



**HAL**  
open science

# A Study of Measures for Contour-based Recognition and Localization of Known Objects in Digital Images

Baptiste Magnier, Hasan Abdulrahman

► **To cite this version:**

Baptiste Magnier, Hasan Abdulrahman. A Study of Measures for Contour-based Recognition and Localization of Known Objects in Digital Images. IPTA 2018 - Image Processing Theory, Tools and Applications, Nov 2018, Xi'an, China. 10.1109/IPTA.2018.8608165 . hal-01942291

**HAL Id: hal-01942291**

**<https://hal.science/hal-01942291>**

Submitted on 17 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A Study of Measures for Contour-based Recognition and Localization of Known Objects in Digital Images

Baptiste Magnier<sup>1</sup> and Hasan Abdulrahman<sup>2</sup>

<sup>1</sup> IMT Mines d'Alès, LGI2P, 6. avenue de Clavières 30100 Alès, France

<sup>2</sup>Northern Technical University, Department of Technical Computer systems, Kirkuk 36001, Iraq  
e-mail: baptiste.magnier@mines-ales.fr ; hasan.abdulrahman@mines-ales.fr

**Abstract**—Usually, the most important structures in an image are extracted by an edge detector. Once extracted edges are binarized, they represent the shape boundary information of an object. For the edge-based localization/matching process, the differences between a reference edge map and a candidate image are quantified by computing a performance measure. This study investigates supervised contour measures for determining the degree to which an object shape differs from a desired position. Therefore, several distance measures are evaluated for different shape alterations: translation, rotation and scale change. Experiments on both synthetic and real images exhibit which measures are accurate enough for an object pose or matching estimation, useful for robot task as to refine the object pose.

**Keywords**—Distance transform, similarity measures, contours.

## I. INTRODUCTION AND CONTEXT

Object detection and recognition are used in many computer vision applications. In the past, several methods have been developed to achieve this in digital images, involving, for example, directly shape context [1], patches [2], points of interest [3] or region based-methods [4]. The shape representation of an object is particularly useful for accurate industrial inspection tasks, where this shape must be aligned with a reference model of the interested object [5]. A contour-based representation remains a class of methods and exploits only the shape boundary information. In this paper, we focus on distance measures between the acquired features (contours [6]) in a candidate image and an ideal contour map model. This represents a supervised evaluation of the shape representation. As the shape of the sought object is already known (by a reference pose, for example), a learning stage is not necessary to estimate the object pose. Consequently, supervised edge detection evaluations compute a score between a ground truth edge map ( $G_t$ ) and a candidate image ( $D_c$ ) to achieve this task. Originally, they are used to assess edge detection methods [7],[8],[9],[11]; here, a contour-based localization evaluation of these measures is investigated. Thus, they compute a score of coherence which qualifies the correct object pose possibility.

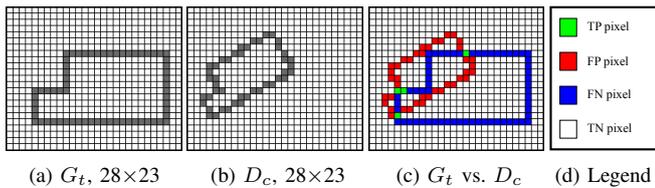


Fig. 1. Example of translation, rotation and scale change of  $D_c$ .

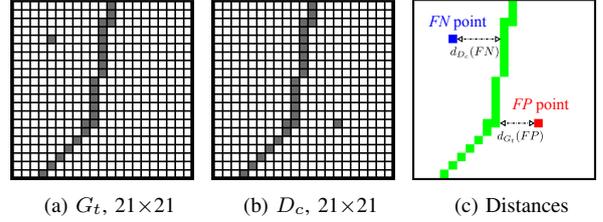


Fig. 2. Example of ground truth ( $G_t$ ) versus (vs.) a candidate contour ( $D_c$ ).

## II. SUPERVISED EDGE DETECTION EVALUATIONS

Several alterations may interfere and disturb the object pose estimation, as translation, rotation or scale change of the interest object. Both their own shape(s) and their contours are altered, as illustrated in Fig.1. Supervised contour measures are presented below. Then, an evaluation process is fulfilled, determining the degree to which an object shape differs from a desired position as a function of different alterations.

Various evaluation methods have been proposed in the literature to assess different shapes of edges using pixel-based ground truth (see reviews in [12],[7],[8],[9],[11]). Indeed, a supervised evaluation criterion computes a dissimilarity measure between a ground truth ( $G_t$ ) and a detected contour map ( $D_c$ ) of an original image  $I$ . In this paper, the closer to 0 the score of the evaluation is, the more the segmentation is qualified as suitable. To assess an edge detector, the confusion matrix remains a cornerstone in boundary detection evaluation methods. Comparing pixel per pixel  $G_t$  and  $D_c$ , the 1st criterion to be assessed is the common presence of edge/non-edge points. A basic evaluation is composed of statistics by combining  $G_t$  and  $D_c$ . Afterwards, denoting  $|\cdot|$  as the cardinality of a set (e.g.  $|G_t|$  represents the number of edge pixels in  $G_t$ ), all points are divided into four sets (cf. Figs.1):

- True Positive points (TPs):  $TP = |G_t \cap D_c|$ ,
- False Positive points (FPs):  $FP = |\neg G_t \cap D_c|$ ,
- False Negative points (FNs):  $FN = |G_t \cap \neg D_c|$ ,
- True Negative points (TNs):  $TN = |\neg G_t \cap \neg D_c|$ .

Various edge detection evaluations involving confusion matrices have been developed, cf. [8][9][11]. The *Dice* measure [13] is one example:  $Dice^*(G_t, D_c) = 1 - \frac{2 \cdot TP}{2 \cdot TP + FN + FP}$ . This type of assessment is useful for region segmentation evaluation [16], but, a reference-based edge map quality measure requires that a displaced edge should be penalized in function of FPs and/or FNs and of the distance from the position where it should be located [9][11], as shown in Fig. 2.

Table 1. List of error measures involving distances, generally:  $k = 1$  or  $k = 2$ , and,  $\kappa = 0.1$  or  $\kappa = 1/9$ .

