

Detection of the Objects With Given Shape on the Grey-Valued Pictures

I.Aksak¹, Ch. Feist², V.Kijko¹, R.Knoefel², V.Matsello¹,
V.Oganovskij³, M.Schlesinger¹, D.Schlesinger³, G.Stanke²

¹Institute of Cybernetics, Ukrainian Academy of Sciences,
40, Prospect Akademika Glushkova, 252022 Kiev, Ukraine,
Phone: (44) 266 25 69, email: schles%image.kiev.ua@ts.kiev.ua

²Gesellschaft zur Foerderung angewandter Informatik
Rudower Chaussee 5, Gebaeude 13.7, D-12484 Berlin, Germany

³National Technical University (Kiev Polytechnical Institute)
252056 Kiev, Ukraine, Prospekt Peremohy, 37

Abstract. The results of investigations of a problem of object detection on various backgrounds are described. New methods for estimation of dissimilarity between pictures and new methods of the pictures matching are developed as well as new procedures of picture pre-processing. Methods of fast computation of dissimilarity criterion are proposed.

1. Introduction

The detection is formulated as a special case of classical best-matching problem. Let e be a picture of some ideal object. The picture depends on its geometrical parameters: scale μ , displacements ξ and η and rotation φ , so that $e = e(\xi, \eta, \mu, \varphi)$. Let ρ be the picture under analysis. Let $R(\rho, e)$ be a function of dissimilarity that shows how well the real picture ρ can be matched with the ideal picture e . The problem consists in finding such displacements ξ^*, η^* , scale μ^* and rotation φ^* of the ideal picture, that provide the best matching of real picture with ideal one in accordance with R . The task consists in calculation of

$$(\xi^*, \eta^*, \mu^*, \varphi^*) = \arg \min_{\xi, \eta, \mu, \varphi} R(\rho, e(\xi, \eta, \mu, \varphi)) \quad (1)$$

The problem of reasonable dissimilarity function has already a long history (see, for example [1] and referred investigations). In this work the problem is considered in connection with thresholding and noise removing on the thresholded picture.

It is well-known that the threshold for picture binarization can be rarely pre-defined so, that a single value is suitable for every picture. Because of that a lot of adaptive thresholding procedures were developed [2,3]. The new approach, presented here, consists in that a threshold depends on both pictures, whose similarity is to be calculated. It is also necessary to remove some noise in the binary pictures before their similarity is calculated. The most popular methods of such noise removal are based on the detecting of connected components or on the ideas of dilation-erosion [4,5]. This operations must be fulfilled after binarization. As threshold's value cannot be pre-defined, such noise removal must be done for several values of threshold and

computational complexity becomes large enough. We propose such procedure that the “noise removing” with subsequent thresholding gives the same result as a thresholding with subsequent commonly used noise removing in the binary picture. This equivalency holds for every threshold’s value.

The dissimilarity function is described in section 2. In section 3 the main ideas of the solution of optimisation task are described and then discussed in sections 4,5,6.

2. The Function of Dissimilarity of Ideal and Real Pictures

Let T be a set of pixels (field of vision) and t be a pixel. An ideal picture e is meant as a pair of two subsets T_O - a set of pixels which belong to the object, and T_B - a set of background pixels. A real picture ρ is a function $\rho: T \rightarrow \{0,1,\dots,255\}$. The dissimilarity function is defined on the base of some assumptions about pictures properties using the following considerations.

2.1. The first step is based on the assumption, that a background is, as a rule, darker than an object. In this case for every two pixels t' and t'' , such that the first one belongs to the background and the second - to the object, the inequality

$$\forall t' \in T_B \forall t'' \in T_O (\rho(t') < \rho(t'')) . \quad (2)$$

would be valid. This event could be surely interpreted that the object $e = (T_B, T_O)$ is present on the picture. While the set T is finite, the event (2) is equivalent to

$$\exists \theta (\forall t' \in T_B (\rho(t') < \theta)) \& (\forall t'' \in T_O (\rho(t'') > \theta)) . \quad (3)$$

Event (3) means, that there exists such threshold θ , that thresholding of the real picture ρ with θ transforms it into the ideal picture e . It is clear that, firstly, the event (3) gives a good reason for the decision that the object e is present on the picture and, secondly, that this event occurs rather rarely. To weaken this criterion it is assumed that one can neglect the small noisy distortions of the picture.

