Scaled-distance-transforms and monotonicity of autocorrelations

Jean-Luc Bouchot and Frédéric Morain-Nicolier

Abstract—**In this paper, we investigate the use of distance transforms in a scale space domain for the image (and signal) misalignment problem. We show that it is possible to build an autocorrelation function that is monotonic with respect to the amount of translation. This creates a new paradigm for image comparison and gives yet a new generalization of distance transforms to grey-level images. Its behavior is analyzed on a natural scene image and its robustness against noise is verified numerically.**

Index Terms—**autocorrelation, distance transforms, scale-space, misalignment, Hausdorff measure**

I. Introduction

According to some research in the neuroscience (see for instance Cadieu et.al. [1]), the human brain manages to recognize some objects by a cascade of different detectors with increasing complexity. It is now clear that the neurons are activated in presence of certain fixed stimuli. In particular it seems that edges play an important role in the detection process. Due, however to the scale invariance nature of the human visual system, it is beforehand impossible to tell the scale of an object activating a neuron at a given moment.

On the other side, computer vision scientists have been working intensively on structural similarity measures. A major piece of work is dealing with image quality estimation and yields one of the best measures so far [2]. The authors compared images by separating their resemblances into three components. First a luminance normalization, second a contrast normalization, while the third component seems to have the most impact and is described as being a structural component of the image. More recently, some work have been done on monogenic representations and phase-based imaging for image comparison [3], [4]. It is clear since the original work of [5] that structural information is contained in phases.

It is also understood that edges play a great role in machine recognition. The HOG descriptor [6] gives a description of local neighborhoods only based on weighted edges. Based on this, Shrivastava et. al. [7] recently described how considering the whole environment (i.e. the whole image) for comparison, even though one uses local descriptors, can drastically improve some results.

In this paper, we investigate a new way to look at multivariate real-valued functions by defining a distance transform in a scale space. This approach is particularly suited for the misalignment problem, e.g. for image registration and stereo vision. Unlike most of the similarities used in image comparison and pattern matching [8], our scaled distance transform yields a monotonic behavior when dealing with autocorrelations.

J.-L. Bouchot is with the Department of Mathematics, Drexel University, Philadelphia, PA 19104, e-mail: jean-luc.bouchot@drexel.edu

F. Morain-Nicolier is with the Centre de Recherche en Sciences et Technologies d'Information et Communiations, University of Reims - Champagne-Ardennes, IUT Troyes, France, e-mail: frederic.nicolier@univ-reims.fr

Manuscript received April 1, 2014;

A. Monotonicity of autocorrelation and misalignment

In this note, we are particularly interested in similarity measures designed for the following misalignment problem:

$$t^* := \operatorname{argmin}_{t \in \mathbb{R}^n} J(f_t, g) \tag{1}$$

where

- f and g are two multivariate signals: $f, g : \mathbb{R}^n \to \mathbb{R}$,
- f_t corresponds to the translation of the signal f by a vector t, and
- J corresponds to an objective function. It can be understood as a distance measure or a similarity measure (in which case we would want t^* to maximize this equation).

When $f = g$, we are dealing with an autocorrelation minimization. In order to use local minimization techniques, it is essential for the objective function f to show some kind of monotonicity with respect to the amount of misalignment such as $J(f_{\lambda_1 t}, f) \leq J(f_{\lambda_2 t}, f)$ for all $0 \leq \lambda_1 \leq \lambda_2$. However, as described in [8], most of the usual similarity measures or distance metrics, such as any f-divergences and Minkowski distances, do not fulfil this property. This has led researchers to increase the work on local feature selections. Unfortunately, while SIFT [9] and HOG [6] like features appear robust for natural images, they behave poorly, for instance, in interferometric imaging (such as those obtained by Optical Coherence Tomography). In such cases, purely structure based measures show stronger potential. To the best of our knowledge, only measures based on the discrepancy norm [10] and Hausdorff measures [11].

B. Organization and contribution

In this paper we investigate a new similarity measure in a scale-space that fulfills the above monotonicity criterion. It is based on a combination of distance transforms at different scales. We start in Section II by re-introducing the mandatory background. In particular basics about distance-transforms and scale space are given. Then, we introduce our novel approch in Section III. We derive our *scaled-distance-transform* (SDT) for one dimensional signals and analyze its theoretical properties. Section III-B extends our results to higher dimensional signals. Finally Sections IV and V give some numerical results and draw conclusions.