Error measure name	Formulation	Parameters
Pratt's FoM [14]	$FoM(G_t, D_c) = 1 - \frac{1}{\max( G_t ,  D_c )} \cdot \sum_{p \in D_c} \frac{1}{1 + \kappa \cdot d_{G_t}^2(p)}$	$\kappa \in ]0; 1]$
FoM revisited [15]	$F(G_t, D_c) = 1 - \frac{1}{ G_t  + \beta \cdot FP} \cdot \sum_{p \in G_t} \frac{1}{1 + \kappa \cdot d_{D_c}^2(p)}$	$\kappa \in ]0; 1]$ and $\beta \in \mathbb{R}^+$
Combination of FoM and statistics [17]	$d_4(G_t, D_c) = \frac{1}{2} \cdot \sqrt{\frac{(TP - \max( G_t ,  D_c ))^2 + FN^2 + FP^2}{(\max( G_t ,  D_c ))^2} + FoM^2(G_t, D_c)}$	$\kappa \in ]0; 1]$
Edge map quality measure [18]	$D_p(G_t, D_c) = \frac{1/2}{ I - G_t } \cdot \sum_{p \in FP} \left(1 - \frac{1}{1 + \kappa \cdot d_{G_t}^2(p)}\right) + \frac{1/2}{ G_t } \cdot \sum_{p \in FN} \left(1 - \frac{1}{1 + \kappa \cdot d_{TP}^2(p)}\right)$	$\kappa \in ]0; 1]$
Symmetric FoM [9]	$SFoM(G_t, D_c) = \frac{1}{2} \cdot FoM(G_t, D_c) + \frac{1}{2} \cdot FoM(D_c, G_t)$	$\kappa \in ]0; 1]$
Maximum FoM [9]	$MFoM(G_t, D_c) = \max(FoM(G_t, D_c), FoM(D_c, G_t))$	$\kappa \in ]0; 1]$
Yasnoff measure [19]	$\Upsilon(G_t, D_c) = \frac{100}{ I } \cdot \sqrt{\sum_{p \in D_c} d_{G_t}^2(p)}$	None
Hausdorff distance [20]	$H(G_t, D_c) = \max\left(\max_{p \in D_c}(d_{G_t}(p)), \max_{p \in G_t}(d_{D_c}(p))\right)$	None
Maximum distance [12]	$f_2d_6(G_t, D_c) = \max\left(\frac{1}{ D_c } \cdot \sum_{p \in D_c} d_{G_t}(p), \frac{1}{ G_t } \cdot \sum_{p \in G_t} d_{D_c}(p)\right)$	None
Distance to $G_t$ [21][12][8]	$D^k(G_t, D_c) = \frac{1}{ D_c } \cdot \sqrt[k]{\sum_{p \in D_c} d_{G_t}^k(p)}$ , $k = 1$ for [21] and [12]	$k \in \mathbb{R}^+$
Oversegmentation measure [22]	$\Theta(G_t, D_c) = \frac{1}{FP} \cdot \sum_{p \in D_c} \left(\frac{d_{G_t}(p)}{\delta_{TH}}\right)^k$	for [22]: $k \in \mathbb{R}^+$ and $\delta_{TH} \in \mathbb{R}_*^+$
Undersegmentation measure [22]	$\Omega(G_t, D_c) = \frac{1}{FN} \cdot \sum_{p \in G_t} \left(\frac{d_{D_c}(p)}{\delta_{TH}}\right)^k$	for [22]: $k \in \mathbb{R}^+$ and $\delta_{TH} \in \mathbb{R}_*^+$
Relative Distance Error [12], [23], [24]	$RDE_k(G_t, D_c) = \sqrt[k]{\frac{1}{ D_c } \cdot \sum_{p \in D_c} d_{G_t}^k(p)} + \sqrt[k]{\frac{1}{ G_t } \cdot \sum_{p \in G_t} d_{D_c}^k(p)}$ ,	$k \in \mathbb{R}^+$ , $k = 1$ for [12], $k = 2$ for [23], [24]
Symmetric distance [12][8]	$S^k(G_t, D_c) = \sqrt[k]{\frac{\sum_{p \in D_c} d_{G_t}^k(p) + \sum_{p \in G_t} d_{D_c}^k(p)}{ D_c \cup G_t }}$ , $k = 1$ for [12]	$k \in \mathbb{R}^+$
Baddeley's Delta Metric [25]	$\Delta^k(G_t, D_c) = \sqrt[k]{\frac{1}{ I } \cdot \sum_{p \in I}  w(d_{G_t}(p)) - w(d_{D_c}(p)) ^k}$	$k \in \mathbb{R}^+$ and a convex function $w: \mathbb{R} \rightarrow \mathbb{R}$
Magnier <i>et al.</i> measure [26]	$\Gamma(G_t, D_c) = \frac{FP+FN}{ G_t ^2} \cdot \sqrt{\sum_{p \in D_c} d_{G_t}^2(p)}$	None
Complete distance measure [9]	$\Psi(G_t, D_c) = \frac{FP+FN}{ G_t ^2} \cdot \sqrt{\sum_{p \in G_t} d_{D_c}^2(p) + \sum_{p \in D_c} d_{G_t}^2(p)}$	None
$\lambda$ measure [28]	$\lambda(G_t, D_c) = \frac{FP+FN}{ G_t ^2} \cdot \sqrt{\sum_{p \in D_c} d_{G_t}^2(p) + \min( G_t ^2, \frac{ G_t ^2}{TP^2}) \cdot \sum_{p \in G_t} d_{D_c}^2(p)}$	None
Edge Mismatch Measure (EMM) [27]	$EMM(G_t, D_c) = 1 - \frac{TP}{TP + \omega \cdot \left[\sum_{p \in FN} \delta_{D_c}(p) + \epsilon \cdot \sum_{p \in FP} \delta_{G_t}(p)\right]}$ $\delta_{D_c}(p) = \begin{cases} d_{D_c}(p), & \text{if } d_{D_c}(p) < M_{dist} \\ D_{max}, & \text{otherwise} \end{cases}$ and $\delta_{G_t}(p) = \begin{cases} d_{G_t}(p), & \text{if } d_{G_t}(p) < M_{dist} \\ D_{max}, & \text{otherwise.} \end{cases}$	$M_{dist} \in \mathbb{R}^+$ , $D_{max} \in \mathbb{R}^+$ , $\omega \in \mathbb{R}^+$ , $\epsilon \in \mathbb{R}^+$ . In [27]: $M_{dist} = 0.025 \cdot  I $ , $D_{max} =  I /10$ , $\omega = 10/ I $ , $\epsilon = 2$ .

Table 1 reviews the most relevant measures involving distances. Thus, for a pixel  $p$  belonging to the candidate contour  $D_c$ ,  $d_{G_t}(p)$  represents the minimal Euclidian distance between  $p$  and  $G_t$ . These types of distance measures play an important role in image matching and may be used to determine the degree of resemblance between two objects [20][12]. To achieve this, if  $p$  belongs to  $G_t$ ,  $d_{D_c}(p)$  corresponds to the minimal distance between  $p$  and  $D_c$ , Fig. 2 illustrates the difference between  $d_{G_t}(p)$  and  $d_{D_c}(p)$ . Mathematically, denoting  $(x_p, y_p)$  and  $(x_t, y_t)$  the pixel coordinates of two points  $p$  and  $t$  respectively, thus  $d_{G_t}(p)$  and  $d_{D_c}(p)$  are described by:

$$\begin{cases} \text{for } p \in D_c: d_{G_t}(p) = \inf \left\{ \sqrt{(x_p - x_t)^2 + (y_p - y_t)^2}, t \in G_t \right\}, \\ \text{for } p \in G_t: d_{D_c}(p) = \inf \left\{ \sqrt{(x_p - x_t)^2 + (y_p - y_t)^2}, t \in D_c \right\}. \end{cases}$$

These distance functions refer to the Euclidean distance.

