MATCHING CRITERIA IN TEMPLATE LOCALIZING – COMPARATIVE ANALYSIS OF EXPERIMENTAL RESULTS

Yulka PETKOVA*, Dimitar TYANEV*

*Technical University of Varna,
Department of Computer Sciences and Technologies,
1, Studentska Str., 9010, Varna, Bulgaria

Abstract: In the present paper we made an analysis of different matching criteria, which are applied when a task of template matching is solving. It is made an analysis of the experimental results, which are derived when these different criteria are applied on the selected sets of representative points. Comparisons are made about two basic criteria – speed and how they cope with any type of distortions.

Keywords: Template Matching, Matching Criteria, Image Processing.

INTRODUCTION

One of the most important problems in the field of image analysis is the localization of any template in a greater image.

A wide variety of applications need the decision of the pointed problem. Some of them are:

- Wafer alignment, using arbitrary artwork as the pattern to recognize;
- Fiducial recognition for PCB assembly by pick-and-place robot;
- Registration of alignment marks on printed material to be inspected (e.g. wallpaper rolls, currency sheets);
- Robot vision guidance, locating objects on conveyor belts, pallets and trays
- Tracking cars in video sequences.

The problem of template matching is connected with finding of any matching measure and algorithm for an evaluation of the extent of coincidence of the template with the correspondent part of the image. This task has been a research object since 60-th years of the past century, which explains the presence of the great number of matching criteria and searching algorithms [2, 5, 6]. One of the basic purposes is directed to the achievement of high reliability. This means that the matching criterion has to cope successfully with any type of distortions – translation, rotation, scaling, unlinear changes in the intensity, etc. The other important goal is the real time processing. The second purpose is harder to achieve than the first one. This is because of the great data sets, which are to be processed. They are searched ways for a decreasing of the essential information from the template, which will lead to the decreasing of search-time and in the same time the accuracy of localizing has to be high [1, 2, 5, 6].

The present work is devoted to the examination of the abilities of different matching measures and different algorithms for speeding-up their calculations. The high accuracy, derived with high speed is the basic criterion for an estimation of algorithms.

1. MATCHING MEASURES AND ALGORITHMS FOR THEIR REALIZATION

1.1. Full coinciding

The trivial decision of the problem is to pass over all pixels of the image in the area $x \in [0, W_2 - W_1)$ and $y \in [0, H_2 - H_1)$, where [W_2, H_2] and [W_1, H_1] are sizes of the template and of the image respectively.

All positions are passed and the corresponding pixels are compared. When the pixels don't coincide, then the current position rejects. The matching measure is the full coinciding. The disadvantages of this algorithm are: it doesn't work for noised images and its computational cost is $O(N^4)$ and it works too slowly for great images.

1.2. Algorithm which counts coinciding pixels

Like the previous algorithm this one compares pixels, but when it occurs they don't coincide, the work continues to the end and the coinciding pixels are counted. Finally the position of the found template is considered to be this one, which corresponds to the biggest number of coinciding pixels.

This algorithm decides the problem of the first one, but if there is a kind of any little distortion, little change of the brightness, or of the contrast, it will not be able to find the template.

Moreover, in practice it is slower then the first one, because it has to pass every position, independently of how much positions don't coincide or at least while the number of pixels, which don't coincide is not higher then the best one till the moment.

These two algorithms are not used in practice, because the other matching criteria give much better results for almost the same time.

1.3. Algorithm with the differential criterion as a matching estimation

The basic disadvantage of the above two algorithms and criteria is avoided with the evaluating of the difference (or "error") between the template and the corresponding part of the image, which has the same size as a template.

The matching measure is expressed with:

(1) $$\text{Dist(x,y)} = \sum_{i,j} abs\left(I_1(i,j) - I_2(x+i, y+j)\right),$$

where Dist is named as a distance between the template and the part of the image.

The algorithm with the differential criterion for the matching estimation is still with a great computational complexity, but it gives results even if the image is noised.

