Optimum Template Selection for Image Registration Using ICMM

Adnan A. Y. Mustafa
Department of Mechanical and Industrial Engineering
Kuwait University
P. O. Box 5969 - Safat
Kuwait 13060
symymus@kuc01.kuniv.edu.kw

Abstract

In our previous work on image registration we developed a novel image registration method, called the Intensity Combinatorial Minimization Method (*ICMM*), that has many appealing features. Two important features of this method is that ICMM is computationally efficient and has the unique feature of being invariant to the image processed by an injective function. In this paper we extend the use of *ICMM* for template matching. The extraction of an optimum template is investigated. Optimization of both template location and template size are addressed. We introduce the *Intensity Variation Number*, which is an image information measure that is strongly related to entropy. We show that optimization is a function of the Intensity Variation Number. Results of tests conducted on real images with noise are presented that support our theories.

1 Introduction

Image registration and template matching is an area of active research in computer and machine vision. The importance of registration and template matching are not limited to this field but are also important in many other fields because of their frequent occurrences. They arise in the domain of object recognition [1], multi-spectral image analysis [2], aerial image analysis [3], meteorology [4], medical imagery [5], and many more [6]. The main goal of the template matching process (which is at the core of image registration) is to find the translational, rotational and scaling offsets between the template and the image.

In our previous work on image registration [7] [8], we had developed an efficient area-based image registration method, called the *Intensity Combinatorial Minimization Method* (*ICMM*), that has many appealing features. In addition to being computationally efficient, where only simple calculations are required, the method has the unique feature of being invariant to the image (or the template) processed by an injective (one-to-one mapping) function. Tests conducted with this method produced good overall registration results indicating that this method is more efficient and robust than other traditional registration methods. In this paper we investigate the extraction of an optimum template. Optimization of both template location and template size are addressed. We will show

BMVC 1998 doi:10.5244/C.12.81

that optimization is a function of the *Intensity Variation Number,* which is an image information measure that is strongly related to entropy.

This paper is organized into six sections as follows:

1. Introduction
2. Literature Review
3. Description of the Intensity Combinatorial Minimization Method (*ICMM*)
4. Optimum Template Selection
5. Experimental Results
6. Conclusion and Future Work

2 Literature Review

Image registration methods can be classified into two classes; *Area-Based Methods* and *Feature-Based Methods*. In area-based methods, a sub-region of one of the images is taken as the registration template, which is swept over the other image to find it's location by maximizing (or minimizing) a similarity metric. *Area Correlation* is the most common area-based method employed for registration. The similarity metric used here is the area correlation of the two images. The location where the maximum correlation is found is taken as the translational registration offset between the template and the image. Although the *Area Correlation Method* is quite robust for most applications, it's main disadvantage is that it is computationally intensive. Another template matching technique is to subtract the template from the image, as the template is swept over the image, and observe the resulting matrix. The resulting matrix is then searched for the minimum which is taken as the point of registration. This method is called the *Sum of Absolute Difference Method*. Our method, the *Intensity Combinatorial Minimization Method* (*ICMM*) [7] [8], is based on a new metric for detecting similarity between the template and the image (discussed below in more detail).

In feature-based methods (also known as point-based methods), the images are initially processed by some feature extraction process where prominent features are extracted. Correspondences between extracted features of the two images determines the registration point. The method of using fiducials constitutes these methods in the most primitive form [9]. A point based method that matches a pattern of points were described in [10]. Here a simple approach to point matching is performed by measuring the displacement of feature points and counting the number of displacement occurrences. In [11], a relaxation point pattern matching method is used. A figure of merit is assigned to each pair of matched points according to how closely other pairs in the set match. Stockman, et al. [12] used a clustering approach to match image features between the image and the model, based on local evidence. Goshtasby [13] used normalized invariant moments to match images that are offset rotationally. Cox et al. [14] presented an iterative procedure matching line segments to find correspondences between the image and the template. Viola [6] recently aligned images by maximizing mutual information between the image and the model.