On the one hand, some distance measures are specified in the evaluation of over-segmentation (i.e., distances of FPs), for example:  $\Upsilon$ ,  $D^k$ ,  $\Theta$  and  $\Gamma$ ; see also [12], [9], [11]. On the other hand, the  $\Omega$  measure assesses an edge detection by computing only under-segmentation (distances of FNs). Other

edge detection evaluation measures consider both distances of FPs and FNs [10]. A perfect segmentation using an over-segmentation measure could be an image including no edge points. As demonstrated in [11], another limitation of only over- and under-segmentation evaluations are that several binary images can produce the same result. Therefore, a complete and optimum edge detection evaluation measure should combine assessments of both over- and under-segmentation, as  $H$ ,  $\Delta^k$ ,  $f_2d_6$ ,  $S^k$ ,  $RDE_k$ ,  $\Psi$  and  $\lambda$ .

One of the most popular descriptors is named the Figure of Merit (FoM). This distance measure has an advantage because it ranges from 0 to 1, where 0 corresponds to a perfect segmentation [14]. Nonetheless, for FoM, the distance of the FNs is not recorded and contours having a small displacement compared to their desired positions are strongly penalized as statistic measures (detailed in [11]). Several evaluation measures are derived from FoM:  $F$ ,  $d_4$ ,  $EMM$ ,  $MFoM$ ,  $SFoM$  and  $D_p$ . Contrary to FoM, the  $F$  measure computes the distances of FNs but not of the FPs. In addition, the  $d_4$

measure depends particularly on  $TP$ ,  $FP$ ,  $FN$  and  $\approx 1/4$  on  $FoM$ , but  $d_4$  penalizes FNs like the  $FoM$  measure. Note that  $EMM$  computes a score different from 1 if there exists at least one TP (cf. experiments). Otherwise,  $SFoM$  and  $MFoM$  take into account both distances of FNs and FPs, so they can compute a global evaluation of a contour image. However,  $MFoM$  does not consider FPs and FNs at the same time, contrary to  $SFoM$ . Another way to compute a global measure is represented by  $D_p$ . Nevertheless,  $D_p$  is more sensitive to FNs than FPs because of its huge coefficient (cf. Table 1).

A second measure widely computed in matching techniques is represented by the Hausdorff distance  $H$ , measuring the mismatch of two sets of points [20]. This measure is useful in object recognition, the algorithm aims to minimize  $H$ , reporting the mismatch of two shapes [29][30]. It is well known that this distance could be strongly deviated by only one pixel positioned sufficiently far from the pattern, there are several enhancements of  $H$  presented in [20][12][31]. As pointed out in [12], an average distance from the edge pixels in  $D_c$  to those in  $G_t$  is more appropriate, like  $S^k$ ,  $RDE_k$  or  $\Psi$ . Moreover,  $S^k$  takes small values in the presence of low level of outliers, whereas the score becomes large as the level of mistaken points increases [12][8] but is sensitive to remote misplaced points [9]. Otherwise, derived from  $H$ , the Delta Metric ( $\Delta^k$ ) [25] intends to estimate the dissimilarity between each element of two binary images, but is highly sensitive to distances of misplaced points. The  $\lambda$  measure penalizes highly FNs compared to FPs, because the more FNs are present in  $D_c$ , the more  $D_c$  the desirable object becomes unrecognizable.

Finally, an objective assessment of these measures (and others) has been performed by varying the thresholds on contours of images obtained by filtering techniques, see [7][9][10], [28][11]. Theoretically, the minimum score of the measure corresponds to the best edge map, compared to the ground truth. Here, a study is led as a function of the position of the candidate object contours which may be corrupted by undesirable pixels caused by object translation, rotation of scale change (or/and due to noise and/or blur in real images).

### III. EVALUATIONS AND EXPERIMENTAL RESULTS

Object localization, recognition and matching are practical tools in the computer vision domain. Usually, they are designed to meet the requirements of industrial alignment applications (automated optical inspection, industrial automation, target search, visual servoing...). The distance measures detailed above are compared here with the specific purpose of contour-based object localization evaluation. Such measures are tested on synthetic and real image in order to assess if they could be useful and accurate enough to recognize and match objects. The two subsections above present both experiments on synthetic and real images. 22 error measure are tested:  $Dice^*$ ,  $FoM$ ,  $F$ ,  $d_4$ ,  $MFoM$ ,  $SFoM$ ,  $D_p$ ,  $\Upsilon$ ,  $H$ ,  $f_2d_6$ ,  $D_{k=2}^k$ ,  $\Theta$ ,  $\Omega$ ,  $RDE_{k=1}$ ,  $RDE_{k=2}$ ,  $S_{k=1}^k$ ,  $S_{k=2}^k$ ,  $\Delta^k$ ,  $\Gamma$ ,  $\Psi$ ,  $\lambda$  and  $EMM$ , showcasing the advantages and drawbacks of each one. The tests carried out in the experiments are

intended to be as complete as possible, and thus as close as possible to reality. Note that the Matlab code of the distance measures as  $FoM$ ,  $D^k$ ,  $S^k$  and  $\Delta^k$  are available at <http://kermitimagetoolkit.net/library/code/>. Others are available on the mathworks website: <https://fr.mathworks.com/matlabcentral/fileexchange/63326-objective-supervised-edge-detection-evaluation-by-varying-thresholds-of-the-thin-edges>. Firstly, to verify if the measure has the required properties, different alterations are applied to create synthetic localization results simulating real ones.

#### A. Localization and Recognition of Synthetic Objects

In order to quantify the reliability of a measure of dissimilarity, an edge map of a synthetic shape is affected by various alterations: rotation, translation and scale change. Thus, it allows to verify if the evolution of the provided score obtained by a measure is in line with the expected behavior: small error for close shapes and large penalty for shapes that differ. Note that there exist matching techniques imposing grid of specific window on the image, typically in order to search small objects in the image [20][32][33] or line tackers [34]. Here the distance measures consider both full images of the same size: the ground truth (with an isolated object) and the candidate object contour. For the quantification of similarity, special consideration should be given to the ability of a skill measure to penalize important dissimilarities and ignore the unimportant ones; as an example few spurious isolated undesirable pixels must not disturb a measure if the object is close to its desired location. To do that, a synthetic shape is created (see Fig. 3(a)). Candidate shapes are also created to simulate movements, thereupon compared with the desired location (ground truth). For assessment, a validation measure is sought so that it can appropriately penalize differences in the three continuous deformations: 1) rotation, 2) translation and 3) scale. Thus, a curve is obtained for each measure and each alteration type. Only some curves are reported in this paper, curves which are not reported are similar to other ones.