II. Background review

A. Distance transforms and Hausdorff measures

In his novel work, Baddeley [12] introduced a metric based on the Hausdorff measure to compare two binary images. While his idea was first to compare the efficacy of different edge detection algorithms, it has led to many research in the area of binary pattern matching. In particular it introduced the notion of Local Distance Maps that also show a monotonic behavior with respect to the displacement.

In this section we consider two bounded open sets A and B in an ambient metric space $(\mathbb{R}^n, d)$.

Definition 1 (Distance transform to the foreground)**.** *The distance transform is defined as*

$$\forall \mathbf{x} \in \mathbb{R}^n,\ DT_d[A](\mathbf{x}) = d(\mathbf{x}, A) = \min_{\mathbf{y}\in A}(d(\mathbf{x}, \mathbf{y})). \tag{2}$$

Remark 1: The output of the distance transform might be infinite for $\mathbf{x} \to \pm\infty$. We can restrict the domain of definition of the DT to a bounded domain Ω containing A.

Remark 2: This is generally called the distance to the foreground. A similar definition, with the distance transform to the background may be defined. We will not give any further details about this in the rest of this note. The distance transform to the background is finite as long as A is bounded.

From there, it is easy to define the distance transform of binary functions (i.e. indicator functions). As long as these functions have support on a bounded domain we can identify them with their support and consider the distance transform of their support. We will equivalently write

$$DT_d[A] = DT_d[\chi_A] \tag{3}$$

where χ_A represents the indicator function of the set A. Comparing indicator functions can also be done with the Hausdorff metric [11], [13]. It is defined as follows:

Definition 2 (Hausdorff distance)**.** *Given two bounded sets $A, B \subset (\mathbb{R}^n, d)$, the Hausdorff distance is defined as*

$$H(A, B) = \max(\sup_{\mathbf{x}\in A}(d(\mathbf{x}, B)), \sup_{\mathbf{y}\in B}(d(\mathbf{y}, A))) \tag{4}$$

This Hausdorff distance can however be reformulated in a simpler manner [12]

$$\begin{aligned} H(A, B) &= \sup_{x\in\Omega} |d(x, A) - d(x, B)| \\ &= \sup_{x\in\Omega} |DT_d[A](x) - DT_d[B](x)| \end{aligned} \tag{5}$$

which is in this case implemented by running two distance transform algorithm. Some efficient and robust approximations have been developed that makes this idea tractable [14], [15], [16]. The importance of this distance is justified by the following theorem:

Theorem 1. *Given a non-empty finite set $A \subset \mathbb{R}^n$ and a vector $\mathbf{t} \in \mathbb{R}^n$, it holds [11]:*

$$H(A, T_{\mathbf{t}}A) = \|\mathbf{t}\|$$

where the norm depends on the distance used in the definition above.

As a consequence, we have the monotonicity along a line of the autocorrelation function:

Corollary 1. *Let $\mathbf{t} \in \mathbb{R}^n$ be a vector and $A \subset \mathbb{R}^n$ a non-empty finite set. For any two constants $0 \le \lambda_1 \le \lambda_2$, it holds:*

$$H(A, T_{\lambda_1\mathbf{t}}A) \le H(A, T_{\lambda_2\mathbf{t}}A)$$

B. Scale space

As this area can be very detailed if we want an exhaustive description, we only review the topics needed for this letter: the Gaussian scale space. We refere the curious reader to [17] for more details.

Scale space representations have first been introduced and studied in details in [18] and then used in different computer vision algorithms [9], [6]. It allows to give a description of an image based on its spatial coordinates as well as in terms of scale. Given the well known Gaussian smoothing kernel $g(x, y; s) := \frac{1}{2\pi s}e^{-\frac{x^2+y^2}{2s}}$ depending on the continuous scale parameter $s \in \mathbb{R}^+$, one can define the scale space representation of an image I as, $L(x, y; 0) = I$ and for s as above,

$$L(x, y; s) = (g(\cdot, \cdot; s) * I)(x, y) \tag{6}$$

where $*$ denotes the product of convolution. These ideas have been successfully used in the context of feature detection such as SIFT [9] or HOG [6].

