Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 28 (2007) 1957–1968
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Image registration using robust M-estimators

K.V. Arya [a,*], P. Gupta [b], P.K. Kalra [c], P. Mitra [d]

[a] *Department of Computer Science & Information Technology, M. J. P. Rohilkhand University, Bareilly 243006, India*
[b] *Department of Computer Science & Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India*
[c] *Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur 208016, India*
[d] *Department of Computer Science & Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

Received 1 July 2005; received in revised form 28 April 2007
Available online 24 May 2007

Communicated by M. Kamel

Abstract

In this paper, a method for robust image registration based on *M-estimator Correlation Coefficient* (*MCC*) is presented. A real valued correlation mask function is computed using Huber and Tukey's robust statistics and is used as a similarity measure for registering image windows. The mask function suppresses the influence of outlier points and makes the registration algorithm robust to noisy pixels, brightness fluctuations and presence of occluding objects. The superiority of the proposed algorithm, in terms of registration performance and computation time is demonstrated through experimental studies on different types of real world images.

Keywords: Robust statistics; M-estimator; Huber statistics; Tukey statistics; Normalized cross-correlation; Influence function

1. Introduction

Many image processing applications often need to compare or combine information given by multiple images. To perform this task, image registration is one of the fundamental steps. Image registration is the process of determining correspondence between all the points in two images of the same scene. One of the images used for registration is kept unchanged and is referred as reference or template image while the other one can be warped, and is called the sensed or target image.

Template matching is a popular method for registering objects, symbols, characters and faces due to the simplicity of implementation. Template matching is the process of finding location of a sub-image, called template image, inside a given image. It involves determining the similarities between a given template and windows of the same size in the target image, and then identifying the window that produces the highest similarity measure. Registration in real world images involves problems of noisy environment and shadow or occlusion in the image scene. Robustness is an important property required for successful registration in above environments.

Many feature-based robust registration methods have been proposed in literature (Brown, 1992; Dai and Khorram, 1999; Ghaffary and Sawechuk, 1983; Zitova and Flusser, 2003). These methods try to match image features, e.g., lines, corners, contours, between the target and the reference image. Clustering technique presented by Goshtasby et al. (1986) and Stockman et al. (1982) attempt to match points connected by line segments. Barrow et al. (1977) introduces the chamfer matching for image registration, where line features detected in the two image are matched by minimizing the distance between them. Borgefors (1988)

* Corresponding author. Tel.: +91 581 2520003; fax: +91 581 2528384.
E-mail addresses: kvarya@gmail.com (K.V. Arya), pg@iitk.ac.in (P. Gupta), kalra@iitk.ac.in (P.K. Kalra), pabitra@cse.iitkgp.ernet.in (P. Mitra).

0167-8655/$ - see front matter
doi:10.1016/j.patrec.2007.05.006

presented a modified version of Barrow's work by applying the sequential distance transform together with the root mean square error. The angles between the relevant intersecting lines have been used by Zana and Klein (1999). Differential descriptors of the image function in the neighborhood of the detected control points has been presented by Montesinos et al. (2000). In (Freeman, 1974; Li et al., 1995; Saghri and Freeman, 1981) chain code representation of contour was described as an invariant descriptor for registration.

Other methods use the invariant shape descriptors of the image regions. Sester et al. (1998) proposed the use of elongation parameter, compactness, number of holes etc. Flusser and Suk (1994) have described moment based invariant approach for image registration. Registration of images with geometric distortion is presented in (Flusser and Suk, 1994; Goshtasby, 1988). Shekhar et al. (1999) has combined different types of features and their descriptors for image registration. A cross-entropy based similarity measure called mutual information is described in (Wells et al., 1996; Maes et al., 1996; Viola and Wells, 1997; Pluim et al., 2000; Roche et al., 2000; Zhu, 2002).

Feature computation in the above feature-based methods is often computationally demanding. In (Brown, 1992) a family of methods, based on correlation of intensity values in two images, has been proposed for fast image registration. The sum of absolute or squared differences (*SSD*) (Barnea and Silverman, 1972) based methods compute the sum of absolute or squared differences in pixel brightness between reference image and target image. This technique is very sensitive to occlusion. Normalized cross-correlation (*NCC*) between pixel intensities can also be used for image registration. NCC computes the degree of similarity between reference image and target image for window size equal to that of template image and then pickup the position of best match. The NCC (Barnea and Silverman, 1972; Aggarwal et al., 1981) is effectively used for registration of the images having uniform brightness, but this technique does fail to pickup the correct position in case of partial occlusion in the given scene.

Venot et al. (1984) used the sum of number of sign changes between corresponding pixels in a pair of the images as a matching score. This method also cannot handle occlusion. Lai (2000) has given an efficient image matching algorithm for partial occlusion. Kaneko et al. (2002) has presented a similarity measure termed as increment sign correlation (*ISC*), which was shown to be robust for fluctuation in illumination and a class of occlusion. But this method fails if the occluding object is not of uniform brightness. A modified version of ISC called selective correlation coefficient (*SCC*) was proposed by Kaneko et al. (2003), where they define a mask function using increment information of brightness to filter out irrelevant pixel information. The limitation of this method is that in case of occlusion the 50 percent of the pixels in the occluded region still participate in the computation of correlation coefficient, and may lead to registration failure. The masking function used in calculation of SCC is binary in nature and hence does not represent the level of consistency between corresponding pixels in the two images. The efficiency of the method could be improved by defining the real valued mask coefficients.

In this paper an M-estimator correlation-coefficient based image registration algorithm has been proposed. It uses the principles of robust statistics (Black and Rangarajan, 1996; Huber, 1981; Rey, 1983; Hampel et al., 1986). It falls in the category of brightness based registration methods and is fast compared to known feature based methods (Aggarwal et al., 1981; Kaneko et al., 2003). The goal of this work is to circumvent the problem of sensitivity of brightness based methods to fluctuations in image brightness, noise and presence of occlusion. Real valued mask coefficients has been defined for this purpose using two M-estimators, namely, Huber's estimator (Huber, 1981) and Tukey's Bisquare estimator (Tukey, 1977).

M-estimators have been widely used in statistics literature (Huber, 1981; Rey, 1983; Hampel et al., 1986; Tukey, 1977), to separate the outlying noise in the data. M-estimator-based mask coefficient reduces the influence of noisy pixels and those in occluded regions. This leads to improved registration performance in real world images which is demonstrated through experiments on object, face and aerial images under different environmental conditions.

The rest of the paper is organized as follows. Section 2 presents the correlation based methods for image registration. The M-estimator based image registration method developed in this work is discussed in Section 3. The experimental results are analyzed in Section 4. Section 5 concludes the paper.

2. Correlation-based algorithms for image registration

In correlation-based methods, correspondence for a pixel in the template image is achieved by searching over windows of same sizes in the target image. This section discusses some of the existing correlation-based methods.