1) *Rotation*: The test is performed by rotating the control shape incrementally up to a total rotation of  $2\pi$ , as illustrated in Fig. 3. Then, the metric distances are computed between the rotated contour and the reference shape. The shape of the curve of the measure scores is expected to be roughly symmetric around  $\pi$ . It is noteworthy that locations of local extrema of the curve depend on the shape of the object. Usually, a minimum is computed for distance measures at  $180^\circ$ , due the rough symmetry of the shape, as for  $H$ . For example, in Fig. 3(d), some contours of the rotated shape are aligned with the desired location, creating two minima at  $160^\circ$  and  $200^\circ$  for the  $\lambda$  measure. So, this measure penalizes strongly FNs, so it computes an error at  $180^\circ$  larger than before or after, because various details of the shape are missing. However, the  $Dice$  measure computes a pixel per pixel score and does not record distances, so its behavior is unstable. Also,  $FoM$  measure and its derivatives are also unstable and very sensitive to small rotations, minima are not clearly detectable.

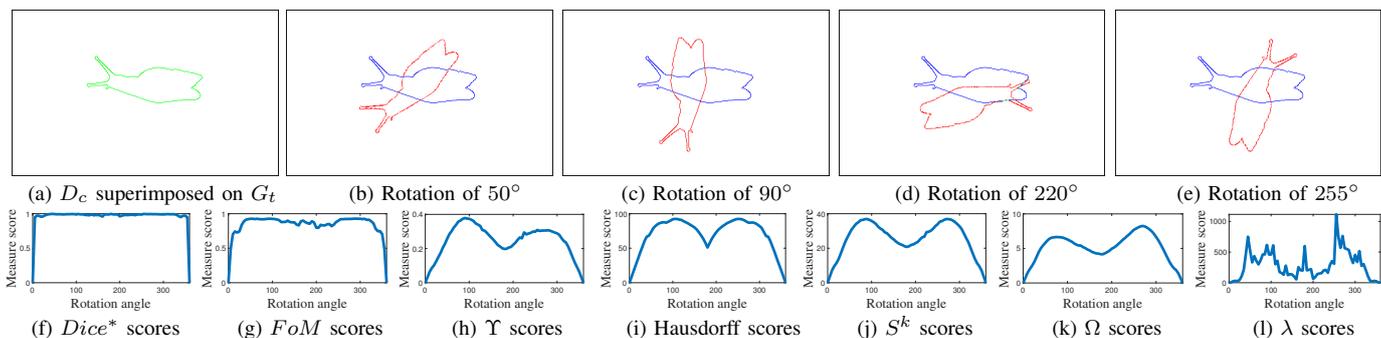


Fig. 3. Examples of localization metrics behaviors for a rotation alteration for each  $5^\circ$ : from  $0^\circ$  to  $360^\circ$ . The legend is available in Fig. 1(d).  $FoM$ ,  $F$ ,  $SFoM$  and  $MFoM$  obtain almost the same scores. Score evolutions for are the same for  $Dice^*$ ,  $d_4$  and  $D_p$ . Also  $\Upsilon$  behaves like  $D^k$ ,  $\Gamma$ ,  $\Psi$  and  $\Theta$  which is the perfect symmetry of  $\Omega$ . Scores tied to  $S^k_{k=\{1,2\}}$ ,  $RDE_{k=\{1,2\}}$ ,  $f_2d_6$ ,  $\Delta^k$  and  $\Psi$  are also similar.  $EMM$  behaves like  $\lambda$ .

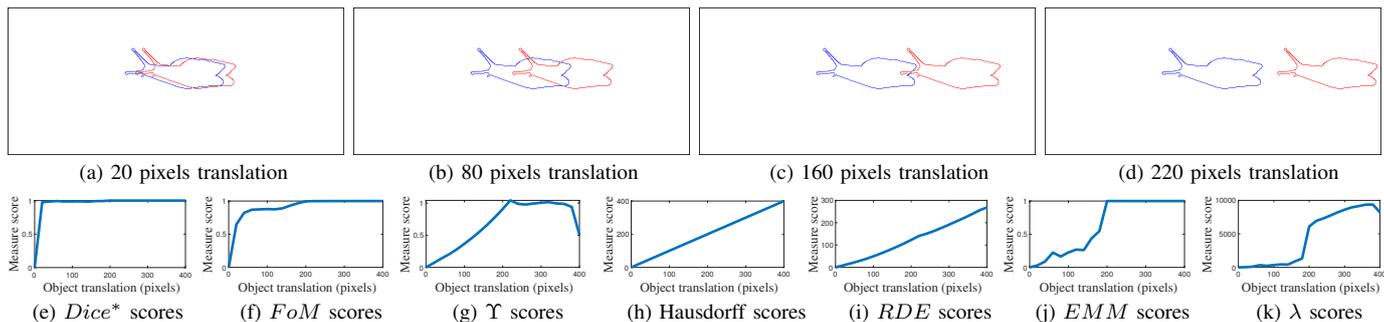


Fig. 4. Examples of localization metrics behaviors for a translation alteration: 20 pixels on the X-axis (horizontal) per translation.  $FoM$ ,  $F$ ,  $SFoM$  and  $MFoM$  obtain almost the same scores. Score evolutions for are the same for  $Dice^*$ ,  $d_4$  and  $D_p$ . Also  $\Upsilon$  behaves like  $D^k$ ,  $\Gamma$  and  $\Psi$ . Scores tied to  $H$ ,  $f_2d_6$ ,  $\Omega$  and  $\Delta^k$  are also similar. Score evolutions for are the same for  $S^k_{k=\{1,2\}}$  and  $RDE_{k=\{1,2\}}$ .

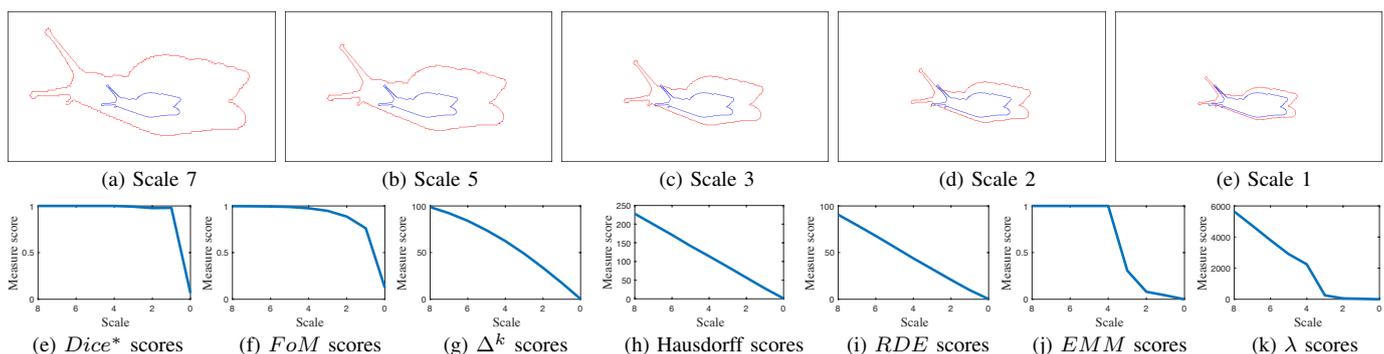


Fig. 5. Examples of localization metrics behaviors for a scale alteration.  $FoM$ ,  $F$ ,  $SFoM$  and  $MFoM$  obtain almost the same scores. Score evolutions for are the same for  $Dice^*$ ,  $d_4$  and  $D_p$ . Scores tied to  $H$ ,  $RDE_{k=\{1,2\}}$ ,  $S^k_{k=\{1,2\}}$ ,  $f_2d_6$ ,  $D^k$ ,  $\Upsilon$ ,  $\Theta$ ,  $\Omega$ ,  $\Gamma$  and  $\Psi$  are also similar.