III. Scaled distance transforms

The whole idea of our new approach is to use the scale-space representation to embed the structures of images in a more robust similarity measure that is monotonic with respect to the amount of translation. We first look at the one-dimensional case and deal with higher-dimensional signals in the following section. On the other hand, our approach can be seen as yet another generalization of the distance transform to continuous, or, at least, non-binary signals. However, unlike other generalizations [19], [20], [21], the autocorrelation function based on the Hausdorff distance in a scale space shows the monotonic behavior expected. Our generalization differs also from others by the fact that it extends the idea of edge comparison, as suggested by the original work.

A. One-dimensional signals

Edges are found as zero-crossing of the Laplacian of the scale-space representation: a point $\mathbf{x}$ is said to be an edge at a given scale $s \ge 0$ if the Gaussian regularized derivative reaches an extremum at $\mathbf{x}$. In other words, we want both a non-vanishing first derivative and a zero-crossing of the second one.

In the one-dimensional case, this translates to:

$$\begin{aligned} (\partial_x g_s * f)(\mathbf{x}) &\neq 0, \\ (\partial_{xx} g_s * f)(\mathbf{x}) &= 0. \end{aligned}$$

We denote by $E(f; s)$ the set of edges found in f at scale s. It is defined as

$$\begin{aligned} E(f; s) := &\{\mathbf{x} \in \mathbb{R} : (\partial_{\mathbf{x}} g_s * f)(\mathbf{x}) \neq 0\} \\ &\cap \{\mathbf{x} \in \mathbb{R} : (\partial_{\mathbf{xx}} g_s * f)(\mathbf{x}) = 0\} \end{aligned} \tag{7}$$

This can equivalently be written as $E(f;s) = \left(\partial_{\mathbf{x}} g_s * f\right)^{-1}(\mathbb{R}\backslash\{0\}) \cap \left(\partial_{\mathbf{xx}} g_s * f\right)^{-1}(\{0\})$, where the inverse should be understood in an algebraic way as

$$\mathbf{x} \in \left(\partial_{\mathbf{xx}} g_s * f\right)^{-1}(\{0\}) \Leftrightarrow \left(\partial_{\mathbf{xx}} g_s * f\right)(\mathbf{x}) = 0$$

Definition 3. *Let f be a function defined on $\Omega \subset \mathbb{R}$, open and bounded, with values in $\mathbb{R}$. Its scaled-distance transform is a positive real-valued mapping from $\Omega \times \mathbb{R}^+$:*

$$SDT[f] : \begin{cases} \Omega \times \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ (\mathbf{x}, s) & \mapsto & SDT[f](\mathbf{x}, s) \end{cases} \tag{8}$$

$$SDT[f](\mathbf{x}, s) = \min\{d(\mathbf{x}, \mathbf{y}), \mathbf{y} \in E(f;s)\} \tag{9}$$

B. Higher dimensional signals

We use a very similar approach to define the scaled-distance transform in a multivariate setting. However, here we calculate the distance transform to an edge, which can appear in any direction. We say that an edge located at $\mathbf{x}$ is oriented in the i^{th} direction, $1 \le i \le n$, if

$$\begin{aligned} \mathbf{x} \in E_i(f;s) := \{\mathbf{x} \in \mathbb{R} : \left(\partial_{x_i} g_s * f\right)(\mathbf{x}) \neq 0\} \\ \{\mathbf{x} \in \mathbb{R} : \left(\partial_{x_i x_i} g_s * f\right)(\mathbf{x}) = 0\}. \end{aligned} \tag{10}$$

The set of all edge locations at scale s is defined as

$$E(f;s) = \bigcup_i E_i(f;s). \tag{11}$$

Hence we get the following definition:

Definition 4. *Let f be a function defined on $\Omega \subset \mathbb{R}^n$, open and bounded, with values in $\mathbb{R}$. Its scaled-distance transform is a positive real-valued mapping from $\Omega \times \mathbb{R}^+$:*

$$SDT[f] : \begin{cases} \Omega \times \mathbb{R}^+ & \rightarrow & \mathbb{R}^+ \\ (\mathbf{x}, s) & \mapsto & SDT[f](\mathbf{x}, s) \end{cases} \tag{12}$$

$$SDT[f](\mathbf{x}, s) = \min\{d(\mathbf{x}, \mathbf{y}), \mathbf{y} \in E(f;s)\} \tag{13}$$