1.4. Algorithm with the minimal square error (MSE)

Minimal square error [3, 4, 5] substitutes operation "calculation of the absolute value" with the square of the difference between the intensities of corresponding pixels:

(2) $$\text{Dist(x,y)} = \sum_{i,j} \left(I_1(i,j) - I_2(x+i, y+j)\right)^2.$$

The result of the squaring is a positive number and it still has a sense of an absolute value of the difference.

An advantage of MSE is that it makes the distance between the differences exponentially greater and when a choice should be made, the coincides can be selected more accurately:

(3) $$abs\left\{abs(a_1 - b_1) - abs(a_2 - b_2)\right\} \le abs\left\{(a_1 - b_1)^2 - (a_2 - b_2)^2\right\}.$$

Another advantage is that sometimes the squaring is computationally faster then the calculation of the absolute value.

1.5. Algorithm with correlation estimation

Commonly the discrete correlation is a ratio between two signals. Many types of correlation are known, but in the field of image processing the normalized cross correlation is used, which is defined as follows:

(4) $$R_{xy} = \frac{\sum_{i,j}\left[\left(I_1(i,j) - I_1^{avg}\right)\left(I_2(x+i, y+j) - I_2^{avg}\right)\right]}{\sqrt{\sum_{i,j}\left(I_1(i,j) - I_1^{avg}\right)^2}\sqrt{\sum_{i,j}\left(I_2(i+x, j+y) - I_2^{avg}\right)^2}},$$

where: $I_1(i,j)$ is the intensity in the point (i,j) from the template;

$I_2(x+i,y+j)$ is the intensity in the point $(x+i,y+j)$ from the image;

I_1^{avg}, I_2^{avg} are the mean values of the template and of the image intensity respectively.

The high computational intensity of this algorithm is a reason for searching ways for its speeding-up, as for example elimination of the constants from the formula or reducing the number of pixels, which take part in the comparison and so on [5].

1.6. Fast discrete correlation

There are many algorithms for speeding-up the correlation. Many of them are based on the statement that the correlation in neighbor positions is connected, i.e. the repeating calculations of neighbor coefficients are making $R_{x,y}$, $R_{x+1,y}$, $R_{x,y+1}$ and so on.

In the spatial domain the method with the optimization of calculations is not so obvious. This is the reason for using transformations in other domains, in particular – in frequency domain (Fourier domain) [3, 4, 5].

The transformations can be calculated much faster using Fast Fourier Transformation (FFT) and Inverse Fast Fourier Transformation (IFFT). These algorithms are published by Cooly and Tukey in 1965 and it employees the computational surplus from the basic formula.

2. SEARCHING ALGORITHMS WORKING WITH A SUBSET OF THE TEMPLATE

The other direction for an optimization of searching algorithm is to choose a subset of the representative points from the template, instead of using the whole set of points. This is equivalent to the looking through a mask with sight-holes, which show only particular pixels and don't give any information about the rest part of the image. Thus, the task consists of comparing only the subset of the selected points. Matching measures are the same, but the problem is which points are to be selected, so the search will be effective.

It is clear, that a random choice of points of taking points with any before settled and fixed coordinates is none too expediently. The basic problems of these methods come from the fact that they don't take into consideration the template or the image structure.

2.1. Points, which belong to the edges

Points, which belong to the edges of the image, bring a great quantity of information, because they outline the different objects in the image and describe its structure. Edge detection is a deeply researched part from the field of image processing and there are many edge detectors, which are used in practice - Canny, Roberts, Sobel, Kirsch, Prewitt, LoG, DoG and many others [2, 3, 4].

2.2. Equipotential method for point extraction

It is known [1], that any procedure for linear dividing, which is based on the D-optimality criterion of the selected subset derives the best linear estimation for the coefficients of the dividing function. This means such choice of elements from the selected subset, which can present the spatial structure of the image in the best way. The analysis of so formed subsets shows that the spatial positions of their elements form the protrusive cover of the object structure.

The image is presented as a three-dimensional object, the third axes is for the intensity changes. Points, which satisfy the criterion of D-optimality, are the outermost from the peripheral cover of the three-dimensional object. Points are derived, when the object is intersected with the equipotential planes [1].

