Implicit Model-Oriented Optimal Thresholding Using the Komolgorov-Smirnov Similarity Measure

Xavier Fernàndez
R&D Dept., Industrias de Óptica, S.A. (INDO)
Sta. Eulàlia 181, 08902 L'Hospitalet (Barcelona - Spain)
xavier@indo.es

Abstract

We analyze the problem of local thresholding of a scene under the constraint of the geometric model of the target to be located, in the scope of the locating search process of an essentially-binary target in a gray-level scene. An optimal threshold is obtained which maximizes the fitting of the thresholded image to the target model template in the binary domain. As a result of this maximizing process the Kolmogorov-Smirnov similarity measure is obtained, which allows target spatial location in the scene with no need of explicit thresholding of the image, and avoids the high computing cost associated to gray-level domain similarity measures, such as normalized correlation. When used as a similarity measure, the nonparametric characteristic of the Kolmogorov-Smirnov statistic yields invariance and normalizing properties, which are shown to improve the properties of commonly-used gray-level domain similarity measures.

Keywords: Optimal thresholding, similarity measure, template matching, Kolmogorov-Smirnov.

1. Introduction

The problem of locating a small, essentially binary target in a gray-valued complex scene, arises frequently in industrial applications. This is the case of fiducial marks, where target geometric characteristics (shape and size) are known in advance [1], thus allowing a model of the target to be used in the recognition and locating process [2].

Three general approaches can be used to detect and locate objects in a scene. The first one is based on image thresholding, labeling and computation of object characteristics [3]; this approach, however, is not precise enough for small targets, and fails when there are incomplete targets or object overlapping. Second approach applies Hough transform (HT) or similar transforms to an edge map of the image [4]. Beside its high computing cost, this kind of procedure has difficulties to cope with small targets, because their edge segments are short and therefore induce peak spreading and uncertainty in the transformed space.

The third approach is template matching [1], implemented as a matched filter or, more generally, as a spatial search using a similarity measure between the image and a target model. Computing cost is proportional to the number of target pixels, making this approach suited to small objects. In addition, it can deal with incomplete, overlapped or low-contrast targets, with minimal sensitivity on locating precision. Consequently, template matching is an appropriate technique to locate fiducial marks and other kinds of small targets.

Normalized correlation (NC) is a similarity measure widely used for template matching on gray-valued images because its property of invariance to linear luminance changes and inherent normalization of the similarity range [5]. Nonetheless, its computation requires $2N^2$ products for a $N \times N$ pixel target, and multiply operations have usually a high computing cost on the majority of processing systems.

Template matching can be simplified if it is used in the binary domain, because computation of a binary similarity measure [6] is based on pixel counting instead of products and other mathematical operations, thus improving computing cost. The disadvantage is the need to binarize

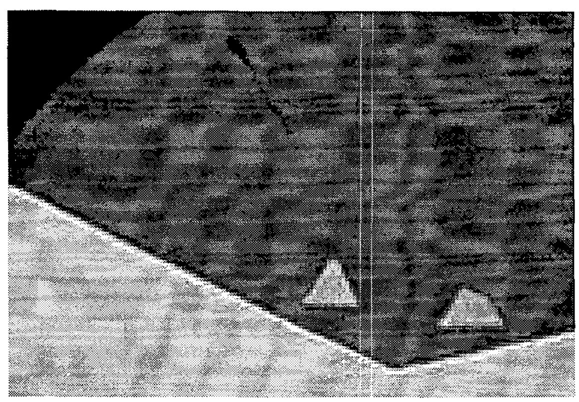

Fig. 1: A complex scene containig two triangular-shaped targets, one of them distorted.

the image prior to the search, and the difficulty to decide a threshold which ensures that target shape is preserved after thresholding [7].

Threshold selection techniques tend to fail when a small target must be segmented from a complex scene [8], because large objects may bias histograms (Fig. 1), yielding sometimes a target embedded in the background of the thresholded image [9]. Spatial-adaptive thresholding algorithms [10] increase computing cost, while in practice their advantages do not overcome drawbacks arising from small objects, because neighboring objects can again bias the local threshold selection.