2) *Translation*: In this test, the contour shape is gradually translated by moving the shape away from its original location along a horizontal straight line, cf. Fig.4 (the results are the same for translations in other directions). The resulting measure scores have a minimum at 0, where the two contours are collocated. Scores obtained for measures as  $H$  and  $RDE$  grows linearly, as a function of the shape displacement. Concerning  $Dice$  or  $EMM$ , the obtained curves have obvious discontinuities, demonstrating the limited sensitivity of these measures to translation.  $EMM$  obtains a score of 1 when no TP exists. Also,  $FoM$  and its derivatives are sensitive to displacements, with a jump after few pixel displacements and with score attaining its maximum (value 1) when no common points between the two shapes. The over segmentation of  $\Upsilon$  is not monotonous when not TP are present. Other measures behave desirably, with a good response: increasing monoton-

ically. Note that the  $\lambda$  measure decreases at the end of this experiment because there exists less pixels in the candidate shape than in the image containing all the object contours.

3) *Scale Chngement*: For this experiment, the shape undergoes a decrease in size with the maximal scale 8th the original size (see Fig.5). Scores from the measures are computed between the ground truth and the rescaled shape. For all shapes, excepted  $Dice$ , each measure exhibits the expected behavior of being strictly decreasing. However, the graph of  $EMM$  has sharp discontinuities that exhibit its unstable response to scaling, because its responses are 1 without TP. The scores of the  $FoM$  and its derivatives are not decreasing enough, excepted where the two contours are collocated. Finally,  $\lambda$  behaves correctly converging through 0 until a scale of 3, and continue to converge less rapidly after.

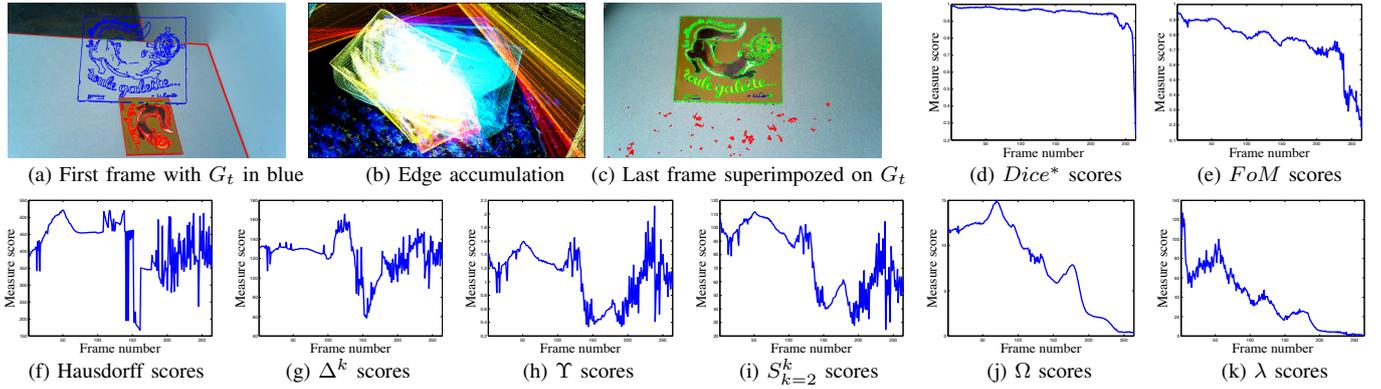


Fig. 6. Localization metrics behaviors on real experiment. The image in (b) represents the edge movements as a function of the time (from blue to red).  $FoM$ ,  $F$ ,  $SFoM$  and  $MFoM$  obtain almost the same scores which are close to  $Dice^*$ ,  $d_4$  and  $D_p$  evolution. Scores tied to  $H$ ,  $\Delta^k$ ,  $\Theta$ ,  $\Gamma$ ,  $\Psi$ ,  $\Upsilon$ ,  $D^k$ ,  $S^k_{k=2}$ ,  $RDE_{k=1}$ ,  $RDE_{k=2}$  and  $f_2d_6$  are similar. Score evolutions are the same for  $S^k_{k=1}$ ,  $EMM$  and  $\lambda$ .

### B. Localization and Recognition of Real Objects

Experiments on color real images were also performed. In Figs. 6 and 8, thin edges are extracted using the Canny edge detector ( $\sigma = 1$ ) [6][36] and a non-maximum suppression: the selected pixels are those having gradient magnitude at a local maximum along the gradient direction, which is perpendicular to the edge orientation [35]. Here, by moving the camera, the aim is to determine when the object is at the desired position in the image using thin binary edges as features, the scores must converge to 0 as summarized in Fig. 7. This desired position corresponds to the object in the last video frame, as shown in Figs. 6(c) and 8(c), each frame may be corrupted by numerous FPs and the candidate object may contain FNs. The first video contains 264 frames, whereas the second video is composed of 116 frames of size  $720 \times 1280$ . The ground truth corresponds to binary boundaries of the desired position of the object, represented in blue pixels in Figs. 6(a) and 8(a). Green pixels represent TPs, red points are FPs whereas blue pixels which are also  $G_t$  are tied to FNs. These features are dilated with a structural element of size  $3 \times 3$  for a better visualization.

For the first video, object contours are easily extracted, with spurious undesirable points at the end of the video, as illustrated in Fig. 6(a)-(c). Consequently, only edge points detected out of the candidate object shape may disturb the contour-based localization. Indeed, curves tied to  $H$ ,  $\Delta^k$ ,  $\Upsilon$  and  $S^k$  behave stochastically (see the caption of Fig. 6 concerning other measures). Also,  $Dice$  and  $FoM$  scores

converge only to lower values when the candidate object is close to the desired location. Only  $\Omega$  and  $\lambda$  measures have the intended behavior for this sequence, as summarized in Tab.2.

The second video is heavily corrupted by a random noise on the each color plane ( $SNR \approx 11dB$ ). These disturbances create spurious pixels in the edge detection process, but especially, the edges of the candidate object are not well localized or not present. Therefore, the majority of measures behaves stochastically, and the last frame does not represent the best score for them (see Fig.8 and its caption). For the  $FoM$  measure, usually its scores decrease, even though the curve converges considerably only for the last frames. On the  $EMM$  measure converges rapidly, but, remains constant after few number of frames. The  $\Omega$  measure is relatively constant, excepted for the last frames where it decreases. Lastly, the  $\lambda$  measure behaves as expected with a minimum at the end.

