(x)$, for all x in a neighbourhood of x^*.
  Otherwise said, a local minimum generates a local maximum in both f_3 and f_4 functions.
- Let $g(x) = f(x) + c$ and let $\hat{g} = (g_1, g_2, g_3, g_4)$ be its four-fold representation. It holds that
  - $g_1(x) = f_1(x) + c$;
  - $g_2(x) = f_2(x) + c$;
  - $g_3(x) = f_3(x)$;
  - $g_4(x) = f_4(x)$.
  Otherwise said, f_1 and f_2 are stable under vertical translations, while f_3 and f_4 are oblivious to them.
- Let $g(x) = c \cdot f(x)$, with $c \in [0,1]$ and let $\hat{g} = (g_1, g_2, g_3, g_4)$ be its four-fold representation. It holds that $g_i(x) = c \cdot f_i(x)$ for any $i \in \{1,\ldots,4\}$ and $x \in X$.
- If f is an increasing function, for any $x \in [\underline{X},\overline{X}]$, it holds that
  - $f_1(x) = f(x)$;
  - $f_2(x) = M$, with $M = f(\overline{X})$;
  - $f_3(x) = 0$;

- $f_4(x) = M - f(x)$.
- If f is a decreasing function, for any $x \in [\underline{X}, \overline{X}]$, it holds that
  - $f_1(x) = M$, with $M = f(\underline{X})$;
  - $f_2(x) = f(x)$;
  - $f_3(x) = M - f(x)$;
  - $f_4(x) = 0$.
- If f is concave, i.e., there exist $x_0 \in [\underline{X}, \overline{X}]$ such that for all $x \in [\underline{X}, x_0]$, $f(x) \leq f(x_0)$ and for all $x \in [x_0, \overline{X}]$, $f(x_0) \geq f(x)$, then the following holds:

  - $f_1(x) = \begin{cases} f(x) & \text{if } x \in [\underline{X}, x_0[\\ f(x_0) & \text{if } x \in [x_0, \overline{X}] \end{cases}$

  - $f_2(x) = \begin{cases} f(x_0) & \text{if } x \in [\underline{X}, x_0[\\ f(x) & \text{if } x \in [x_0, \overline{X}] \end{cases}$

  - $f_3(x) = \begin{cases} 0 & \text{if } x \in [\underline{X}, x_0[\\ f(x_0) - f(x) & \text{if } x \in [x_0, \overline{X}] \end{cases}$

  - $f_4(x) = \begin{cases} f(x_0) - f(x) & \text{if } x \in [\underline{X}, x_0[\\ 0 & \text{if } x \in [x_0, \overline{X}]. \end{cases}$

3.3 A similarity measure for functions

We present two different ways of measuring the similarity focusing on two interpretations of the four-fold representation in Section 3.2. The first one focuses on the comparison of all the available information, while the second one restricts the analysis to the study of the left and right subtracted cumulative functions.

Let H denote the set of all functions from $[\underline{X}, \overline{X}]$ to $[0, 1]$, i.e.,

$$H = \{f \mid f : [\underline{X}, \overline{X}] \to [0, 1]\}.$$

Definition 4 The similarity function $q_1 : H \times H \to [0, \infty)$ is defined as

$$q_1(f, g) = \sum_{x \in [\underline{X}, \overline{X}]} \sum_{i \in \{1, \dots, 4\}} |f_i(x) - g_i(x)|, \quad \text{for any } f, g \in H. \tag{4}$$

This operator takes into account the differences between the four functions of the four-fold representation. The advantage is that the generated function is a metric.

Proposition 2 *The similarity q_1 satisfies that*

(i) $q_1(f, g) = 0$ iff $f(x) = g(x)$, for all $x \in X$;
(ii) $q_1(f, g) = q_1(g, f)$, for any $f, g \in H$;
(iii) $q_1(f, h) \leq q_1(f, g) + q_1(g, h)$, for any $f, g, h \in H$.

Proof (i) The similarity $q_1(f, g) = 0$ if and only if $f_1(x) = g_1(x)$, $f_2(x) = g_2(x)$, $f_3(x) = g_3(x)$ and $f_4(x) = g_4(x)$, for all $x \in [\underline{X}, \overline{X}]$. Since, $f_1(x) = g_1(x)$ means that $f^L(x) = g^L(x)$, for all $x \in [\underline{X}, \overline{X}]$, it holds that $f_3(x) = g_3(x)$ for all $x \in [\underline{X}, \overline{X}]$ if and only if $f^L(x) - f(x) = g^L(x) - g(x)$ for all $x \in [\underline{X}, \overline{X}]$ if and only if $f^L(x) - f(x) = f^L(x) - g(x)$ for all $x \in [\underline{X}, \overline{X}]$ if and only if $f(x) = g(x)$ for all $x \in [\underline{X}, \overline{X}]$.

(*ii*) Direct.

(*iii*) Straight, taking into account the triangle inequality of the absolute value.

Note that $f_3(x) - g_3(x)$ can be rewritten as $f^L(x) - f(x) - (g^L(x) - g(x)) = (f^L(x) - g^L(x)) - (f(x) - g(x))$. Hence, big differences in the cumulative functions (which are counted also in $f_1(x) - g_1(x)$) may make the punctual differences insignificant. Analogously, the same problem arises with the right cumulative and right subtracted cumulative functions. In order to solve this obstacle, we propose another way of measuring the similarity.

Definition 5 The similarity function $q_2 : H \times H \to [0, \infty)$ is defined as

$$q_2(f, g) = \sum_{x \in [\underline{X}, \overline{X}]} \left(|f_3(x) - g_3(x)| + |f_4(x) - g_4(x)| \right). \tag{5}$$

Proposition 3 *The similarity q_2 satisfies that*

(i) $q_2(f, g) = 0$ *if and only if* $f(x) = g(x) + c$ *for all* $x \in X$, *with some fixed* $c \in [-\min_{x \in [\underline{X}, \overline{X}]}(f(x), g(x)), 1 - \max_{x \in [\underline{X}, \overline{X}]}(f(x), g(x))]$;
(ii) $q_2(f, g) = q_2(g, f)$, *for any* $f, g \in H$;
(iii) $q_2(f, h) \leq q_2(f, h) + q_2(h, g)$, *for any* $f, g, h \in H$.