2.1. Sum of squared differences (SSD)

Let $F = \{F_1, F_2, \ldots, F_n\}$ and $f = \{f_1, f_2, \ldots, f_n\}$ be the one-dimensional lists of brightness values for same sized windows in the target and template image respectively. In (Barnea and Silverman, 1972) the sum of square of Euclidean distance between the corresponding pixels in the two images F and f is defined by SSD, where the size of template image window is n,

$$\mathrm{SSD} = \sum_{i=1}^{n} (F_i - f_i)^2.$$

Note that the value of SSD close to zero indicates the best match. The computation of SSD is efficient but is very sensitive to changes in image brightness due to occlusion and shading.

2.2. Normalized cross-correlation coefficient (NCC)

Normalized cross-correlation coefficient (NCC) (Barnea and Silverman, 1972; Aggarwal et al., 1981) is defined as

$$\mathrm{NCC} = \frac{\sum_{i=1}^{n}(F_i - \bar{F})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n}(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{n}(f_i - \bar{f})^2}},$$

where $\bar{F}$ and $\bar{f}$ are the mean brightness values in the target image and template image, respectively.

The SSD and NCC both estimate the degree of linear dependence between the corresponding pixel brightness values being compared. The absolute value of NCC lies between 0 and 1 where value closer to 1 indicates the better match. The NCC is preferred over SSD because it is invariant to linear brightness and contrast variations between matching windows. The SSD and NCC are equivalent for Normalized images. It is observed that the SSD and NCC both perform badly in the presence of non-linear pixel brightness variation due to illumination variation, occlusion and shadow.

2.3. Increment sign correlation coefficient (ISC)

Increment sign correlation algorithm (Kaneko et al., 2002) first converts the list of pixel brightness values to a list of corresponding binary codes $B = \{b_1, b_2, \ldots, b_{n-1}\}$ based on the brightness increment information. For the target image the binary codes (b_i^F) is defined as

$$b_i^F = \begin{cases} 1 & \text{if } F_{i+1} \geqslant F_i, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Similarly, the binary codes (b_i^f) for template image is defined as

$$b_i^f = \begin{cases} 1 & \text{if } f_{i+1} \geqslant f_i, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The Increment Sign Correlation Coefficient (ISC), between b_i^F and b_i^f is defined as follows:

$$\mathrm{ISC} = \frac{1}{n}\sum_{i=1}^{n}\{b_i^F b_i^f + (1 - b_i^F)(1 - b_i^f)\}.$$

2.4. Selective correlation coefficient (SCC)

Selective correlation coefficient SCC method (Kaneko et al., 2003) is an extension of NCC method with a masking function for corresponding pixels in both the images. Thus only some selected pixels contribute to similarity computation. Irrelevant pixels (e.g., in occluded region) are masked. SCC is defined as

$$\mathrm{SCC} = \frac{\sum_{i=1}^{n}c_i(F_i - \bar{F})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n}c_i(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{n}c_i(f_i - \bar{f})^2}},$$

where $\bar{F}$ and $\bar{f}$ have the same meaning as in Section 2.2, and the mask coefficient c_i represents the similarity of sign increment in the adjacent pixels in both the images. The mask coefficient is defined as the bitwise exclusive NOR of corresponding binary codes of the template and target image. Mask coefficient is represented as follows:

$$c_i = \begin{cases} 1 - |b_i^F - b_i^f| & \text{if } i = 0 \text{ or even}, \\ c_{i-1} & \text{if } i = \text{odd}, \end{cases}$$

where b_i^F and b_i^f are defined in (1) and (2) respectively.

3. Image registration using M-estimators

This section presents the definition of M-estimators followed by the robust image registration algorithms developed in this work. In Section 3.1, general definition of M-estimators has been given. The two M-estimators, viz Huber and Tukey, used in this study have also been described in Section 3.1. The two robust image registration algorithms using Huber's and Tukey's M-estimators have been given in Section 3.2.

3.1. M-estimators

M-estimators are generalizations of the usual maximum likelihood estimates (Huber, 1981; Rey, 1983; Hampel et al., 1986; Black and Rangarajan, 1996). Classically a parameter V is obtained by maximizing the likelihood function L, i.e. if x_i is the residual of ith data point, optimal parameter V^* is given by

$$V^* = \operatorname{argmax}\left(L = \prod_i f(x_i|V)\right)$$

or equivalently

$$V^* = \operatorname{argmax}\left(-\ln L = -\sum_i \ln f(x_i|V)\right).$$

The estimators of type M are solutions of the more general structure

$$V^* = \operatorname{argmin}\left(M = \sum \rho(x_i, V)\right),$$

where the function $\rho(\cdot)$ is a symmetric positive definite function with a unique minimum at zero and is chosen to be increasing slower than quadratically.

Instead of solving this problem directly, we can reformulate it as an iterated weighted least-square problem. That is, for estimating a parameter vector $V = [v_1, v_2, \ldots, v_n]^{\mathrm{T}}$, the M-estimator of V based on the function $\rho(x_i)$, is the solution of following n equations:

$$\sum \psi(x_i, V)\frac{\partial x_i}{\partial j} = 0 \quad \text{for } j = 1, 2, \ldots, n.$$

The derivative $\psi(x) = \mathrm{d}\rho(x)/\mathrm{d}x$ is called the *influence function*, which measures the influence of data point on the value of parameter estimate. For a robust estimator the influence of any single data point does not introduce any significant error (Rey, 1983). This makes it less sensitive to outliers. The M-estimator should fulfill the following constraints:

(1) Influence function should be bounded.
(2) Robust estimator should be unique i.e. the objective function of parameter vector V should have unique minimum. This requires that the individual ρ-function is convex in variable V.
(3) The gradient $\frac{\partial\rho(.)}{\partial V} \neq 0$, whenever $\frac{\partial\rho^2(.)}{\partial V^2}$ is zero.

Several M-estimators have been discussed in literature (Huber, 1981; Rey, 1983; Hampel et al., 1986; Tukey, 1977) using different types of influence functions. It is observed from the simulation results that Huber's and Tukey's M-estimators give the best performance and hence, these two M-estimators are considered for this work.

3.1.1. Huber's M-estimator

Huber's function is a parabola in the vicinity of zero, and increases linearly at a given level $|x| > k$. This helps in restricting influence of outliers. Using this estimator an asymptotic efficiency of 95% on the standard normal distribution is obtained with the tuning constant $k = 1.345\sigma$, where σ is estimated standard deviation of errors. From the simulation results, we observe that for the same value of tuning parameter k it is equally efficient on many non-normal distributions also. This estimator is found to be superior to most of the other estimators for a wide range of data. The $\rho(\cdot)$ function and the influence function $\psi(\cdot)$ for this M-estimator are given below

$$\begin{aligned}\rho(x) &= \begin{cases} x^2/2 & \text{if } |x| < k, \\ k(|x| - k/2) & \text{if } |x| \geqslant k, \end{cases}\\ \psi(x) &= \begin{cases} x & \text{if } |x| < k, \\ k\,\mathrm{sgn}(x) & \text{if } |x| \geqslant k. \end{cases}\end{aligned} \tag{3}$$

The nature of the ρ and the influence function ψ defined in the above equations are shown in Fig. 1.

