An Algorithm for Fast and Robust Template Matching Using M-estimators

Yulka Petkova[1], Nuri Nuri[2] and Milena Karova[3]

Abstract – **In this paper we propose a fast and reliable template matching algorithm. We use M-estimators to evaluate the similarity between compared images. Sets of representative points are used in computations. A coarse-to-fine algorithm is also used for speeding up the procedure. The received experimental results show high reliability and precision of template localization.**

Keywords – **Template Matching, Image Processing, M-estimators, Representative Points**

I. INTRODUCTION

Image Registration is a process, which compares two or more images of one and the same scene, received at different moments and from different points of view and/or from different devices. It is a very important step in all tasks for image analysis and it is necessary to:

- integrate information, received from different devices;
- find differences between images of one and the same scene (object), received at different moments and different conditions;
- extract three-dimensional information from images, where the object are disordered or images, received by sensors (cameras), which are at different positions, i. e. from different points of view;
- template (model) based object recognition.

Registration should solve different problems, including noise from sensors, different sensors, different kind of transformations - translation, rotation, scaling, and their combination also, changes in atmosphere conditions - clouds, mist, etc. and combinations of them.

One of the most frequently solved tasks from this area is the task for template matching. This task consists of finding approximately or fully coinsidence of the template T^M in the image I^N, i.e. coinsidence with any of sub-images I_i^M, as it is shown in fig. 1.

Two basic problems, which should be solved, are to achieve great precision and reliability of algorithms and in the same time this algorithms have to be fast. The best way to achieve a great precision and reliability is to use robust similarity measures and detailed search or to use so called "brute-force" method.

In generally algorithms, based on the "brute-force" method are with high computational complexity and respectively - they work too slowly. Thus they cannot be used in real-time applications.

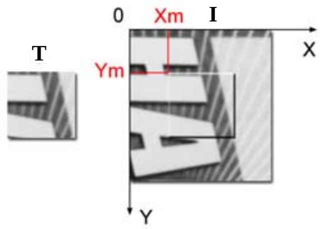

Fig. 1. Illustration of template matching problem

Many approaches are published for speeding up the algorithms for template matching. Some of them are: making computations in the frequency domain [4], [5], search around the most probable position [10], reducing the number of sub-windows for making computations [14], skipping the non-promising positions [8], [14], reducing sets of points taking part in comparisons [13], [17], using pyramids of images [6], [7], "coarse-to-fine" search [9], [11], [12], and many others.

These approaches are good enough for some applications, but they are not satisfactory for others. If we consider, for example the pyramidal approach, which is preferred when the goal is a high precision of template localization. When the best coincidence is found in the finest resolution, then it is possible to search for a subpixel accuracy of localization [7]. Another great advantage of the pyramidal approach is the fast search in the already built pyramids. The disadvantages are the slow building of pyramids and the probability to skip the position of the best coincidence.

Different similarity measures are used to achieve a great precision of template matching. Some of them are correlation measures - Sum of Absolute Differences, Sum of Squared Differences, Cross Correlation and Normalized Cross Correlation, measures, based on distances between sets of representative points - chamfer distance, Hausdorff distance, Frechet distance, etc.

In [1] for example authors use so called "sum of robust differences - SRD", which are computed, using M-estimators. The basic idea of M-estimators is to reduce the influence of outliers over the error of comparison [1], [2], [3]. M-estimators are generalizations of the usual maximum likelihood estimates.

In this paper we propose to combine using of M-estimators together with computations over sets of representative points and with algorithms for speeding-up the process of template matching.

[1]Yulka Petkova is with the Department of Computer Sciences and Technologies at Technical University of Varna, 1 Studentska Str., Varna 9010, Bulgaria, E-mail: jppet@abv.bg.

[2]Nuri Nuri is with the Department of Computer Sciences and Technologies at Technical University of Varna, 1 Studentska Str., Varna 9010, Bulgaria, E-mail: nuri.ismail@gmail.bg.

[3]Milena Karova is with the Department of Computer Sciences and Technologies at Technical University of Varna, 1 Studentska Str., Varna 9010, Bulgaria, E-mail: milena.karova@gmail.com.

The rest of the paper is organized as follows: Section II presents our proposition for representative points' selection. Section III presents some M-estimators, we have used in our work. In section IV the algorithm for fast template matching is described. Section V is about experimental results. We make conclusions in section VI.

II. SETS OF REPRESENTATIVE POINTS

A. Equipotential points

In [16] we propose a method for important points extraction, based on the criterion on D-optimality. According to this criterion the most informative points lie on the protruded peripheral wrapper of the object. If we interpret the image as a three-dimensional object, we propose the choice to be made by equipotential planes, which are parallel to the plane *xOy* and which cut the three-dimensional image profile (relief) on proper intensity levels. Thus, the extracted points outline the horizontal contours of the local "hollows" and "hills" from the three-dimensional image profile.