Proof (i) Note that $q_2(f, g) = 0$ if and only if $f_3(x) = g_3(x)$ and $f_4(x) = g_4(x)$ for all $x \in [\underline{X}, \overline{X}]$ if and only if $f^L(x) - f(x) = g^L(x) - g(x)$ and $f^R(x) - f(x) = g^R(x) - g(x)$ for all $x \in [\underline{X}, \overline{X}]$ if and only if $f^L(x) - g^L(x) = f^R(x) - g^R(x) = f(x) - g(x)$ for all $x \in [\underline{X}, \overline{X}]$.
Let $x^* \in \mathrm{argmax}_{x \in [\underline{X}, \overline{X}]} f(x)$. It holds that

$$
\begin{aligned}
f^L(x^*) - g^L(x^*) = f(x^*) - g(x^*) = f^R(x^*) - g^R(x^*) \text{ if and only if} \\
f(x^*) - g^L(x^*) = f(x^*) - g(x^*) = f(x^*) - g^R(x^*) \text{ if and only if} \\
g^L(x^*) = g(x^*) = g^R(x^*).
\end{aligned}
\tag{6}
$$

Due to Proposition 1, this means that $x^* \in \mathrm{argmax}_{x \in [\underline{X}, \overline{X}]} g(x)$. Let $c = f(x^*) - g(x^*)$. For any, $x_1 \in [\underline{X}, \overline{X}]$ such that $x_1 \leq x^*$, it holds that

$$
\begin{aligned}
f^R(x_1) - g^R(x_1) = f(x_1) - g(x_1) \text{ if and only if} \\
f(x^*) - g(x^*) = f(x_1) - g(x_1) = c.
\end{aligned}
\tag{7}
$$

Similarly, for any, $x_2 \in [\underline{X}, \overline{X}]$ such that $x_2 \geq x^*$, it holds that

$$
\begin{aligned}
f^L(x_2) - g^L(x_2) = f(x_2) - g(x_2) \text{ if and only if} \\
f(x^*) - g(x^*) = f(x_2) - g(x_2) = c.
\end{aligned}
\tag{8}
$$

Hence, $f(x) = g(x) + c$ for any $x \in [\underline{X}, \overline{X}]$.

(*ii*) Direct.

(*iii*) Straight, taking into account the triangle inequality of the absolute value.

It is worth mentioning that even q_2 overcomes the limitations of the metric q_1, in which big differences in the cumulative functions may degenerate the metric, the new similarity is not a metric any more. Indeed, it yields 0 for any pair of functions satisfying that one is a vertical translation of the other. Anyway, in situations in which the importance is given to the form and size and not to the height of the functions (as it is the case of the proposed application in hyperspectral images) this need not be a disadvantage, but an advantage instead.

3.4 Practical analysis of the comparison measure

The measures q_1 and q_2 are meant to have a behaviour more logical than that of the point-to-point measurements. In order to illustrate that, Fig. 3 displays the results of the comparison of two functions f and g against distorted versions of itself. In that figure, the leftmost column includes the original signals in the experiment. Then, each of the three rightmost columns displays the dissimilarity of the original version of the signal w.r.t a modified version of itself. Such modifications consist of a lateral displacement of the function d positions to the right (column (b)), a point-wise scaling by a constant factor k (column (c)) or uniform white noise contamination of $\pm n\%$ of its value (column (d))[1]. In all cases, the disimilarity is measured using q_1, q_2 and the absolute difference $s(f,g) = \sum_{x \in [\underline{X}, \overline{X}]} |f(x) - g(x)|$.

In Fig. 3 we can observe that the measures q_1 and q_2, as well as s, quantify linearly the distortions due to pointwise alterations of the signal (columns (c) and (d)). Note-worthingly, q_1 and q_2 have a better behaviour when the modification of the function affects its structure, as in the case of the displacements. In Fig. 3(b) we observe how the dissimilarity measured with q_1 is a quasi-linear w.r.t. the displacement. While q_2 does not manage to keep such linearity, it still performs better than s, which performs poorly in such cases.

4 Hyperspectral imaging

Functional representation of information is not uncommon in science, specially when linear measurement parameters are considered in the data gathering process. Examples of such can be time series, in which scalar data becomes a function of time, or scale-spaces [43], for which dedicated parameters are introduced with the same goal. Hyperspectral image processing is a suitable field of application of the measurements in Section 3 for a list of reasons. Among them, two reasons gain special relevance. First, the fact that hyperspectral imaging involves all the factors that make image processing an evident target for fuzzy set theory, as analyzed in Section 1. Second, the absence of widely-accepted low level operations on top of which more complex image processing methods can be built.

Hyperspectral imaging generally refers to the measurement of light in wavelengths beyond those visible to human beings. Specifically, such measurements are

---

[1] In the case of column (d) the measurements have been repeated, then averaged, 1000 times.

(a) Original Signal   (b) Lateral displacement   (c) Scaling   (d) White noise
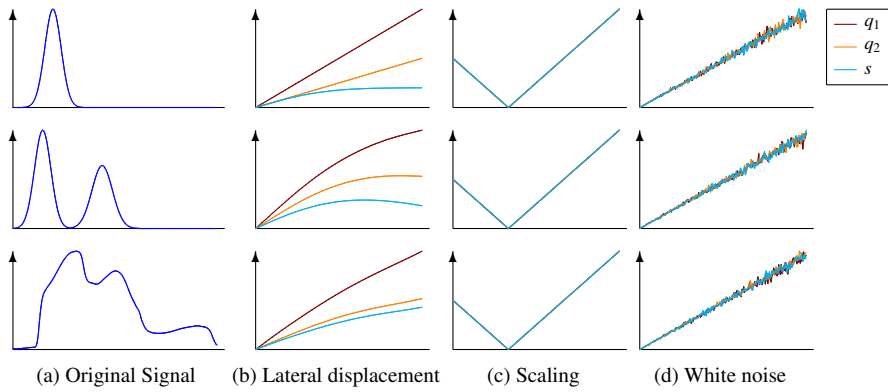
Fig. 3 Analysis of the responses yielded by the measures q_1, q_2 and *absolute difference (AD)* in the comparison of a function with a distorted version of itself. (a) Original functions with 200 sample points, (b) dissimilarity measured w.r.t. lateral displacement of d positions to the right, (c) dissimilarity measured w.r.t. multiplication of the signal by a factor k, (d) average dissimilarity measured w.r.t. 1000 random contaminations of the signal with uniform, white noise in the range $[-n, n]$, n expressing a percentage of the maximum value of the original signal.