3 Description of the Intensity Combinatorial Minimization Method (*ICMM*)

The *Intensity Combinatorial Minimization Method* (*ICMM*) was described in [7] [8]. A brief description of the method and its features is given here. ICMM is a registration method based on a new metric for detecting similarity between images. The similarity measure is based on the simple fact that images of the same scene should appear similar. This implies that if an intensity of a point of an object (i.e. its grayscale value) in one scene changes to a different value in another scene, then all points that are at the same intensity (or most of them) should also map into the same new grayscale value (or a finite set of possible grayscale values). In other words, the variation of the intensity of all points with a certain grayscale value in the first image to that of the second image should be small.

3.1 ICMM Definition

ICMM locates a template match by sweeping the template over the image, and examining the overlaid pixels to calculate the registration matrix R. This is done by counting the number of distinct combination pairs of the intensity of the template and the image window. R is then searched for the minimum value which corresponds to the registration and match point. *ICMM* is formally defined by,

$$R(\mathbf{x}) = \sum_{\mathbf{y}} \Gamma\left(T(\mathbf{y}), I(\mathbf{x}+\mathbf{y}), S\right) \tag{1}$$

where $T(\mathbf{y})$ is the template intensity at location $\mathbf{y}$, $I(\mathbf{x})$ is the image intensity at location $\mathbf{x}$, and initially $S = \phi$. Γ is defined by,

$$\Gamma(u,v,S) = \begin{cases} 1 \text{ and } S = S \cup \{(u,v)\}, & if\ (u,v) \notin S \\ 0, & if\ (u,v) \in S \end{cases} \tag{2}$$

The registration point is then found at the point $\mathbf{x} = \{\mathbf{x} \mid \min\{R(\mathbf{x})\}\}$.

3.2 Features of the ICMM

ICMM exhibits several appealing features for registration applications: low noise sensitivity, invariance to injective functions processing, computationally efficient, good results when small templates are employed, a performance increase with an increase in the number of image intensities, as well as other features. The method's invariance to the image being processed by an injective function is unique.

4 Optimum Template Selection

In this paper we investigate two important aspects of template optimization; size and location. The problem of template size selection for template matching (and image registration) is of major importance as it dictates the time required for matching (and registration). Matching may fail if a small template is employed due to insufficient matching information. Matching may also fail if a large template is employed, where the template would not fully fit in the image except at a few number of locations. While a template ½ the size of the image might seem to be a suitable compromise, it really is not, as it results in the most computations required for registration.

Given a template size of $k \times l$ and an image size of $m \times n$, the number of operations (N) required for sweeping the template across the image is: $N = kl(m-k+1)(n-l+1)$. N attains a minimum at the edges, ($k = l = 0$, $k = m+1$ and $l = n+1$). The maximum value of N occurs at the midpoint ($k = ½(m+1)$, $l = ½(m+1)$). Since there is more freedom and control in selecting a small template size than a large template size, the optimum template size is thus the smallest possible template size that will produce adequate matching results. Template location selection is critical since the template must be taken in a region of the image that contains "a lot of information", as this will produce more accurate results. The definition of "a lot of information" is a template that contains sufficient detail for it to be found (uniquely if possible) in the other image.

4.1 The Intensity Variation Number

The *image intensity variation number* (Π) is a measure of image information that -we show below- is strongly related to entropy. For grayscale images, Π is measured by counting the number of distinct intensity values in the image[‡],

$$\Pi = \sum_{\mathbf{x}} \Gamma_o\big(I(\mathbf{x}), S\big) \tag{3}$$

where $I(\mathbf{x})$ represents the image intensity value at location $\mathbf{x}$. Initially $S = \phi$, and

$$\Gamma_o(u,S) = \begin{cases} 1 \text{ and } S = S \cup \{u\}\,, & \textit{if } u \notin S \\ 0, & \textit{else} \end{cases} \tag{4}$$

Π takes on values between 0 and L-1, where $L = 2^b$ is the maximum number of intensity levels in the image and b is the number of bits used to represent the image. Thus for 8-bit grayscale images, $255 \geq \Pi \geq 0$. The normalized *intensity variation number* ($\overline{\Pi}$) is defined by,

$$\overline{\Pi} = \frac{1}{L-1}\Pi \tag{5}$$

Values of $\overline{\Pi}$ lie between 0 and unity. The significance of $\overline{\Pi}$ is that as $\overline{\Pi}$ becomes large and approaches unity, indicating the image has more distinct gray levels appearing in the image, the possibility of more information contained in the image increases.