2.2. Let $B(\rho, \theta)$ be the result of thresholding of picture ρ with θ . Then the event (3) may be represented as

$$\exists \theta (B(\rho, \theta) = e) \quad (4)$$

Let us define noise removal operation. Let ρ be a binary picture, T_0 be the set of pixels for which $\rho(t) = 0$ and T_1 be the set of pixels with $\rho(t) = 1$. The connected subsets of small size from T_0 and T_1 are considered as noise. The removing of noise consists in inverting of its brightness. The result of noise removing will be denoted as $S(\rho)$. Matching criterion:

$$\exists \theta (S(B(\rho, \theta)) = e) \quad (5)$$

is more weak, than (3). This criterion means that there exists such threshold θ that after thresholding of the picture ρ and subsequent noise removing the binary picture is obtained, that is equal to the picture e . One can see, that if some picture satisfies (4) then it satisfies (5). Of course, opposite statement is not valid and, consequently, criterion (4) is stronger than (5). At the first sight the computational complexity of (5) is much greater than complexity of (4). Really, for testing of (5) the real picture must be thresholded with every possible value of threshold and then noisy spots must be removed on every obtained binary picture. As for criterion (4), it requires only the

testing of simple inequality $\min_{t \in T_0} \rho(t) > \max_{t \in T_B} \rho(t)$. Nevertheless testing of (5) can

be essentially accelerated due to the following statement. There exists such transformation S' , that for every ρ and θ $B(S'(\rho), \theta) = S(B(\rho, \theta))$ (6)

is valid. By virtue, (5) can be represented as $\exists \theta (B(S'(\rho), \theta) = e)$ (7)

Testing of (7) requires one-fold transformation of picture ρ into the picture $\rho_1 = S'(\rho)$ and subsequent testing of $\min_{t \in T_0} \rho_1(t) > \max_{t \in T_B} \rho_1(t)$ (8)

The concrete transformation S' that satisfies (6) is described in section 4.

2.3. The next step is based on the assumption that thin lines (“hairs”) can also be considered as a noise. Let t be a pixel for which $\rho(t) = 1$ and there exists such pixel t' ,($\rho(t') = 0$), that the distance between t and t' is less then some predefined value. Such pixel t will be referred to as a black border pixel. The dilation consists in inverting of black border pixels. Similarly the erosion is defined. Removing of black lines, whose thickness is not greater than ε pixels, is defined as a sequence of ε erosions and ε dilations. Removing of thin white lines is defined as a sequence of ε dilations and ε erosions.

Let us denote by T the removing of thin distortions and weaken (5) with an account that such distortions can be neglected. As a result the following criterion

$$\exists \theta (T(S(B(\rho, \theta))) = e) \quad (9)$$

is obtained. For T the same statement is valid as the statement (6). Namely, there exists such transformation T' , that $T(B(\rho, \theta)) = B(T'(\rho, \theta))$ (10)

holds for every threshold θ and every picture ρ . Due to (10) the criterion (9) can be represented as $\exists \theta (B(T'(S'(\rho)), \theta) = e)$ (11)

Using (11) one can test the criterion computationally much more effectively. The picture transformation T' , that satisfies (10), is described in the section 5.

$$\text{Binary criterion} \quad \exists \theta (B(\rho, \theta) = e) , \quad (12)$$

can be replaced by the integer-valued function that represents a natural fuzzy variant of (12) and the dissimilarity function $R(\rho, e) = \min_{\theta} d(B(\rho, \theta), e)$, (13)

where d is a distance by Hemming between two binary pictures. It is evident that if $R(\rho, e) = 0$ then (12) holds. The value $R(\rho, e) \neq 0$ is considered as a degree of mismatching of ρ and e .