acquired at different wavelengths in a certain spectral range, normally a subset between the visible and near infrared regions of the electromagnetic spectrum. Given a spectral range and a vector of equally-distant wavelengths within it $\mathbb{M} = (\mu_1, \ldots, \mu_n)$, a hyperspectral image is a mapping $f : \mathscr{R} \times \mathscr{C} \times \mathbb{M} \mapsto [0, 1]$, so that $f(r, c, i)$ represents the reflectance of energy at position (r, c) for a wave at μ_i nanometers. Hence, it can be figured out as an image in which the information at each row r and column c is a spectrum, *i.e.*, a function.

Recent years have brought a drastic increase in the use of hyperspectral imaging. The reasons partially lie on the decay of the prizes of the machinery, but also on the attractive applications hyperspectral imagery has. Specifically, on the fact that the reflectance at certain wavelengths in the non-visible spectra is related to material properties of the visible objects. Hence, the reflectance at certain wavelengths is dependent upon physical facts such as humidity or presence of certain components. This makes such facts *visible* under hyperspectral imaging systems, even if they do not manifest for the human-visual system. This has found a list of applications in different fields, specially remote sensing, drone-based surveillance and control, and agro-bio product analysis [23].

The benefits and possibilities of hyperspectral imaging are evident, but they come at the cost of greater data complexity [24]. To compare, regular images have 3 or 4 values per pixel, while hyperspectral images contain functions with $256 - 1024$ measured wavelengths, whose interpretation (individually, or globally) is far from evident. When it comes to hyperspectal imagery even visualization is non-evident [35]. The low interpretability of hyperspectral imagery is indeed a significant problem, more so if we consider that most image processing tasks are based on human interpretation of visual artefacts, sometimes even attempting to mimic human cognitive processes. We believe that most of this complexity would unravel by developing low-level operators that enable the implementation of more complex image processing

algorithms. In this context, we focus on spectra comparison, one of the key low-level operations in data (also, image) processing.

The comparison of pixel information is ubiquitous in image processing. It affects most of the processes, from the definition of the basic operators (e.g. differential operators [62]) to the very definition of color spaces (e.g CIELab). When it comes to hyperspectal comparison, the complexity of the data makes it arduous to define reliable, interpretable operators. In our opinion, operators should make full use of the fact that spectra are not vectors, but functions. In this sense, we understand that spectra comparison must be tackled from a deep understanding of the conditions of spectra. Specifically, we focus on agro-bio applications, one the most prolific use of hyperspectral imagery.

5 Experiments on in-lab hyperspectral imagery

5.1 Edge detection on hyperspectral images based on the USAN principle

There exists a certain divergence on the way in which object boundaries are perceived by humans, at least when it comes to the directionality of the process. This divergence roots on the experiments carried out in the context of neurology and psychotropics, and has been matter of debate for a long time. Although the contributions to the debate are multiple, the contradiction of modern theories (e.g. [34, 10]) have been frequently echoed in the image processing literature [49, 48].

A large portion of the literature in image processing considers images to be scalar-valued functions, so that the boundaries manifesting as intensity changes shall be detected throughout the analysis of their derivatives. This has lead to a prolific field of study on optimal filters or kernels for gradient characterization, including the well-known Sobel [67], Prewitt [62] or Canny [11] operators, but also very recent proposals as the anisotropic Gaussian kernels [65, 75]. In case images are vector-valued (e.g. color images), the required analysis is that of the Jacobian of the image [22]. Despite being a mainstream option, very relevant authors have taken a totally different approach. Within their framework for the *raw primal sketch*, Marr and Hildreth considered that boundary manifestations, and hence boundary detection filters, *need not be orientation dependent* [49]. Other authors have also elaborated on similar ideas to produce orientation independent filters to model object boundaries. One of such efforts is the SUSAN edge detector, which relies on the relative heterogeneity of the pixels in a small image region to model the probability of the presence of a boundary [66].

The USAN edge detector is based on the so-called Univalue Segment Assimilating Nucleus (USAN) principle, which refers to the fact that most of the values in an area are normally similar to the pixel in its center (its *nucleus*). Any divergence from such trend is thought to be due to the presence of multiple objects in the region, and hence indicator of boundary of multiple objects. Mathematically, it renders in a weighted summation of the dissimilarities between each pixel and the pixel in its surrounding area.

Let Ω be the positions of an image I, and let $N : \Omega \mapsto \mathscr{P}(\Omega)$ be a mapping such that $N(x,y)$ yields the positions in a certain neighbourhood around (x,y). The edge value at each pixel is computed from the difference of the pixels and its neighbours. In [66], Smith and Brady proposed to compute such differences with a sigmoid-like function of the pairwise differences. Specifically, Smith and Brady propose to measure the uniformity of the area around each pixel as

$$U(x,y) = \sum_{(x',y') \in N(x,y)} e^{-\left(\frac{|I(x,y)-I(x',y')|}{t}\right)^{\alpha}}, \tag{9}$$

where t corresponds to the maximum expected dissimilarity between two pixels in the same object, and α is a coefficient influencing the slope of the sigmoid-like function. In [66], α is set to 6, although in our experiments better results were reached with $\alpha \in [2,4]$. For any $(x,y) \in \Omega$, $U(x,y) \in [0, \mathrm{card}(N(x,y))]$, with 0 measuring the maximal heterogeneity in the neighbourhood and $\mathrm{card}(A)$ denoting the cardinal of a set A. Hence, an *edginess* value can be obtained as

$$E(x,y) = 1 - \frac{U(x,y)}{N(x,y)}. \tag{10}$$

A key for the application of the USAN is the comparison of individual elements within the local context of the image. This seems to be a relatively evident question for scalar valued images, but becomes a complicated problem when the amount of information at each pixel increases. In fact, relevant color spaces as the CIELab have been created with the goal of easing pixel comparison; also, some authors propose custom-made kernels to produce distance-aware color spaces on demand. Nonetheless, provided a sensitive comparison operator, the SUSAN operator can be applied to any type of images.