The *intensity variation neighborhood number* (π) is a local image information measure. For a $n \times n$ image neighborhood size, π_n is defined by counting the number of distinct intensity values in the neighborhood. Thus, at any image point,

$$\pi_n(\mathbf{x}) = \sum_{\mathbf{y}} \Gamma_o\big(I(\mathbf{x}+\mathbf{y}), S\big) \tag{6}$$

The maximum π_n value found in the image, for a $n \times n$ image neighborhood size, is defined as the *image intensity variation neighborhood number* (Π_n),

$$\Pi_n = \max(\pi_n) \tag{7}$$

For grayscale images, Π_n represents the maximum number of distinct intensity values in any $n \times n$ image neighbourhood.

Similarly the normalized intensity variation neighborhood number ($\overline{\pi}_n$) and the normalized image intensity variation neighborhood number ($\overline{\Pi}_n$) are defined, respectively by,

$$\overline{\pi}_n = \frac{1}{L-1}\pi_n \quad \text{and} \quad \overline{\Pi}_n = \frac{1}{L-1}\Pi_n \tag{8}$$

Note that for any image, a value of $\overline{\Pi} = 1.0$ indicates that all possible intensity levels are present in the image, and

$$\Pi \geq \Pi_n \geq \pi_n \text{ and } \overline{\Pi} \geq \overline{\Pi}_n \geq \overline{\pi}_n \tag{9}$$

[‡] For color images, Π is measured by counting the number of distinct colors in the image.

4.2 The Intensity Variation Number and Entropy

Image entropy (H) is a measure of the information embedded in a given image and is measured by quantifying the image intensity histogram,

$$H = -\sum_{i=0}^{L-1} h_i \log_2(h_i) \text{ bits/image} \tag{10}$$

where h denotes the image intensity histogram and L is the number of gray levels in the image. The *Intensity Variation Number* (Π) can also be described as a function of the histogram,

$$\Pi = \sum_{i=0}^{L-1} \Gamma_1(h_i) \quad \text{where} \quad \Gamma_1(u) = \begin{cases} 0, & \textit{if } u = 0 \\ 1, & \textit{else} \end{cases} \tag{11}$$

Hence, a great correlation exists between image entropy and the image intensity variation number. However, using the image intensity variation number has the advantage of being easier to calculate than entropy.

A similarity between local entropy and the intensity variation neighborhood number also exists. Local entropy is a measure that measures the amount of information present in an image sub-region. For a $n \times n$ image neighborhood size, the local entropy (H_n) at location $\mathbf{x}$ is measured by,

$$H_n(\mathbf{x}) = -\sum_{i=0}^{L-1} h_i^* \log_2(h_i^*) \quad \text{bits/image} \tag{12}$$

where, h^* denotes the image intensity histogram in the neighborhood of $\mathbf{x}$. Similarly, the intensity variation neighborhood number (π) can also be described in terms of the local image histogram,

$$\pi_n(\mathbf{x}) = \sum_{i=0}^{L-1} \Gamma_1(h_i^*) \tag{13}$$

4.3 Optimum Template Location and Size

Hence from above we arrive at the following conclusions:

i. For a given template size ($n \times n$), the optimum template location is at the image location that maximizes the number of distinct intensity values, i.e. at $\overline{\Pi}_n$.
ii. The optimum $n \times n$-template size for a given image is the smallest template size that maximizes $\overline{\Pi}_n$.

