Information Sciences 569 (2021) 479–495

Contents lists available at ScienceDirect

Information Sciences

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

On the role of distance transformations in Baddeley's Delta Metric

C. Lopez-Molina [a,b,c,*], S. Iglesias-Rey [a,b], H. Bustince [a], B. De Baets [c]

[a] Dept. of Estadistica, Informatica y Matematicas, Universidad Publica de Navarra, 31006 Pamplona, Spain
[b] NavarraBiomed, Complejo Hospitalario de Navarra, 31006 Pamplona, Spain
[c] KERMIT, Dept. of Data Analysis and Mathematical Modelling, Ghent University, 9000 Ghent, Belgium

ARTICLE INFO

Article history:
Received 19 November 2020
Received in revised form 29 March 2021
Accepted 16 May 2021
Available online 19 May 2021

Keywords:
Image comparison
Binary image
Baddeley's Delta Metric
Distance transform
Generalized distance transform

ABSTRACT

Comparison and similarity measurement have been a key topic in computer vision for a long time. There is, indeed, an extensive list of algorithms and measures for image or subimage comparison. The superiority or inferiority of different measures is hard to scrutinize, especially considering the dimensionality of their parameter space and their many different configurations. In this work, we focus on the comparison of binary images, and study different variations of Baddeley's Delta Metric, a popular metric for such images. We study the possible parameterizations of the metric, stressing the numerical and behavioural impact of different settings. Specifically, we consider the parameter settings proposed by the original author, as well as the substitution of distance transformations by regularized distance transformations, as recently presented by Brunet and Sills. We take a qualitative perspective on the effects of the settings, and also perform quantitative experiments on separability of datasets for boundary evaluation.
© 2021 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Binary images are one of the most common visual representations in modern computer vision. Although not very useful for information acquisition, they are ubiquitous in the representation of intermediate or final results in image processing. In fact, most of the low-level or mid-level image processing tasks represent their results as binary images, e.g. object tracking [1], boundary detection or object detection/segmentation [2]. This led to a vast literature on binary image processing, including very successful theories such as mathematical morphology [3] or tools such as local binary patterns [4,5]. One of the most active subfields of binary image processing is that of image comparison. This relates to the fact that performance evaluation, in many image processing tasks, is carried out by comparing the binary images produced by automated methods to gold standard (ground truth) images, boosting the need for strategies and measures for binary image comparison.

The nature of the pixel information in binary images makes them eligible for a large variety of comparison measures. For example, we find strategies inspired by the literature in classification (each pixel is an element to be classified) or based on set theory (an image is the set of the featured, *i.e.*, 1-valued, pixels). Also, less popular alternatives can be found based on trigonometry [6] or even dedicated signal representation frameworks [7,8].

* Corresponding author.
E-mail addresses: carlos.lopez@unavarra.es (C. Lopez-Molina), sara.iglesias@unavarra.es (S. Iglesias-Rey), bustince@unavarra.es (H. Bustince), bernard.debaets@ugent.be (B. De Baets).

https://doi.org/10.1016/j.ins.2021.05.034
0020-0255/© 2021 The Authors. Published by Elsevier Inc.
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

A prominent class of measures consists of the classification-based measures, which ground the comparison of two images in creating a binary confusion matrix [9]. Classification-based comparison measures have gained momentum in recent years, partly because of the simplistic generation and interpretation of the confusion matrices, which allow for the quantification of the convergent and divergent information in two images through pixel counting. Normally, counting coincident and divergent pixels leads to the generation of a binary confusion matrix, which is further processed to produce scalar measurements which represent the overall similarity of the images (see, e.g. [9] or [10], for a list of such measures). The underlying challenge in the generation of such confusion matrices is precisely how pixels are counted, or how images are matched prior to the counting. Generally, binary image comparison requires complex strategies which allow pixels in two images to be categorized as coincident (*true positives*, in confusion matrix jargon) if they are located near one another, even if not completely concurrent in position [11]. This led to a series of methods that allow for displacement-tolerant matching between binary images in the generation of the confusion matrix [12,13]. Displacement-tolerant matching algorithms still face a problem due to the potentially different number of 1-valued pixels in each image. If matching is performed on a 1-to-1 basis (as with the Cost Scaling Assignment algorithm [14], or with variants of the Hungarian/Munkres algorithm [15]), visually similar images might result in a large number of false positives; alternatively, 1-to-N or N-to-M matching schemes face the reverse problem: being unable to identify false positives, especially if images are cluttered or contain a large number of noisy pixels/regions.

In order to avoid problems in pixel counting, an alternative class of comparison measures for binary images was developed based on set theory and/or distance transformations. Distance transformations generate an alternative representation of binary images aimed at representing the structural (binary) features of the image. This representation can be used to compare images regardless of the number of featured pixels they have, or their distribution over the image. Early examples of measures partially based on distance transformations are Pratt's FoM (PFoM) [16] and Haralick's measure [17]. These ideas further evolved into full-fledged metrics on the binary image universe. Since featured (1-valued) pixels constitute a subset of positions in the image, some geometric measures for multidimensional information objects were translated to the context of binary image comparison. The most popular alternative is the Hausdorff Metric (HM), which has been further modified to produce task-specific operators [18,19].

A very relevant metric for binary images, heavily inspired by HM, is Baddeley's Delta Metric (BDM) [20,21]. In his works, Baddeley advocates for metrics over other types of comparison measures, but at the same time aims at correcting the instability of HM against subtle changes in the images to be compared. While both goals seem to be fairly accomplished in [20,21], BDM still suffers from some drawbacks, mostly due to the significant number of parameters in its formulation. In this work we study, from different perspectives, the role of the parameters in BDM. Within this study, we include the parameters in the original proposal by Baddeley, but also further contributions in the literature, especially the use of alternatives for distance transformations [22]. We consider the qualitative comparison of the results for each parameter setting, while also performing large-scale quantitative evaluations on real datasets, aiming at the best-performing configuration strategies for the metric.

The remainder of this paper is organized as follows. Section 2 recaps some notions on distance transformations for binary images. Section 3 presents BDM, together with an analysis of its behaviour depending on the underlying distance transformation. Section 4 contains two experiments contrasting the behaviour of different configurations of BDM in the comparison of images in the BSDS500 [23], a popular dataset for boundary detection. Finally, Section 5 includes the discussion on our developments.

2. Distance transforms of binary images

This section recaps some common concepts used in upcoming sections, specifically those related to distance transformations and their different configurations.

Definition 1. Let $q \in [-\infty, +\infty]$ and $n \in \mathbb{N}$. The q-power mean is the mapping $(\mathbb{R}^+)^n \to \mathbb{R}^+$ defined by

$$\mathrm{G}_q(\mathbf{x}) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^q\right)^{1/q}.$$

The value of q has a critical effect on the behaviour of G_q. For example, it holds that $G_{-\infty}$ (resp. G_{∞}) coincides with the minimum (resp. maximum) operator, while G_0 coincides with the geommetric mean. Also, if $q \leq 0$, then $G_q(\mathbf{x}) = 0$ if and only if $x_i = 0$ for some $i \in \{1, \ldots, n\}$; this means in particular that 0 is an annihilator of G_q when $q \leq 0$ [24]. This observation will prove relevant in this work.

Definition 2. An image is a function $f : \Omega \mapsto T$, with $\Omega = \{1, \ldots, R\} \times \{1, \ldots, C\}$ the set of positions and T the set of tones.

Assuming some fixed Ω, we refer with $\mathbb{I}_T$ to the set of all images with tonal palette T. This work mostly elaborates on the use of binary images, *i.e.*, images in $\mathbb{I}_{\{0,1\}}$. Binary images can be seen as mappings $\Omega \mapsto \{0, 1\}$ or as subsets of Ω, so that $p \in I$ if and only if $I(p) = 1$, with $p \in \Omega$ and $I \in \mathbb{I}_{\{0,1\}}$. This dual notation allows to combine mathematical tools from both interpretations, so as to simplify the manuscript and to ease the understanding of the underlying concepts.

Definition 3. A function $g : U \times U \mapsto \mathbb{R}^+$ is called a pseudo-metric on a universe U if and only if it satisfies the following properties, for any $a, b, c \in U$:

(i) $g(a, a) = 0$ for all $a \in U$;
(ii) Symmetry: $g(a, b) = g(b, a)$;
(iii) Triangle inequality: $g(a, c) \leqslant g(a, b) + g(b, c)$.

Definition 4. A function $g : U \times U \mapsto \mathbb{R}^+$ is called a metric (or distance function) on a universe U if and only if it satisfies the following properties, for any $a, b, c \in U$:

(i) Identity of indiscernibles: $g(a, b) = 0$ iff $a = b$;
(ii) Symmetry: $g(a, b) = g(b, a)$;
(iii) Triangle inequality: $g(a, c) \leqslant g(a, b) + g(b, c)$.

Definition 5. The Distance Transform (DT) of a binary image $I \in \mathbb{1}_{\{0,1\}}$, for some metric m on Ω, is the image $\mathscr{T}_m[I] \in \mathbb{1}_{\mathbb{R}^+}$ so that

$$\mathscr{T}_m[I](p) = \min_{p' \in I} m(p, p') \tag{1}$$

for all $p \in \Omega$.

Distance transformations have a natural connection to other mathematical models with application to image processing, e.g., Voronoi surfaces [25,26]. Also, they relate to other well-studied image processing paradigms, such as mathematical morphology [3].

In [22], Brunet and Sills introduced a generalized version of the distance transformation which replaces the minimum operator by the power mean of the distances to all pixels in the image.

Definition 6. [22] The Generalized Distance Transform of a binary image $I \in \mathbb{1}_{\{0,1\}}$, for some metric m on Ω, is the image $\mathscr{P}^q_m[I] \in \mathbb{1}_{\mathbb{R}^+}$ so that

$$\mathscr{P}^q_m[I](p) = \mathsf{G}_q\Big((w(m(p, p'))_{p' \in I}\Big) = \left(\frac{1}{|I|}\sum_{p' \in I} w(m(p, p'))^q\right)^{1/q} \tag{2}$$

for all $p \in \Omega$, with $q \in [-\infty, \infty]$ and w a concave, increasing function such that $w(0) = 0$.

Brunet and Sills [22] impose $q < 0$ in their original definition. The reason relates to the properties of the power mean. If $q \leq 0$, we have that $\mathscr{P}^q_m[I](p) = 0$ if and only if $p \in I$. This ensures $\mathscr{P}^q_m$ to be invertible as long as $q \leq 0$, which has an impact on the properties of subsequent operators.

It is relevant to note that the name *generalized distance transform* was used prior to [22], with a purpose different from the one in that work. Specifically, it was used as an attempt to use distance transforms for non-binary images. Early efforts in this direction are [27,28], although it was Felzenszwalb and Huttenlocher who properly coined the term [29]. In this work, in order to avoid any potential conflict in the naming, we refer to the transformation proposed by Brunet and Sills as *Power Distance Transformation* (PDT).