As a result, the following formulation of the problem of detection of pre-defined ideal picture on the real picture is obtained. It is necessary to remove some distortions on the real picture ρ . Then transformed picture ρ_1 must be obtained and for real picture such threshold θ^* , such displacements ξ^* and η^* , scale μ^* and rotation φ^* of ideal picture must be found, that the distance by Hemming between

thresholded picture $B(\rho_1, \theta^*)$ and transformed ideal picture $e(\xi^*, \eta^*, \mu^*, \varphi^*)$ would be minimal. It means, that it is necessary to calculate

$$(\xi^*, \eta^*, \mu^*, \varphi^*) = \arg \min_{\xi, \eta, \mu, \varphi} \min_{\theta} d(B(\rho_1, \theta), e(\xi, \eta, \mu, \varphi)) \quad (14)$$

3. Methods for Acceleration of the Matching

3.1. Let us analyse the innermost part of (14). Number of operations for calculation the distance d under fixed values of $\theta, \xi, \eta, \mu, \varphi$, is of order n - number of pixels in the field of vision. To calculate $\min_q d(B(r_1, q), e(x, h, m, j))$

the distance d must be calculated for each threshold value. The amount of this values will be denoted by ω . However the total amount of necessary operations is not of order $\omega \cdot n$, as one can suppose at the first glance, but is of order $\omega + n$.

3.2. Computation of (14) requires that the value $\min_{\theta} d(B(\rho_1, \theta), e(\xi, \eta, \mu, \varphi))$ must be obtained for every four-tuple ξ, η, μ, φ . Let g be the amount of these values. The total amount of computation is not of order $g \cdot (\omega + n)$, as it could be suspected, but is of essentially less order $g \cdot (\omega + n')$, where n' is the number of pixels of ideal picture at the border of subsets T_B and T_O .

4. Noise Removing in the Grey-Valued Picture

The algorithm for noise removing consists of four steps: 1) sorting of pixels in increasing order of their brightness; 2) "white" noise removing; 3) sorting of pixels in decreasing order of their brightness; 4) "black" noise removing.

Ordering of pixels is fulfilled during two-fold scanning of the pixels using additional array with the length equalled to the number of different brightness.

"White" noise removing is carried out during one-fold scanning of ordered pixels. Let t_i be a pixel having brightness $v(t_i)$. Processing of the pixel consists in checking whether there exists "white" noise region G that contains this pixel. Initially the region contains only t_i . Then the region grows so that. some pixel t' is included into it if: a). t' is a neighbour of the region G ; b). $v(t') \geq v(t_i)$. The region grows until one of the following two conditions are satisfied: 1) the size of the obtained region G exceeds the pre-defined size. In this case the region G is not considered as a noise and the next pixel t_{i+1} is processed; 2) no pixel can be included into G . In this case G is a noise and the brightness of all its pixels are equalled to maximal brightness of the neighbouring pixels. Removing of "black" noise is similar to "white" noise removing.

5. Removing of Thin Distortions on the Multilevel Pictures

Let ρ be a multilevel picture, $B(\rho, \theta)$ be a result of thresholding of ρ using θ , $T(B(\rho, \theta))$ be a result of removing of "hairs" on B . It is necessary to construct such transformation T' that the equality $B(T'(\rho), \theta) = T(B(\rho, \theta))$ (15)

is valid for every multilevel picture ρ and every threshold θ . For the removal of white "hairs" the pair "erosion-dilation" must be fulfilled. So it is necessary to find out such transformation T' that $B(T'(\rho), \theta) = D(E(E(D(B(\rho, \theta))))))$ (16)

is valid for every ρ and every θ , where D and E mean dilation and erosion. Let us consider only D and find the transformation TD , that satisfy the equality

$$B(TD(\rho), \theta) = D(B(\rho, \theta)) \quad (17)$$

Let $T(t)$ be the d -neighbourhood of pixel t , $\rho \max$ be the maximal brightness of pixel from $T(t)$. Let us define the picture $TD(\rho)$ such that the brightness of t is $\rho \max$. Then $B(TD(\rho), \theta)$ is equal 1 if and only if $\theta \leq \rho \max$. (18)

The brightness of picture $D(B(\rho, \theta))$ is equal 1 if and only if there exists a pixel t' in $T(t)$ for which $B(\rho, \theta)$ is equal 1. It is equivalent to the condition that there exists a pixel t' for which $\rho(t') \geq \theta$ and in its turn is equivalent to $\rho \max \geq \theta$. (19)