### IV. CONCLUSION

This study presents an assessment of measures for contour-based recognition and localization of known objects. Thus, the  $Dice$  and 20 distance measures are evaluated through different shape alterations: translation, rotation and scale change. Experiments on real images exhibit which measures are accurate enough for an object pose or shape matching estimation. Clearly, a measure involving false negative distances remains more accurate than other techniques, as the under-segmentation measure  $\Omega$ . However, when some parts of the candidate image are missing, but detected close to their desired positions, they are not taken into account by  $\Omega$ . Alternatively, the  $\lambda$  measure remains well defined for this assessment. Missing edges are penalized higher than spurious

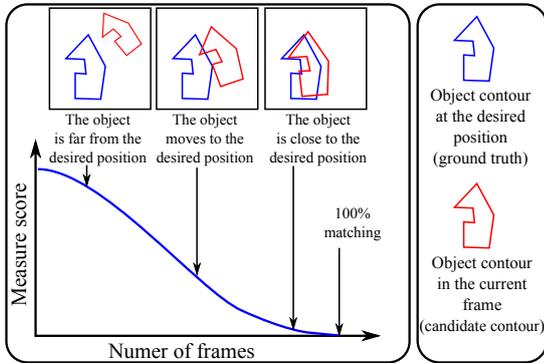


Fig. 7. Expected behavior of a measure scores.

Table 2. Reliability of the reviewed edge detection evaluation measures.

Meas.	Tr.	Rot.	Sc.	Real	Meas.	Tr.	Rot.	Sc.	Real
$Dice^*$	X	X	X	X	$D^k$	$\approx$	$\checkmark$	$\checkmark$	X
$FoM$	X	X	X	X	$\Theta$	$\checkmark$	$\checkmark$	$\checkmark$	X
$F$	X	X	X	X	$\Omega$	$\checkmark$	$\checkmark$	$\checkmark$	$\approx$
$d_4$	X	X	X	X	$RDE$	$\checkmark$	$\checkmark$	$\checkmark$	X
$SFoM$	X	X	X	X	$S^k$	$\checkmark$	$\checkmark$	$\checkmark$	X
$MFoM$	X	X	X	X	$\Delta^k$	$\checkmark$	$\checkmark$	$\checkmark$	X
$D_p$	X	X	X	X	$\Gamma$	$\approx$	$\checkmark$	$\checkmark$	X
$\Upsilon$	$\approx$	$\checkmark$	$\checkmark$	X	$\Psi$	$\approx$	$\checkmark$	$\checkmark$	X
$H$	$\checkmark$	$\checkmark$	$\checkmark$	X	$\lambda$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$
$f_2d_6$	$\checkmark$	$\checkmark$	$\checkmark$	X	$EMM$	$\approx$	$\checkmark$	X	X

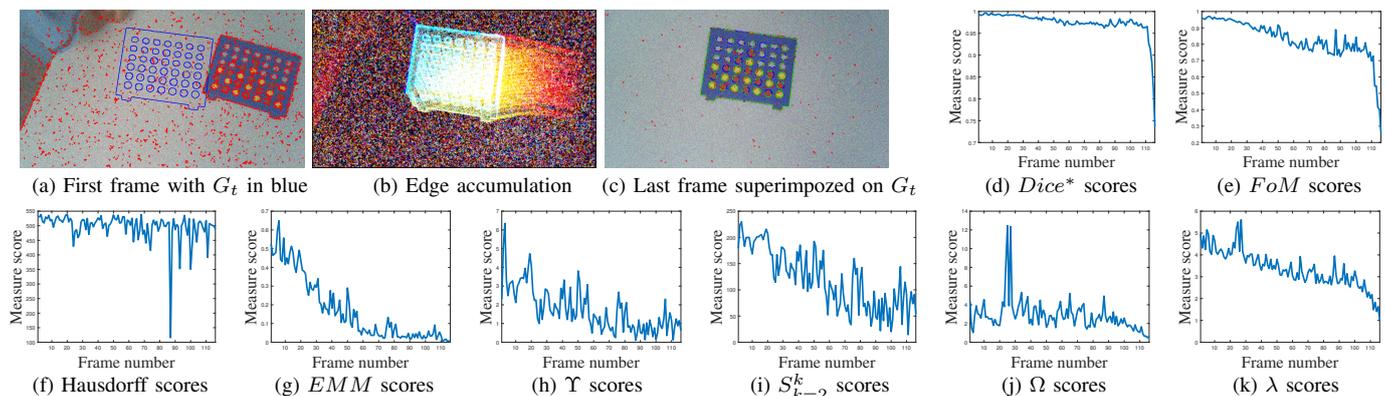


Fig. 8. Localization metrics behaviors on real experiment with noisy images. The image in (b) represents the edge movements as a function of the time (from blue to red), showing the huge number of noise pixels along the video.  $FoM$ ,  $F$ ,  $SFoM$  and  $MFoM$  obtain almost the same scores which are close to  $Dice^*$ ,  $d_4$  and  $D_p$  evolution. The measure  $\Delta^k$  behaves like the Hausdorff distance, stochastically along the video without convergence. Also  $\Upsilon$  behaves like  $D^k$ ,  $\Theta$ ,  $S_{k=2}^k$ ,  $RDE_{k=1}$ ,  $RDE_{k=2}$  and  $f_2d_6$ . Scores tied to  $\Psi$ ,  $\Gamma$ ,  $EMM$  and  $S_{k=1}^k$  are also similar.

ones because some isolated points disturb the shape location, as the majority of the measures, cf. Tab.2. Finally,  $\lambda$  does not need any tuning parameter or specific window and may be useful to validate a visual servoing process.

The studied measures are tied to edge binarized images (edge detection methods are not discussed, see [9], [10], [11]). In practice, illumination changes modify seriously the image contrast and, additionally to the noise and the blur, may be an obstacle for shape recognition. To overcome this problem, the combination of edge-based measures and nonlinear illumination changes measure [37] could be deeply investigate.