The PDT is related to some other proposals in literature. For example, it is somehow similar to the k-distances by Öfverstedt et al. [30], which replace the minimum operation in Eq. (1) by the 2-power mean of the k lowest distances. In this regard, the k-distances are a combination of the PDTs in [22] with OWA-like operators [31].

The original formulation of the PDT [22] keeps the use of a concave function w, $w(x) = 0$ if and only if $x = 0$, as in the original works by Baddeley [20,21]. This function is ignored in the present manuscript because any such concave mapping of a metric m is still a metric. Hence, removing the function w saves a parameter while losing no flexibility at all in the configuration of the operator.

The formulation in Eq. (2) is the result of replacing the minimum operator in Eq. (1) by a power mean, whose behaviour is controlled by the parameter q. For $q = -\infty$, such power mean becomes the minimum operator, and hence Eq. (2) is equivalent to Eq. (1) Alternatively, $q = \infty$ would render the power mean in Eq. (2) into the maximum operator. Different values of q can turn into other well-studied means, as the arithmetic mean ($q = 1$) or the harmonic mean ($q = -1$). As said before, and unlike the DT, the PDT is not always an invertible operation. While $\mathscr{T}_m[I](p) = 0$ iff $p \in I$, making the DT trivially invertible, this only holds for the PDT if $q \leq 0$. We refer to [24] for a detailed review on aggregation operators, including averaging operators and means.

Distance transformations can be based on any metric on the set of positions Ω. Let m be a metric on U, then the bounded version of m, on a universe U, is given by

$$m_t(a,b) = \min(t, m(a,b)), \tag{3}$$

for any $a, b \in U$, with $t \in \mathbb{R}^+, t > 0$, an arbitrary parameter. We refer with *bounded distance transformation* to the distance transformation based on a bounded metric. An unbounded metric, *i.e.* m_∞, is simply referred to as m. In the remainder of this work, generic metrics will be referred to as m, while the Euclidean metric will be represented as d.

Fig. 1 includes the colourized versions of the distance transforms of fives sample images. The results by each transformation are normalized row-wise for the purpose of clarity; a homogeneous normalization of all transforms is impractical due to the very different range of each DT and PDT, as is evident from the rightmost scales of Fig. 1. In this figure, as well as in the remainder of this work, DTs and PDTs are based on the Euclidean metric (d), which is the most recurrent choice in image processing.

There is a list of facts to be remarked from Fig. 1. The behaviour of $\mathscr{T}_d$ is as expected, with large values being generated far from the featured pixels in the image I. As opposed to $\mathscr{T}_d$, the bounded transformation $\mathscr{T}_{d_{20}}$ manages to create flat areas that represent regions farther than 20 pixels away from any featured pixel. The behaviour of $\mathscr{P}_d^q$ is significantly more complex to understand. PDT computes, at each pixel, a mean of the distances to all featured pixels in I. This brings two consequences. Firstly, we have that the value of $\mathscr{P}_d^q[I]$ at some pixel $p \in \Omega$ is not based on the specific position of a singular pixel in I, as in $\mathscr{T}_d[I]$ or $\mathscr{T}_{d_{20}}[I]$, but on the overall distribution of featured pixels over the image. Of course, this is problematic when the featured pixels are spread across the image, as is the case of the images sketching a starfish; in these cases, the transform barely represents the original image I. Secondly, the minimal values yielded by $\mathscr{P}_d^q$ might be greater than zero, depending on the specific value of q. If $q > 0$, then there are no zero values in the distance transform. In fact, there is no guarantee that values at pixels $p \in I$ will be lower than those at $p \notin I$. If $q \leq 0$, then the distance transform will be zero for all $p \in I$. However, pixels nearby (even adjacent to) those $p \in I$ might produce values that are significantly greater than zero. This is due to the sharp decrease of $G_q(\mathbf{x})$, for $q \leq 0$, when $x_i \to 0$ for some x_i.

The parameter q in Eq. (2) is, hence, both relevant and non-evident. While $q = -\infty$ would recover the original DT, any value $q \approx 0$ will lead to severe regularization effects (except for $p \in I$ when $q < 0$, as explained before). The regularization effect can be effectively, yet not totally, reduced by using a bounded metric d_t. The reason is that, by using a bounded metric, pixels in $p \in \Omega$ farther than t pixels from the closest $p' \in I$ will be t-valued in $\mathscr{P}_{d_t}^q[I]$. This can be observed in images $\mathscr{P}_{d_{20}}^1[I]$, where the shape of the original objects in I can be recognized, as opposed to the images $\mathscr{P}_d^1[I]$. In any case, and even if the regularization effect is obviously reduced, the results by $\mathscr{P}_{d_{20}}^1$ appear as regularized versions of those by $\mathscr{T}_{d_{20}}$. The regularization effect of $\mathscr{P}_{d_t}^q$ will be dependent upon q, with $q = -\infty$ inducing no regularization.

3. Baddeley's Delta Metric for binary image comparison

This section recaps the literature and history of Baddeley's Delta Metric (Section 3.1), prior to a description of the impact of its parameters (Section 3.2).

3.1. Definition of Baddeley's Delta Metric

Binary images can be analogously seen as functions or as subsets of positions (in Ω). Elaborating on this dual interpretation, a frequent choice for binary image comparison is the Hausdorff metric [32]. Let $\mathscr{T}_m$ be a distance transformation on binary images based on some metric m. The Hausdorff distance between two images, $A, B \in \mathbb{I}_{\{0,1\}}$ is given by

$$HM(A,B) = \max\left(\max_{a \in A} \mathscr{T}_m[B](a),\ \max_{b \in B} \mathscr{T}_m[A](b)\right). \tag{4}$$

The main concern with the Hausdorff metric is the fact that the value yielded for any two images is dependent upon two pixels only, namely those producing the maximum value in one of the distance transforms. While this makes the metric adequate for certain tasks, especially those in which noise is unexpected or undesired, its behaviour is rather unstable for most applications.

The literature contains a large number of proposals tuning the Hausdorff metric, with different, specific goals. Very relevant in this context is the work of Dubuisson and Jain [18], focused on the performance of different image-to-image, asymmetric similarity measures,[1] in combination with bidirectional aggregation operators. In total, they put to the test 24 different Hausdorff-metric-inspired comparison operators in the context of delineated object recognition. In follow-up works, Takacs and Wechsler [19,33] considered a slightly different tuning to adapt the results in [18] to face recognition through the comparison of boundaries. Note that the robustness against outliers and noise has not been the only area of improvement in Hausdorff-metric-based comparison measures. For example, Baudrier et al. combined the ideas in [18] with a local, sliding window-based, analysis of subregions of the image. It is worth mentioning that all these modifications of the Hausdorff metric often render into non-

[1] While the authors in [18] present some operators as *directed distance measures*, it is well established that metrics should be positive and symmetric. Hence, we prefer to refer to those operators as asymmetric similarity measures, avoiding any conflict with metric axioms.

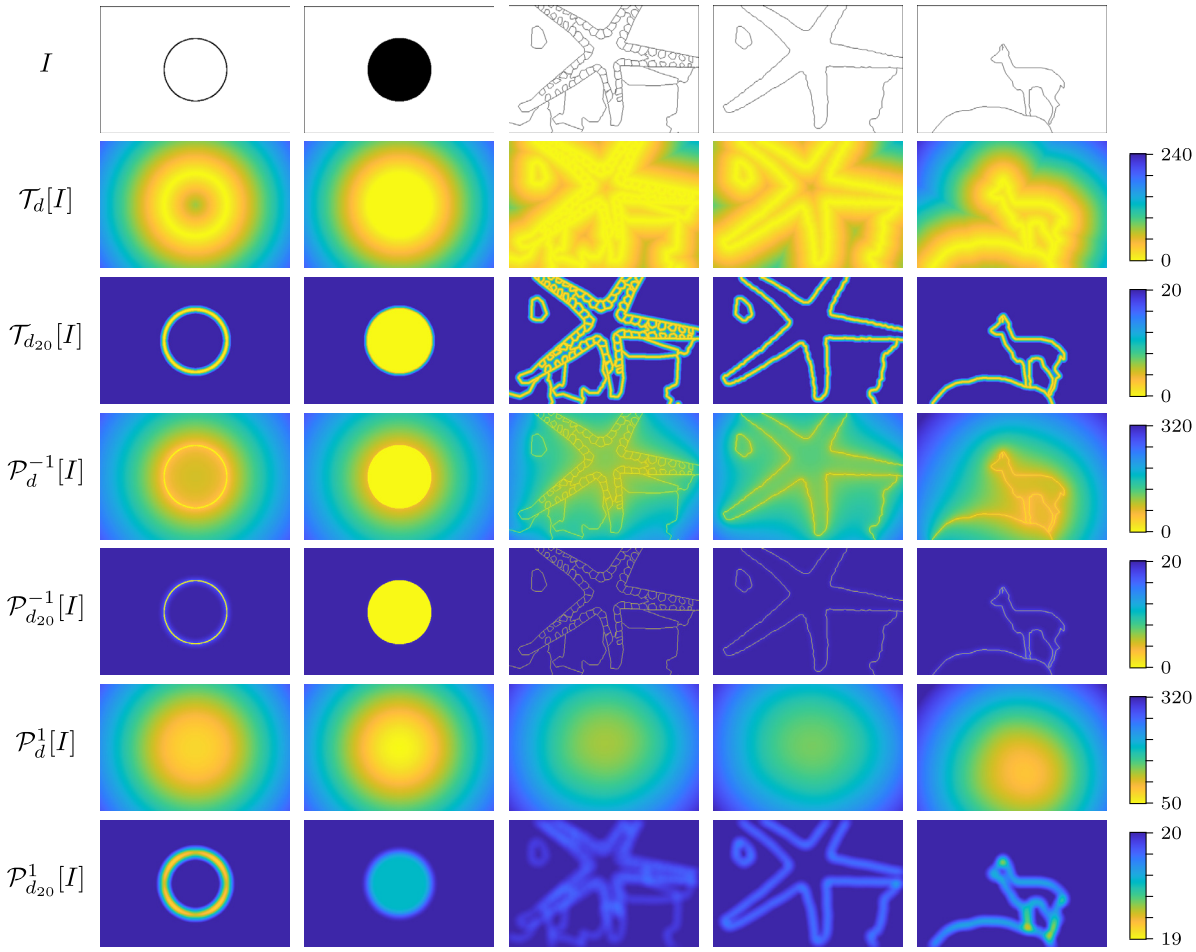

Fig. 1. Distance transforms of different binary images. The two leftmost images are synthetic examples, while the three rightmost images are taken from the BSDS500 [23]. The upper row displays the binary images, with 1s represented in black for better visualization. The lower rows display the results by different distance transformations, each of them expressed following the notation in Section 2. Note that each distance transformation uses its own numerical scale.

metric comparison operators, since the authors focus on increasing the flexibility of the operators rather than on preserving the metric axioms.