One can see, that the condition under which brightness $B(TD(\rho), \theta)$ of the pixel t equals 1 is the same as the condition, under which brightness $D(B(\rho, \theta))$ is 1. Consequently, these brightness are always equal. So, transformation TD satisfies (17). Similarly, one can construct a transformation TE such that:

$$B(TE(\rho), \theta) = E(B(\rho, \theta)). \quad (20)$$

This transformation assigns for a pixel t the brightness $\rho \min$ -minimal brightness of picture ρ from $T(t)$. The properties (17) and (20) cause that the equality $B(TD(TE(TE(TD(\rho))))), \theta) = D(E(E(D(B(\rho, \theta))))))$ holds also for every θ . The last expression defines explicitly the operator T' that satisfies (16).

6. Fast Matching of Grey-Valued Pictures

6.1. Statement of the Problem

Let $\rho: T \rightarrow \{0, 1, \dots, 255\}$ be a picture under analysis; $\rho(t)$ be a brightness of pixel t ; Θ be a threshold; $B(\rho, \Theta)$ be a result of thresholding ρ with Θ ; $e = (T_o, T_b)$ be an ideal pattern, T_o be an object, T_b be a background, $T_o \wedge T_b = \emptyset$; ξ, η, μ, φ be values of horizontal and vertical translations, scale and rotation respectively (this four-tuple will be referred to as a variant); $e(\xi, \eta, \mu, \varphi)$ be the ideal pattern under translation ξ and η , scale μ and rotation φ ; $d(v, e), e = (T_o, T_b)$ be a dissimilarity function between v and e . Its value is defined as a total quantity of pixels for which $((v(t) = 1) \wedge (t \in T_b)) \vee ((v(t) = 0) \wedge (t \in T_o))$. The problem is formulated as the optimisation task: $(\xi^*, \eta^*, \mu^*, \varphi^*) = \arg \min_{\xi, \eta, \mu, \varphi} \min_{\Theta} d(B(\rho, \Theta), e(\xi, \eta, \mu, \varphi))$.

6.2. The Problem Analysis and Elimination of Calculations

6.2.1. Let us consider dissimilarity minimisation under fixed variant, i.e. calculation of $F(\rho, \xi, \eta, \mu, \varphi) = \min_{\Theta} d(B(\rho, \Theta), e(\xi, \eta, \mu, \varphi))$. It is fulfilled by three steps.

a. Two histograms must be calculated: $G_o(\Theta)$ (respectively $G_b(\Theta)$) being an amount of pixels, for which $(t \in T_o) \wedge (\rho(t) = \Theta)$ (respectively $(t \in T_b) \wedge (\rho(t) = \Theta)$).

b. Then the cumulative histograms $G'_o(\Theta) = \sum_{i=0}^{\Theta} G_o(i)$, $0 \leq \Theta < \Theta_{\max}$ and

$G'_b(\Theta) = \sum_{i=0}^{\Theta} G_b(i)$, $0 \leq \Theta < \Theta_{\max}$ must be calculated using recurrent expressions:

$$G'_o(0) = G_o(0), G'_o(\Theta) = G'_o(\Theta - 1) + G_o(\Theta), \Theta = 1, 2, \dots, \Theta_{\max} - 1;$$

$$G'_b(\Theta_{\max} - 1) = G_b(\Theta_{\max}), G'_b(\Theta) = G'_b(\Theta + 1) + G_b(\Theta + 1), \Theta = \Theta_{\max} - 2, \dots, 1, 0$$

$G'_o(\Theta)$ means the amount of pixels from subset T_o , which would be black after thresholding of picture ρ with Θ . The function $G'_b(\Theta)$ has similar meaning.

c. $F(\rho, \xi, \eta, \mu, \varphi)$ is being calculated. This value is equal $\min_{\Theta} (G'_o(\Theta) + G'_b(\Theta))$.

The amount of operations that are necessary to obtain functions $G_o(\Theta), G_b(\Theta)$, is of order n - the amount of pixels in $T_o \cup T_b$. All other calculations are of order Θ_{\max} - the amount of brightness values.