In this experiment we attempt to produce fuzzy edge images from hyperspectral images using the comparison operators introduced in Section 3. In order to do so, we just need to reformulate Eq. (9) as

$$U(x,y) = \sum_{(x',y') \in N(x,y)} e^{-\left(\frac{q_i(I(x,y),I(x',y'))}{t}\right)^{\alpha}}, \tag{11}$$

for any of the q_i presented in Section 3 (also, the absolute difference s).

In order for the values yield by q_i to be normalized (in $[0,1]$), simple arithmetic can be used. Assuming that spectra have normalized reflectance values (percentage of maximum reflection), the results of the measures q_i just need to be divided by the number of wavelengths in the spectra.

5.2 Image segmentation using the USAN principle

Image segmentation is a mid level image processing task aimed at discriminating the areas of the image occupied by each visible object. Since the description of the

objects in an image is normally hierarchical, some authors have proposed segmentation tasks to produce hierarchical descriptions of the scene (see, e.g. the discussion on ultrametric contour maps in [2]). Due to the simplicity of the scenes in this experiment, we restrict to scalar labelling of individual pixels. Hence, a segmentation algorithm can be seen as a mapping from the universe of visual images to the universe of labelled images. Let $\Omega = \{1,\ldots,R\} \times \{1,\ldots,C\}$ be the positions of an image, we denote by $\mathbb{I}_X$ to the set of images whose pixels take values in X. Hence, a hyperspectral segmentation algorithm is a mapping $T : \mathbb{I}_{\mathbb{M}\mapsto[0,1]} \mapsto \mathbb{I}_{\mathbb{N}}$ such that, for any two positions $(x,y),(x',y') \in \Omega$, $T(I)(x,y) = T(I)(x',y')$ if and only if (x,y) and (x',y') are occupied by the same object in I. Alternative descriptions of segmentation can be found in, e.g., [28].

In the context of object segmentation in hyperspectral imagery, we intend to automatically discriminate two regions: the area in which vegetable material appears and the background area. Hence, only two classes (equiv. labels) are considered.

Let $H \in \mathbb{I}_{\mathbb{M}\mapsto[0,1]}$ be a hyperspectral image. Our algorithm, created as a modification of that in [44], is as represented in Fig. 4. A stepwise description is the following:

1. Create a *Peak energy image* of the hyperspectral image ($G \in \mathbb{I}_{\mathbb{R}}$), given by

$$G(x,y) = \bigvee_{\mu \in \mathbb{M}} I(x,y,\mu).$$

2. Compute a *fuzzy edge image* using the operator in Eq. (10). We refer to this image as $E \in \mathbb{I}_{\mathbb{R}}$.
3. Perform the watershed transform on E, so that the image is partitioned in regions. This process is likely to oversegment the image, creating highly partitioned regions.
4. Create tiled representations of the peak energy and the centrality of each region in C. These representations (namely G' and C') can be seen as spatially-granular interpretations of the image or, equivalently, as superpixel-based representations of different image features.
5. Since G' and C' represent semantic facts (energy and centrality), they can be combined using a logical operator. In this case, we are interested in tiles (regions) being in the center of the image and having high energy, and so we chose a t-norm. For this experiment, the product has been chosen.
6. The resulting image is thresholded using hysteresis [12]. The upper threshold for the hysteresis is computed according to the Otsu method [59], while the lower threshold is set as 70% of the upper one.

5.3 Experimental Dataset

In this study, we used a collection of 40 images of different food and vegetables to build our dataset. All images were of different size and shape, but each having 320 columns and a variable number of rows. The reflectance of each pixel was measured at 256 different wavelengths using a setup for hyperspectral imaging consisted of a
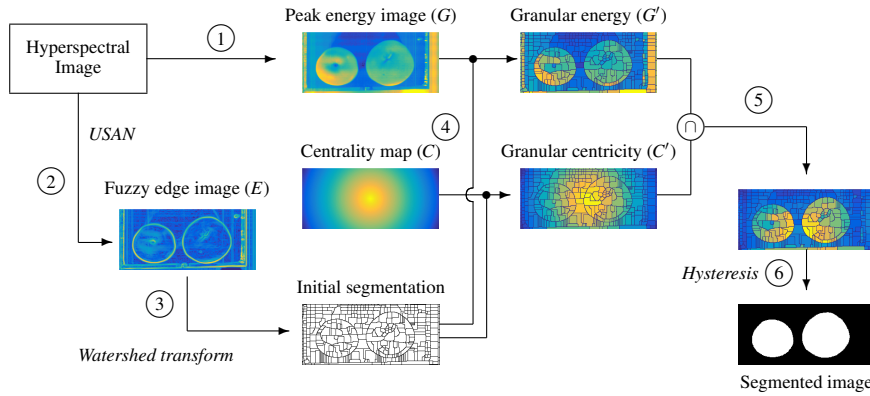
Fig. 4 Schematic representation of the image segmentation algorithm based on local contrast measurements, as explained in Section 5.2.

computer-controlled translation stage; four halogen lamps (Lexman 46 W) to illuminate the samples; an Indium gallium arsenide (InGaAs) CCD camera with a resolution of 320x256 pixels (Xenics, Model Xeva-1.7-320, Leuven, Belgium) coupled to a spectrograph (ImSpector N17E, Specim, Spectral Imaging Ltd., Oulu, Finland), and to a lens (Xenics, Model OPT-000107, SWIR 16 mm f/1.4, Leuven, Belgium) with a focal length of 16 mm; and, a computer fitted with the specific software Xeneth 2.5. The exposure time was optimized at 2 ms in order to maximize the spectral signal-to-noise ratio while avoiding saturation of specular reflective regions.

This image acquisition system has a line scanning or pushbroom configuration which acquires a whole line of an image each time requiring to move the sample under the field of view of the camera [41]. Therefore, images of one spatial and one spectral dimension are acquired by the detector of the camera and the third dimension is achieved by means of sample movement under the camera. Images are measured in the 900-1700 nm wavelength range with 3 nm resolution.