At the root of the problem there is the fact that threshold selection techniques do not take into account the geometry of the objects we are intending to segment, as was suggested by Rosenfeld [11]. But in the scope of a template matching process, target geometry is known in advance, at least in its theoretical form: the model or template.

In the present work we thus adopt the strategy of constraining threshold selection to maximize the fitting of the segmented target to the model. As a result, the Kolmogorov-Smirnov (KS) similarity measure is obtained [12], allowing spatial target location with no need of explicit thresholding of the image. KS shows the low computing cost associated to the binary-domain similarity measures; whilst complete gray-level information of the image is preserved and used [9].

2. Optimization of threshold selection

Taking into account that the ideal target is essentially binary, it can be efficiently represented by a model or template having two regions (foreground T_f and background T_b, fig.2), each of them having uniform luminance, and separated by the definite shape of the ideal target.

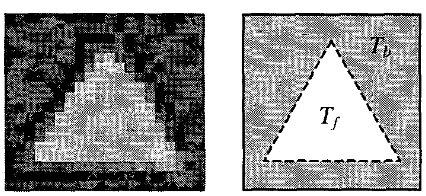

Fig. 2: Background and foreground regions of the template (right) which models the target (left).

Once the image has been thresholded, computation of the similarity between the model and the contents of each window of the image, both binary, is based on pixel match and mismatch counting. A proper way to obtain from them a value for the similarity is the modified form of the Yule [6] binary similarity measure:

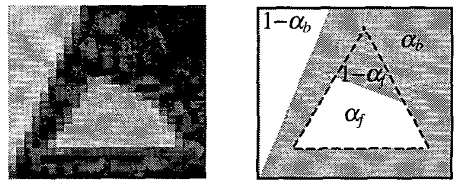

Fig. 3: Definition of area fractions (right) for a distorted target and an inhomogeneous background (left).

$$J = \left|\alpha_f + \alpha_b - 1\right| \tag{1}$$

where α_f accounts for the fraction of the model foreground area which matches corresponding pixels of the image, and similarly α_b accounts for the matching area fraction of the model background (fig.3). It is easily seen that expression (1) yields a maximum value of 1 for a perfect match, and a 0 value when image and model are uncorrelated, independently of model foreground and background pixel number, which is a desirable normalizing feature for a similarity measure. Absolute value accounts for both bright and dark foreground target types.

Clearly, area fraction values α_f and α_b are strongly dependent on the particular threshold τ used to binarize the image; therefore eq. (1) cab be expressed in terms of this threshold:

$$J(\tau) = \left|\sum_{g=0}^{\tau-1} h_b(g) + \sum_{g=\tau}^{M-1} h_f(g) - 1\right| \tag{2}$$

where $h_f(g)$ and $h_b(g)$ are the gray-level frequency histograms of the image regions corresponding respectively to the model foreground and model background (fig.4), and M being the number of gray-level quantization steps.

For some threshold value τ_0 expression (2) reaches its maximum value:

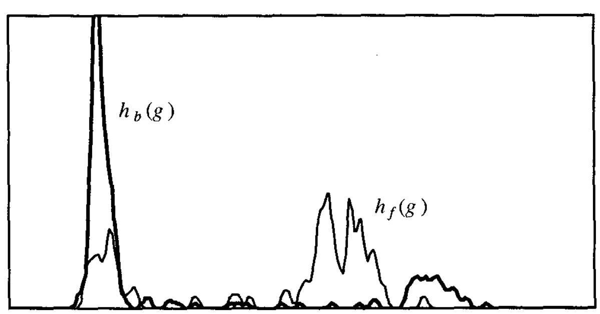

Fig. 4: Gray-level histograms of model-defined background and foreground regions of target on fig. 3.

$$\tau_0 = \underset{0<\tau<M}{\operatorname{argmax}} \left| \sum_{g=0}^{\tau-1} h_b(g) + \sum_{g=\tau}^{M-1} h_f(g) - 1 \right| \qquad (3)$$

thus τ_0 is the optimum threshold value conditioned to the geometry of the ideal target because it maximizes the similarity measure value between the image and the model.