C. Comparing SDTs and monotonicity

Given a finite measure ν on $\mathbb{R}^+$ and a distance d on $\mathbb{R}^n$, we suggest to compare two SDT representations with

$$\delta_{SDT}(f, g) = \int_{s \ge 0} \delta\left(SDT[f](\cdot, s), SDT[g](\cdot, s)\right) \mathrm{d}\nu(s). \tag{14}$$

Proposition 1. *Given a function f in $(\Omega \subset \mathbb{R}^n, d)$. Its autocorrelation function $J_H[f](\mathbf{t}) := \delta_{SDT}(f, f_{\mathbf{t}})$ the Hausdorff distance is used as δ is monotonic with the respect to the displacement. For a given $\mathbf{t} \in \mathbb{R}^n$ and to constants $0 \le \lambda_1 \le \lambda_2$, it holds*

$$J_H[f](\lambda_1 \mathbf{t}) \le J_H[f](\lambda_2 \mathbf{t}) \tag{15}$$

Note that this property is stated for functions defined on Ω which means that the objective function J_H needs not to be defined for all $\mathbf{t}$'s. Fortunately, a simple completion of the function with 0's will solve any problems. The proof relies on the following lemma.

Fig. 1. Mansion image with the reference patch

Lemma 1. *Given a function f, the set of its edges at a given scale s is anti-invariant with respect to translations:*

$$E(T_{\mathbf{t}} f; s) = T_{-\mathbf{t}} E(f;s)$$

Proof. To prove the lemma it is sufficient to show that for any $\mathbf{x} \in E(f;s)$, $T_{-\mathbf{t}}\mathbf{x} \in E(T_{\mathbf{t}} f; s)$. The other part of the inequality is obtained with $f = T_{-\mathbf{t}} g$.

Equations (10) and (11) state that $\mathbf{x} \in E(f;s)$ if there exists an $i \in \{1, \ldots, n\}$ such that $\left(\partial_{x_i} g_s * f\right)(\mathbf{x}) \neq 0$ and $\left(\partial_{x_i x_i} g_s * f\right)(\mathbf{x}) = 0$. In particular we have

$$\begin{aligned} \int_{\tau \in \mathbb{R}^n} \partial_{x_i x_i} g_s(\mathbf{x} - \tau) f(\tau) \mathrm{d}\tau = 0, \\ \int_{\mathbf{u} \in \mathbb{R}^n} \partial_{x_i x_i} g_s(\mathbf{x} + \mathbf{t} - \mathbf{u}) f(\mathbf{u} - \mathbf{t}) \mathrm{d}\mathbf{u} = \\ \int_{\mathbf{u} \in \mathbb{R}^n} \partial_{x_i x_i} g_s(T_{-\mathbf{t}}\mathbf{x} - \mathbf{u}) T_{\mathbf{t}} f(\mathbf{u}) \mathrm{d}\mathbf{u} = 0. \end{aligned}$$

A similar result is obtained regarding the first derivative. Hence $T_{-\mathbf{t}}\mathbf{x} \in E(T_{\mathbf{t}}; s)$ which concludes the proof.

□

Proof. (of the proposition)

With the use of the previous lemma, the proof of Proposition 1 is rather straightforward. Indeed, for a fixed $s \in \mathbb{R}^+$, for two positive constants $0 \le \lambda_1 \le \lambda_2$ and a translation vector $\mathbf{t} \in \mathbb{R}^n$, we have that $H(SDT[f](\cdot;s), SDT[T_{\lambda_1 \mathbf{t}} f](\cdot;s)) \le H(SDT[f](\cdot;s), SDT[T_{\lambda_2 \mathbf{t}} f](\cdot;s))$; due to Corrollary 1 and the definition 4 of the SDT combined with the previous lemma.