5.4 Quantification of the results

The results of our segmentation algorithm can be seen as binary results, so that the background takes the label 0 and the featured biological items take label 1. In parallel, human experts were required to manually label the images by pointing out the areas in which the biological items appeared. Let $I_{GT} \in \mathbb{I}_{\{0,1\}}$ be the ground truth image created by a human and let $I_C \in \mathbb{I}_{\{0,1\}}$ be a candidate image produced by our algorithm. The quality of I_C w.r.t. to I_{GT} is quantified as follows.

- *Statistical interpretation:* If segmentation is regarded as a classification task, quantification shall be done in terms of a confusion matrix. From the standard quantities in the confusion matrix (TP, FP, FN and TN) , we consider the following

quality indicators:

$$\mathrm{TP_R} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}, \ \mathrm{TN_R} = \frac{\mathrm{TN}}{\mathrm{TN+FP}}, \ \text{and Acc} = \frac{\mathrm{TN+TP}}{\mathrm{card}(\Omega)},$$

with $\mathrm{card}(\Omega)$ representing the cardinality of the set Ω.

Note that, due to the conditions of the dataset, most of the pixels are expected to be classified as *background*. Hence, TN is significantly greater than the other quantities. In such cases, it is recommended to use Precision/Recall measures, which avoid using TN. Precision, recall and F_α are defined as

$$\mathrm{Prec} = \frac{\mathrm{TP}}{\mathrm{TP+FP}}, \quad \mathrm{Rec} = \frac{\mathrm{TP}}{\mathrm{TP+FN}}, \ \text{and } F_\alpha = \frac{\mathrm{Prec+Rec}}{\alpha\,\mathrm{Prec}+(1-\alpha)\,\mathrm{Rec}},$$

respectively.

– *Topological interpretation:* This interpretation considers the distribution of the featured object (in this case, the biological item) in the image. We consider two measures for this interpretation. First, the symmetric distance which, as included in [46,45], is given by

$$\mathrm{SD}^k(I_\mathrm{C}, I_\mathrm{GT}) = \left[\frac{\sum_{p \in I_\mathrm{GT}} d^k(p, I_\mathrm{C}) + \sum_{p \in I_\mathrm{C}} d^k(p, I_\mathrm{GT})}{\mathrm{card}(I_\mathrm{GT} \cup I_\mathrm{C})} \right]^{\frac{1}{k}}$$

where d is a metric in $\mathbb{R}^2$ (in our experiments, the Euclidean metric) and k is a regulation exponent, which is set to 1 in our experiments. Second, Baddeley's Delta metric [4,5], given by

$$\Delta^\beta(I_\mathrm{C}, I_\mathrm{GT}) = \left[\frac{1}{\mathrm{card}(\Omega)} \sum_{p \in \Omega} |D(p, I_\mathrm{C}) - D(p, I_\mathrm{GT})|^\beta \right]^{\frac{1}{\beta}},$$

where $D(p,S)$ is the point-to-set distance defined as $D(p,S) = \min_{s \in S} d(p,s)$, d is a metric on Ω and $\beta \in \mathbb{R}^+$ is an exponent[2].

Note that, although both interpretations would yield maximal results in case of optimal solutions, they differ in how they model and penalize errors [46].

5.5 Experimental results

The results of the experiments are included in Table 1. Note that they are discriminated according to the operator N used for the configuration of local neighbourhoods at each pixel in the image. We consider, separately, radial neighbourhoods of radius 5 (Table 1(a)) and radial neighbourhoods of radius 7 (Table 1(b)).

The facts to be observed in Table 1 are diverse. At first glance our algorithm shows good performance regardless of the measure used for comparison and the USAN

---

[2] The original formulation by Baddeley [4,5] included a convex function w, but its effect can be also reached by means of a sensible choice of d and D. In fact, follow-up works have omitted such function [17, 16,77].

Measure		**Quality measures**								
		TP	FP	FN	*Acc*	Prec	Rec	$F_{0.5}$	SD[1]	Δ^2
AD (*s*)	$t = 0.05$	.13	.00	.04	.96	**.97**	.78	.87	.013	7.75
	$t = 0.10$	.14	.01	.03	.96	.96	.80	.87	.012	7.74
	$t = 0.15$	.14	.01	.03	.96	.96	.80	.87	.013	7.85
	$t = 0.20$	.14	**.01**	.03	.96	.96	.80	.87	.013	7.85
q_1	$t = 0.05$	.13	.01	**.04**	.96	.96	.78	.86	.013	7.80
	$t = 0.10$	.14	.01	.04	.96	.96	.79	.87	.013	7.68
	$t = 0.15$	.14	.01	.04	.96	.96	.79	.87	.013	7.67
	$t = 0.20$	.14	.01	.03	.96	.96	.80	.87	.013	7.63
q_2	$t = 0.05$	.14	.01	.03	.96	.96	.80	.87	.012	7.61
	$t = 0.10$	**.14**	.01	.03	**.96**	.96	**.80**	**.87**	**.012**	7.55
	$t = 0.15$	.14	.01	.03	.96	.96	.80	.87	.012	**7.55**
	$t = 0.20$	.14	.01	.03	.96	.96	.80	.87	.012	7.55

(a) Results with circular USAN area of radius $r = 5$.

Measure		**Quality measures**								
		TP	FP	FN	*Acc*	Prec	Rec	$F_{0.5}$	SD[1]	Δ^2
AD (*s*)	$t = 0.05$	.13	.00	**.04**	.96	**.97**	.78	.87	.013	7.81
	$t = 0.10$	.14	.00	.04	**.96**	.97	.79	**.87**	**.012**	**7.46**
	$t = 0.15$	.14	.01	.04	.96	.96	.79	.87	.013	7.73
	$t = 0.20$	.14	.01	.04	.96	.95	.80	.87	.013	7.89
q_1	$t = 0.05$	.13	.01	.04	.96	.96	.79	.87	.013	7.71
	$t = 0.10$	.14	.01	.04	.96	.96	.79	.87	.012	7.87
	$t = 0.15$	.14	.01	.04	.96	.95	.79	.86	.013	8.11
	$t = 0.20$	.14	**.01**	.04	.96	.95	.79	.86	.013	8.09
q_2	$t = 0.05$	.14	.01	.03	.96	.96	.80	.87	.012	7.61
	$t = 0.10$	**.14**	.01	.03	.96	.96	**.80**	.87	.012	7.66
	$t = 0.15$	.14	.01	.03	.96	.96	.80	.87	.013	7.79
	$t = 0.20$	.14	.01	.04	.96	.96	.79	.87	.013	7.84

(b) Results with circular USAN area of radius $r = 7$.