3. HAUSDORFF DISTANCE

Hausdorff distance [5] is a matching measure between two sets of geometric points P and Q. It is expressed with the next formula:

(5) $$H(P,Q)=\max_{a\in P}\left[\min_{b\in Q}\left(d\left(a,b\right)\right)\right],$$

where $d\left(a,b\right)$ is the basic distance between points.

Images are considered as sets of points in the three-dimensional space (wide, length and intensity). Sometimes in practice the k-th maximal value is taken, instead of the absolute maximal value. The reason of this is to provide a greater tolerance to the disturbed images:

(6) $$H_k(P,Q)=\mathrm{k}^{\mathrm{th}}\max_{a\in P}\left[\min_{b\in Q}\left(d\left(a,b\right)\right)\right].$$

4. EXPERIMENTAL RESULTS

4.1. Analytical comparison of algorithms

The computational complexity of the considered algorithms is shown in table 1.

It is seen that the greatest computational cost have searching algorithms, which work with the whole sets of points and which calculate absolute difference, square error and correlation as a measure of matching.

Table 1

Algorithm	**Computational compexity**
Searching over the whole set of points; matching criteria are absolute difference, square error, correlation.	$O(N^4)$
Searching over subset of selected points; matching criteria are absolute difference, square error and correlation.	$O(N^2M)$
Searching over subset of selected points; matching criteria is Hausdorff distance	$O(N^2M^2)$
Fast correlation in the frequency domain	$O(N^2\log(N))$

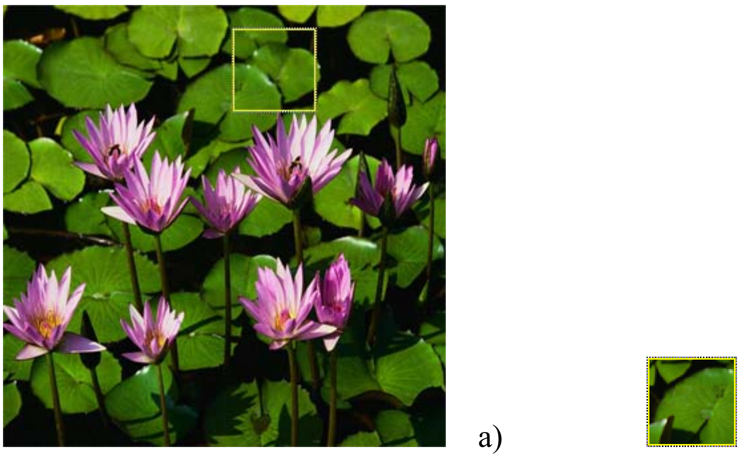
N is the high/wide of image. M is the number of selected points.

4.2. Experimental comparison of algorithms

All experiments are made with the following system: **Motherboard:** ASUS A8N-E; **Processor:** AMD Athlon64 3000+; **RAM:** 1GB DDR400 Dual Channel; **Video adapter:** ASUS nVIDIA GeForce 6600 PCI-E c 256MB DDR500 RAM; **Operation System:** MS Windows XP Pro SP2. We have made efforts all of the tests to be carry out in the same conditions.

4.2.1. The accuracy of algorithms working with the subset of selected points

It is chosen a template from the image Flowers (figure 1a). The position of the template (figure 1b) is (125, 36) and its size is (64x64):



a) b)

Fig. 1. a) – Test image Flowers, **b)** – Chosen template (twice increased)

The number of selected points is 100 and the image is noised with 20% uniform noise. It is valid for all of the tests.

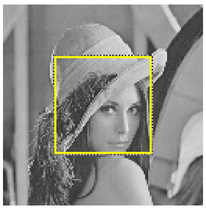
Table 2

Matching criterion	Deviation from the correct position [pixels]							
	Sobel – 20% noise		Roberts – 20% noise		Equipotential points			
					20% noise		25% noise	
	DX	DY	DX	DY	DX	DY	DX	DY
Absolute value of difference	109	228	0	0	0	0	-	-
Minimal square errror	109	228	0	0	0	0	0	0
Cross correlation	14	23	0	0	0	0	9	165
Hausdorff distance	122	216	118	62	20	226	-	-

The results, derived from these tests are shown in table 2. It is visible, that the subset of Sobel edge points does not find the correct position of the template. All of the distances of matching except Hausdorff distance give good results, if they work with the points extracted with Roberts edge detector or with the method of equipotential planes.