The above criteria applies in the absence of noise. However, in the presence of noise, template size selection is dictated by the amount of noise present in the image, as there is no precise method which can measure the amount of true information present in the image.

5 Experimental Results

Tests were conducted on the ten image scenes shown in Figure 1. Some important properties of these images are summarised in Table 1. The images differ in size and in the amount of information available. Entropy values range from 5.684 bits up to 7.862 bits. Π values vary from 124 distinct intensity levels to 253 intensity levels. Figure 2 shows a line chart displaying the images at their corresponding Π values.

5.1 Optimum Template Location

Figure 3 shows $\overline{\Pi}$ plotted against image entropy for the ten images. The correlation coefficient between the two measures are very high (0.81) indicating that $\overline{\Pi}$ can indeed be used as an information measure. This can be exemplified further by calculating Π_n for different n values and calculating the maximum neighborhood entropy values for these points and examining the relation between the two measures. Figure 4 which is a plot of the maximum neighborhood entropy vs. neighborhood entropy at points of Π_n for n = {5, 7, 9, 11, 13, 15, 18, 20, 24, 28, 32}. We see that locations that maximize the number of distinct intensity levels for a given sub-image size (i.e. Π_n) do indeed correspond to locations of maximum entropy. Hence, the optimum template location corresponds to the location of Π_n as this maximizes the amount of information present.

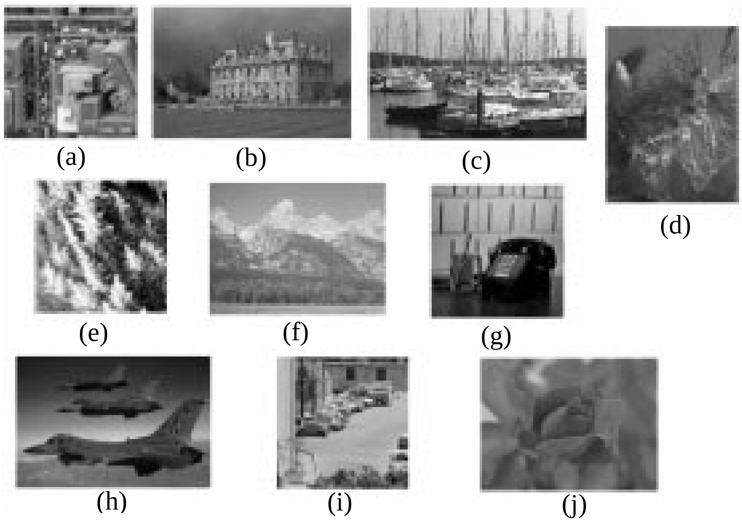

Figure 1: Test images: (a) *Air*, (b) *Building*, (c) *Boat*, (d) *Flower*, (e) *Forest*, (f) *Mountain*, (g) *Telephone*, (h) *Airplane*, (i) *Park*, (j) *Rose*.

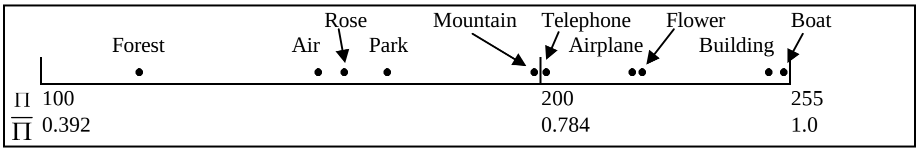

Figure 2: A line chart displaying the test images at their corresponding Π values.

Table 1: Test Image properties					
Image	Size	Entropy	Π	$\overline{\Pi}$	$\log_2(\Pi)$
Air	128×128	6.872	159	0.624	7.313
Building	384×256	7.054	251	0.984	7.972
Boat	420×256	7.862	253	0.992	7.983
Flower	256×341	6.812	222	0.871	7.794
Forest	128×128	5.684	124	0.486	6.954
Mountain	341×256	7.161	199	0.780	7.637
Telephone	128×128	7.366	202	0.792	7.658
Airplane	376×256	7.381	219	0.859	7.775
Park	128×128	6.753	172	0.675	7.426
Rose	341×256	6.391	163	0.639	7.349

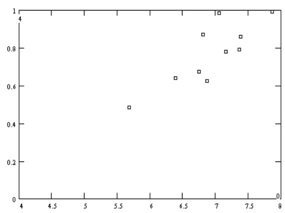

Figure 3: $\overline{\Pi}$ vs. image entropy for the ten test images.