In [20,21], Baddeley introduced a metric for binary images that is more flexible than the Hausdorff metric but, at the same time, preserves the metric axioms. Baddeley defends the use of metrics for a list of reasons, including the fact that metrics generate topologies, which define *notions of continuity and convergence* [20]. Hence, his proposal for binary image comparison is a metric, which has been thereafter referred to as Baddeley's Delta Metric (BDM).

Baddeley's reasoning starts by reformulating the Hausdorff metric as

$$HM(A,B) = \max_{p\in\Omega} |\mathscr{T}_m[A](p) - \mathscr{T}_m[B](p)| \tag{5}$$

which is equivalent to Eq. (4). The original intention of Baddeley is to replace the maximum operator in order for the newborn metric not to be exclusively dependent on a pair of pixels.

Let $A, B \in \mathbb{I}_{\{0,1\}}$ be two binary images on Ω, and let m be some metric on Ω. The distance between them, in terms of BDM is given by

$$\Delta^k(A,B) = \left[\frac{1}{|\Omega|}\sum_{p\in\Omega} |w(\mathscr{T}_m[A](p)) - w(\mathscr{T}_m[B](p))|^k\right]^{\frac{1}{k}}, \tag{6}$$

where $w : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a concave function with $w(x) = 0$ iff $x = 0$, $\mathscr{T}_m$ is a distance transformation and $k \in \mathbb{R}^+$.

The role of the function w is, by far, the most intriguing in Eq. (6). In his original work [20], Baddeley mentions the *problem with over-sensitivity to large error distances* in the Hausdorff metric and then proposes to tune w as a solution. However, it is

clear that any valid mapping w of the distance transform is equivalent to use a w-affected metric within that transform. Otherwise said, any effect in the metric that could be reached by setting w, can be equivalently inferred by modifying the metric m. The function w is, hence, redundant. In this work, we set $w(x) = x$, focusing our attention on the configuration of m, specifically by using bounded metrics. Also, we consider replacing DTs by PDTs, as proposed in [22].

It is relevant to note that, although designed to compare binary images, BDM has been used for many different purposes other than its original goal. Since Δ^k can be used to compare any two subsets of a (discrete) metric space, its uses have surpassed the context of binary images. For example, follow-up works aimed at comparing grayscale images [27,28,34], or hyperspectral signatures [35].

3.2. Configuration of Baddeley's Delta Metric

Baddeley's Delta Metric is known to have certain problems in balancing the different distortions or discrepancies that can occur in binary images. Distortions in binary images can be typically broken down to missing information (false negatives), spurious information (false positives) and spatially-inaccurate information (displacements). Baddeley's Delta Metric normally incurs in overweighting the presence of spurious information over the remaining two distortions, especially when such spurious information is far from the featured pixels in the image.

In order to test the sensitivity of BDM to different types of noise, we have induced progressive distortions to the image leftmost image in the upper row of Fig. 1. Each distortion undergone by the image is modeled as follows:

- *Missing Information*: A segment is removed from the upper part of the circumference. The removed information is specified as a percentage of the total number of pixels in the circumference.
- *Spurious Information*: A spurious response (3×3 block of 1-valued pixels) is added to the image. This response is placed on the circumference (North-East side), then displaced to the east according to a distance-to-object parameter.
- *Displacements*: The radius of the circle is progressively increased, yet preserving the center position of the circumference and the line width.

A visual interpretation of the experiment is displayed in Fig. 2. This figure includes different distorted versions ($I\prime$) of the original image (I) in the leftmost column of Fig. 1. Also, Fig. 2 contains their distance transforms using the unbounded Euclidean metric d, and the absolute pixelwise difference w.r.t. the distance transform of the original image in Fig. 1.

In Fig. 2 we observe that a relatively small number of false positive responses have a significant impact on the results by BDM. Specifically, we see how the inclusion of a small false positive in the image has a severe impact in $\mathscr{T}_d[I']$ (central columns of Fig. 2). The impact of placing a 3×3 patch 80 pixels away from the circumference is, visually, very light. However, we observe in Fig. 2 how the impact in the DT is comparable to that achieved by removing 35% of the circumference. Subsequently, as seen in the lowest row of the figure, such impact produces a large divergence in terms of $|\mathscr{T}_d[I] - \mathscr{T}_d[I']|$. As a final consequence, an almost imperceptible discrepancy between I (in Fig. 1) and I' renders in a large difference in $|\mathscr{T}_d[I] - \mathscr{T}_d[I']|$, comparable to that suffered when removing a significant part of the circumference (leftmost column of Fig. 2) or when dramatically increasing its size (rightmost column of Fig. 2).

Fig. 3 offers a quantitative insight into the results in Fig. 2. In Fig. 3 we display the distance, as measured by BDM, between the image in the leftmost column of Fig. 1 and a distorted version of itself. We use the unbounded Euclidean metric d and $k \in \{1, 1.5, 2\}$. As we can observe in Fig. 3, BDM is much more sensitive to false positives than to other distortions, especially when false positives occur far from other featured pixels. In Fig. 3, columns (a) and (c) display distortions of the original image I much more visually salient than those in column (b). However, in Fig. 3 the distortion measured when adding a block of 3×3 pixels in the far right part of the image is equivalent to that measured when removing 35% of the circumference, or to that measured when the radius of the circumference is increased by 15–20 pixels. The reason lies in the interpretation of the image through its distance transform.

We propose, as Brunet and Sills [22], to replace the distance transform ($\mathscr{T}_m[I]$) in BDM by the PDT ($\mathscr{P}^q_m[I]$). Let $A, B \in \mathbb{1}_{\{0,1\}}$ be two binary images on Ω, and let m be some metric on Ω. The distance between A and B is measured as

$$\Delta^k(A,B) = \left[\frac{1}{|\Omega|}\sum_{p\in\Omega}\left|w(\mathscr{P}^q_m[A](p)) - w(\mathscr{P}^q_m[B](p))\right|^k\right]^{\frac{1}{k}}, \tag{7}$$

A relevant question relates to the preservation of the metric axioms when DTs are replaced by PDTs in Eq. (6). If $q \leq 0$, then Δ^k is a metric (for a proof, see [22]). However, if $q > 0$, then Δ^k is a pseudo-metric, since the identity of indiscernibles no longer holds. However, this is not a significant problem. The role of the identity of indiscernibles is practical from the mathematical point of view, especially when metrics are used for optimization. But, from the perspective of mathematical psychology, it is unclear whether humans have a behaviour coherent with the identity of indiscernibles. Humans often identify as *equal* or *equivalent* things that are, strictly speaking, non-equal. In fact, such human behaviour is partially responsible for the emergence of mathematical theories such as Fuzzy Set Theory [36,37].

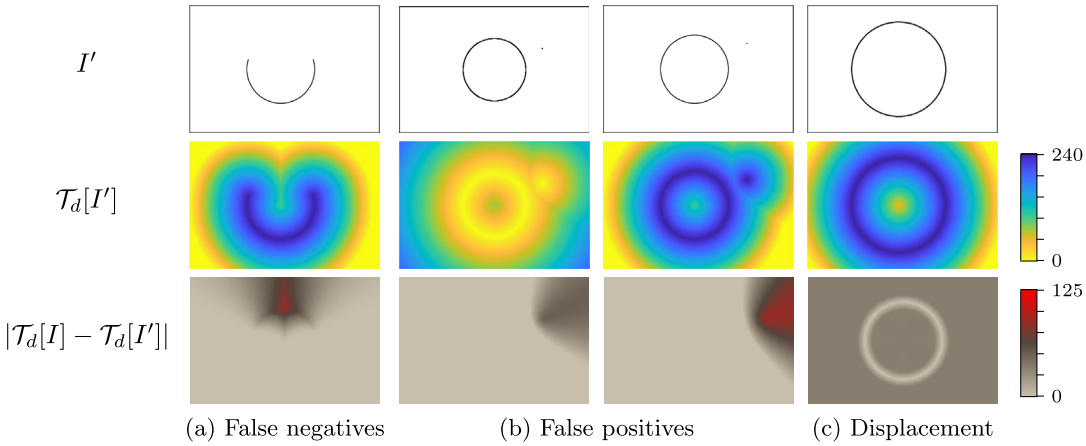

Fig. 2. Visualization of the distance transforms of distorted versions of the image (I) in the leftmost column of Fig. 1. The distortions are: Leftmost column, removal of a segment with 35% of the pixels; middle columns, addition of a 3×3 pixel spot 40 and 80 pixels away from the circumference, respectively; rightmost column, increase of the radius of the circumference in 35 pixels. The lowest row presents the absolute difference between the transform and the original image.

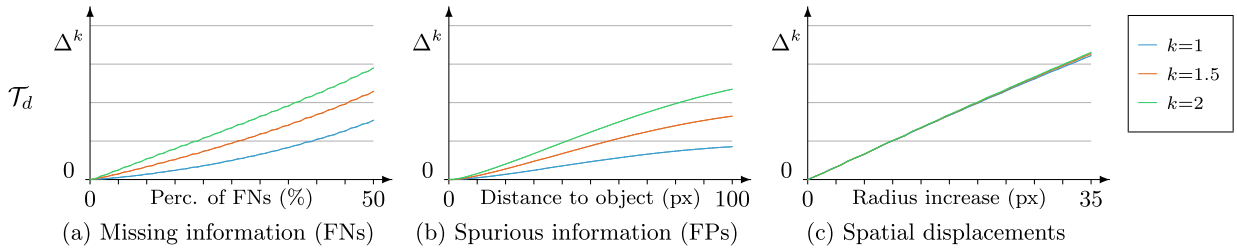

Fig. 3. Quantification, by means of BDM, of different types of distortions of the binary image in the leftmost column of Fig. 1. Each column is dedicated to a different type of distortion. The leftmost column represents the progressive, random elimination of featured pixels. The center column represents the inclusion of a 3×3 patch of 1s at a variable distance of the circumference. The rightmost column represents the increase of the radius of the circumference. All plots are homogeneously scaled on the vertical axis for better comparison.

In the remainder of this work, we keep the naming *Baddeley's Delta Metric* for any value of q, taking into account the fact that it renders into either a metric (for $q \leq 0$) or a pseudo-metric (otherwise). Also, we keep the notation Δ^k for the different parameterizations, in an attempt not to relabel existing operators.