3.1.2. Tukey's Bisquare M-estimator

The Tukey's Bisquare function (Tukey, 1977) can suppress the outliers even further. Using Tukey's Bisquare function, 95% asymptotic efficiency on the standard normal and non-normal distributions is obtained with the tuning constant $c = 4.6851\sigma$. The $\rho(\cdot)$ function and the influence function $\psi(\cdot)$ for this M-estimator are given below

$$\begin{aligned}\rho(x) &= \begin{cases} \frac{c^2}{6}(1 - [1 - (x/c)^2]^3) & \text{if } |x| \leqslant c, \\ \frac{c^2}{6} & \text{if } |x| > c, \end{cases}\\ \psi(x) &= \begin{cases} x[1 - (x/c)^2]^2 & \text{if } |x| \leqslant c, \\ 0 & \text{if } |x| > c. \end{cases}\end{aligned} \tag{4}$$

The nature of the ρ and the influence function ψ for Tukey's Bisquare estimator are shown in Fig. 2.

It can be seen from Figs. 1 and 2 that the ρ function for Tukey saturates for large values of x, whereas for Huber it increases linearly for all x. Through simulation it has been found that the best choice for the tuning constants k for Huber's M-estimator and c of Tukey's Bisquare M-estimator are $k = 1.345\sigma$ and $c = 4.685\sigma$ respectively for the estimated standard deviation σ. These choices of tuning constants matched with that given in (Rey, 1983).

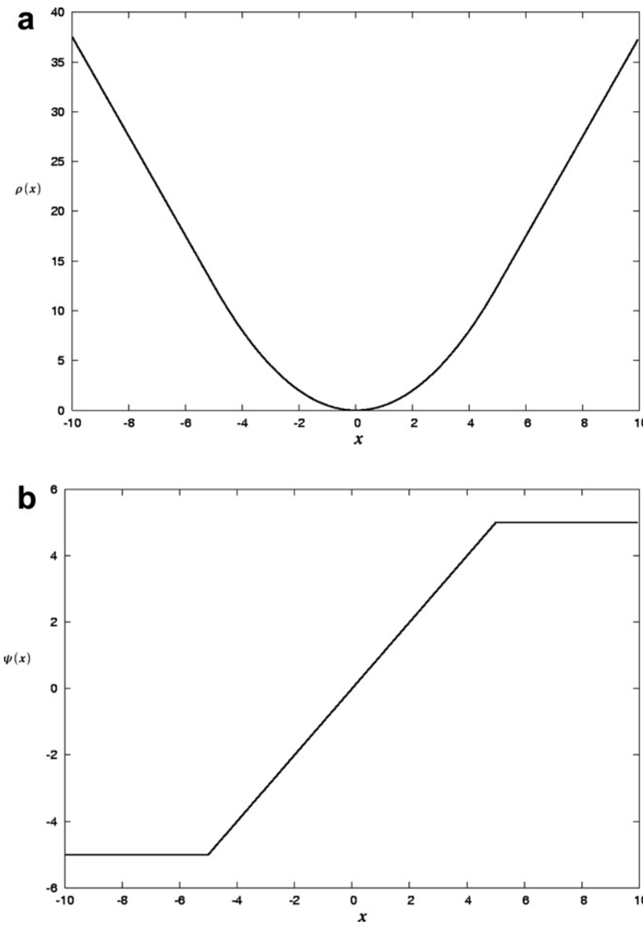

Fig. 1. Huber's M-estimator: (a) ρ-function, (b) influence function ψ.

3.2. Registration algorithms using M-estimators

Based on the two M-estimators, viz. Huber and Tukey, two registration algorithms, "ROBUST_IMAGE_MATCH_HUB" and "ROBUST_IMAGE_MATCH_TUK" have been designed. The algorithms choose windows of certain size in both the template and target images. Let $f = \{f_i\}_{i=1,2,\ldots,n}$ represents a one-dimensional list of intensity values of n pixel in the template image and $F = \{F_i\}_{i=1,\ 2,\ldots,n}$ be the similar list in the target image of the same size n, which is drawn from the scene. These images are represented in one-dimensional list for simplicity without loss of generality. In this paper the pixels intensity values are read row wise and then stacked in the natural order of the rows.

An *M-estimator correlation coefficient* (MCC) is computed between the pixel intensity values of these two windows using Huber's and Tukey's statistics. The masking function thus obtained is real valued and suppresses the influence of outliers i.e., noise pixels and pixel in occluded

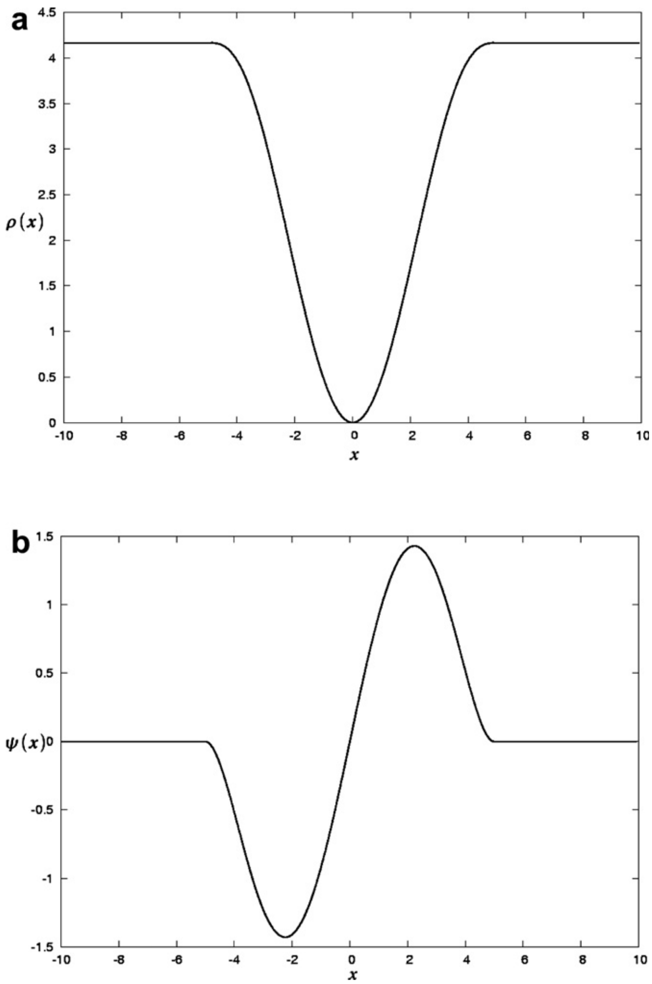

Fig. 2. Tukey's Bisquare M-estimator: (a) ρ-function, (b) influence function ψ.

regions to achieve robustness. The standard deviation σ in the proposed algorithms for computation of MCC is a critical parameter. In our experiments it is computed on-line during the search process adaptively for the pair of template image and equal portion of target image by considering the residual image. In residual image each pixel brightness value is replaced by the corresponding difference of original brightness value and mean brightness value. The M-estimator correlation coefficients give a measure of the degree of similarity between the two windows and the best matching window in the target image is returned and considered to be registered with the corresponding window in the template image. The occlusion has been taken care by using the concept of M-estimator on the residual image where the residual information is computed on-line during correlation computation. This process is detailed in Algorithms 3.1 and 3.2.