Hence, by definition of a positive finite measure, we have that

$$\begin{aligned} \int_{s \ge 0} H(SDT[f](\cdot;s), SDT[T_{\lambda_1 \mathbf{t}} f](\cdot;s)) \mathrm{d}\nu(s) \le \\ \int_{s \ge 0} H(SDT[f](\cdot;s), SDT[T_{\lambda_2 \mathbf{t}} f](\cdot;s)) \mathrm{d}\nu(s), \end{aligned}$$

which is the claimed result. □

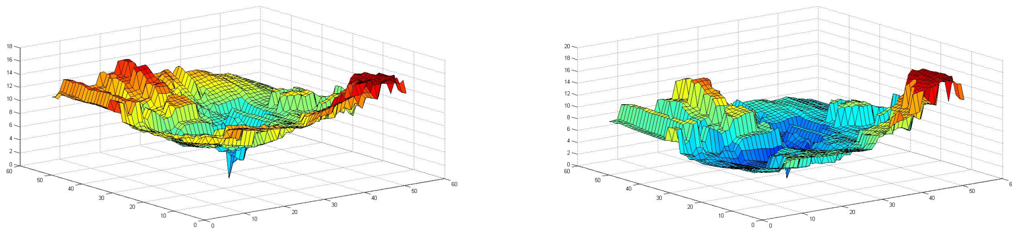

Fig. 2. Discrepancies of the translated pattern in the noise free case. x and y axis corresponds to the displacement of the window that is compared to the black box from Fig. 1.

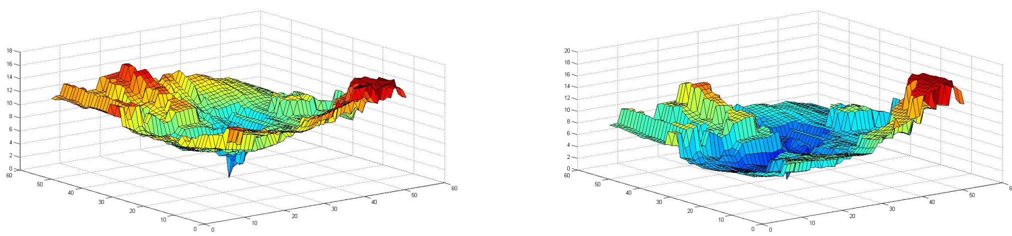

Fig. 3. Discrepancies of the translated pattern in the noisy case. x and y axis corresponds to the displacement of the window that is compared to the black box from Fig. 1.

IV. Numerical results

A. Monotonicity

To analyze the monotonicity property of the SDT, we compare a sliding patch around the black box from the mansion pattern, Figure 1. This patch corresponds to a 51×51 pixel area. We compare its SDT with other patches in its neighborhoord in the image with translations from -25 pixels to $+25$ pixels in all directions. [1]

For the numerical experiments we need to specify how to deal with the finite measure specified in the combination of the Hausdorff distances at every level. We suggest three ideas but it is evident that prior knowledge on the geometry of the scene may help the design of better ones. We consider:

- $\nu = \chi_S$, $S \subset \mathbb{R}^+$; in this case, only a subset of scales is given importance (we do not use this approach further)
- $\nu = 1$; in this case we give the same emphasis on all scales without further considerations
- $\nu(s) = Cs$; here we give more emphasize on objects at coarser scales. It is motivated by the fact that noise and unwanted details will most likely appear in fine scales.

Moreover we apply the SDT on discrete images and therefore use the usual Gaussian pyramid as a scale space. The results of the local discrepancy can be seen in Figure 2. It is important to notice that the monotonicity along a line is not valid due to the fact that we consider "real" patches or a real image and thus we are not dealing with exact autocorrelation. However, it is still locally true and the surface shows few local minima.

B. Noise robustness

It is well known that the Hausdorff distance is however rather unstable when facing noise. This is due to the $\max$ in its definition. To overcome this problem, it has been suggested to average the contribution of all the pixels via the use of a p-mean:

$$H(A,B) \approx \left(\frac{\sum_{x \in X} |d(x,A) - d(x,B)|^p}{|X|} \right)^{1/p} \tag{16}$$

And as p gets bigger this approximation gets better but more sensitive to noise. For this series of tests we have considered the black box reference window as noise free and have added Gaussian noise to the sliding window. The dissimilarity maps can be seen in Figure 3, where we have $p = 16$.