The value of q in Eq. (7) does not only impact the mathematical properties of Δ^k. It also affects how PDTs model the binary features in the images. Since $\mathscr{P}_m^q[I](p) = 0$ for all $p \in I$ if and only if $q < 0$ (we assume that I has two or more 1-valued pixels, since the opposite would be a rather aberrant situation), the distance transform can be unstable near such $p \in I$. This fact can be observed in Fig. 4. In this figure, we display, for three different images I, $\mathscr{P}_m^q[I]$ computed with $q \in \{-1, 1\}$ and $m \in \{d, d_{20}\}$. Also, the figure displays the values of the PDTs at the 190th row of the images, which is marked with a dotted line at the top row of the figure.

The main observation to be made in Fig. 4 is the sudden change of the distance transform near the pixels $p \in I$, when $q = -1$. While it is clear that PDTs with $q < 0$ will be zero at those pixels, we can also observe in Fig. 4 how nearby pixels produce relatively high values. This happens because the power mean of the distances at each pixel can greatly change depending on whether the vector of distances contains a zero or not. This is not a problem *per se*, but might have a negative influence on the behaviour of Δ^k, which is heavily dependent on the pixelwise comparison of distance transforms. Slight displacements of solid shapes (as the deer in Fig. 4(c)) should not be a significant problem; as long as the shapes are mostly overlapping, the distance transforms would remain similar. However, small variations in the position of thin features (as the deer silhouette in Image A, Fig. 4(a)) will lead to significant divergences in the resulting distance transform. If the features are thicker (as in Images B and C, Fig. 4(b) and (c)), the problem will be lighter, since small displacements would still incur in a significant overlapping of such features. According to these facts, positive values of q seem to be more adequate to compare images with thin, linear structures.

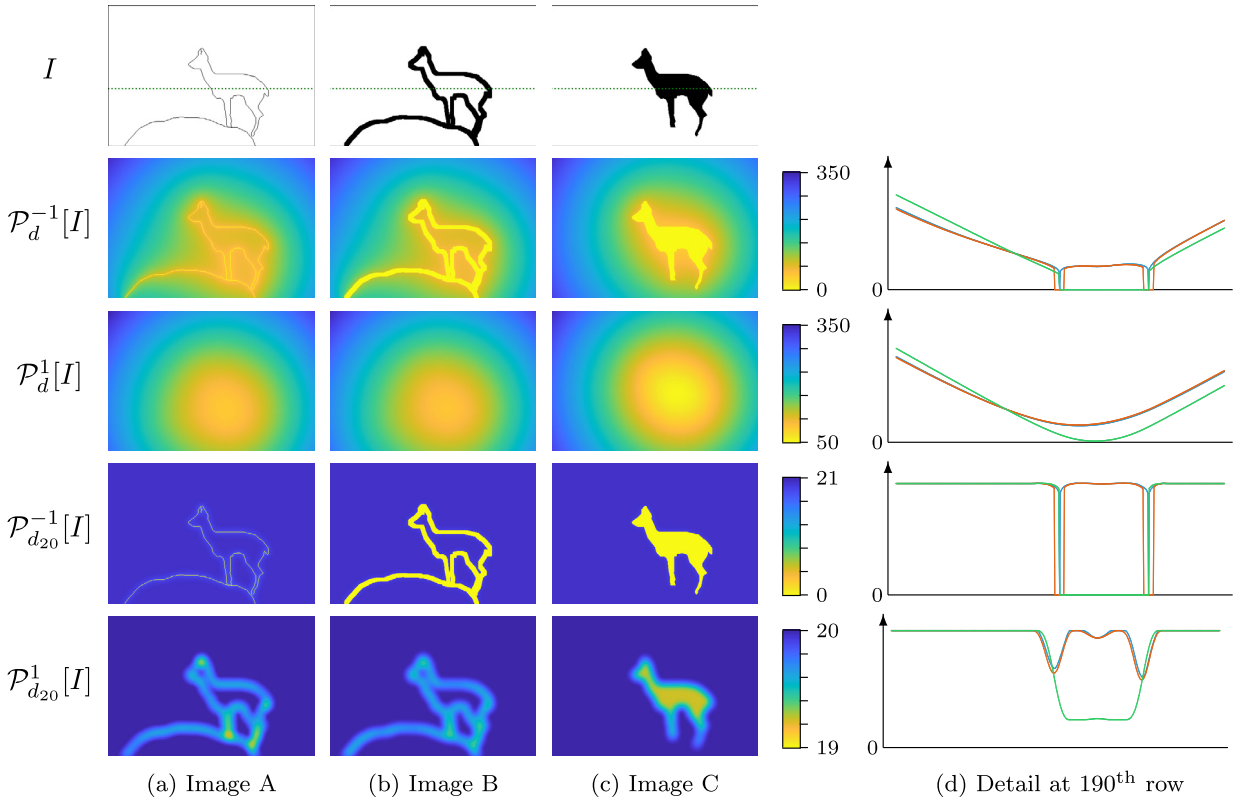

Fig. 4. Visualization of distance transform on three versions of the same image. For each image, we display the distance transform under different combinations of q and the metric d. Also, for each of such combinations, we display on the rightmost column the detail of the 190th line of the image. In this plot, *Image A*, *Image B* and *Image C* take the colors blue, orange and green, respectively.

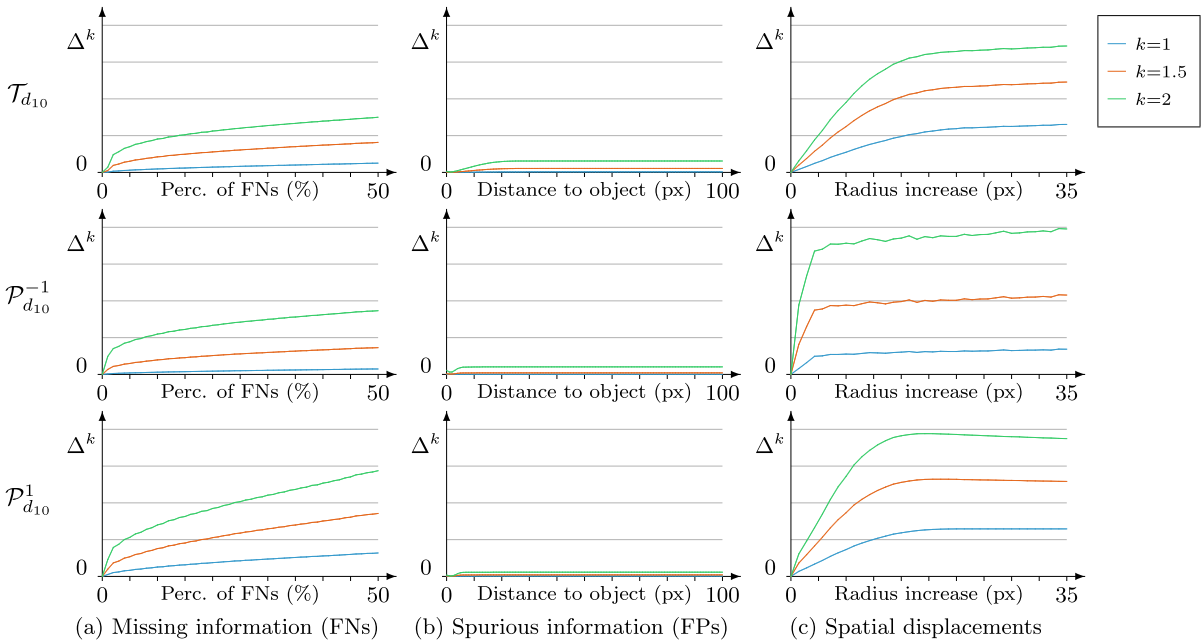

Fig. 5. Replication of the experiment in Fig. 3 using different distance transformations. All plots at each row are homogeneously scaled on the vertical axis for better comparison.

The experiment in Fig. 3 has been repeated using DTs and PDTs other than $\mathscr{T}_d$. The results are displayed in Fig. 5. In the upper row, we consider $\mathscr{T}_{d_{10}}$, while in the two lower rows we display the results using $\mathscr{P}^{-1}_{d_{10}}$ and $\mathscr{P}^{1}_{d_{10}}$. Note that the scale is different for each version of BDM, for the sake of visibility, but is kept constant for each row (*i.e.*, for each distance transformation). In all cases, we observe how the sensitivity to subtle false positives is significantly lower than that to the remaining distortions. Also, we observe that the sensitivity of the image metrics to false negatives and displacements takes a log-like evolution w.r.t. the amount of distortion in the images. For example, in the case of displacements, there is a fast increase for small radius increments, followed by a quasi-flat evolution thereafter.[2] This makes sense, since for small radius increases the distorted image is recognized as *similar* to the original one; then, upon reaching a certain level of distortion, the image is simply *different*, regardless of the specific amount of distortion. Otherwise said, the dissimilarity quantified with a radius increase of 15 pixels and that of 20 pixels can be both perceived as similarly distant from the original image. The influence of the parameter k is as expected in the formulation, with Δ^k increasing along with k.

From the experiments in Figs. 3 and 5, we can infer that bounded distance metrics d_t manage to solve some of the problems in d-based BDM, either applied to DTs or to PDTs. However, this improvement shall be quantitatively confirmed with different images and configurations. The next section is devoted to such task.

4. Large-scale quantitative experiments

This section is focused on understanding the impact of PDTs in Baddeley's Delta Metric, in a context as complex as boundary detection comparison. Boundary images contain relatively little information (in terms of number of pixels), but represent complex shapes and objects. This further extends the initial experiments in [22], in which the authors tested BDM (only for $q < 0$) with synthetic imagery, as well as with meteorological images featuring solid binary objects/regions.

4.1. Metaquality evaluation for boundary detection

Methods for binary image comparison, either metrics, dissimilarity measures, or any other class of functions, can be used for different tasks. For example, they have applications in object tracking [38], object recognition [7] or biometric ID [39] identification. We focus on the context of quality evaluation, specifically restricting to boundary detection and image segmentation. This process is normally carried out by comparing the results by automated methods with gold standards (often produced by humans), making the comparison operator critical in the process [40,41]. Comparison for quality evaluation was, in fact, one of the primary goals Baddeley in his original work [20], citing *deriving optimal algorithms*, the *numerical benchmark for quantifying the performance of an algorithm* and the *measurement of achieved quality* as potential uses of error measures.

The comparison of boundary images has been rather prolific in past years, and holds strong similarities with related tasks (e.g. silhouette matching and/or tracking). Given the variety of alternatives to produce such comparison, it is relevant to question, and eventually quantify, how good comparison methods perform. This study concerns the so-called metaquality: strategies to measure or enforce the quality of the evaluation methods.