3.3. Search method

Exhaustive search with branch-and-bound approach is used for reducing the redundant MCC values. The basic procedure is based on an estimation and comparison of correlation value. The maximal value of MCC reports of the candidate position registered through scanning. At any new potential position value of MCC is ruled out or branched on the way if it is less than the maximal MCC value achieved in the previous iterations. This helps in reducing the number of correlation values to be compared finally to get the maximum value which indicates the best match position.

The method MCC presented in this paper exhibits better time complexity as compared to SCC. In MCC the pixels in the occluded, saturated or noisy region of the image contribute negligibly toward the computation cost because their influence on the correlation computation is reduced by the M-estimators. Therefore, compared to NCC less number of pixels participate in correlation computation. Hence, it has a lower time complexity. In SCC also half of the pixels in the occluded region do not participate in the correlation computation. But it involves an additional preprocessing step of converting the gray-scale image into a binary one. Preprocessing has to be done for both the template and target images and this preprocessing step takes a significant amount of time. As MCC does not undergo any such preprocessing, a lower computation time as compared to that of SCC is obvious.

Algorithm 3.1 ROBUST_IMAGE_MATCH_HUB

Let m_1^H and m_2^H represent the masks for template and target images respectively and f_i, F_i are defined as the pixel brightness values in template and target images respectively.

Step 1: Using the influence function of Huber's M-estimator defined in (3), compute the mask $m_1^H(i)$ $\forall$ residues $x_i = f_i - \bar{f}$ for each pixel $i = 1, 2, \ldots, n$ and mean brightness value $\bar{f}$ of the template image as follows:

$$m_1^H(i) = \begin{cases} x_i & \text{if } x_i < k_1, \\ k_1 \operatorname{sgn}(x_i) & \text{if } x_i \geqslant k_1, \end{cases}$$

where $k_1 = 1.345\sigma_x$ and σ_x is the standard deviation of template image brightness residues.

Step 2: Using the influence function in (3) compute the mask $m_2^H(i)$ $\forall$ residues $y_i = F_i - \bar{F}$ for each pixel $i = 1, 2, \ldots, n$ and mean brightness value $\bar{F}$ in the target image windows (of same size as template image) as follows:

$$m_2^H(i) = \begin{cases} y_i & \text{if } y_i < k_2, \\ k_2 \operatorname{sgn}(y_i) & \text{if } y_i \geqslant k_2, \end{cases}$$

where $k_2 = 1.345\sigma_y$ and σ_y is the standard deviation of target image brightness residues.

Step 3: Compute the *M-estimator correlation coefficient* MCC^H between the masked template and target image as follows:

$$\mathrm{MCC}^H = \frac{\sum_{i=1}^{n} m^H(i)(F_i - \bar{F})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n} m_2^H(i)(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{n} m_1^H(i)(f_i - \bar{f})^2}},$$

where $m^H(i) = \sqrt{m_1^H(i) * m_2^H(i)}$.

Step 4: Return the window position having the maximum value of MCC^H, indicating the best match.

Algorithm 3.2 ROBUST_IMAGE_MATCH_TUK

Let m_1^T and m_2^T represent the masks for template and target images respectively and f_i, F_i are defined as the pixel brightness values in template and target images respectively.

Step 1: Using the influence function of Tukey's M-estimator defined by (4), compute the mask $m_1^\text{T}(i)$ $\forall$ residues $x_i = f_i - \bar{f}$ for each pixel $i = 1, 2, \ldots, n$ and

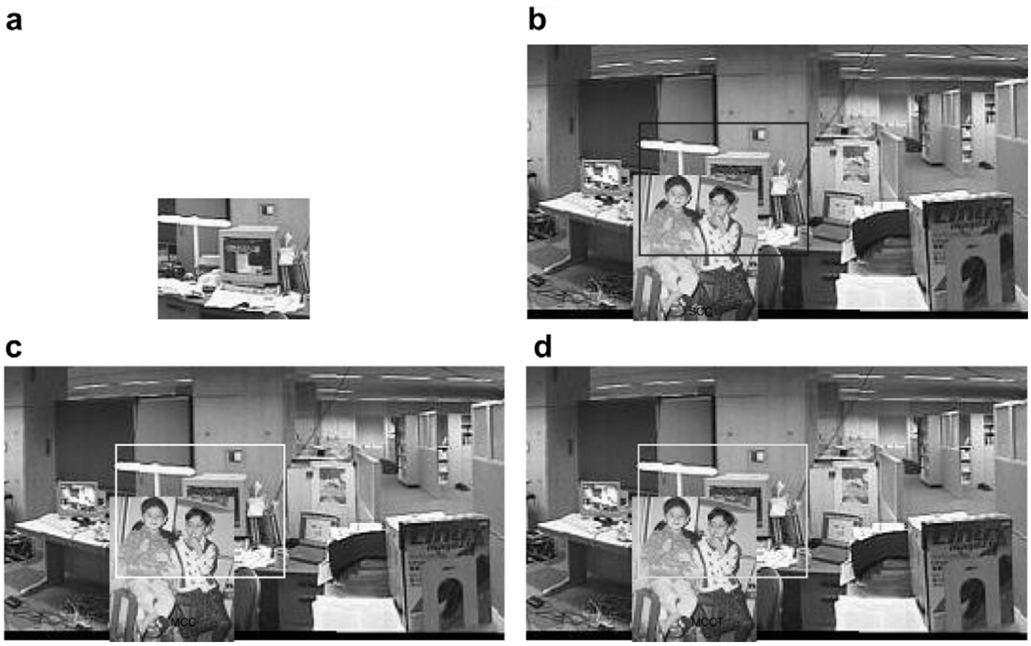

Fig. 3. Object image: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber estimator, (d) target image registered using MCC with Tukey's estimator. The occlusion (40%) has been artificially generated.