[1]Note that all the data for this note can be obtained from the authors' websites: http://www.math.drexel.edu/ jb3455/publi.html or http://pixel-shaker.fr/

V. Conclusion

We have introduced an approach that generalizes the classical distance transform to non-binary signals based on a scale space representation. Compared to other generalizations, the autocorrelation function based on our scaled distance transform shows a monotonic behavior with respect to the misalignment. This generates new ideas for signal and image registration.

The idea to use the scale space for edge detection is not new, it can be understood as the building blocks of wavelet analysis, for instance. However, the monotoncity property of the autocorrelation function described above may motivate new research in image alignments based on wavelet representations and/or multiresolution analysis.

Acknowledgment

This work was initiated while J.-L. B was with the Department of Knowledge-based Mathematical Systems, at the Johannes Kepler university of Linz, supported by the Austrian Science Fund (FWF) grant number P21496 N23. He is now supported by the National Science Foundation under the agreement DMS 1120622.

REFERENCES

[1] C. Cadieu, M. Kouh, A. Pasupathy, C. E. Connor, M. Riesenhuber, and T. Poggio, "A model of V4 shape selectivity and invariance," *Journal of Neurophysiology*, vol. 98, no. 3, pp. 1733–1750, 2007.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.

[3] S. Held, M. Storath, P. Massopust, and B. Forster, "Steerable wavelet frames based on the Riesz transform," *Image Processing, IEEE Transactions on*, vol. 19, no. 3, pp. 653–667, 2010.

[4] S. Bernstein, J.-L. Bouchot, M. Reinhardt, and B. Heise, "Generalized analytic signals in image processing: Comparison, theory and applications," in *Quaternion and Clifford Fourier transform and Wavelets*, ser. Trends in Mathematics. Springer Basel, 2013, pp. 221–246.

[5] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, pp. 886–893.

[7] A. Shrivastava, T. Malisiewicz, A. Gupta, and A. A. Efros, "Data-driven visual similarity for cross-domain image matching," in *ACM Transactions on Graphics (TOG)*, 2011, p. 154.

[8] B. A. Moser, J.-L. Bouchot, and G. Stübl, "On a non-monotonicity effect of similarity measures," in *Similarity-Based Pattern Recognition*. Springer, 2011, pp. 46–60.

[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[10] B. A. Moser, "A similarity measure for image and volumetric data based on Hermann Weyl's discrepancy," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 11, pp. 2321–2329, 2011.

[11] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 9, pp. 850–863, 1993.

[12] A. J. Baddeley, "An error metric for binary images," in *Proceedings, International Workshop on Robust Computer Vision*. Wichmann, 1992, pp. 59–78.

[13] D. P. Huttenlocher and W. J. Rucklidge, "A multi-resolution technique for comparing images using the Hausdorff distance," in *Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93., 1993 IEEE Computer Society Conference on*, 1993, pp. 705–706.

[14] U. Montanari, "A method for obtaining skeletons using a quasi-Euclidean distance," *Journal of the ACM (JACM)*, vol. 15, no. 4, pp. 600–624, 1968.

[15] G. Borgefors, "Distance transformations in digital images," *Computer vision, graphics, and image processing*, vol. 34, no. 3, pp. 344–371, 1986.

[16] ——, "On digital distance transforms in three dimensions," *Computer Vision and Image Understanding*, vol. 64, no. 3, pp. 368–376, 1996.

[17] T. Lindeberg, "Scale-space for discrete signals," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 12, no. 3, pp. 234–254, 1990.

[18] ——, "Scale-space theory: A basic tool for analyzing structures at different scales," *Journal of Applied Statistics*, vol. 21, pp. 225–270, 1994.

[19] D. L. Wilson, A. J. Baddeley, and R. A. Owens, "A new metric for grey-scale image comparison," *International Journal of Computer Vision*, vol. 24, pp. 5–17, 1997.

[20] D. Coquin and P. Bolon, "Application of Baddeley's distance to dissimilarity measurement between gray scale images," *Pattern Recognition Letters*, vol. 22, pp. 1483–1502, 2001.

[21] I. S. Molchanov and P. Terán, "Distance transforms for real-valued functions," *Journal of Mathematical Analysis and Applications*, vol. 278, pp. 472–484, 2003.