If the noise is higher, the correlation makes errors, while the minimal square error does not make. This is seen from the result of the test with 25% noise, selection of 89 points, extracted with the method of equipotential planes. The template is a part of image Flowers with upper left coordinates (141, 15). The only advantage of correlation is seen when the brightness of the image is changed. Minimal square error cannot find the template in such conditions – see table 3.

Table 3

Original image	Matching with MSE	Matching with cross correlation
		
Coordinates of the template (32, 32)	Coordinates of the found template (33, 32)	Coordinates of the found template (32, 32)

4.2.2. Speed of searching algorithms, working with selected important points

We should pay attention to the next fact: if the number of points and the matching criteria are the same, then searching time is one and the same as well, independently of the method for their selection. It is due to the identical data structures and procedures for the choice f points. It doesn't matter what kind of imaged are used as well. All of them just make a list of important points. We comment this to explain the only one method for point extraction we have used in the tests – the method of equipotential planes.

Moreover, for all of tests below noise in the image is 10%, which provides good identification with all of the matching measures. The used template is in the position (163, 58) from the image Flowers. Results are shown in table 4.

Table 4

Matching criterion	Searching with the subset of selected points - time [ms]	Searching with the whole set of points - time [ms]	Speeding-up [times]
Absolute value of differences	128	1735	13.55469
MSE	**94**	**1313**	**13.96809**
Cross correlation	95	1528	16.08421
Hausdorff distance	21391	-	-

The Hausdorff distance calculates too slowly. Minimal square error is the fastest. The absolute value of difference calculates slowly, which is due to the fact, that the compiler doesn't optimize the function **abs()** correctly.

The derived acceleration is satisfactory, especially considering that 200 points are far more then necessary for the template with a size (64x64). In many cases 100 points are fully enough for a correct identification.

The received experimental results show that minimal square error is the best matching distance.

4.2.3. Minimal square error with points, selected with the method of equipotential planes, across the fast correlation in the frequency domain

4.2.3.1. Speed

The test is carried out with the image Flowers (333x333) – see figure 1a, and the results are presented in table 5.

Table 5

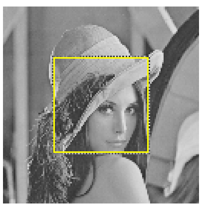
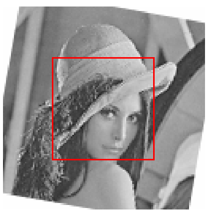
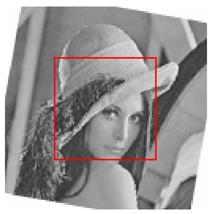
Number of selected points	Template size	Time for searching with correlation in the frequency domain [ms]	Time for searching with a selected subset of points		
			Time for subset selection [ms]	Time for points selection [ms]	Common time for searching [ms]
200	64x64	218	114	17	**131**
200	128x128	217	79	70	**149**
500	128x128	**217**	168	70	238
500	256x256	**203**	28	287	315

The results from table 5 show, that the methods, working with the subset of 200 points are faster even the correlation in the frequency domain.

4.2.3.2. Rotation stability

We consider the rotation as a kind of noise. The possible rotation in the scanned or printed images is not much. The test is provided with the image Lena (128x128), rotated on 10 degree. There is not any additional noise. Template size is (64x64). The derived results are shown in table 6.