Summations in eq. (3) can be expressed in terms of the cumulative distributions $H_b(g)$ and $H_f(g)$:

$$\tau_0 = \underset{0<\tau<M}{\operatorname{argmax}} \left| H_b(\tau) - H_f(\tau) \right| \qquad (4)$$

3. The Kolmogorov-Smirnov similarity measure

The maximum value reached by the binary similarity measure J is thus:

$$J(\tau_0) = \max_{0<\tau<M} \left| H_b(\tau) - H_f(\tau) \right| \qquad (5)$$

This expression corresponds to the definition of the Kolmogorov-Smirnov (KS) statistic, which is used in the homonymous nonparametric test to check the hypothesis that two empirical distributions come from the same ground population [13]. KS statistic is a figure of separability between the histograms $h_f(g)$ and $h_b(g)$, that is to say, it accounts for the overlap between the gray-level distributions of the image regions corresponding respectively to the model foreground and model background (fig.5).

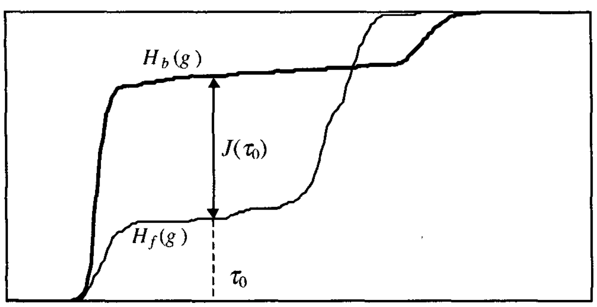

Fig. 5: Computation of KS similarity measure between model on fig. 2 and target on fig. 3, through the maximum vertical distance between cumulative distributions.

Therefore, if there is no target in the image, both distributions are very similar and expression (5) yields a value near 0; but when there is a target exactly shaped as the model, foreground gray-level distribution is completely apart from background distribution, and a perfect match of value 1 is obtained. A partial match results from a mixture of populations in the foreground or background regions, yielding a value between 0 and 1 according to the overlapping degree.

To summarize, the optimal threshold value τ_0 computed in (3) for each image window maximizes the binary-domain similarity measure defined in (1). The value of this binary similarity measure, however, can be obtained directly from the image in the gray-value domain, using the KS statistic (eq. 5); thus there is no need to actually threshold the image. KS statistic acts in practice as a similarity measure between the gray-level image and the binary model, and in opposition to binary domain similarity measures, KS uses the complete gray-level information contained in the image.

4. Properties of the KS similarity measure, compared to normalized correlation

Normalized correlation (NC) is usually the preferred similarity measure to compute similarity in the gray-level domain, because its properties of normalization and invariance to gray-level linear transformations [5], although its computing cost is high. Experiments have shown the performance of KS in comparison to NC [9,14]; in the following sections the properties of both similarity measures are summarized and compared.

4.1. Normalization

KS yields an abstract similarity value, ranging from 0 for a null match, to 1 for a perfect match, irrespectively of the probability law of the gray-level distributions, thanks to the KS nonparametric behavior. Unity value is obtained for a perfect match between target an model shapes, even if gray levels are not uniform in the foreground or the background, while NC requires uniformity to reach unity when the model is binary.

4.2. Invariance

NC is invariant to linear gray-level transformations, while KS is invariant to any monotonically increasing transformation of gray-level values, also due to KS nonparametric nature. Nonlinear gray-level transformations, such as gamma compensation, are frequently found in some image acquisition devices [14].