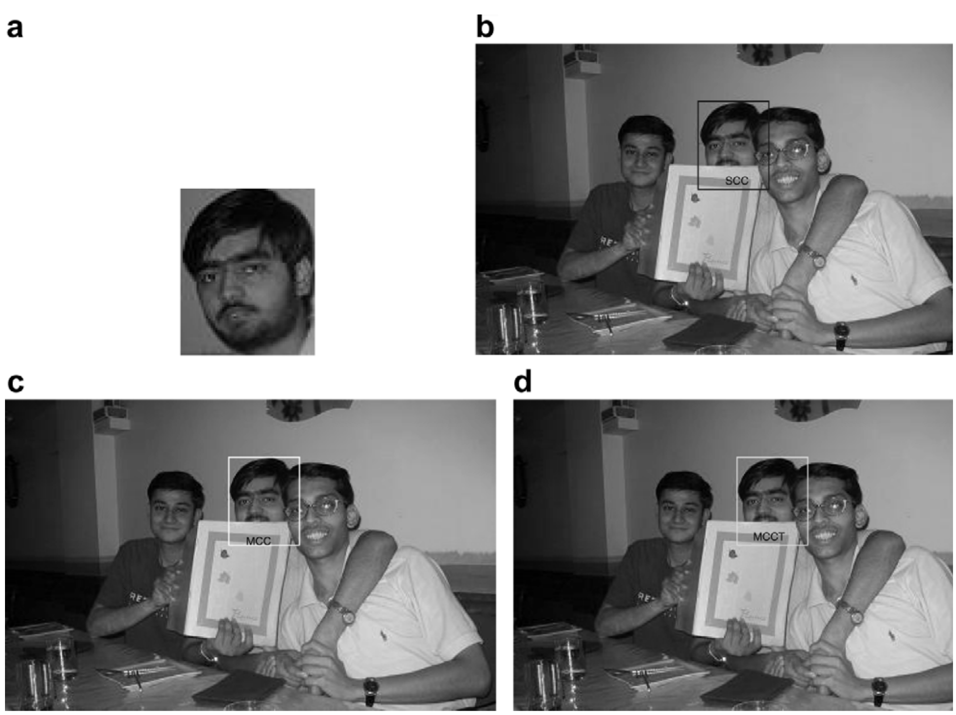

Fig. 4. Face image: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber estimator, (d) target image registered using MCC with Tukey's estimator. The occlusion (40%) is natural.

mean brightness value $\bar{f}$ of the template image as follows:

$$m_1^{\mathrm{T}}(i) = \begin{cases} x_i[1-(x_i/c_1)^2]^2 & \text{if } x_i \leqslant c_1, \\ 0 & \text{if } x_i > c_1, \end{cases}$$

where $c_1 = 4.685\sigma_x$ and σ_x is the standard deviation of template image brightness residues.

Step 2: Using the influence function in (4), compute the mask $m_2^{\mathrm{T}}(i)$ $\forall$ residues $y_i = F_i - \bar{F}$ for each pixel $i = 1, 2, \ldots, n$ and mean brightness value $\bar{F}$ in the target image windows (of same size as template image) using following expression:

$$m_2^{\mathrm{T}}(i) = \begin{cases} y_i[1-(y_i/c_2)^2]^2 & \text{if } y_i \leqslant c_2, \\ 0 & \text{if } y_i > c_2, \end{cases}$$

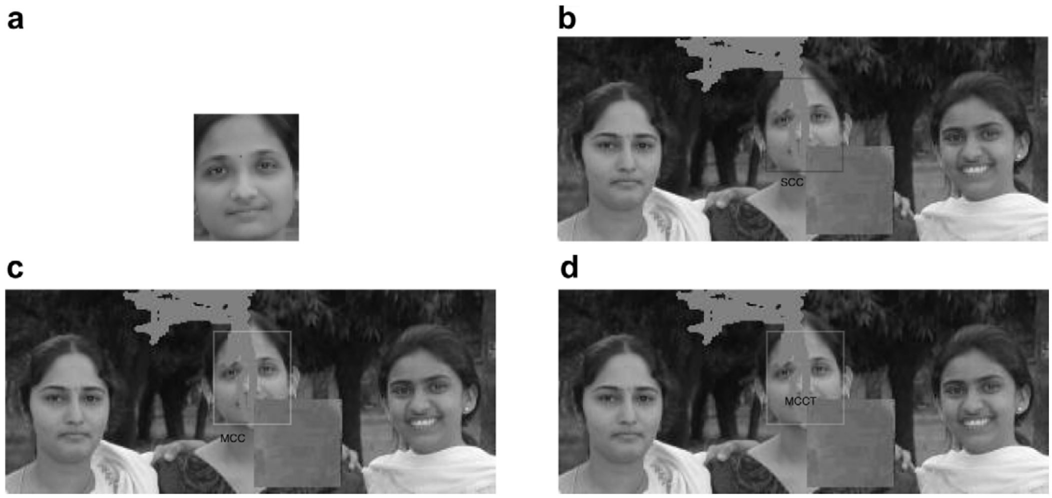

Fig. 5. Face image: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber's estimator, (d) target image registered using MCC with Tukey's estimator. The occlusion (30%) has been artificially generated.

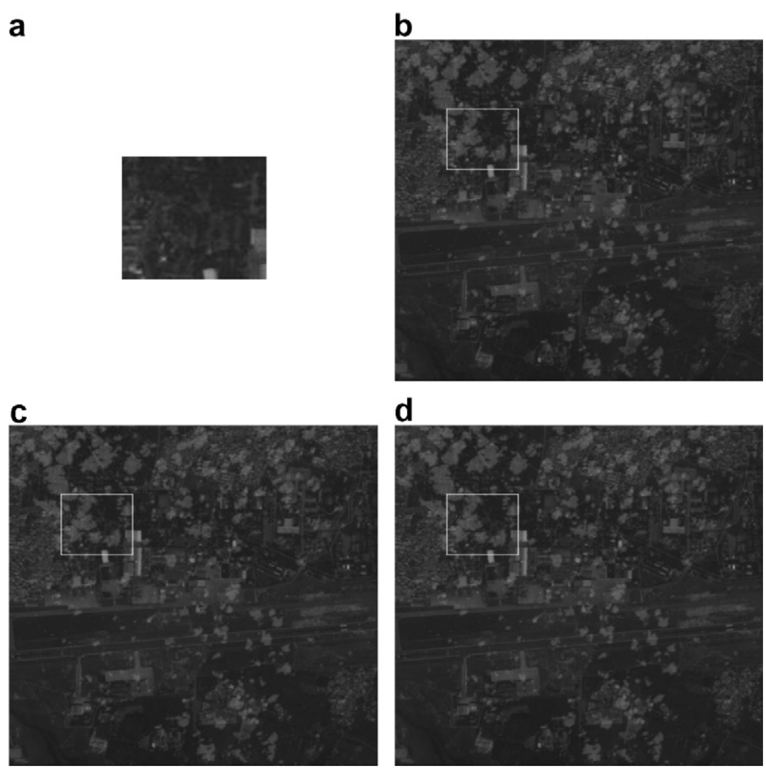

Fig. 6. Satellite image: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber's estimator, (d) target image registered using MCC with Tukey's estimator. The target image contains noise patches.

where $c_2 = 4.685\sigma_y$ and σ_y is the standard deviation of target image brightness residues.

Step 3: Compute the *M-estimator correlation coefficient* MCC^T between the masked template and target images as follows:

$$\mathrm{MCC}^\mathrm{T} = \frac{\sum_{i=1}^{n} m^\mathrm{T}(i)(F_i - \bar{F})(f_i - \bar{f})}{\sqrt{\sum_{i=1}^{n} m_2^\mathrm{T}(i)(F_i - \bar{F})^2}\sqrt{\sum_{i=1}^{n} m_1^\mathrm{T}(i)(f_i - \bar{f})^2}},$$

where $m^\mathrm{T}(i) = \sqrt{m_1^\mathrm{T}(i) * m_2^\mathrm{T}(i)}$.

Step 4: Return the window position having the maximum value of MCC^T, indicating the best match.