Metaquality is complex to design and interpret; we refer to [42] or [43] for more details on this subject. In our experiments, we focus on a specific study within the BSDS500 [23], a popular dataset for image segmentation and boundary detection which has been used for both benchmarking and metaquality evaluation [44]. This dataset contains, per image (481×321 pixels), a list of 5 to 7 hand-made ground truth images. The images, produced by different humans, are normally divergent. Despite this divergence, any human would be able to cluster all boundary images according to the image after which they were produced. This is illustrated in Fig. 6, which lists the ground truth images associated to three different images in the BSDS500. Within the pool of images in Fig. 6, and despite the large intra-class divergences, humans are able to cluster images in the same class, and discriminate images in different classes. We expect binary image comparison measures to replicate this behaviour. That is, we expect binary image dissimilarity measures (and metrics) to yield larger values when comparing inter-class pairs of images than when comparing intra-class ones. This section intends to measure how BDM performs in this regard, under different configurations, to evaluate the impact of such configuration.

4.2. A global analysis of separability

Our first approach to intra- and inter-class separability is based on the analysis of the distributions of the values yielded in the comparison. In this regard, we have used different versions of BDM to compare all pairs of images in the BSDS500 Test set. This set contains 1063 images in 200 classes, accounting for more than 10^6 comparisons.

For this experiment we have selected $q \in \{-1, 1\}$, so as to represent both the metric and the pseudometric versions of BDM. Also, we consider three different metrics on Ω. First, we take $m \in \{d_5, d_{15}\}$, since 5 and 15 correspond to (approx.) 1% and 3% of the image diagonal. These values cover the typical range in boundary comparison, since they represent the distance beyond which a pixel is considered as *far away from a boundary*, hence having its value at the DT/PDT set to t. Also, we

[2] In Fig. 5 some plots in the rightmost column show a slight decrease of the quantified error w.r.t. the radius increase. This fact is due to the bounded nature of Ω, as the enlarged circumference approaches the image limits.

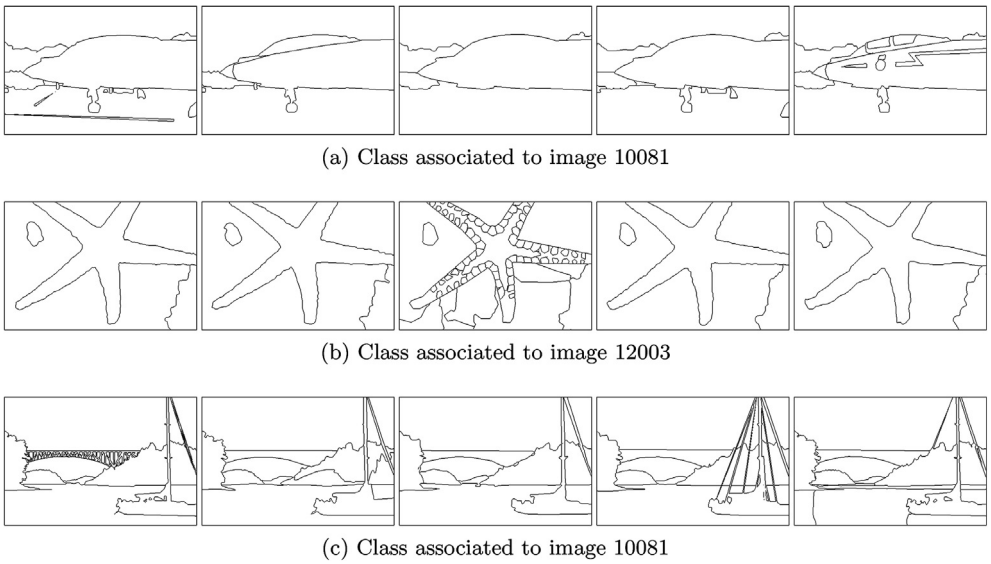

Fig. 6. Different hand-made (ground truth) boundary images extracted from the BSDS500 Test set. Each row corresponds to a different cluster (same-class images).

consider $m = d$ because, in Figs. 1 and 4, the unbounded Euclidean metric seemed to produce a regularization effect on $\mathscr{P}_d^{-1}$ that is worth investigating.

In Fig. 7 we display the distribution of values yielded by different versions of BDM on inter- and intra-class pairs of images. Together with the distributions, we also display on the right axis the resulting accuracy (Acc) in the discrimination of both distributions with each possible threshold. Ideally, if the distributions were non-overlapping, there would be at least one threshold value producing Acc = 1. For the experiment, we have considered different configurations of BDM. In the upper row, we display the results with the original BDM (using $m \in \{d_5, d_{15}, d\}$). Then, the middle and lower rows contain the analogous results when replacing $\mathscr{T}_m$ by $\mathscr{P}_m^{-1}$ and $\mathscr{P}_m^1$.

We recall that distributions in Fig. 7 are displayed in percentual terms, since the number of inter-class comparisons is orders of magnitude greater than those of intra-class ones. In this experiment, there are circa 4500 intra-class comparisons, and over 10^6 inter-class ones. Also, the accuracy (Acc) is computed according to normalized distributions for each class to avoid problems due to class imbalance.

A noticeable fact in Fig. 7 is that the distributions of inter- and intra-class comparisons are fairly separable in the upper row. In a general manner, intra-class comparisons yield significantly lower values than inter-class ones, which indicates a relatively good behaviour of all configurations of BDM. The situation differs when replacing $\mathscr{T}_m$ by $\mathscr{P}_m^q$. As expected from the visual results in Fig. 4, PDTs computed with $q < 0$ and bounded metrics d_t are poor representations of boundary images. Hence, the intra- and inter-class distributions are hardly separable. However, the increasing smoothing effect by $\mathscr{P}_d^{-1}$ partially solves this situation. Regarding the separability with $\mathscr{P}_m^1$, the results are rather opposite to those with $\mathscr{P}_m^{-1}$. The classes are highly separable for $m \in \{d_5, d_{15}\}$. However, the smoothing effect in using d renders in too unspecific distance transforms, leading to poor silhouette recognition.

According to the results in Fig. 7, we have that $\mathscr{P}_m^{-1}$ is a less suitable alternative for the task than $\mathscr{T}_m$ and $\mathscr{P}_m^1$. We relate these results directly to the need for preservation of the property $\mathscr{P}_m^{-1}[I](p) = 0$ if and only if $p \in I$. While this might have a different effect in the recognition of solid shapes, it proves unfit for boundary images. The main problem is the presence of sharp variations in the distance transform at boundary pixels, which severely hamper a sensible comparison by BDM. For boundary comparison, BDM based on $\mathscr{P}_m^{-1}$ hardly reached the performance (Acc) by BDM based on $\mathscr{T}_m$, and did not reach that by $\mathscr{P}_m^1$. Note that, while the peak accuracy can be seen as a scalar evaluation of class separability, it is also important to observe the width and the area under the curve of Acc. In this sense, $\mathscr{P}_m^1$ still produces the best results.

The experiment has been repeated with $k = 2$ (see Fig. 8). While the results and conclusions are analogous to those extracted from Fig. 7, it is also evident that $k = 1$ leads to better results, especially in the settings using PDT.

4.3. A detailed analysis of separability

A more detailed analysis can be made on the basis of the quantitative criteria for class separability in [45]. These criteria can be applied to any dataset with multiple ground truth solutions, as is the case for the BSDS500. The four separability criteria are referred to as weak, moderate, strong and total separability.

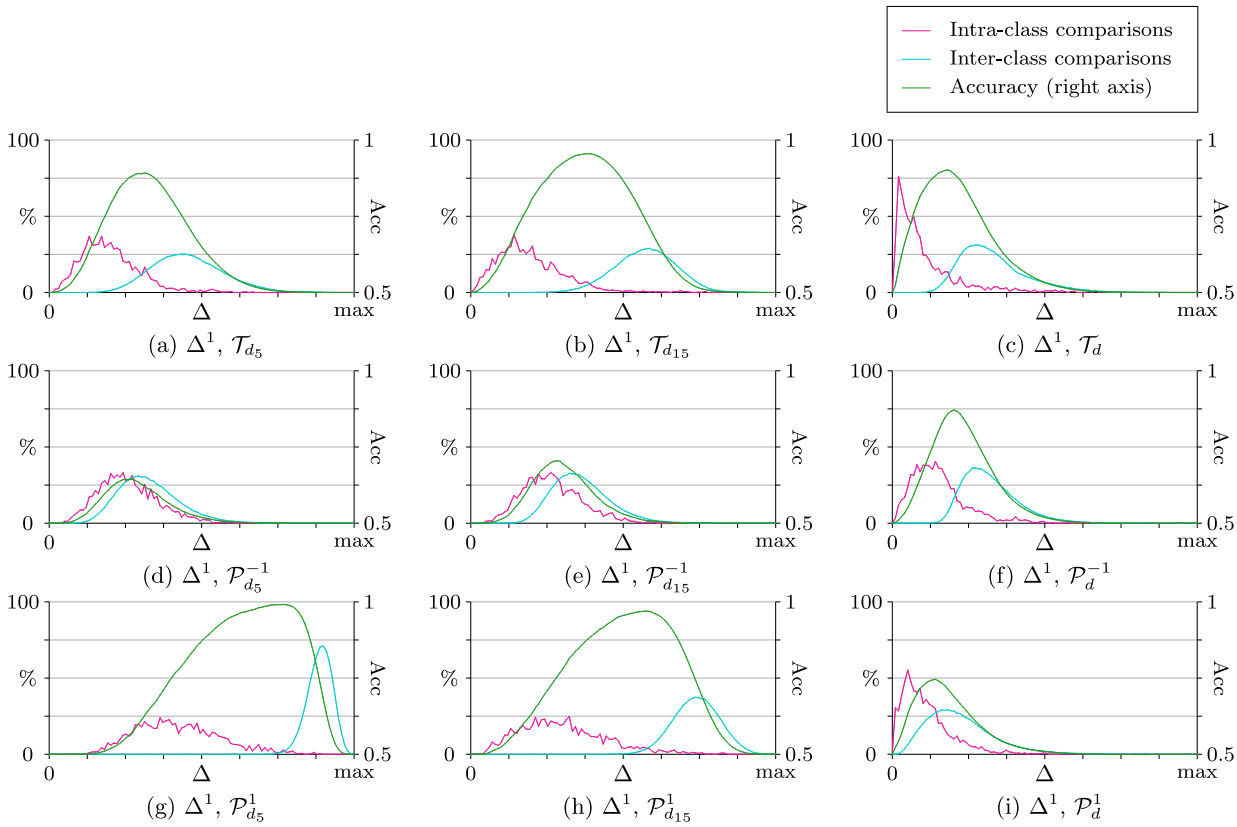

Fig. 7. Distributions of distances for inter- and intra-class pairs of ground truth images in the BSDS500 Test set, using different configurations of BDM. The first rows represents the results by BDM with DTs, while the second and third rows represent the results by BDM when combined with PDTs. Each column features a different base metric for the DT or PDT (d_5, d_{15} or $d = d_\infty$). In all cases, we set $k = 1$. The distributions have been configured with 100 bins.