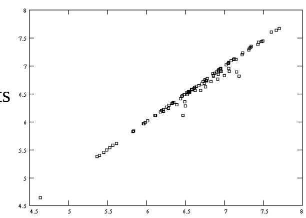

Figure 4: Variation of the maximum neighborhood entropy vs. neighborhood entropy at points of Π_n.

5.2 Optimum Template Size

In this section we study the extraction of the optimum template location and size under two conditions; when no noise is present and in the presence of noise.

5.2.1 Noise Free Images

To investigate the optimum template size, templates of different sizes were studied. Eleven $n \times n$ templates of sizes $n = \{5, 7, 9, 11, 13, 15, 18, 20, 24, 28, 32\}$ were extracted at points of Π_n for each image. These templates were matched to the original images to see if they matched correctly to the original images. The results obtained show that correct match was not found for any template when $n = 5$, and all templates of sizes $n \geq 9$ matched correctly. These results can be explained by examining Figure 5, a plot of the ratio of the intensity variation neighborhood number to the image intensity variation number (Π_n/Π). Correlating this plot with the above results, indicates that templates matched correctly when values of Π_n/Π were larger than 0.25, while

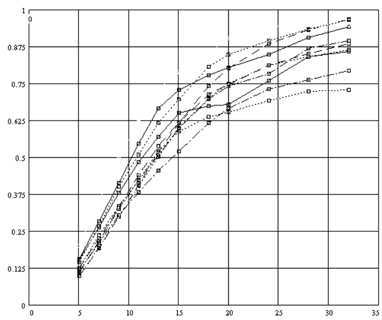

Figure 5: A plot of Π_n/Π for different template sizes.

templates failed to matched correctly when Π_η/Π were smaller than 0.25. Hence, sufficient data for matching is possible only when 25% or more of the actual number of intensity levels of the original image are present in the template.

5.2.2 Noise Images

To investigate the optimum template size in the presence of noise, tests were conducted on noisy variation of the original test images. White noise was added to the original images in noise quantities from 10% to 90% in 10% increments[†]. At each noise level 10 different noise images were generated for each original image, producing 90 noise variations. The templates employed in section 5.2.1 above were then matched to the noise images using *ICMM*.

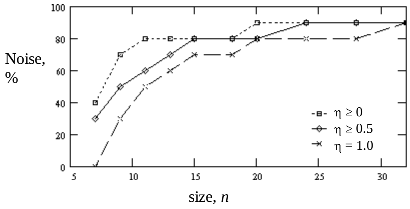

Figure 6: Curves of constant η for the *Air* image.

To measure the goodness of matching at each noise level, a matching performance metric (η) is employed to measure the matching efficiency at each level. η is defined as the ratio of the number of noise images with correct template matches to the total number of noise images. The matching results for the *Air* image are shown in Figure 6 where curves of constant η, for $\eta \geq 0$, $\eta \geq 0.5$ and $\eta = 1.0$, are displayed. At a template size of $n = 5$ the template did not match correctly. At $n = 7$, a correct template match was possible when no noise was present, but when noise was present, performance degraded with increase noise level: at 10% and 20% noise level, the matching performance (η) was 0.8 (i.e. the correct match was found in 8 of the 10 image noise variations at this noise level). At a 30% noise level η degraded further to 0.5, and at a 40% noise level $\eta = 0.2$. At noise levels above 50%, matching was not possible (η=0.0). At $n = 9$, η has better performance: $\eta = 1.0$ at noise levels up to 30%, decreases gradually at noise levels between 40% and 70%, and matching fails at noise levels of 80% and beyond. η continues to increase as n increases, and at a value of n=20, template matching success is possible at a noise level of 90% ($\eta = 0.1$). At $n = 24$, further improvement in η is evident: at a noise level of 90% η has increased to 0.8. Finally at a value of $n = 32$, correct template matching is possible at all noise levels.