6.2.2. Let us consider now dissimilarity minimisation task under fixed scale and rotation, i.e. calculation of $F(\rho, \mu, \varphi) = \min_{\xi, \eta} F(\rho, \xi, \eta, \mu, \varphi)$, which is necessary to

obtain the values of $F(\rho, \xi, \eta, \mu, \varphi)$ for every ξ and η . It requires to fulfil steps a, b and c. For the first value of ξ and η step a is fulfilled by the following algorithm.

For every Θ $G_o(\Theta) = G_b(\Theta) = 0$; For every $t \in T_o$ $G_o(\rho(t)) = G_o(\rho(t)) + 1$;

For every $t \in T_b$ $G_b(\rho(t)) = G_b(\rho(t)) + 1$;

Calculation time is proportional to the amount of pixels in $T_o \cup T_b$. However, the processing time for all other values of ξ and η can be reduced essentially. Really, let ξ_0 and η_0 , ξ_1 and η_1 be two different translations, $e_0 = (T_{o0}, T_{b0})$ and $e_1 = (T_{o1}, T_{b1})$ be corresponding ideal patterns. Let us suppose that histograms $G_{o0}(\Theta)$ and $G_{b0}(\Theta)$ for the ideal pattern e_0 are already available. Then histograms $G_{o1}(\Theta)$ and $G_{b1}(\Theta)$ must be constructed by recalculation of histograms G_{o0} and G_{b0} using the following algorithm.

a. The four-tuple of difference sets

$T_{o01} = T_{o1} \setminus T_{o0}, T_{o10} = T_{o0} \setminus T_{o1}, T_{b01} = T_{b1} \setminus T_{b0}, T_{b10} = T_{b0} \setminus T_{b1}$ must be constructed.

b. $\forall \Theta: \{ G_{o1}(\Theta) = G_{o0}(\Theta); G_{b1}(\Theta) = G_{b0}(\Theta); \}$

$$\begin{aligned} \forall t \in T_{o01} \quad G_{o1}(\rho(t)) &= G_{o1}(\rho(t)) + 1; \quad \forall t \in T_{o10} \quad G_{o1}(\rho(t)) = G_{o1}(\rho(t)) - 1; \\ \forall t \in T_{b01} \quad G_{b1}(\rho(t)) &= G_{b1}(\rho(t)) + 1; \quad \forall t \in T_{b10} \quad G_{b1}(\rho(t)) = G_{b1}(\rho(t)) - 1. \end{aligned}$$

As the lists of pixels from the difference sets can be constructed in advance, the time of computations on the step b may be proportional to the volume of difference sets. When the difference between the ideal patterns e_0 and e_1 is small, the volume of difference sets is considerably less than the volume of $T_o \cup T_b$. The computation of step a may be accelerated. Let the ideal pattern e_1 is obtained by the translation $(\Delta i, \Delta j)$ of the ideal pattern e_0 . Let the ideal pattern e_2 is obtained by the same translation of the ideal pattern e_1 . It is obvious that the difference sets for e_2 and e_1 can be obtained by the translation $(\Delta i, \Delta j)$ of the difference sets for e_1 and e_0 . The complexity of such recalculation is also proportional to the volume of difference sets. Due to this the construction of difference sets is made only once for the first position of ideal pattern. For every other position the difference sets are only recalculated.

7. Application

The developed best-matching method was implemented for searching of geometrical objects in the pictures of shoeprints with the aim of subsequent search of the given shoeprint in the database. Two examples of detecting some geometrical objects are shown in Fig. 1.

Acknowledgement

This research was sponsored the Ministry of Education and Investigations of Germany which is gratefully acknowledged.

References

1. V.Starovoitov, Towards a Measure of Diversity between Grey-Scale Images. In *Comp. anal. of images and patterns: 6th int. conf; proc./ CAIP'95, Prague*, 214-221, 1995.
2. W.Tao, H.Burkhardt, An Effective Image Thresholding Method Using a Fuzzy Compactness Measure. In *Proc. of 12th IAPR Int. conf. on pattern recogn.*, **1**, 47-51, 1994.
3. P.K.Sahoo, S.Soltani, A.K.C.Wong, Y.C.Chen, A Survey of Thresholding Techniques. In *Comput. Vision Graphics Image Process.*, **41**, 233-260, 1988.
4. Haralick R. M., S. R. Sternberg, and X. Zhuang, 1987. Image analysis using mathematical morphology, *IEEE Trans. on Pattern Anal. and Machine Intell.* **9**(4), 532-550.
5. Serra, J. In *Image Analysis and Mathematical Morphology*, Acad.Press, NY, 1982.