4. Experimental results

Experimental results conducted in several real life images and some of them are presented in this section. Three important categories of images where registration is usually required are considered. Theses are face, object and satellite images. To study the robustness property of the proposed image registration method the experiments have been run on noisy images, color images, multimodal images and images containing objects with varying degrees of occlusion.

Both artificial and natural occlusion are considered. The performance of proposed method is compared with the best correlation based image registration method reported in literature, namely selective correlation coefficient (SCC) based method (Kaneko et al., 2003). The few representative template images and target images with the registered region using the proposed (MCC) and SCC methods are shown in Figs. 3–11. The Canon *A*70, 3.2*Megapixel*,

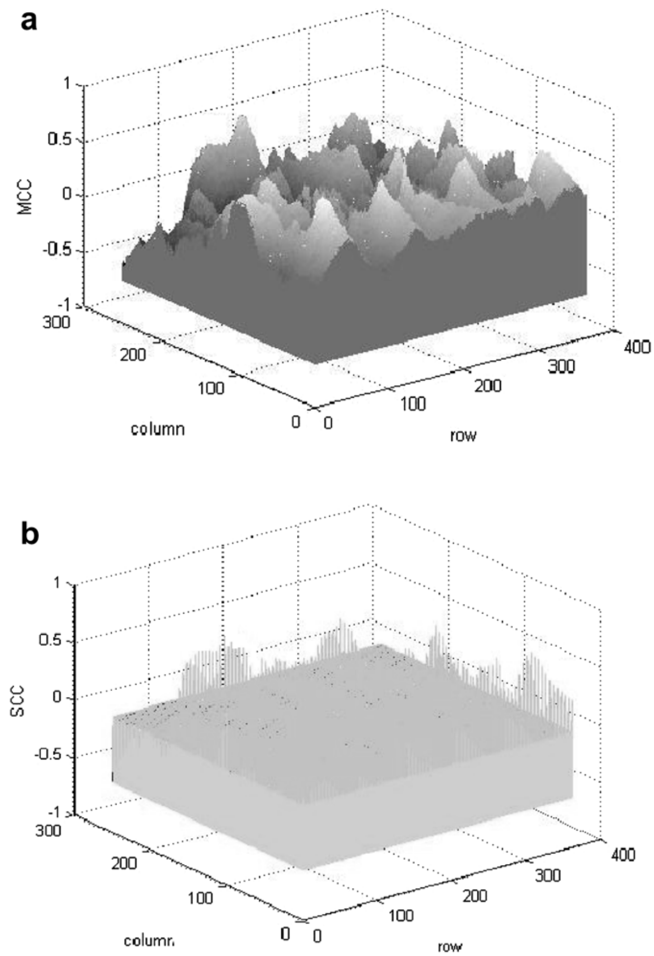

Fig. 8. Profiles of correlation values: (a) MCC, (b) SCC.

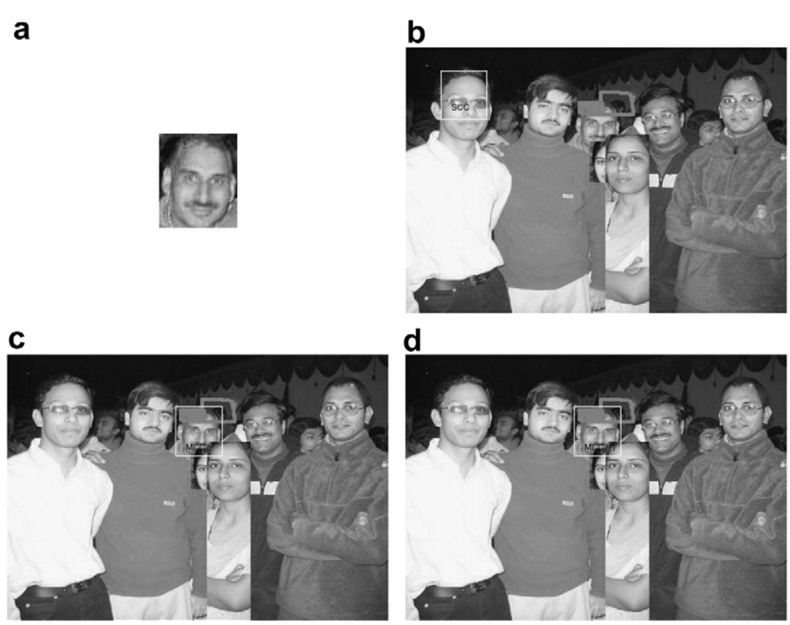

Fig. 7. Face image: (a) template, (b) target image registered u sing SCC, (c) target image registered using MCC with Huber's estimator, (d) target image registered using MCC with Tukey's estimator. The target image contains high degree (60%) of occlusion.

ccd camera was used for taking the images of faces and objects. The satellite images obtained from Indian Remote-sensing Satellite $IRS - P6(RESOURCE\ SAT - 1)$ with AWiFS camera of spatial resolution 56 m. Table 1 shows the experimental specifications.

Target images in Figs. 3 and 4 contain moderate amount of artificial and natural occlusion (about 40%). Both SCC and MCC successfully register the template image in this case, although the time required by MCC is much less as compared to SCC (Table 2). The image in Fig. 5 contains relatively high amount of artificial occlusion, but still both MCC and SCC are successful. Here also MCC take less time compared to SCC. Note that in all the above images the pose of the object in target image is different from that in template image. In the satellite image (Fig. 6) patches of noise are present in significant portion of target image. In this case also both MCC and SCC performed satisfactorily but MCC takes less time compared to SCC. It is seen from Fig. 7, that SCC registered incorrectly whereas MCC is still successful. In this image the amount of occlusion is very high (about 60%), and is of complex intensity variation.

Fig. 8 shows the profiles of correlation values of MCC and SCC for Fig. 7. In Fig. 8a the highest peak indicates the correct match position. It is evident from Fig. 8b that SCC also has a peak at the same location but it is not the highest over the entire scene and hence, SCC has failed