Similar analysis can be derived for the remaining images. Table 2 shows common values of η for all images. From this table, several important observations can be stated about the overall matching results:

1. The matching performance indicates two distinct image groups. The first group consists of the *Air*, *Forest*, *Park* and *Rose* images. Successful template matching for these images start at a template size $n = 7$, and by $n = 32$, $\eta = 1.0$ at all noise levels. The second group consists of the remaining six images (*Building*, *Boat*, *Flower*, *Mountain*, *Telephone*, and *Airplane*). The matching performance for the second group lags behind the first group where successful template matching starts at size n=9, and by $n = 32$, $\eta = 1.0$ at noise levels up to 80%. This difference in performance of the two image groups can be explained by referring back to Figure 2

[†] The noise percentage refers to the percentage of image pixels that had noise added to it.

and examining $\overline{\Pi}$ for the various images. All images of the first group have 0.675 $\geq \overline{\Pi}$ while all images in the second group have $\overline{\Pi} \geq 0.780$. Hence, a template of a given size, *n*, from the first group will have a greater $\overline{\pi}$ value than those of the second group and a higher rate of successful matching.

2. Template sizes of $n \geq 18$ have $\eta = 1.0$ for all noise levels up to 50%.
3. $\eta > 0$ for all noise levels when $n \geq 20$, for all images except the boat image (where for the boat image $\eta > 0$ for all noise levels only when $n \geq 28$. This is explained by noting that this image has the largest $\overline{\Pi}$, where $\overline{\Pi} = 0.992$).
4. It is obvious that the matching performance is sensitive to the noise level content. By categorising the noise content into categories based on noise, we can arrive at general conclusions (see Figure 7):
   - No noise (noise level 0%): The optimum template size for this level is $n = 9$, as this size assures correct matching every time ($\eta = 1.0$).
   - Low noise content (30% ≥ noise level > 0%): The optimum template size for this level is $n = 15$, as this size assures correct matching every time ($\eta = 1.0$).
   - Medium noise content (60% ≥ noise level > 30%): The optimum template size for this level is $n = 20$, as this size assures correct matching every time (η=1.0).
   - High noise content (80% ≥ noise level > 60%): The optimum template size for this level is $n = 24$, as this size assures correct matching every time ($\eta = 1.0$).
   - Ultra-high noise content (noise level > 80%): The optimum template size for this level is $n > 32$. A template size larger than 32 will have $\eta \geq 0.7$ and will most likely be close to $\eta = 1.0$.

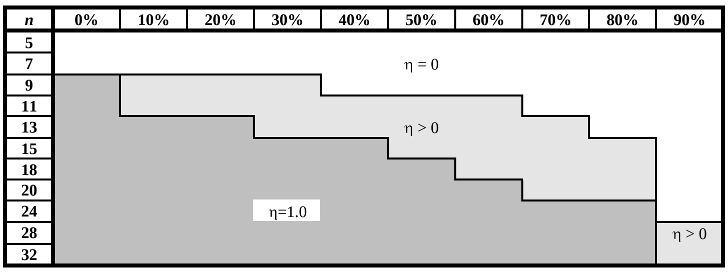

Figure 7: Overall Matching Results for all Test Images

Table 2: Optimum Template Size for Various Noise Contents					
Noise Content	**None (%0)**	**Low (>0%–30%)**	**Medium (>30%–60%)**	**High (>60%–80%)**	**Ultra-high (> 80%)**
Template size, *n*	9	15	20	24	32

6 Conclusions and Future Work

The *image intensity variation number* (Π) is a measure of image information that is strongly related to entropy. For grayscale images, Π is measured by counting the number of distinct intensity values in the image. Similarly, the *intensity variation neighborhood number* (π) is a local image information measure. For a $n \times n$ image neighborhood size, π_n is defined by counting the number of distinct intensity values in the neighborhood.