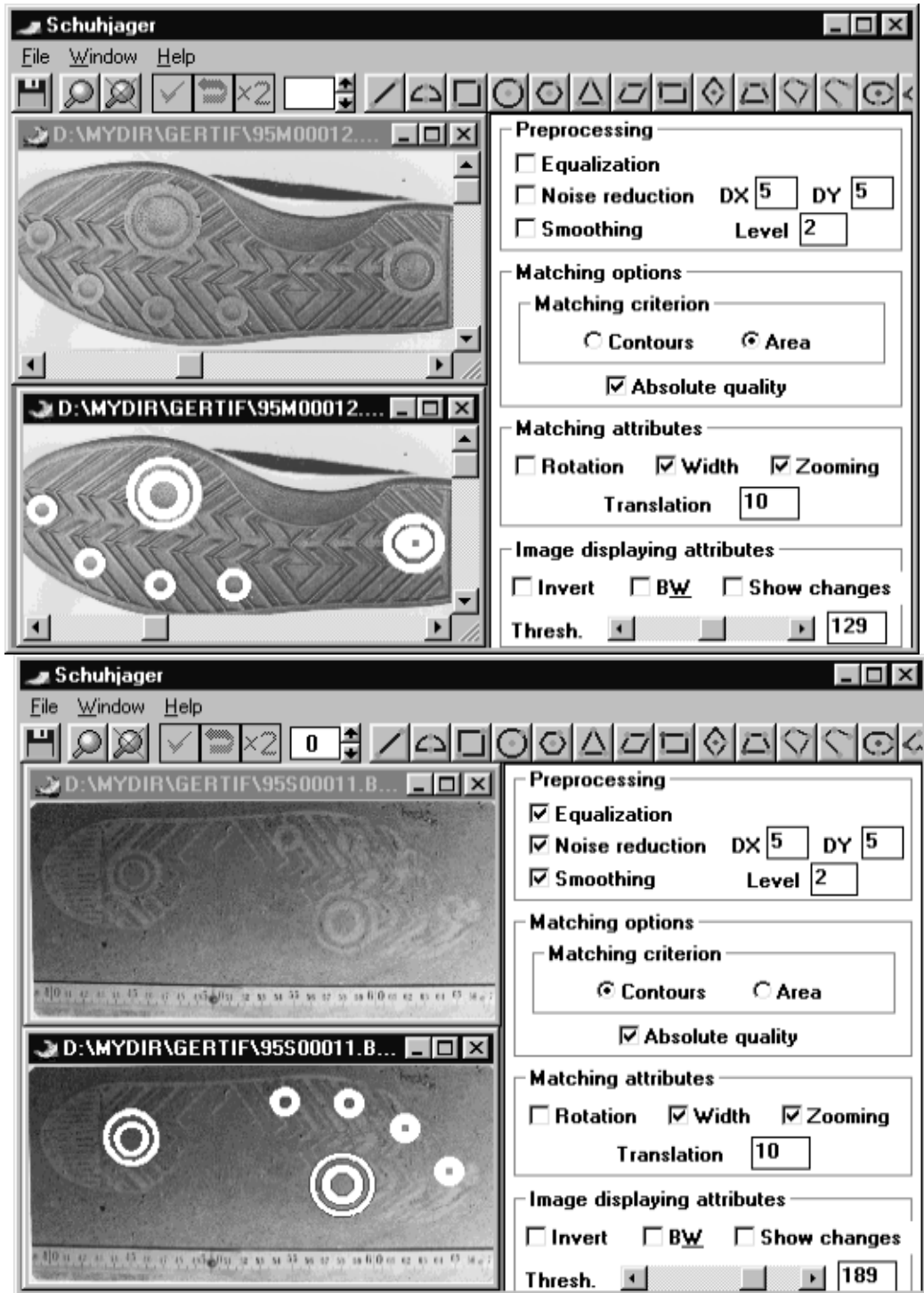


Fig. 1. Detecting of circles on shoeprint images.