A dataset for image processing can be modelled as a triplet $\mathbb{D} = (\mathbf{I}, \mathbf{E}, \lambda)$ such that.

- $\mathbf{I} = \{I_1, \ldots, I_k\} \subseteq \mathbb{G}$ is the set of original images in the dataset;
- $\mathbf{E} = \{E_1, \ldots, E_n\} \subseteq \mathbb{B}$ is the set of ground truth images in the dataset;
- $\lambda : \{1, \ldots, n\} \to \{1, \ldots, k\}$ is a mapping such that $\lambda(i) = j$ if the image E_i was created by a human from image I_j.

Let $\mathbb{D}$ be a dataset and q be a metric or dissimilarity measure used to compare the binary images. The four separability criteria are defined as follows.

(S_1) *Weak separability:* The pair $(\mathbb{D}, q)$ is weakly separable if

$$\min_{\substack{\lambda(i) = \lambda(j) \\ i \neq j}} q(E_i, E_j) \leqslant \min_{\lambda(i) \neq \lambda(r)} q(E_i, E_r),$$

for all $i \in \{1, \ldots, n\}$.

(S_2) *Moderate separability:* The pair $(\mathbb{D}, q)$ is moderately separable if

$$\max_{\lambda(i) = \lambda(j)} q(E_i, E_j) \leqslant \min_{\lambda(i) \neq \lambda(r)} q(E_i, E_r),$$

for all $i \in \{1, \ldots, n\}$.

(S_3) *Strong separability:* The pair $(\mathbb{D}, q)$ is strongly separable if

$$\max_{\lambda(i) = \lambda(j) = m} q(E_i, E_j) \leqslant \min_{\substack{\lambda(r) = m \\ \lambda(s) \neq m}} q(E_r, E_s),$$

for all $m \in \{1, \ldots, k\}$.

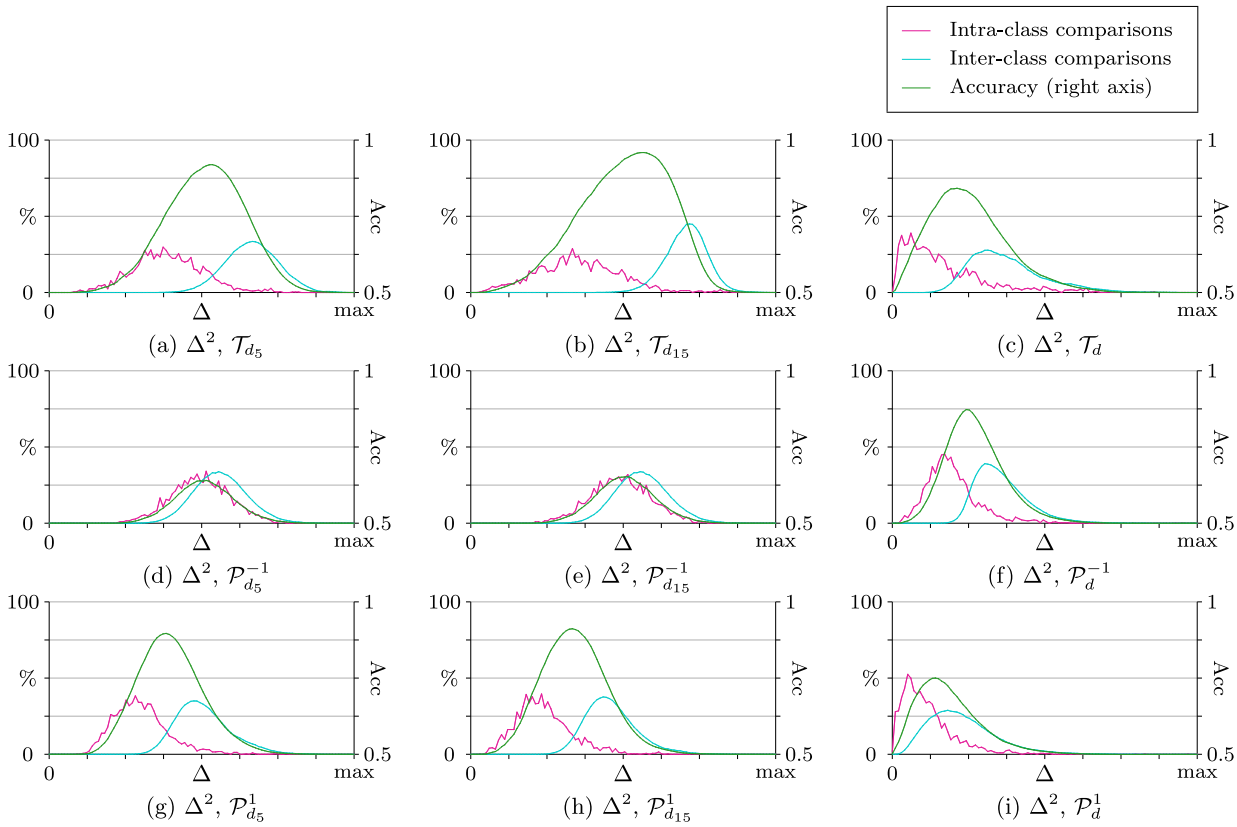

Fig. 8. Distributions of distances for inter- and intra-class pairs of ground truth images in the BSDS500 Test set, using different configurations of BDM. The first rows represents the results by BDM with DTs, while the second and third rows represent the results by BDM when combined with PDTs. Each column features a different base metric for the DT or PDT (d_5, d_{15} or $d = d_\infty$). In all cases, we set $k = 2$. The distributions have been configured with 100 bins.

(S_4) *Total separability:* The pair $(\mathbb{D}, q)$ is totally separable if

$$\max_{\lambda(i)=\lambda(j)} q(E_i, E_j) \leqslant \min_{\lambda(r)\neq\lambda(s)} q(E_r, E_s).$$

The criteria are presented in increasing severity. In this way, for any pair $(\mathbb{D}, q)$, it holds that

$$S_4 \Rightarrow S_3 \Rightarrow S_2 \Rightarrow S_1.$$

The interpretation of the four criteria can be made from the point of view of class recognition. (S_1) checks whether, for each image in a dataset, the *closest* image belongs to the same class. (S_2) discriminates whether, for each image, all intra-class distances are smaller than inter-class ones. (S_3) checks that, for each class, all intra-class distances are smaller than the inter-class ones. Finally, (S_4) is satisfied if and only if all intra-class distances in the dataset are smaller than the inter-class ones.

While S_4 is a Boolean criterion, S_1–S_3 can be expressed as ratios. For example, for S_1 we have the ratio R_1 given by:

$$R_1(\mathbb{D}, q) = \frac{1}{n}\left|\left\{ i \in \{1, \ldots, n\} \mid \min_{\substack{\lambda(i)=\lambda(j)\\ i\neq j}} q(E_i, E_j) \leqslant \min_{\lambda(i)\neq\lambda(r)} q(E_i, E_r) \right\}\right|, \tag{8}$$

which represents the ratio of images $E_i \in \mathbb{B}$ that comply with the criterion. Obviously, $R_i(\mathbb{D}, q) = 1$ if and only if $S_i(\mathbb{D}, q)$.

It is worth noting that any metric or dissimilarity measure reaching an Acc of 1 in the experiments in Section 4.2, would comply with total separability (S_4), and consequently with all other criteria (S_1–S_3).

We have computed the separability ratios R_i on the BSDS500 using BDM combined with the distance transformations considered in Section 4.3. The results are displayed in Fig. 9, presented for $k \in \{1, 2\}$.

In Fig. 9, we observe that all R_i decrease as i increases, which is natural considering the increasing restrictivity of the ratios and the criteria according to which they are computed. The behaviour of BDM, under the various configurations, is in line

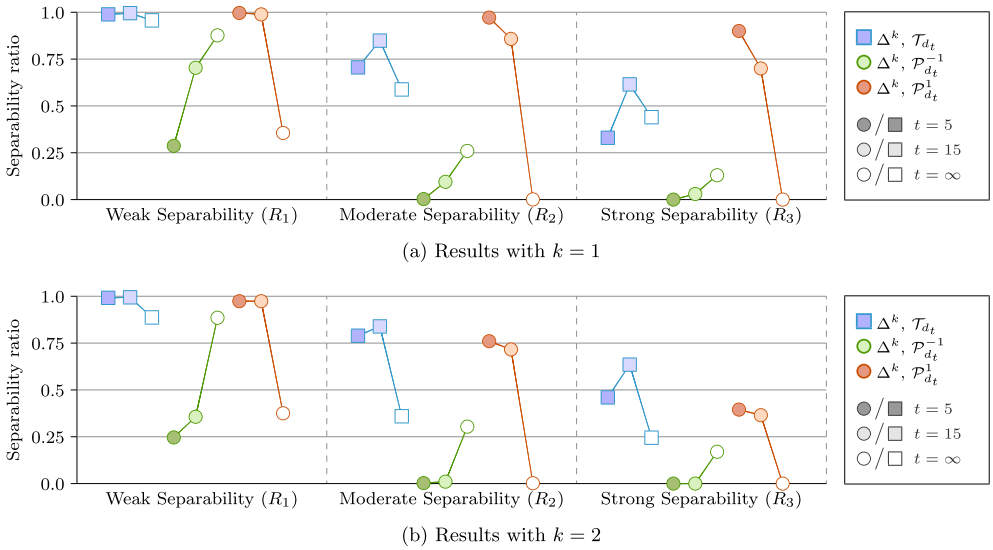

Fig. 9. Separability ratios of different configurations of BDM on the BSDS500 Test set. The configurations of BDM are the ones depicted in Fig. 7 (for $k = 1$) and Fig. 8 (for $k = 2$).

with the results in Figs. 7 and 8. For example, we observe how $\mathscr{T}_m$ and $\mathscr{P}^1_m$ perform better with bounded than unbounded metrics. Also, in terms of overall performance, we observe that $\mathscr{P}^1_m$ reaches a greater separability than $\mathscr{T}_m$ and, especially than $\mathscr{P}^{-1}_m$. Finally, we have that $k = 1$ generally produces a better separability than $k = 2$, which is consistent with the results in Section 4.2.

4.4. A detailed view into inter- and intra-class separation

According to the results in Sections 4.2 and 4.3, BDM based on $\mathscr{P}^1_{d_t}$, with $t \in \{5, 15\}$ offers, in general, a better separability than the other alternatives. We shall analyze an example in which this transformation outperforms its counterparts to understand this fact.