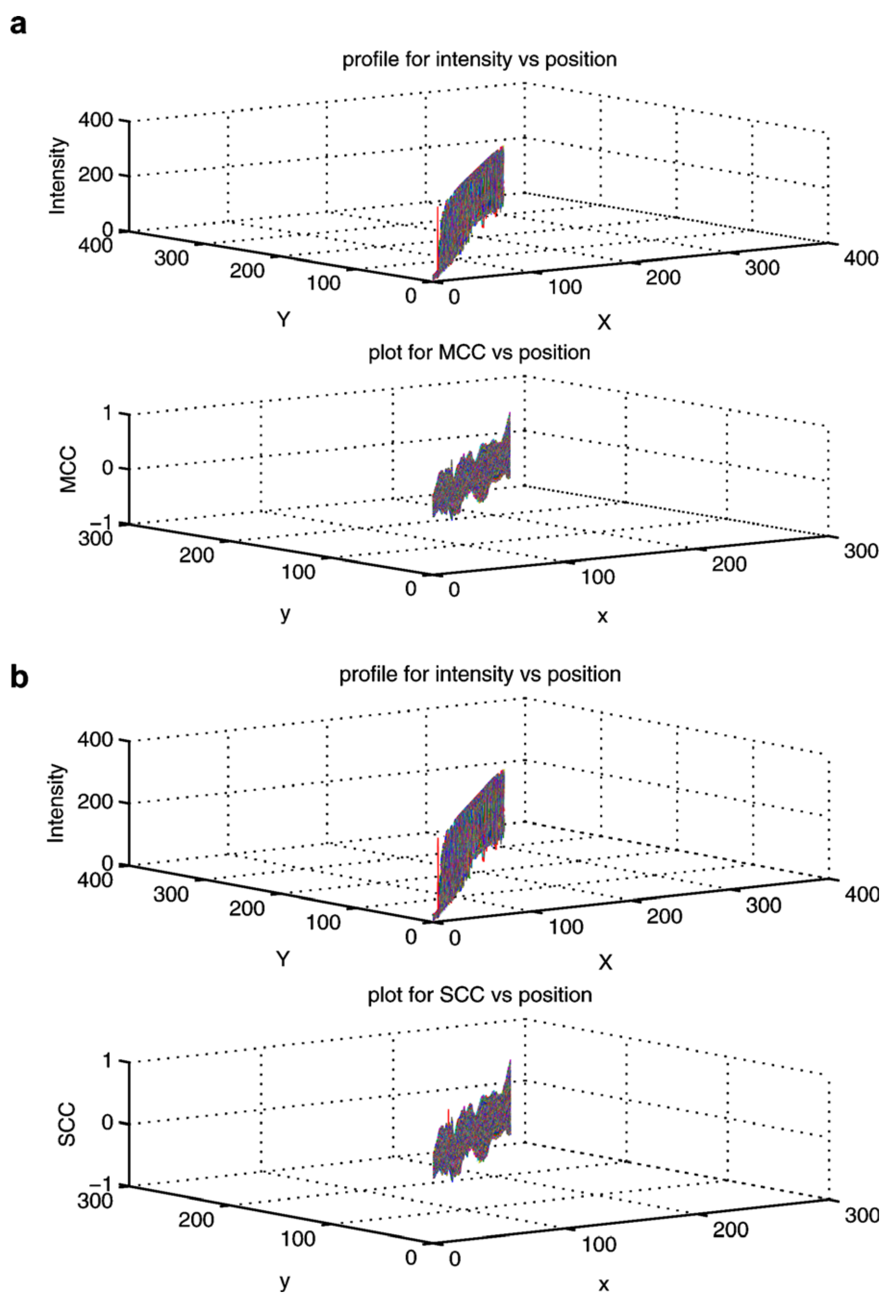

Fig. 9. Profiles of brightness and correlation values: (a) MCC, (b) SCC.

to locate the correct match position. Fig. 9a and 9b show the profile of brightness values and variation of MCC and SCC respectively.

The registration of multimodal satellite images is demonstrated in Fig. 10. In this case the template is taken from Band2 and target image is a Band3 image which contains noise patches. It is clearly evident from Fig. 10 and Table 2 that MCC performed satisfactorily but the registration performance of SCC for multimodal images is unsatisfactory. The experiments are also carried out on several color

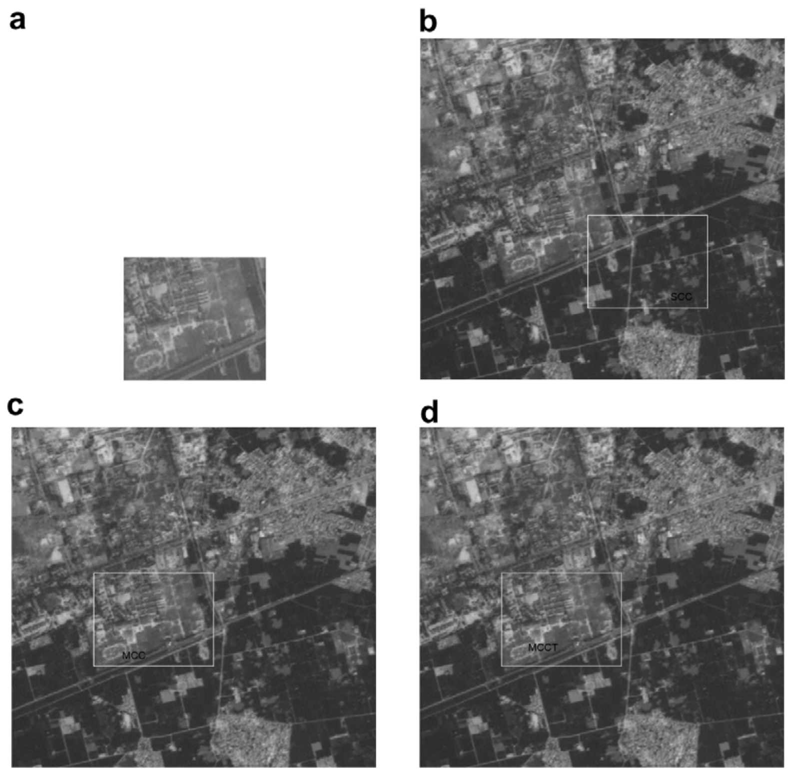

Fig. 10. Multimodal Satellite images: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber's estimator, (d) target image registered using MCC with Tukey's estimator. The target image contains noise patches.

Fig. 11. Colored image: (a) template, (b) target image registered using SCC, (c) target image registered using MCC with Huber's estimator, (d) target image registered using MCC with Tukey's estimator. The target image contains occlusion. (For interpretation of the references in colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Experimental specifications

Object image (computer)	
Target	265 × 146
Template	90 × 71
Face image (3 boys)	
Target	396 × 251
Template	72 × 78
Face image (3 girls)	
Target	300 × 125
Template	48 × 58
Satellite image	
Target	512 × 512
Template	90 × 71
Face image (party picture)	
Target	410 × 308
Template	50 × 57
Colored image	
Target	200 × 150 × 3
Template	32 × 110 × 3
Satellite image (multimodal)	
Target	512 × 512
Template	120 × 92

Table 2
Comparison of computational time (in seconds) required by registration algorithms

Image	MCC-Huber	MCC- Tukey	SCC
Object image *Computer* (Natural occlusion)	19.602	22.563	38.797
Face image *3 Girls* (Artificial occlusion)	20.779	41.000	55.250
Face image *3 Boys* (Natural occlusion)	58.457	81.079	136.156
Satellite image (with Noise Patches)	397.567	440.981	872.050
Face image *Party picture* (≈60% occlusion)	107.082	120.338	231.596 Incorrect match
Satellite image *Multimodal* (with Noise Patches)	477.667	759.955	1441.568 Incorrect match
Colored image (Artificial occlusion)	171.332	208.889	144.177 Incorrect match

images and an example for demonstration purpose is shown in Fig. 11. The color picture used here is downloaded from internet and occlusion is created artificially. Finally, it is clear from Fig. 11 that both Huber's and Tukey's MCC performed the accurate registration whereas the SCC gave incorrect match. However, it is to be noted here that on the gray scale version of the same set of images the SCC also gave the correct match. It is to be noted that we have tried the normalized correlation coefficient (NCC) method in all the cases and it has been found that it failed to register the template image. The success of MCC in case of gray scale, color and multimodal satellite images demonstrates that the proposed method is highly robust.