Table 1 The results gathered in the experiments. Results are discriminated according to the radius of the neighbourhood considered in the SUSAN edge detector. Appart from that parameter, we have considered three different spectral comparison measures (*s*, q_1 and q_2), each of them in combination with four different thresholds in the computation of the USAN area (*t*). Comparison measures are explained in Section 5.4. The best performers for each quantitative measure and USAN area are boldfaced.

threshold t. The algorithm is specially good in avoiding of FPs (normally featuring under 1% of FPs), although this comes at the cost of a significant number of FNs. In-lab agro-bio analysis normally intends to extract one or many spectral profiles from each item. In this regard, misclassifying parts of the object (FNs) is a problem much lighter than misclassifying regions of the background (FPs), whose spectral profiles might severely modify the spectral information from the item. Also, it is remarkable that the methodology presented here works better than those in previous approaches, e.g. than the one we presented in [44] on the same dataset.

With respect to the comparison of the results by different measures, we find very little differences in performance. The reasons lie in the treatment of the fuzzy edge images. Although each version of the SUSAN edge detector produces rather different images E, the watershed transform (which mostly uses the position of the local maxima in E, but not their intensity) equalizes their result. Still, differences arise in the topological interpretation of the results. This is to be expected, again. In the statistical interpretation of the results, most of the methods coincide in segmenting the objects, what creates a very significant number of TPs. The differences between the results mostly appear at the item boundaries. Such differences are visually noticeable, but account for a relatively small number of pixels. For the statistical measurements the impact is subtle, but for the topological measures they create a noticeable impact.

From the results in Table 1, we understand that the segmentation algorithm reaches its goal, but leaves few room for the interpretation of the differences between comparison measures. A potential solution would have been to test the quality of the edge images E in a pure edge detection dataset, but up to our knowledge there is no available dataset for edge detection in the context of in-lab hyperspectral imaging. In any case, we undestand that the results gathered in the experiment fully depict the validity of our proposals (q_1 and q_2), to perform competitively with s in real data scenarios.

6 Conclusions

This work has elaborated on the notions and literature to produce two comparison measures for functions in discrete domains. One of the functions (q_1) is a metric in the universe of such functions, while the other one (q_2) is a measure such that any two functions satisfying that one is the vertical transformation of the other, i.e., $f = g + c$, with $c \in \mathbb{R}$, are equally similar and they yield the minimum value. Such measures can be further use to compare functions coming from membership degrees in T2FSs, but also as spectral signatures in hyperspectral images. In fact, they have been put to the test in the context of in-lab hyperspectral image segmentation. Their results indicate that the new measures perform similarly to (eventually, slightly better than) the absolute difference, but hold more interesting theoretical properties. Still, it seems evident that the measures shall be used in contexts or applications in which their impact is better observed. Since item comparison is a very low image processing operation, those applications might as well be in the fields of data clustering or classification.

References

1. Allison, D.B., Cui, X., Page, G.P., Sabripour, M.: Microarray data analysis: from disarray to consolidation and consensus. Nature reviews genetics **7**(1), 55–65 (2006)
2. Arbelaez, P.A.: Une approche métrique pour la segmentation d'images. Ph.D. thesis, Université Paris-Dauphine (2005)
3. Attneave, F.: Dimensions of similarity. The American Journal of Psychology **63**(4), 516–556 (1950)
4. Baddeley, A.J.: An error metric for binary images. In: W. Förstner, S. Ruwiedel (eds.) Robust Computer Vision: Quality of Vision Algorithms, pp. 59–78. Wichmann Verlag, Karlsruhe (1992)
5. Baddeley, A.J.: Errors in binary images and an L^p version of the Hausdorff metric. Nieuw Archief voor Wiskunde **10**, 157–183 (1992)
6. Benson, E.S.: Trackable life: Data, sequence, and organism in movement ecology. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences **57**, 137–147 (2016)
7. Bibiloni, P., Gonzalez-Hidalgo, M., Massanet, S.: Soft color morphology. In: Proc. of the IEEE Intl. Conf. on Fuzzy Systems. IEEE (2017)
8. Bustince, H., Barrenechea, E., Pagola, M.: Restricted equivalence functions. Fuzzy Sets and Systems **157**(17), 2333–2346 (2006)
9. Bustince, H., Barrenechea, E., Pagola, M., Fernandez, J., Xu, Z., Bedregal, B., Montero, J., Hagras, H., Herrera, F., De Baets, B.: A historical account of types of fuzzy sets and their relationships. IEEE Transactions on Fuzzy Systems **24**(1), 179–194 (2016)
10. Campbell, F.W., Robson, J.: Application of fourier analysis to the visibility of gratings. The Journal of physiology **197**(3), 551–566 (1968)
11. Canny, J.: Finding edges and lines in images. Tech. rep., Massachussets Institute of Technology (1983)
12. Canny, J.: A computational approach to edge detection. IEEE Trans. on Pattern Analysis and Machine Intelligence **8**(6), 679–698 (1986)
13. Chamorro-Martínez, J., Martínez-Jiménez, P.M., Soto-Hidalgo, J.M., Prados-Suárez, B.: Perception-based fuzzy sets for visual texture modelling. Soft Computing **18**(12), 2485–2499 (2014)
14. Chamorro-Martínez, J., Soto-Hidalgo, J.M., Martínez-Jiménez, P.M., Sánchez, D.: Fuzzy color spaces: A conceptual approach to color vision. IEEE Trans. on Fuzzy Systems **25**(5), 1264–1280 (2017)
15. Chanussot, J., Mauris, G., Lambert, P.: Fuzzy fusion techniques for linear features detection in multi-temporal sar images. IEEE Trans. on Geoscience and Remote Sensing **37**(3), 1292–1305 (1999)
16. Coquin, D., Bolon, P.: Application of Baddeley's distance to dissimilarity measurement between gray scale images. Pattern Recognition Letters **22**(14), 1483–1502 (2001)
17. Coquin, D., Bolon, P., Onea, A.: Objective metric for colour image comparison. In: Proc. of the European Signal Processing Conference (2000)
18. De Baets, B., De Meyer, H.: Transitivity-preserving fuzzification schemes for cardinality-based similarity measures. European Journal of Operational Research **160**(3), 726–740 (2005)
19. De Baets, B., De Meyer, H.: On the cycle-transitive comparison of artificially coupled random variables. International Journal of Approximate Reasoning **47**(3), 306–322 (2008)
20. De Baets, B., De Meyer, H., Naessens, H.: A class of rational cardinality-based similarity measures. Journal of Computational and Applied Mathematics **132**(1), 51–69 (2001)
21. De Miguel, L., Campión, M., Candeal, J., Induráin, E., Paternain, D.: Pointwise aggregation of maps: Its structural functional equation and some applications to social choice theory. Fuzzy Sets and Systems **325**(Supplement C), 137–151 (2017)
22. Di Zenzo, S.: A note on the gradient of a multi-image. Computer Vision, Graphics, and Image Processing **33**(1), 116–125 (1986)
23. ElMasry, G., Kamruzzaman, M., Sun, D.W., Allen, P.: Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. Critical reviews in food science and nutrition **52**(11), 999–1023 (2012)
24. ElMasry, G., Sun, D.W.: Hyperspectral imaging for food quality analysis and control, chap. Principles of hyperspectral imaging technology, pp. 3–43. Academic Press: San Diego, CA, USA (2010)
25. González-Hidalgo, M., Massanet, S.: A fuzzy mathematical morphology based on discrete t-norms: fundamentals and applications to image processing. Soft Computing pp. 1–15 (2013)
26. Gregori, V., Miñana, J.J., Morillas, S., Sapena, A.: Characterizing a class of completable fuzzy metric spaces. Topology and its Applications **203**, 3–11 (2016)