The optimum template location corresponds to the location that maximizes the amount of image information for a given sub-image size. This location is the location of the *image intensity variation neighborhood number* (Π_n). The optimum template size is dependent on the amount of noise present in the image. This was verified by conducting tests on 910 images from ten real image. The results indicated that when no noise was present a template size of $n = 9$ (i.e. 9×9), was sufficient for correct template matching and registration. When the noise content in the image was low, a template size of $n = 15$ was necessary for matching. When the noise content in the image was at a medium level, a template size of $n = 20$ was necessary for correct matching. Images with higher noise contents required a template size of $n = 24$ for correct template matching, and images with ultra-high noise content required a template size of at least $n = 32$. Our future work will concentrate on extending the *ICMM* method further by concentrating on an important aspect: scale and rotation invariance.

Acknowledgment

The author would like to thank the *Kuwait University Research Administration* for supporting this research through grant EM-099.

References

[1] Robert Haralick and Linda G. Shapiro. *Computer and Robot Vision*, Addison Wesley, Reading, Massachusetts, 1993.

[2] Paul Anuta. "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques", *IEEE Transaction on Geoscience Electronics,* GE-8 N 4, Oct., 1970.

[3] A. Goshtasby. "Template Matching in Rotated Images", *IEEE Trans. Pattern Analysis and Machine Intelligence,* V7, N3, pp. 338-344, 1985.

[4] J. Leese and C. Novak. "An Automated Technique for Obtaining Cloud Motion from Geosynchronous Satellite Data Using Cross Correlation", *Journal of Applied Meteorology*, V 10, February, 1971.

[5] Paul Viola and William Wells III. "Alignment by Maximization of Mutual Information". *International Journal of Computer Vision*, 24 (2), pp. 137-154, 1997.

[6] Lisa Brown. "A Survey of Image Registration Methods", *ACM Computing Surveys*, V 24, N 4, pp. 325-376, 1992.

[7] Adnan A. Mustafa and Mark Ganter. "Efficient Image Registration by Intensity Combinatorial Minimization". In *Vision Interface '94*, pp. 163-171, Banff, Alberta, Canada, May 18-20, 1994.

[8] Adnan A. Mustafa and Mark Ganter. "An Efficient Image Registration Method by Minimizing Intensity Combinations". In *Research in Computer and Robot Vision,* Archibald, C. and Kwok, P. (Eds.), World Scientific Press, Singapore, 1995.

[9] C. Bose and I. Amir. "Design of Fiducials for Accurate Registration Using Machine Vision", *IEEE Transaction on Pattern Analysis and Machine Intelligence,* V-12, N 12, December, pp. 1196-2010, 1990.

[10] D. Kahl, A. Rosenfeld and A. Danker. "Some Experiments in Point Pattern Matching", *IEEE Trans. Syst., Man, and Cybern.,* V10, N2, pp. 0.15-116, 1980.

[11] S. Ranade and A. Rosenfeld. “Point Pattern Matching in Relaxation”, *Pattern Recognition*, V12, pp. 269-75, 1980.

[12] G. Stockman, S. Kopstein, and S. Benett. “Matching Images to Models for Registration and Object Detection via Clustering”, *IEEE Trans. Pattern Analysis and Machine Intelligence*, V4, N3, pp. 229-241, 1982.

[13] A. Goshtasby, G. Stockman and C. Page. “A Region Based Approach to Digital Image Registration with Subpixel Accuracy”, *IEEE Trans. Geoscience and Remote Sensing,* V24, N3, pp. 390-399, 1986.

[14] I. Cox, J. Krusal and D. Wallach. “Predicting and Estimating the Accuracy of a Subpixel Registration Algorithm”, *IEEE Transaction on Pattern Analysis and Machine Intelligence*, V-12, N 8, August, pp. 721-734, 1990.