We define the following rule (Eq. 1) to extract representative points:

$$If \left(\begin{array}{l} \left(\begin{array}{l} [(PI(i,j) \le Pt_{max}) \cap (PI(i+1,j) > Pt_{max})] \cup \\ \cup [(PI(i,j) \ge Pt_{max}) \cap (PI(i+1,j) < Pt_{max})] \end{array} \right) \cup \\ \cup \left(\begin{array}{l} [(PI(i,j) \le Pt_{min}) \cap (PI(i+1,j) > Pt_{min})] \cup \\ \cup [(PI(i,j) \ge Pt_{min}) \cap (PI(i+1,j) < Pt_{min})] \end{array} \right) \cup \\ \cup \left(\begin{array}{l} [(PI(i,j) \le Pt_{max}) \cap (PI(i,j+1) > Pt_{max})] \cup \\ \cup [(PI(i,j) \ge Pt_{max}) \cap (PI(i,j+1) < Pt_{max})] \end{array} \right) \cup \\ \cup \left(\begin{array}{l} [(PI(i,j) \le Pt_{min}) \cap (PI(i,j+1) > Pt_{min})] \cup \\ \cup [(PI(i,j) \ge Pt_{min}) \cap (PI(i,j+1) < Pt_{min})] \end{array} \right) \end{array} \right) Then \quad (i,j) \in EP \tag{1}$$

In this equation $PI(i,j)$ is the intensity at point (i,j); Pt_{min} and Pt_{max} are intensity thresholds - low and high respectively; *EP* is the extracted set of representative points.

We use the three-dimensional image profile only to extract coordinates of representative points. In the next computations (for template localization) these points take part with their intensities.

B. Edge points

In [15] we propose a new edge definition, based on the discontinued first derivative of the intensity function. In practice at the points of discontinuance second derivative has local extreme values. The rule for determining if a point belongs to an edge (for one-dimensional case) is:

$$\text{If} \quad \left| \underset{i}{extr} \left(\frac{d^2 I(x)}{dx^2} \right) \right| > \theta, \quad \text{then} \quad i \in E, \tag{2}$$

where $I(x)$ is the intensity function; E is a set of points i, which are edge points; θ is a before settled threshold.

Te correspondence between edges and local extremes of the second derivative of intensity function is illustrated in Fig. 2.

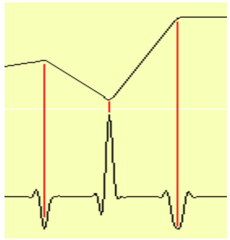

Fig.2 Correspondence between edges and extremas

The basic advantage of the proposed edge detector (we have named it TyPe ED) is that extracted edge points are closer to the human acceptance than these, selected with other edge detectors.

III. M-ESTIMATORS

A lot of M-estimators have been discussed in literature - Huber's (Eq. 3), Tukey's (Eq. 4), Geman's and McClure's (Eq. 5), Lorentzian's (Eq. 6), etc. [1], [2]. They use different types of influence functions.

In all the following equations ρ is the robust error measure, r is the difference between intensities of compared points and σ is a threshold, which can vary from 0 to 255.

- Huber's estimator:

$$\rho_1(r,\sigma) = \begin{cases} \dfrac{r^2}{2} & r \le \sigma \\ \sigma\left(r - \dfrac{\sigma}{2}\right) & r > \sigma \end{cases} \tag{3}$$

- Tukey's estimator:

$$\rho_2(r,\sigma) = \begin{cases} \dfrac{\sigma^2}{6}\left[1 - \left(1 - \left(\dfrac{r}{\sigma}\right)^2\right)^3\right] & r \le \sigma \\ \dfrac{\sigma^2}{6} & r > \sigma \end{cases} \tag{4}$$

- Geman and McClure estimator:

$$\rho_3(r,\sigma) = \frac{r^2}{r^2 + \sigma^2} \tag{5}$$

- Lorentzian's estimator:

$$\rho_4(r,\sigma) = \log\left(1 + \frac{1}{2}\left(\frac{r}{\sigma}\right)^2\right) \tag{6}$$

The similarity measure is "sum of robust differences - SRD" - Eq. 7, as it is in [11]:

$$SRD_{\rho,\sigma}(T,I) = \sum_{i=0}^{n}\sum_{j=0}^{n} \rho(r,\sigma) \tag{7}$$

IV. ACCELERATED ALGORITHM FOR TEMPLATE MATCHING

In this study we propose to combine specially extracted representative points and robust similarity measure together with an algorithm for fast template matching. This algorithm is earlier proposed by us [17] and it is based on the principle of coarse-to-fine search of the most probable position of the template into the examined image.