27. Gregori, V., Morillas, S., Sapena, A.: Examples of fuzzy metrics and applications. Fuzzy Sets and Systems **170**(1), 95–111 (2011)
28. Guada, C., Gómez, D., Rodríguez, J., Yáñez, J., Montero, J.: Classifying image analysis techniques from their output. International Journal of Computational Intelligence Systems **9**, 43–68 (2016)
29. Haralick, R.M., Sternberg, S.R., Zhuang, X.: Image analysis using mathematical morphology. IEEE Trans. on Pattern Analysis and Machine Intelligence **9**(4), 532–550 (1987)
30. Harding, J., Walker, C., Walker, E.: Lattices of convex normal functions. Fuzzy Sets and Systems **159**(9), 1061–1071 (2008)
31. Harding, J., Walker, E.A., Walker, C.L.: The Truth Value Algebra of Type-2 Fuzzy Sets. CRC Press, Taylor & Francis Group (2016)
32. Hernández, P., Cubillo, S., Torres-Blanc, C.: Negations on type-2-fuzzy sets. Fuzzy Sets and Systems **252**, 111–124 (2014)
33. Hernández, P., Cubillo, S., Torres-Blanc, C.: On t-norms for type-2 fuzzy sets. IEEE Transactions on Fuzzy Systems **23**(4), 1155–1163 (2015)
34. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. The Journal of physiology **160**(1), 106–154 (1962)
35. Jacobson, N.P., Gupta, M.R.: Design goals and solutions for display of hyperspectral images. IEEE Trans. on Geoscience and Remote Sensing **43**(11), 2684–2692 (2005)
36. Jurio, A., Pagola, M., Paternain, D., Lopez-Molina, C., Melo-Pinto, P.: Interval-valued restricted equivalence functions applied on clustering techniques. In: 13th International Fuzzy Systems Association World Congress and 6th European Society for Fuzzy Logic and Technology Conference, Portugal, (2009)
37. Karmakar, G.C., Dooley, L.S.: A generic fuzzy rule based image segmentation algorithm. Pattern Recognition Letters **23**(10), 1215–1227 (2002)
38. Karnik, N.N., Mendel, J.M.: Operations on type-2 fuzzy sets. Fuzzy Sets and Systems **122**(2), 327 – 348 (2001)
39. Kramosil, I., Michálek, J.: Fuzzy metrics and statistical metric spaces. Kybernetika **11**(5), 336–344 (1975)
40. Law, T., Itoh, H., Seki, H.: Image filtering, edge detection, and edge tracing using fuzzy reasoning. IEEE Trans. on Pattern Analysis and Machine Intelligence **18**(5), 481–491 (1996)
41. Lawrence, K.C., Park, B., Windham, W.R., Mao, C.: Calibration of a pushbroom hyperspectral imaging system for agricultural inspection. Transactions of the ASAE **46**(2), 513–521 (2003)
42. Liang, Q., Mendel, J.M.: Interval type-2 fuzzy logic systems: theory and design. IEEE Transactions on Fuzzy Systems **8**(5), 535–550 (2000)
43. Lindeberg, T.: Edge detection and ridge detection with automatic scale selection. International Journal of Computer Vision **30**(2), 117–156 (1998)
44. Lopez-Molina, C., Ayala-Martini, D., Lopez-Maestresalas, A., Bustince, H.: Baddeleys delta metric for local contrast computation in hyperspectral imagery. Progress in Artificial Intelligence **6**(2), 121–132 (2017)
45. Lopez-Molina, C., Bustince, H., De Baets, B.: Separability criteria for the evaluation of boundary detection benchmarks. IEEE Trans. on Image Processing **25**(3), 1047–1055 (2016)
46. Lopez-Molina, C., De Baets, B., Bustince, H.: Quantitative error measures for edge detection. Pattern Recognition **46**(4), 1125–1139 (2013)
47. Lopez-Molina, C., Vidal-Diez de Ulzurrun, G., Baetens, J., Van den Bulcke, J., De Baets, B.: Unsupervised ridge detection using second order anisotropic Gaussian kernels. Signal Processing (2015)
48. Marr, D.: Vision. MIT Press (1982)
49. Marr, D., Hildreth, E.: Theory of edge detection. Proceedings of the Royal Society of London **207**(1167), 187–217 (1980)
50. Martín, J., Mayor, G., Valero, O.: Functionally expressible multidistances. In: Proc. of the Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT), pp. 41–46 (2011)
51. Masuzzo, P.: Towards an open data exchange ecosystem: forging a new path for cell migration data analysis and mining. Ph.D. thesis, Ghent University (2016)
52. McCulloch, J., Wagner, C., Aickelin, U.: Extending similarity measures of interval type-2 fuzzy sets to general type-2 fuzzy sets. In: 2013 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), pp. 1–8 (2013)
53. Mendel, J.M., John, R.I., Liu, F.: Interval type-2 fuzzy logic systems made simple. IEEE Transactions on Fuzzy Systems **14**(6), 808–821 (2006)
54. Mitchell, H., Estrakh, D.: A modified owa operator and its use in lossless dpcm image compression. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **5**(04), 429–436 (1997)