4.3. Computing cost

Both NC and KS have a computing cost of order $o(N^2)$ for a $N{\times}N$ pixel target; nonetheless NC requires products where KS uses only operations of increment/decrement of frequency counters, so KS is usually faster than NC. KS computing speed can be further improved by reducing the number M of gray-level quantization steps. Some experiments have sown that KS computation can be more

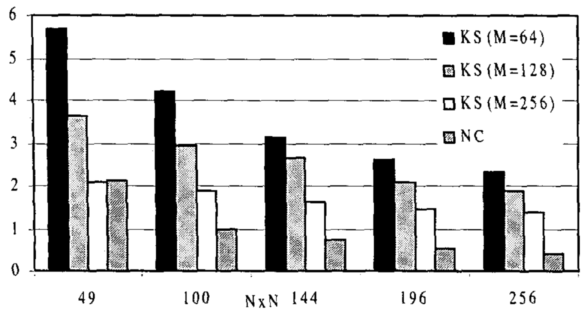

Fig. 6: NC and KC relative computing speed, depending on target size. For KS, different *M* values of gray-level quantization steps are considered.

than three times faster than NC for targets of 15×15 pixels and 256 gray levels (fig 6).

5. Conclusions

A cost-saving approach to template matching is to search the target in the binary domain after binarizing the image, and this solution seems to be naturally suited to locate essentially-binary targets such as fiducial marks. Threshold selection is, however, a critical step, specially for small targets. Furthermore, image binarization neglects an important part of the information contained in the gray values.

We has shown that a good solution is to constrain local thresholding to maximize fitting of the thresholded image to the target model. From this process the Kolmogorov-Smirnov statistic has been obtained, acting as a similarity measure between the gray-level image and the binary model. In fact, image binarization is not necessary at all when the aim is to obtain a similarity value to detect the target. Therefore, optimal thresholding remains implicit within the process.

KS statistic shows a computing simplicity similar to that of binary-domain similarity measures, but presents the properties of normalization and gray-level transformation invariance, which are important for real-world target locating applications.

These KS properties even exceed those of the cost-computing, extensively used Normalized Correlation. Thus, KS can be regarded as an advantageous, efficient alternative to Normalized Correlation.

6. References

[1] C.B. Bose, I. Amir, "Design of fiducials for accurate registration using machine vision", *IEEE Trans.PAMI,* Vol.12, nº12, pp. 1196-1200 (1990).

[2] D.H. Ballard, C.M. Brown, *Computer Vision*, Prentice Hall, New Jersey (1982).

[3] A. Rosenfeld et al., "Scene labelling by relaxation operations", *IEEE Trans.SMC* **6**, 420-433 (1976).

[4] E.R. Davies, *Machine Vision*, Academic Press, London (1990).

[5] W.M.Silver, "Normalized correlation search in alignment, gauging and inspection", *Proc. SPIE* Vol 755, pp. 23-34 (1983).

[6] J. D. Tubbs, "A note on binary template matching", *Pattern Recognition*, **22**, 359-365 (1989).

[7] J. Kittler and J. Illingworth, "Minimum error thresholding", *Pattern Recognition*, **19**, 41-47 (1986).

[8] G. Wang et al., "Efficient method for multiscale small target detection from a natural scene", *Optical Engineering* **35**(3), pp.761-768 (1996).

[9] X. Fernàndez, "Template matching of binary targets in grey-scale images: a nonparametric approach", *Pattern Recognition* **30**(7),1175-1182 (1997).

[10] Y. Nakagawa and A. Rosenfeld, "Some experiments on variable thresholding", *Pattern Recognition*, **11**, 191-204 (1979).

[11] A. Rosenfeld, "Image analysis: problems, progress and prospects", *Pattern Recognition*, **17**, 3-12, (1984).

[12] H.C. Schau, "Kolmogorov-Smirnov test in image processing", *Optical Engineering*, **20**(2), 275-280 (1981).

[13] T. Hettmansperger, *Statistical Inference Based on Ranks*, Wiley, New York (1984).

[14] X. Fernàndez, J. Ferré-Borrull, S. Bosch, "Definition of a robust measure of similarity for the localization of small shapes in scenes", *Proc. SPIE* vol. 3405, pp.1056-1062 (1998).