It consists of three steps. The first one is a fast determination of the region in which the probability for finding the template is the greatest. It is achieved by using a regular set of spread sub-windows into the image. The searched image is considered as a regular set of sub-windows, every one of them with a size $M/2^k$ * $M/2^k$, where M*M is the size of the template. The similarity measure is computed for these sub-windows. The position with coordinates of the upper left corner (x,y) with the greatest value of the similarity measure is taken to be the base for the next comparisons.

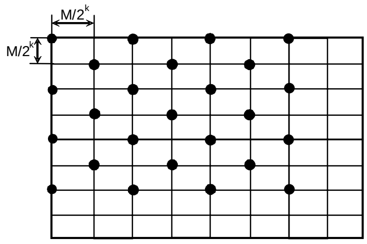

Fig.3 Sub-windows of the first step

The goal of the second step is the precise localization of the template. The search continues in a limited region around the determined during the first step the most probable position, in order to find the precise location of the template. This step is an iterative process which is executed for $p = k;\; p = p + 1;\; p < \mathrm{int}(\log_2 M)$. The similarity measure is computed at eight sub-windows with coordinates of the upper left corner, as it is shown in fig. 4. The position of the best coincidence is determined again and the coordinates of its upper left corner are (x, y).

$(x-M/2^p, y-M/2^p)$	…	$(x, y-M/2^p)$	…	$(x+M/2^p, y-M/2^p)$
…	…	…	…	…
$(x-M/2^p, y)$	…	-	…	$(x+M/2^p, y)$
…	…	…	…	
$(x-M/2^p, y+M/2^p)$	…	$(x, y+M/2^p)$	…	$(x+M/2^p, y+M/2^p)$

Fig.4 Positions of the left upper corners of sub-windows at which the similarity measure is computed

The third step is the final exact determination of the template location. The computations from the third step continue while the new determined coordinates coincide with those of the previous computations.

V. EXPERIMENTAL RESULTS

A lot of experiments have been made. They examine the precision of template localization and number of successful localizations as a part of all trials. The experiments are made in different conditions - different levels of noise and different transformations - rotation and scaling. We compare results, received by using equipotential points, edge points, extracted by TyPe edge detector and Sobel edge points and using M-estimators with the results, received by using the classical normalized cross correlation (NCC), computed over all possible points. All the examined images are 128 x 128 pixels and templates are 64 x 64 pixels in size. We examine the reliability presuming the accessible deviation of +/- 1 pixel from the right position.

We have examined all four M-estimators, encountered in section III. Huber's and Tukey's M-estimators have given the best performance and hence, the results from these two M-estimators are summarized and considered here.

The results in Table I present the successive template localizations in the image.

Table I
RELIABILITY IN THE PRESENCE OF NOISE [%]

Noise [%]	5	10	30	50	70
M-estimators with Equipotential Points	100	100	100	99	99
M-estimators with TyPe edge points	100	100	100	100	99
M-estimators with Sobel edge points	100	100	99	98	98
NCC over all points	100	100	100	100	100

It is seen that results, received by M-estimators, computed over sets of representative points and using the accelerated algorithm for template matching do not decrease the reliability of localization, even in the presence of noise.

Table II is about the precision of template matching. The same conditions of experiments have been kept.

Table II
TEMPLATE MATCHING IN THE PRESENCE OF NOISE - PRECISION

Mean Square Error [$*10^{-6}$]

Noise [%]	5	10	30	50	70
M-estimators with Equipotential Points	0,02	0,02	0,03	0,03	0,05
M-estimators with TyPe edge points	0,02	0,02	0,02	0,04	0,06
M-estimators with Sobel edge points	0,02	0,02	0,02	0,04	0,05
NCC over all points	0,005	0,01	0,01	0,02	0,02

We can make the same conclusion, considering these results - the derived precision is very close to the received by NCC, computed over all points.

We have examined the reliability of our approach over rotated images also. The received results are presented in Table III.

Table III
RELIABILITY IN THE PRESENCE OF ROTATION [%]

Rotation [°]	1	3	5	10	15	20	30
M-estimators with Equipotential Points	100	100	100	100	91	85	64
M-estimators with TyPe edge points	100	100	100	100	99	89	61
M-estimators with Sobel edge points	100	100	100	100	98	84	63
NCC over all points	100	100	100	100	95	88	42

The reliability of our approach over scaled images has been also examined. Table IV represents these results.

Table IV
RELIABILITY IN THE PRESENCE OF SCALING [%]

Scale factor	1,1	1,2	1,3	1,5	2	0,9	0,8
M-estimators with Equipotential Points	100	100	84	52	33	92	40
M-estimators with TyPe edge points	100	91	58	47	41	91	50
M-estimators with Sobel edge points	100	88	62	56	26	92	80
NCC over all points	96	82	48	34	22	90	40

For rotation and scaling we can summarize that the reliability, derived by M-estimators and proposed sets of representative points is better than this, received by Normalized Cross Correlation, computed over all points. It is due the fact, that the used sets of points represent objects in the best way, according to the mentioned criteria.