#### REFERENCES

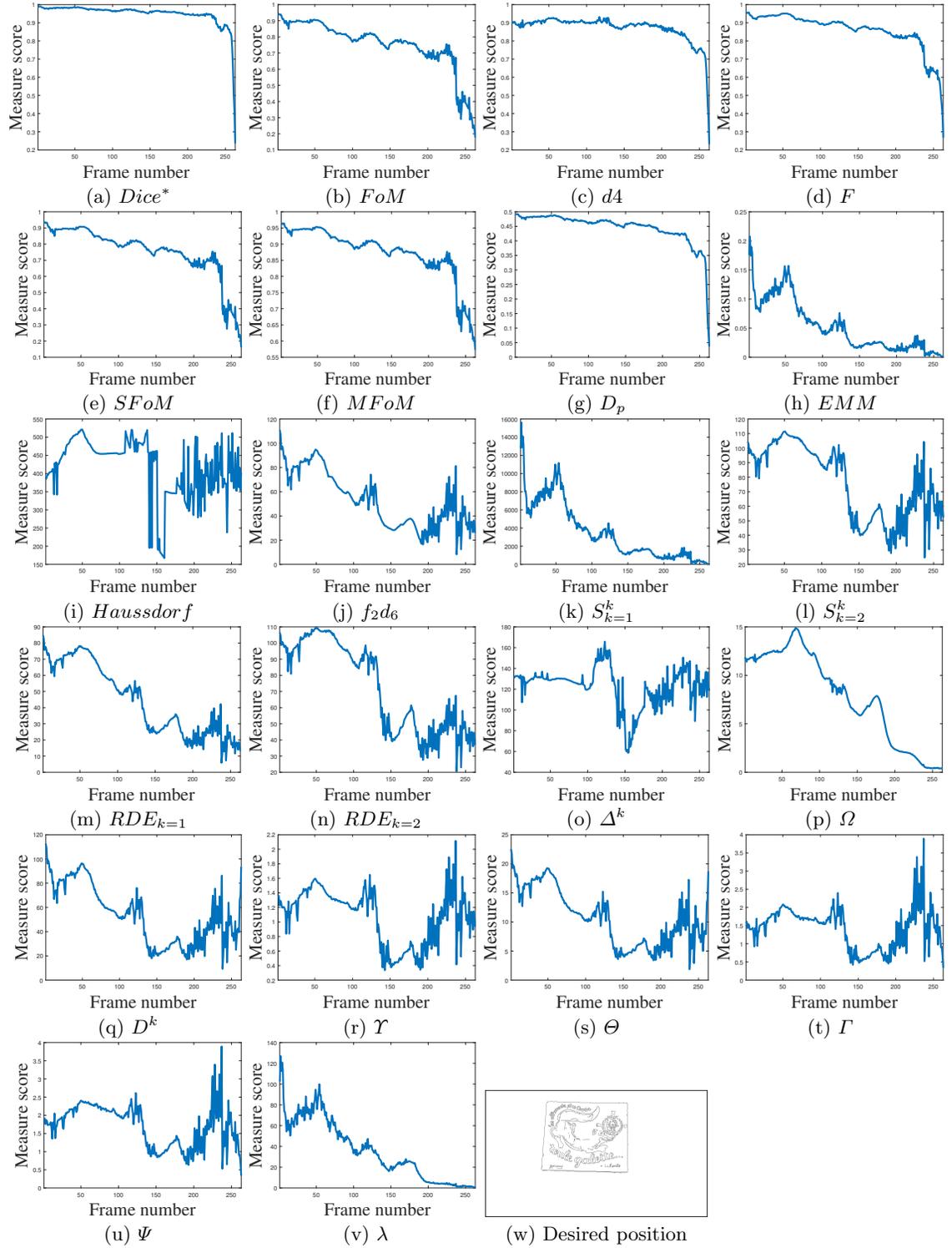
- [1] S. Belongie, J. Malik and J. Puzicha. Shape context: A new descriptor for shape matching and object recognition. In *Advances in neural information processing systems*, 831–837, 2001.
- [2] A. Crivellaro, M. Rad, Y., Verdie, K.M. Yi, P. Fua, and V. Lepetit. Robust 3d object tracking from monocular images using stable parts. *IEEE TPAMI*, 1465–1479, 2017.
- [3] C. Schmid, R. Mohr, and C. Bauckhage. Evaluation of interest point detectors. *International Journal of computer vision*, 37(2):151–172, 2000.
- [4] D. Zhang and G. Lu. Review of shape representation and description techniques. *Pattern recognition*, 37(1):1–19, 2004.
- [5] C. Steger, M. Ulrich and C. Wiedemann. Machine vision algorithms and applications/ *John Wiley & Sons*, 2018.
- [6] J. Canny. A computational approach to edge detection. *IEEE TPAMI*, 6:679–698, 1986.
- [7] S. Chabrier, H. Laurent, C. Rosenberger and B. Emile. Comparative study of contour detection evaluation criteria based on dissimilarity measures. *EURASIP J. Image Video Process*, 2008:2, 2008.
- [8] C. Lopez-Molina, B., De Baets and H. Bustince. Quantitative error measures for edge detection. *Patt. Rec.*, 46:1125–1139, 2013.
- [9] Magnier, B. Edge detection: A review of dissimilarity evaluations and a proposed normalized measure. *Multimed. Tools Appl.*, 77:1–45, 2017.
- [10] H. Abdulrahman, B. Magnier, and P. Montesinos. From contours to ground truth: How to evaluate edge detectors by filtering. *Journal of WSCG*, 25:133–142, 2017.
- [11] B. Magnier, H. Abdulrahman and P. Montesinos A Review of Supervised Edge Detection Evaluation Methods and an Objective Comparison of Filtering Gradient Computations Using Hysteresis Thresholds *J. Imaging*, 4(6):74, 2018.
- [12] M.P. Dubuisson and A.K. Jain. A modified Hausdorff distance for object matching. *IEEE ICPR*, 1:566–568, 1994.
- [13] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26:297–302, 1945.
- [14] I.E. Abdou, and W.K. Pratt. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proc. IEEE 1979*, 67:753–763.
- [15] A.J. Pinho, and L.B. Almeida. Edge detection filters based on artificial neural networks. In *ICIAP*, 159–164, 1995.
- [16] W.R. Crum, O. Camara, Oscar and D.L.G. Hill. Generalized overlap measures for evaluation and validation in medical image analysis. *IEEE Trans. on Medical Imaging*, 25(11):1451–1461, 2006.
- [17] A.G. Boaventura, and A. Gonzaga. Method to evaluate the performance of edge detector. In *Brazilian Symp. on Comput. Graph. Image Process; Citeseer: State College, PA, USA*, 234–236, 2006.
- [18] K. Panetta, C. Gao, S. Agaian, and S. Nercessian. A New Reference-Based Edge Map Quality Measure. *IEEE Trans. Syst. Man Cybern. Syst.*: 46:1505–1517, 2016.
- [19] W. Yasnoff, W. Galbraith, and J. Bacus. Error measures for objective assessment of scene segmentation algorithms. *Anal. Quant. Cytol.*, 1:107–121, 1978.
- [20] D. Huttenlocher, and W. Rucklidge. A multi-resolution technique for comparing images using the hausdorff distance. In *IEEE CVPR*, 705–706, 1993.
- [21] T. Peli, and D. Malah. A Study of Edge Detection Algorithms. *CGIP: Indianapolis, IN, USA*, 20:1–21, 1982.
- [22] C. Odet, B. Belaroussi, and H. Benoit-Cattin. Scalable discrepancy measures for segmentation evaluation. In *ICIP*, 1:785–788, 2002.
- [23] S.F. Yang-Mao, Y.K. Chan, and Y.P. Chu. Edge enhancement nucleus and cytoplasm contour detector of cervical smear images. *IEEE Trans. Syst. Man Cybern. Part B*, 38:353–366, 2008.
- [24] B. Magnier. An objective evaluation of edge detection methods based on oriented half kernels. In *ICISP*, 80–89, 2018.
- [25] A.J. Baddeley. An error metric for binary images. In *Robust Computer Vision: Quality of Vision Algorithms; Wichmann*, 59–78, 1992.
- [26] B. Magnier, A. Le, and A. Zogo. A Quantitative Error Measure for the Evaluation of Roof Edge Detectors. In *IEEE IST*, 429–434, 2016.
- [27] M. Sezgin, and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *J. Elec. Im.*, 13:146–166, 2004.
- [28] H. Abdulrahman, B. Magnier, and P. Montesinos. A New Objective Supervised Edge Detection Assessment using Hysteresis Thresholds. In *ICIAP*, pp 3–14, 2017.
- [29] Hemery, B.; Laurent, H.; Emile, B.; Rosenberger, C. Comparative study of localization metrics for the evaluation of image interpretation systems. *J. Elec. Im.*, 19:023017, 2010.
- [30] J. Paumard. Robust comparison of binary images. *Patt. Rec. Lett.*, 18:1057–1063, 1997.
- [31] E. Baudrier, F. Nicolier, G. Millon, S. Ruan. Binary-image comparison with local-dissimilarity quantification. *Patt. Rec.*, 41:1461–1478, 2008.
- [32] W.J. Rucklidge. Efficiently locating objects using the Hausdorff distance. *International Journal of computer vision*, 24(3):251–270, 1997.
- [33] N. Boughnim, J. Marot, C. Fossati, S. Bourennane. Hand posture classification by means of a new contour signature. *ACIVS*, 384–394, 2012.
- [34] R. Deriche and O. Faugeras. Tracking line segments. *Image and vision computing*, 8(4):261–270, 1990.
- [35] D. Ziou and S. Tabbone. Edge detection techniques: An overview. *Int. J. on Patt. Rec. and Image Anal.*, 8:537–559, 1998.
- [36] S. Di Zenzo. A Note on the Gradient of a Multi image. *J. Computer Vision, Graphics, and Image Proc.*, 33:116–125, 1986.
- [37] C. Steger. Similarity measures for occlusion, clutter, and illumination invariant object recognition. *Joint Patt. Rec. Symp.*, 148–154, 2001.