Let (A_1, A_2, B) be a triplet of ground truth images in the BSDS500 Test set, so that A_1 and A_2 belong to the same class, and B belongs to a different one. Ideally, all versions of BDM should yield smaller values for the comparison of A_1 and A_2 than for the comparison of A_1 and B. In Fig. 10, we show an example of such triplet for which, among the configurations in Fig. 9, only BDM based on $\mathscr{P}^1_{d_t}$ succeeds to yield the expected values. Note that, despite the differences in the images, any human would be able to recognize A_1 and A_2 being much closer than A_1 and B.

Fig. 10 displays a triplet of images in the BSDS500, together with some visualizations of distance transforms and the differences between them. As in Figs. 1 and 4, scales have been customized for each DT and PDT, in order to make the results row-wise comparable and interpretable. Note that $t = 25$ is used for a better display of the distance transforms.

In Fig. 10, we firstly observe the problems in the use of $\mathscr{T}_d$, due to the unbounded nature of d. As a consequence, this transformation generates very high values at the upper corners of some images, which subsequently leads to large divergences in columns (d) and (e). Since image B has a spatial distribution similar to image A_1, $\mathscr{T}_d[A_1]$ is in fact more similar to $\mathscr{T}_d[B]$ than to $\mathscr{T}_d[A_2]$. This problem is partially corrected by using $\mathscr{T}_{d_{25}}$ instead of $\mathscr{T}_d$. However, the large number of featured pixels in A_2 leads to large divergences in the comparison of the transforms of A_1 and A_2, as seen in column (d).

In the case of $\mathscr{P}^{-1}_{d_{25}}$, we have a situation similar to that in Fig. 4. Since the distance transform is 0 at the boundary pixels, it results in a crisp interpretation of the images. As can be seen in columns (a)–(c), transforms based on $\mathscr{P}^{-1}_{d_{25}}$ contain little information other than the position of the boundaries. Under these conditions, the comparison of distance transforms is mostly based on exact coincidences of boundaries in the images.

In Fig. 10, the only transformation able to properly model divergences between images is $\mathscr{P}^1_{d_{25}}$. This, however, is not due to a proper modelling of the coincidences and divergences of the images. As we can observe in the lower row of Fig. 10, the key lies in the smoothness of $\mathscr{P}^1_{d_{25}}[A_2]$. Due to the large number of binary pixels all over the image, $\mathscr{P}^1_{d_{25}}$ yields large values even for those pixels that actually contain featured pixels (as those representing trees in the upper region of the image). This happens because, even at those featured positions, the 1-power mean of the distances to all other featured pixels is high. This can be seen in the low values associated with the contour regions in the transforms (columns (a)–(c)). Hence, in the comparison of the transforms of A_1 and A_2 (see column (d)), the discrepancies in the upper area of the image are not as significant as with the other distance transformations. This indicates that $\mathscr{P}^1_{d_{25}}$ succeeds in recognizing A_1 as being closer to A_2 than to B;

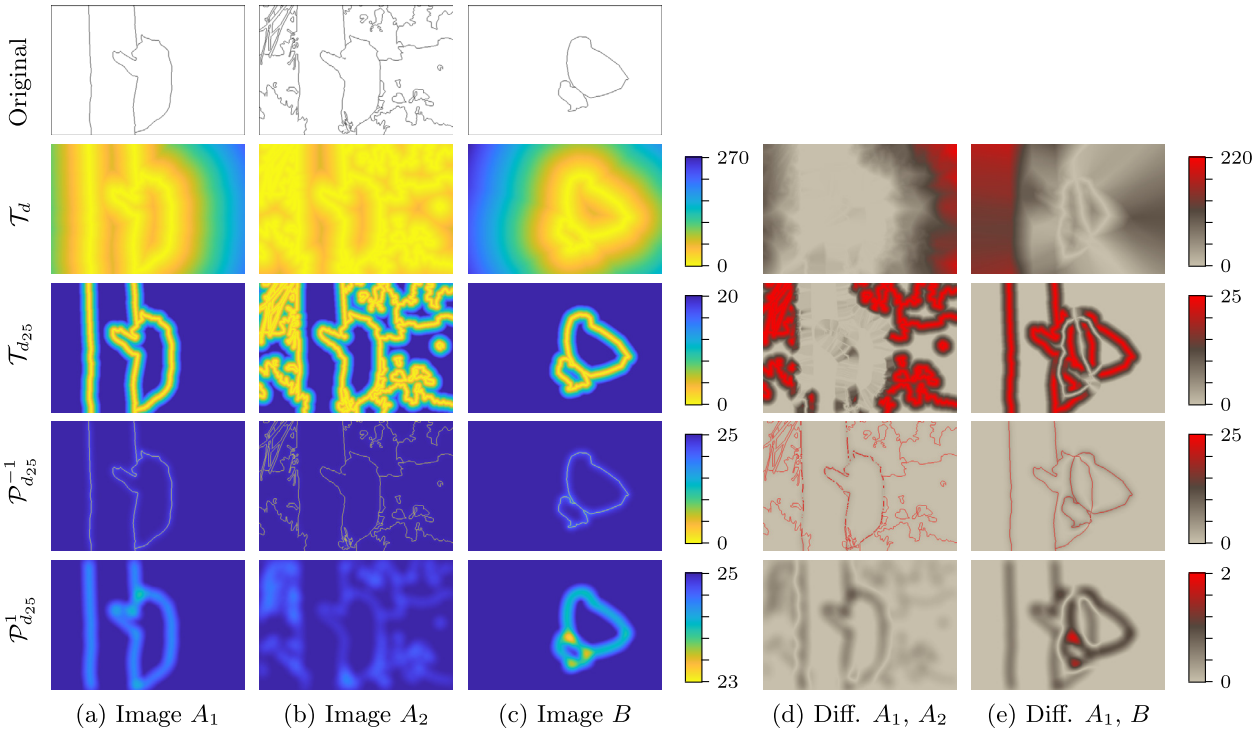

Fig. 10. Example of distance transforms of three images in the BSDS500 Test set. The first row includes the original images. For four different distance transformations, we display (a–c) the distance transforms and the absolute difference between the distance transforms of (d) A_1 and A_2 and (e) A_1 and B. The images have been homogeneously normalized in each row.

also, it indicates that the mechanisms powering such recognition are based on a bare, unspecific representation of the image in the distance transforms, and not on a good modelling of the coincidences. Hence, yet yielding the right results, these transforms need to be carefully understood.

5. Discussion

In this work we have considered the possible parametrizations of Baddeley's Delta Metric (BDM), devoting special attention to the configuration of the underlying distance transformation; in this regard, we have considered both bounded and unbounded distance metrics, as well as the Power Distance Transformation (PDT) by Brunet and Sills [22]. We have performed quantitative and qualitative experiments testing the behaviour, properties and performance of different configurations of the metric.

As a conclusion, we can state that the most relevant setting for BDM is the use of bounded distance transformations. The use of PDTs can be of great use as well, although the use of the parameter q, and the regularization it induces, need to be carefully analyzed and studied in specific applications. In our case, *i.e.* the context of boundary detection, PDT (with $q > 0$) actually outperforms other configurations. Notably, it outperforms PDT with $q < 0$. However, we can also state that PDT (a) involves the fine setting of parameters, specifically the powers k and q, (b) yields transforms whose interpretation is not as straightforward as those by the standard DT, and (c) might require a configuration related to the morphological characteristics of the binary features present in the images to be compared. Although PDT is a fully competitive alternative to DT, its use in a specific context needs to be carefully considered and, eventually, finely configured.

CRediT authorship contribution statement

C. Lopez-Molina: Conceptualization, Methodology, Validation, Formal analysis. **S. Iglesias-Rey:** Software, Validation. **H. Bustince:** Methodology, Formal analysis. **B. De Baets:** Conceptualization, Formal analysis, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the financial support by the Spanish Ministry of Science (project PID2019-108392GB-I00 AEI/FEDER, UE), as well as that by Navarra Servicios y Tecnologías S.A. (NASERTIC).

Appendix A. Efficient implementation of a Power Distance Transformation

The fast implementation of distance transformations has been heavily studied [46,47], and is normally supported by standard pre-coded libraries. However, the implementation of a PDT cannot be done, efficiently, on the basis of such algorithms. Instead, we propose the following implementation.

The PDT of an image I using a bounded metric m_t is given by:

$$\mathscr{P}^q_{m_t}[I](p) = \left(\frac{1}{|I|}\sum_{p'\in I}(m_t(p,p'))^q\right)^{1/q}.$$

Let an image I be understood as the set of its 1-valued positions, $I \subseteq \Omega$. For every given position $p \in \Omega$, we can write $I = I_{\text{near}}(p) \cup I_{\text{far}}(p)$, with $I_{\text{near}}(p)$ representing the pixels in I at a distance strictly smaller than t (in terms of m_t), and I_{far} representing the pixels at distance t.

$$\begin{aligned}
\mathscr{P}^q_{m_t}[I](p) &= \left(\tfrac{1}{|I|}\left(\sum_{p'\in I_{\text{near}}(p)}(m_t(p,p'))^q + \sum_{p'\in I_{\text{far}}(p)}(m_t(p,p'))^q\right)\right)^{1/q} \\
&= \left(\tfrac{1}{|I|}\left(\sum_{p'\in I_{\text{near}}(p)}(m_t(p,p'))^q + \sum_{p'\in I_{\text{far}}(p)} t^q\right)\right)^{1/q} \\
&= \left(\tfrac{1}{|I|}\left(\sum_{p'\in I_{\text{near}}(p)}(m_t(p,p'))^q + \left(|I| - |I_{\text{near}(p)}|\right)\cdot t^q\right)\right)^{1/q}.
\end{aligned}$$

In order to implement this reformulation of the PDT, we define a kernel k_{in}, centered at the origin. Let $\|\cdot\|$ be the norm induced by m_t. At each position (x,y), k_{in} is defined as:

$$k_{\text{in}}(x,y) = \begin{cases} \|(x,y)\|^q & \text{, if } \|(x,y)\| < t \\ 0 & \text{, otherwise} \end{cases}.$$

Also, we define a kernel k_{count} as

$$k_{\text{count}}(x,y) = \begin{cases} 1 & \text{, if } \|(x,y)\| < t \\ 0 & \text{, otherwise} \end{cases}.$$

The PDT of an image I, given the bounded metric m_t, at each pixel $p \in \Omega$ is given by:

$$\mathscr{P}^q_{m_t}[I](p) = \left(\frac{1}{|I|}(\alpha+\beta)\right)^{1/q}$$

with

$$\alpha = (I * k_{\text{in}})(p) \quad \text{and} \quad \beta = (|I| - (I * k_{\text{count}})(p))\cdot t^q.$$

With this implementation, we discriminate the sum of the distances to the nearby pixels (α) and the sum of the distances to the far away pixels (β). The procedure is hence completed with two kernel convolutions.

There are two parameter configurations to be analyzed separately in this strategy.