The comparison of computational time required by the MCC and SCC registration methods is given in Table 2. The values are for a Pentium-IV PC with 256 MB RAM in a MATLAB environment. It is evident from Table 2 that MCC requires significantly less time as compared to SCC for all the images. The reason for MCC to take less time is that it does not convert the gray scale images in to binary images whereas, in SCC the masking function used is binary and an additional time is required for preprocessing to convert the gray scale/color or multimodal images in to the corresponding binary codes.

5. Conclusions

This paper has proposed a method, namely MCC, to robustly register images in the presence of noise and occlusion up to 60%. This method is based on the concept of M-estimators. The Huber's and Tukey's M-estimators are used in this work. It has been shown that the proposed algorithm performs efficiently in occluded and noisy images whereas other correlation based methods produce the incorrect registration. MCC method is faster and robust as compared to the selective correlation coefficient (SCC) based method. It has been shown through experimental results that MCC can efficiently register face, object, satellite images, multimodal images and color images; thereby signifying its suitability in a wide range of applications.

Acknowledgements

The authors are thankful to anonymous reviewers for their valuable suggestions and comments on the original manuscript. It has helped us to improve the quality of the paper.

References

Aggarwal, J.K., Davis, L.S., Martin, W.N., 1981. Correspondence processes in dynamic scene analysis. Proc. IEEE 69 (5), 562–571.

Barnea, D., Silverman, H., 1972. A class of algorithms of fast digital image registration. IEEE Trans. Comput. 21 (2), 179–186.

Barrow, H.G., Tanenbaum, J.M., Balles, R.C., Wolf, H.C., 1977. Parametric correspondence and chamfer matching: Two new techniques for image matching. In: Proc. Fifth Internat. Joint Conf. on Artificial Intelligence, pp. 659–663.

Black, M., Rangarajan, A., 1996. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. Internat. J. Comput. Vision 19 (1), 57–91.

Borgefors, G., 1988. Hierarchical chamfer matching: A parametric edge matching algorithm. IEEE Trans. Pattern Anal. Machine Intell. 10 (6), 849–865.

Brown, L.G., 1992. A survey of image registration techniques. ACM Comput. Surveys 24 (4), 326–376.

Dai, X., Khorram, S., 1999. A feature-based image registration algorithm using improved chain-code representation combined with invariant moments. IEEE Trans. Geosci. Remote Sens. 37 (5), 2351–2362.

Flusser, J., Suk, T., 1994. A moment-based approach to registration of images with affine geometric distortion. IEEE Trans. Geosci. Remote Sens. 32 (2), 382–387.

Freeman, H., 1974. computer processing of line drawing images. Comput. Surveys 6, 57–98.

Ghaffary, B.K., Sawechuk, A.A., 1983. A survey of new techniques for image registration and mapping. In: Proc. SPIE: Applications of Digital Image Processing 432, pp. 222–239.

Goshtasby, A., 1988. Registration of images with geometric distortions. IEEE Trans. Geosci. Remote Sensing 26 (1), 60–64.

Goshtasby, A., Stockman, G.C., Page, C.V., 1986. A region-based approach to digital image registration with subpixel accuracy. IEEE Trans. Geosci. Remote Sensing 24 (3), 390–399.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W.A., 1986. Robust Statistics: The Approached Based on Influence Function. John Wiley & Sons, NY.

Huber, P.J., 1981. Robust Statistics. John Wiley & Sons, NY.

Kaneko, S., Murase, I., Igarashi, S., 2002. Robust image registration by increment sign correlation. Pattern Recognition 35 (10), 2223–2234.

Kaneko, S., Satoh, Y., Igarashi, S., 2003. Using selective correlation coefficient for robust image registration. Pattern Recognition 36 (5), 1165–1173.

Lai, S.-H., 2000. Robust image alignment under partial occlusion and spatially varying illumination change. Comput. Vision Image Understanding 78, 84–98.

Li, H., Manjunath, B.S., Mitra, S., 1995. A contour-based approach to, multisensor image registration. IEEE Trans. Image Process. 4 (3), 320–334.

Maes, F., Collignon, A., Vanderneulen, D., Suetens, P., 1996. Multimodality image registration by maximization of mutual information. Proc. Workshop on Mathematical Methods in Biomedical Image Analysis. IEEE Computer Society Press, pp. 14–22.

Montesinos, P., Gouet, V., Deriche, R., Pele, D., 2000. Matching color uncalibrated image using differential invariants. Image Vision Comput. 18 (9), 659–671.

Pluim, J.P.W., Maintz, J.B.A., Viergever, M.A., 2000. Image registration by maximization of combined mutual information and gradient information. IEEE Trans. Med. Imaging 19 (8), 809–814.

Rey, W.J.J., 1983. Introduction to Robust and Quasi-robust Statistical Methods. Springer-Verlag, Berlin Heidelberg.

Roche, A., Malandain, G., Ayache, N., 2000. Unified maximum likelihood approaches in medical image registration. Internat. J. Imaging Systems Technol. 11 (1), 71–80.

Saghri, J.A., Freeman, H., 1981. Analysis of the precision of generalized chain codes for the representation of planar curves. IEEE Trans. Pattern Anal. Machine Intell. 3 (5), 533–539.

Sester, M., Hild, H., Fritsch, D., 1998. Definition of ground control feature for image registration using gis data. In: Proc. ISPRS Commission III Symp. on Object Recognition and Scene Classification from Multispectral and Multisensor Pixels, OH, pp. 537–543.

Shekhar, C., Govindu, V., Chellapa, R., 1999. Multisensor image registration by feature consensus. Pattern Recognition 32 (1), 39–52.

Stockman, G., Kopstein, F., Benett, S., 1982. Matching image to models for registration and object detection via clustering. IEEE Trans. Pattern Anal. Machine Intell. 4 (3), 229–241.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley Publishers, Cy.

Venot, A., Lebruchec, J.F., Roucayrol, J.C., 1984. A new class of similarity measure for robust image registration. In: Proc. CVGIP 28, pp. 176–184.

Viola, P., Wells, W.M., 1997. Alignment by maximization of mutual information. Internat. J. Comput. Vision 24 (2), 137–154.

Wells, W., Viola, P., Atsumi, H., Nakajima, S., Kikinis, R., 1996. Multimodal volume registration by maximization of mutual information. Med. Image Anal. 1 (1), 35–51.

Zana, F., Klein, J.C., 1999. A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. IEEE Trans. Med. Imaging 18 (5), 419–428.

Zhu, M., 2002. Volume image registration by cross-entropy optimization. IEEE Trans. Med. Imaging 21 (2), 174–180.

Zitova, B., Flusser, J., 2003. Image registration methods: A survey. Image Vision Comput. 21 (11), 977–1000.