55. Mizumoto, M., Tanaka, K.: Some properties of fuzzy sets of type 2. Information and Control **31**(4), 312–340 (1976)
56. Mizumoto, M., Tanaka, K.: Fuzzy sets and type 2 under algebraic product and algebraic sum. Fuzzy Sets and Systems **5**(3), 277–290 (1981)
57. Molinari, F.: About a new family of multidistances. Fuzzy Sets and Systems **195**, 118–122 (2012)
58. Odet, C., Belaroussi, B., Benoit-Cattin, H.: Scalable discrepancy measures for segmentation evaluation. In: Proc. of the International Conf. on Image Processing, vol. 1, pp. 785–788 (2002)
59. Otsu, N.: Threshold selection method for gray-level histograms. IEEE Trans. on Systems, Man, and Cybernetics **9**(1), 62–66 (1979)
60. Palmeira, E., Bedregal, B., Bustince, H.: Applying two different methods to extend restricted dissimilarity functions. In: Proceedings of the 2015 Conference of the International Fuzzy Systems Association and the European Society for Fuzzy Logic and Technology (2015)
61. Palmeira, E.S., Bedregal, B.: Restricted Equivalence Function on L([0,1]), pp. 410–420. Springer International Publishing, Cham (2018)
62. Prewitt, J.M.S.: Object enhancement and extraction, pp. 75–149. Picture Processing and Psychopictorics. Academic Press (1970)
63. Ruzon, M.A., Tomasi, C.: Edge, junction, and corner detection using color distributions. IEEE Trans. on Pattern Analysis and Machine Intelligence **23**(11), 1281–1295 (2001)
64. Santini, S., Jain, R.: Similarity measures. IEEE Trans. on Pattern Analysis Machine Intelligence **21**(9), 871–883 (1999)
65. Shui, P.L., Zhang, W.C.: Noise-robust edge detector combining isotropic and anisotropic Gaussian kernels. Pattern Recognition **45**(2), 806–820 (2012)
66. Smith, S.M., Brady, J.M.: SUSAN- a new approach to low level image processing. International Journal of Computer Vision **23**, 45–78 (1997)
67. Sobel, I., Feldman, G.: A 3x3 isotropic gradient operator for image processing (1968). Presented at a talk at the Stanford Artificial Intelligence Project
68. Strasser, B.J.: Data-driven sciences: From wonder cabinets to electronic databases. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences **43**(1), 85–87 (2012)
69. Tversky, A.: Features of similarity. Psychological Review **84**(4), 327–352 (1977)
70. Tversky, A., Gati, I.: Studies of similarity. Cognition and Categorization **1**, 79–98 (1978)
71. Vanegas, M.C.: Spatial relations and spatial reasoning for the interpretation of earth observation images using a structural model. Ph.D. thesis (2011)
72. Waegeman, W., Pahikkala, T., Airola, A., Salakoski, T., Stock, M., De Baets, B.: A kernel-based framework for learning graded relations from data. IEEE Trans. on Fuzzy Systems **20**(6), 1090–1101 (2012)
73. Walker, C., Walker, E.: Type-2 operations on finite chains. Fuzzy Sets and Systems **236**, 33–49 (2014)
74. Walker, C.L., Walker, E.A.: The algebra of fuzzy truth values. Fuzzy Sets and Systems **149**(2), 309–347 (2005)
75. Wang, G., De Baets, B.: Edge detection based on the fusion of multiscale anisotropic edge strength measurements. In: Advances in Fuzzy Logic and Technology 2017, pp. 530–536. Springer (2017)
76. Wilkin, T., Beliakov, G.: Robust image denoising and smoothing with generalised spatial-tonal averages. In: Proc. of the IEEE International Conf. on Fuzzy Systems, pp. 1–7 (2017)
77. Wilson, D.L., Baddeley, A.J., Owens, R.A.: A new metric for grey-scale image comparison. International Journal of Computer Vision **24**, 5–17 (1997)
78. Wu, D., Mendel, J.M.: A vector similarity measure for linguistic approximation: Interval type-2 and type-1 fuzzy sets. Information Sciences **178**(2), 381–402 (2008)
79. Wu, D., Mendel, J.M.: A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets. Information Sciences **179**(8), 1169–1192 (2009)
80. Wu, H., Mendel, J.M.: Uncertainty bounds and their use in the design of interval type-2 fuzzy logic systems. IEEE Transactions on Fuzzy Systems **10**(5), 622–639 (2002)
81. Yearsley, J.M., Barque-Duran, A., Scerrati, E., Hampton, J.A., Pothos, E.M.: The triangle inequality constraint in similarity judgments. Progress in Biophysics and Molecular Biology **130**, 26 – 32 (2017)
82. Zadeh, L.: The concept of a linguistic variable and its application to approximate reasoning (I). Information sciences **8**(1), 199–249 (1975)
83. Zadeh, L.A.: Similarity relations and fuzzy orderings. Information Sciences **3**(2), 177–200 (1971)
84. Zadeh, L.A.: Fuzzy logic= computing with words. IEEE Trans. on Fuzzy Systems **4**(2), 103–111 (1996)
85. Zeng, W., Li, D., Yin, Q.: Distance and similarity measures between hesitant fuzzy sets and their application in pattern recognition. Pattern Recognition Letters **84**(Supplement C), 267 – 271 (2016)