- In case $q = 0$, the kernel k_{in} cannot be computed. In this case, the q-power mean becomes the geometric mean, but its computation following the above schema will prove erroneous. Two alternatives are available; firstly, since the geometric mean is continuous and monotone w.r.t. q, the 0-power mean can be approximated as the average of the $(-\epsilon)$-power mean and the ϵ-power mean, with ϵ a small enough number. In terms of PDT, we could approximate

  $$\mathscr{P}^0_{m_t}[I] = \frac{1}{2}(\mathscr{P}^{\epsilon}_{m_t}[I] + \mathscr{P}^{-\epsilon}_{m_t}[I]).$$

  Note that this could generate a loss of precision due to the powers being near-zero values. Alternatively, $\mathscr{P}^0_{m_t}[I]$ can be implemented as a geometric mean:

$$\mathscr{P}^0_{m_t}[I](p) = \left(\frac{1}{|I|}\prod_{p'\in I} m_t(p,p')\right)^{1/|I|},$$

which is straightforward using a simplistic, yet inefficient, iteration-based implementation.

- In case $q < 0$, the center position of k_{in} will inevitably yield $0^q = \infty$. This can be solved from two perspectives. Firstly, according to the definition of the power mean, that value shall produce no problem: if the center position is featured in the image ($p \in I$), the power mean will be zero; if it is not featured ($p \notin I$), then the value at the center of k_{in} will be ignored, avoiding any further problem. This was the alternative adopted in our experiments, so as to be faithful to the original proposal by Brunet and Sills [22]. Secondly, if opting for an approximate solution, one can proceed in three manners. The first option is to ignore that value setting it to 0, presuming it is unimportant in the mean of $|I|$ values. A second option is to replace the 0-value produced by the metric m_t in the origin of k_{in} by some small enough ϵ. This can lead to extraordinarily large, unstable values as $\epsilon \to 0$. A third option is to find a numerically sensible solution using simple interpolation in the kernel. For a fixed metric, the partial derivatives can be computed and implemented. In a general manner, a first-order discrete approximation at $(x, y) = (0, 0)$ can be computed as

$$k_{\text{in}}(0,0) = k_{\text{in}}(0,-1) + k_{\text{in}}(0,-1) - k_{\text{in}}(0,-2)$$

or

$$k_{\text{in}}(0,0) = \frac{1}{2}(k_{\text{in}}(0,-1)) + k_{\text{in}}(0,1)).$$

Discrete interpolation and differentation, yet not mathematically sound, have been used in a myriad of works in image processing, e.g. for gradient computation [48,49].

References

[1] P. Rusnok, P. Hurtik, O. Kaláb, D. Musiolek, P. Kočárek, M. Tomis, Data analysis of multivariate time series of insect tracking, in: Joint International Conference on Soft Computing and Intelligent Systems (SCIS) and International Symposium on Advanced Intelligent Systems (ISIS), 2018, pp. 771–776.
[2] G. Wang, G. Van Stappen, B. De Baets, Automated artemia length measurement using U-shaped fully convolutional networks and second-order anisotropic gaussian kernels, Computers and Electronics in Agriculture 168 (2020) 105102.
[3] J. Serra, Image Analysis and Mathematical Morphology, Academic Press Inc, 1983.
[4] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distributions, Pattern Recognition 29 (1) (1996) 51–59.
[5] H. Zhou, R. Wang, C. Wang, A novel extended local-binary-pattern operator for texture analysis, Information Sciences 178 (22) (2008) 4314–4325.
[6] J. Gimenez, J. Martinez, A.G. Flesia, Unsupervised edge map scoring: A statistical complexity approach, Computer Vision and Image Understanding 122 (2014) 131–142.
[7] F. Mokhtarian, A. Mackworth, Scale-based description and recognition of planar curves and two-dimensional shapes, IEEE Transanction on Pattern Analysis and Machine Intelligence (1) (1986) 34–43..
[8] F. Mokhtarian, A.K. Mackworth, A theory of multiscale, curvature-based shape representation for planar curves, IEEE Transactions on Pattern Analysis and Machine Intelligence (8) (1992) 789–805..
[9] K. Bowyer, C. Kranenburg, S. Dougherty, Edge detector evaluation using empirical ROC curves, Computer Vision and Image Understanding 84 (1) (2001) 77–103.
[10] B. Hemery, H. Laurent, B. Emile, C. Rosenberger, Comparative study of localization metrics for the evaluation of image interpretation systems, Journal of Electronic Imaging 19 (2) (2010), 023017/1–023017/21.
[11] M. Segui Prieto, A. Allen, A similarity metric for edge images, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (10) (2003) 1265–1273.
[12] 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 26 (5) (2004) 530–549.
[13] F.J. Estrada, A.D. Jepson, Benchmarking image segmentation algorithms, International Journal of Computer Vision 85 (2) (2009) 167–181.
[14] A.V. Goldberg, R. Kennedy, An efficient cost scaling algorithm for the assignment problem, Mathematical Programming 71 (1995) 153–177.
[15] G. Liu, R.M. Haralick, Optimal matching problem in detection and recognition performance evaluation, Pattern Recognition 35 (10) (2002) 2125–2139.
[16] I. Abdou, W. Pratt, Quantitative design and evaluation of enhancement/thresholding edge detectors, Proceedings of the IEEE 67 (5) (1979) 753–763.
[17] R.M. Haralick, Digital step edges from zero crossing of second directional derivatives, IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (1) (1984) 58–68.
[18] M.-P. Dubuisson, A.K. Jain, A modified Hausdorff distance for object matching, in: Proc. of the IEEE International Conference on Pattern Recognition, vol. 1, 1994, pp. 566–568..
[19] B. Takacs, Comparing face images using the modified Hausdorff distance, Pattern Recognition 31 (12) (1998) 1873–1881.
[20] A.J. Baddeley, An error metric for binary images, in: W. Förstner, S. Ruwiedel (Eds.), Robust Computer Vision: Quality of Vision Algorithms, Wichmann Verlag, Karlsruhe, 1992, pp. 59–78.
[21] A.J. Baddeley, Errors in binary images and an L^p version of the Hausdorff metric, Nieuw Archief voor Wiskunde 10 (1992) 157–183.
[22] D. Brunet, D. Sills, A generalized distance transform: Theory and applications to weather analysis and forecasting, IEEE Trans. on Geoscience and Remote Sensing 55 (3) (2016) 1752–1764.
[23] P. Arbelaez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (2011) 898–916.
[24] G. Beliakov, A. Pradera, T. Calvo, Aggregation Functions: A Guide for Practitioners, Vol. 221 of Studies in Fuzziness and Soft Computing, Springer, 2007..
[25] D.P. Huttenlocher, W.J. Rucklidge, A multi-resolution technique for comparing images using the Hausdorff distance, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 705–706.
[26] D.P. Huttenlocher, K. Kedem, M. Sharir, The upper envelope of Voronoi surfaces and its applications, Discrete & Computational Geometry 9 (3) (1993) 267–291.
[27] D.L. Wilson, A.J. Baddeley, R.A. Owens, A new metric for grey-scale image comparison, International Journal of Computer Vision 24 (1997) 5–17.

[28] D. Coquin, P. Bolon, Application of Baddeley's distance to dissimilarity measurement between gray scale images, Pattern Recognition Letters 22 (14) (2001) 1483–1502.
[29] P.F. Felzenszwalb, D.P. Huttenlocher, Distance transforms of sampled functions, Theory of Computing 8 (1) (2012) 415–428.
[30] J. Öfverstedt, J. Lindblad, N. Sladoje, Stochastic distance transform: Theory, algorithms and applications, Journal of Mathematical Imaging and Vision 62 (5) (2020) 751–769.
[31] R. Yager, Families of OWA Operators, Fuzzy Sets and Systems 59 (2) (1993) 125–148.
[32] D.-G. Sim, O.-K. Kwon, R.-H. Park, Object matching algorithms using robust Hausdorff distance measures, IEEE Transactions on Image Processing 8 (3) (1999) 425–429.
[33] B. Takacs, H. Wechsler, Face recognition using binary image metrics, in: Proc. of the IEEE International Conference on Automatic Face and Gesture Recognition, 1998, pp. 294–299.
[34] D. Coquin, P. Bolon, A. Onea, Objective metric for colour image comparison, Proc. of the European Signal Processing Conference (2000) 1–4.
[35] C. Lopez-Molina, D. Ayala-Martini, A. Lopez-Maestresalas, H. Bustince, Baddeleyś delta metric for local contrast computation in hyperspectral imagery, Progress in Artificial Intelligence 6 (2) (2017) 121–132.
[36] L. Zadeh, Fuzzy sets, Information and Control 8 (1) (1965) 338–353.
[37] L. Zadeh, The concept of a linguistic variable and its application to approximate reasoning (I), Information Sciences 8 (1) (1975) 199–249.
[38] A. Yilmaz, X. Li, M. Shah, Contour-based object tracking with occlusion handling in video acquired using mobile cameras, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (11) (2004) 1531–1536.
[39] F. Zana, J.-C. Klein, A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform, IEEE Transactions on Medical Imaging 18 (5) (1999) 419–428.
[40] T. Peli, D. Malah, A study of edge detection algorithms, Computer Graphics and Image Processing 20 (1) (1982) 1–21.
[41] C. Lopez-Molina, B. De Baets, H. Bustince, Quantitative error measures for edge detection, Pattern Recognition 46 (4) (2013) 1125–1139.
[42] J. Pont-Tuset, F. Marques, Measures and meta-measures for the supervised evaluation of image segmentation, in: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 2013, pp. 2131–2138..
[43] C. Lopez-Molina, B. De Baets, H. Bustince, Twofold consensus for boundary detection ground truth, Knowledge-Based Systems 98 (2016) 162–171.
[44] B. Peng, L. Zhang, X. Mou, M.-H. Yang, Evaluation of segmentation quality via adaptive composition of reference segmentations, IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (10) (2017) 1929–1941.
[45] C. Lopez-Molina, H. Bustince, B. De Baets, Separability criteria for the evaluation of boundary detection benchmarks, IEEE Transactions on Image Processing 25 (3) (2016) 1047–1055.
[46] L. Boxer, R. Miller, Efficient computation of the Euclidean distance transform, Computer Vision and Image Understanding 80 (3) (2000) 379–383.
[47] J. Wang, Y. Tan, Efficient Euclidean distance transform algorithm of binary images in arbitrary dimensions, Pattern Recognition 46 (1) (2013) 230–242.
[48] I. Sobel, G. Feldman, A 3x3 isotropic gradient operator for image processing, presented at a talk at the Stanford Artificial Intelligence Project, 1968.
[49] O. Laligant, F. Truchetet, A nonlinear derivative scheme applied to edge detection, IEEE Transactions on Pattern Analysis and Machine Intelligence 32 (2) (2010) 242–257.