Table 6

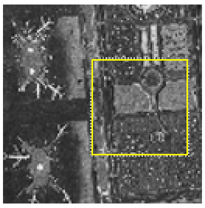
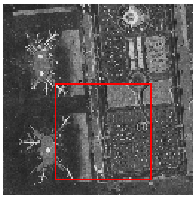
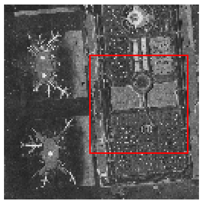
Original and template	MSE over 200 points	Correlation in frequency domain
		
Coordinates of the template (32, 32)	**Coordinates of the found template (31, 32)**	**Coordinates of the found template (31, 32)**

We can say that algorithms work well with such little rotations. They cannot determine the angle of rotation, but they localize the template accurately rough.

4.2.3.3. Scaling stability

On the analogy of the rotation we consider the scaling as a kind of noise and compare the results, derived from both algorithms. We consider a little scaling, which is possible in photo pictures. The tested image is a part of the satellite picture LAX (128x128), The template size is (64x64). Results are shown in table 7.

Table 7

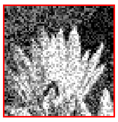
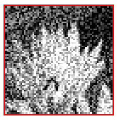
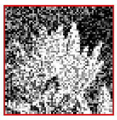
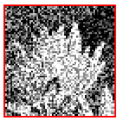
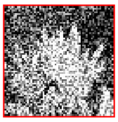
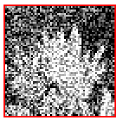
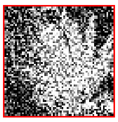
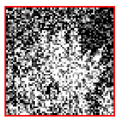
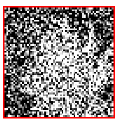
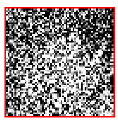
Original and template	**Decreased with 20%, MSE over 118 points**	**Decreased with 20%, correlation in the frequency domain**
		
Coordinates of the template (56, 34)	**Coordinates of the found template (35, 53)**	**Coordinates of the found template (56, 33)**

Here we can see the advantage of the correlation over the whole set of points if the level of noise is high and the templates are hard identified. We can accept that the correlation over the whole set of points finds the template correctly.

4.2.3.4. Noise stability

The test is provided with a subset of 200 points. The results are shown in table 8.

Table 8

Level of noise		**MSE with subset of points**			**Correlation in the frequency domain**		
%	**SNR [dB]**	**Found part of the image**	**Found in position X**	**Found in position Y**	**Found part of the image**	**Found in position X**	**Found in position Y**
30	10.4575		56	61		56	61
40	7.9588		56	61		56	61
50	6.0206		56	61		56	61
60	4.4369		**56**	**62**		56	61
70	3.098		**172**	**188**		56	61
100	0		**172**	**183**		56	61

It is seen, that when the level of noise is 60% or more, the searching with a subset of points cannot find the template. Correlation in the frequency domain keeps accuracy even the noise is high.

CONCLUSION

From the analytical analysis and from the analysis of the derived experimental results we can conclude that the best of the matching measure is the cross correlation. When the correlation works with a selected subset of points, the formed with the method of equipotential planes subset of points gives the best results. It is due to the applied criterion of D-optimality. The minimal square error has almost the same abilities.

Many other experiments can be made. They can study different algorithms, methods and criteria for matching and for selection subsets of representative points, which have to take part in the comparisons. Different searching strategies can be studied as well.

The authors make no claim that their study is exhaustive, but they consider that it is representative enough to make adequate conclusions.

REFERENCES:

[1]. Tyanev, D., Y. Petkova. *"A New Method for Important Points Extraction"*, International Scientific Conference CompSysTech2004, Rousse, Bulgaria, June 2004, pp. III.11-1-11-8.

[2]. Manickam, Swami, Roth, Scott, Bushman, Thomas, *"Intelligent and optimal normalized correlation for high-speed pattern matching"* - Datacube Inc., November, 2000.

[3]. Parker, J.R., "*Algorithms for Image Processing and Computer Vision"*, Wiley Computer Publishing, 1997

[4]. Olson, C.F., "*Maximum-Likelihood Image Matching",* IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(6): 853-857, June 2002.

[5]. Ritter, G. X, J. N. Wilson, *"Handbook of computer vision algorithms in image algebra"*, CRC Press LLC, 2001.

[6]. Lundberg, F., "Maximum Entropy: An Approach to Fast Template Matching"– Oct., 2000.