# A Study of Measures for Contour-based Recognition and Localization of Known Objects in Digital Images: Additional Results

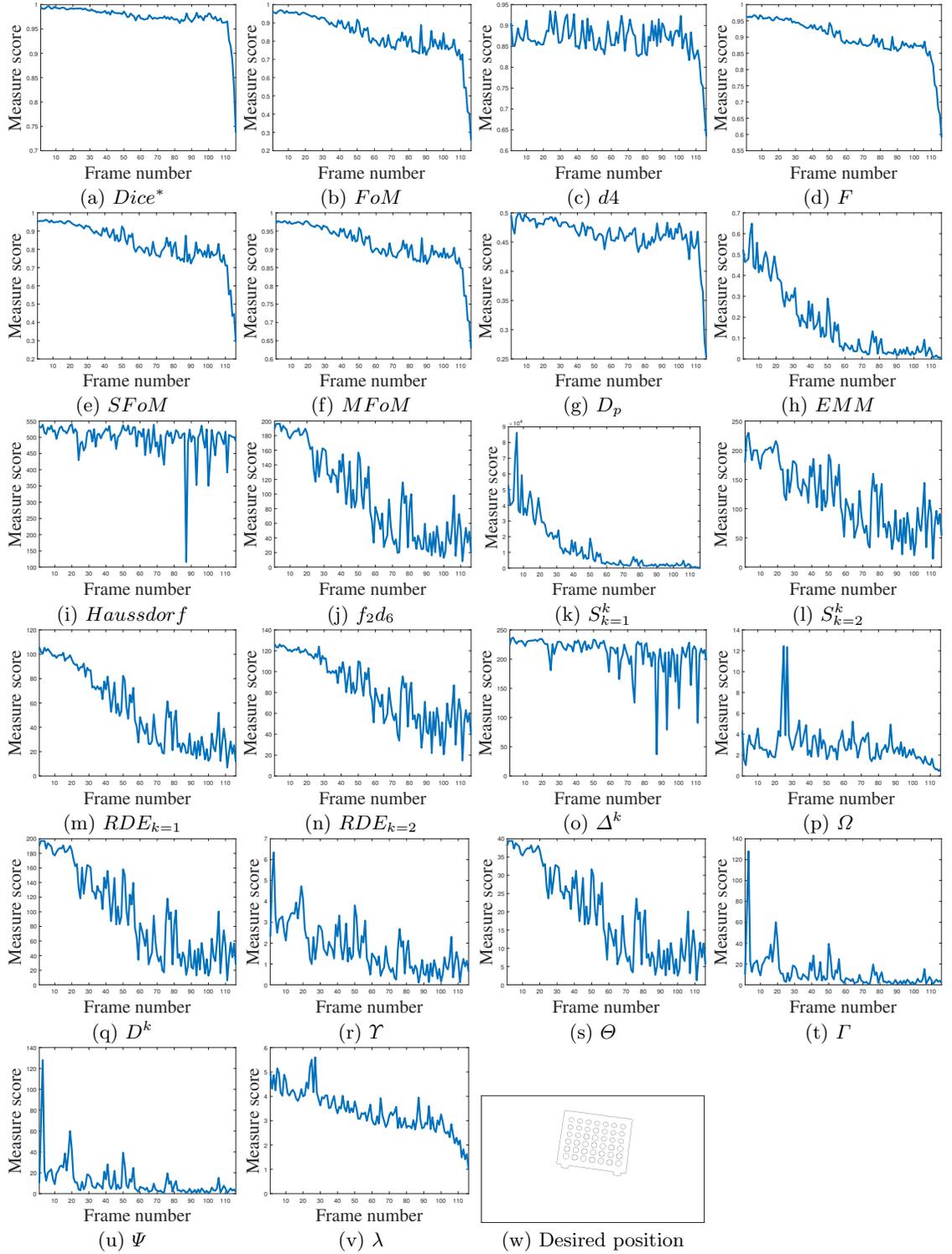
Baptiste Magnier and Hasan Abdulrahman

IMT Mines Alès, LGI2P, 6. avenue de Clavières 30100 Alès, France

**Abstract.** This document presents additional results on real images. The showed curves are tied to Fig. 6 and Fig. 8 in the original paper.



**Fig. 1.** Localization metrics behaviors on real experiment tied to Fig. 6 in the original paper.



**Fig. 2.** Localization metrics behaviors on real experiment with noisy images tied to Fig. 8 in the original paper.