VI. CONCLUSION

To conclude we can point out the advantages of our approach:

- We use M-estimators only on sets of selected points. They can work on points selected by Equipotential, Sobel and TyPe Edge Detector algorithms.
- We do not need any time and memory consuming preprocessing like p-pyramids. We use filtering and gray level color enhancement only.
- M-estimators can be used with fast search strategies like "coarse-to-fine" search and they are still robust to noise and different type of transformations like scaling (up to 25%) and rotation (up to 20 degrees).
- The combination of M-estimators and "coarse-to-fine" search strategy results in a fast template matching algorithm with high level of robustness.
- Easy implementation and low computational complexity.

We can generalize that the received results are good enough, and we can use the proposed method to solve tasks for template matching in applications for which the received precision is satisfactory.

REFERENCES

[1] J-H. Chen, C-S. Chen, Y-S. Chen, "Fast Algorithm for Robust Template Matching with M-estimators", IEEE Transactions on Signal Processing, Vol. 51, Issue 1, Jan. 2003, pp. 230 - 243

[2] R. Dahyot, P. Charbonnier, F. Heitz, "Robust Visual Recognition of Colour Images", CVPR-2000 International conference on Computer Vision and Pattern Recognition - pp. 685-690 - Hilton Head Island, USA

[3] K.V. Arya, P. Gupta, P.K. Kalra, P. Mitra, "Image registration using robust M-estimators", Pattern Recognition Letters 28, pp. 1957–1968, 2007

[4] M. Uenohara, T. Kanade, "Use of Fourier and Karhunen-Loeve Decomposition for Fast Pattern Matching With a Large Set of Templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 19, No8, pp. 891-898, 1997

[5] B. Zitova, J. Flusser, "Image Registration Methods: A Survey", Image and Vision Computing, 21, pp. 977 – 1000, 2003

[6] S. Yoshimura, T. Kanade, "Fast Template Matching Based on the Normalized Correlation by Using Multiresolution Eigenimages", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No 12, pp. 2086-2093, 1993

[7] J. MacLean, J. Tsotsos, "Fast Pattern Recognition Using Gradient-Descent Search in an Image Pyramid", Proceedings of 15th International Conference on Pattern Recognition, Vol. 2, Issue , 2000, pp. 873 - 877

[8] Lee Chang-Hsing, Ling-Hwei Chen, "A Fast Motion Estimation Algorithm Based on the Block Sum Pyramid", IEEE Transactions on Image Processing, Vol. 6, No11, pp. 1587-1591, 1997

[9] Lai Shang-Hong, Ming Fang, "Robust and Efficient Image Alignment with Spatially Varying Illumination Models", IEEE, pp. 1063-6919, 1999

[10] K. Fredriksson, "Rotation Invariant Template Matching", PhD Thesis, Series of Publications A, ReportA-201-3, Helsinki, ISBN1238-645, ISBN952-10-09-0, 2001

[11] T. F. Cootes, G. J. Page, C. b. Jackson, C. J. Taylor, "Statistical Grey-Level Models for Object Location and Identification", Proceeding of British Machine Vision Conference, BMVA Press, pp. 533-542, 1995

[12] L. Chmielewski, D. Kozinska, "Image Registration", Proc. 3rd Polish Conference Computer Pattern Recognition Systems KOSYR 2003, pp. 163-168, 2003

[13] Chang Shih-Hsu, Fang-Hsuan Cheng, Wen-Hsing Hsu, Guo-Zua Wu, "Fast Algorithm for Point Pattern Matching: Invariant to Translations, Rotations and Scale Changes, Pattern Recognition", Vol. 30, No2, pp. 311-320, 1997

[14] C. Castillo, C. Chang, "An Approach to Vision-Based Person Detection in Robotic Applications", In 2nd Iberian Conference on PRIA, Vol. 2, pp. 209-216, 2005

[15] D. S. Tyanev, Y. P. Petkova, "About the Possibilities of a New Edge Definition", 28th ISSE 2005, 19th – 22nd May 2005, Wiener Neustadt, Austria, IEEE Catalog Number 05EX1142C, ISBN: 0-7803-9325-2, pp. 273 - 378

[16] D. Tyanev, Y. P. Petkova, "A New Method for Important Points Extraction", International Conference on Computer Systems and Technologies CompSysTech'2004, 17th – 18th June 2004, Rousse, Bulgaria, ISBN 954-9641-384, pp. IIIA.11-1 – IIIA11-8

[17] Y. Petkova, N. Nuri, "A Strategy for Fast Template Localization in an Image", Fourth International Conference "Computer Science" - 2008, Kavala, Greece, Sep. 2008, pp. 675 - 680.