



Image Filtering and Registration by using Hybrid Area and Feature Based Method

Kiran P. Dange¹, Neha P. Athavale²

^{1,2}Department of Electronics and Communication
Usha Mittal Institute of Technology,
S.N.D.T. University, Mumbai-49, INDIA

Abstract: During acquisition and transmission, noise can be introduced into images. The main problem of image processing is to effectively remove noise from an image, but keep its features intact. The impulse noise is removed by using Gaussian filter. This filtering process assumes a Gaussian spatial profile in the neighborhood of noisy pixel. Image registration can be defined as the process of transformation of various image data sets into one co-ordinate system. The need of image registration is its ability to compare or integrate the data obtained from the different measurements. For any image registration, filtering is necessary to remove the noise. Image registration is classified into image segmentation and Scale invariant feature transformation. The last process is the template matching. This method provides best filtering at higher noise density. It also provides maximum cross correlation coefficient for registered images.

Key points: Gaussian noise, K-means clustering algorithm, Scale invariant Feature transform, Affine Transform, Template matching.

I. INTRODUCTION

Filtering is the process which effectively removes noise or any disturbances while doing the transmission. A fundamental problem of image processing is to effectively remove noise from an image. The nature of the problem depends upon the type of noise added to the image.

Two common types of impulse noise are fixed and random valued noise. Filtering is categorized into linear and non-linear filter. Linear filter could produce very high blurring effect. The non linear filters have been utilized because of improved performance of removal of impulse noise and preserving edge details. The most widely used non-linear filter is standard median (SM) filter which is having tendency to sort the image pixels into descending order and calculate the median. This median replaces the noisy pixel. But the disadvantage of standard median filter is that it does not preserve fine image details. Center weighted and weighted median filters are improved version of median filter. Center weighted median filter (CWM) is a filter which provides weights to the center pixel for the filtering process. Weighted median (WM) filtering provides weights to control the filtering process which preserve image details [10]. Additive Gaussian noise is characterized by adding to each image pixel a value from the zero mean Gaussian distribution. The zero mean property of the distribution allows such noise to be removed by locally averaging pixel values. Ideally, removing Gaussian noise would involve smoothing inside the distinct regions of an image without degrading the sharpness of their edges.

The proposed filtering techniques initially estimates the percentage of corruption and position of impulses in the observed noisy image with the help of an impulse detector. The filtering process estimates the corrupted pixel from the weighted sum of its neighboring non-noisy pixels [10],[11].

It is necessary to form a larger image with a set of overlapping images and so it is required to find the translations to align the images. The process of image registration finds the translations so that two or more overlapping images are aligned properly. It gives the projection from the view point through any position in the sensed images. Image Registration is said to be the process of transforming the different sets of image data into one co-ordinate system [4], [8]. The image which is registered is called as reference image and the another image which has to be matched with the reference image is called as sensed image.

The first step of image registration is image segmentation. Image segmentation is divided into a variety of methods which can be applied to either monochrome or color images [4]. The segmentation can be done by various methods like Segmentation by clustering in color space, edge detection, normalized cuts, gray-level thresholding, surface based segmentation, etc. The thresholding method may be recommended as the simplest and standard method for automatic threshold selection. The key concept behind this method is to obtain an optimal threshold that maximizes a function of threshold level. But the memory required for storage increases. In Edge based method boundaries formed are not necessarily closed. There is no significant improvement for

multi-spectral images. Region-based method segmentation uses closed boundary method. Our method makes use of the feature based method for image segmentation in which the hue saturation and intensity value (HSV) color space is analyzed, then the extraction of features is done. Then K-means algorithm is implemented. Our project focuses on segmentation by k-means clustering algorithm. One of the advantage is that the proposed system has low memory requirement. The throughput is also increased. And so the computational efficiency is increased.

The second step in registration is to describe the image features. The class of descriptors, describing these image features are distribution based descriptor, spatial-frequency technique and differential descriptor. The evaluation of the descriptors is performed in the context of matching and recognition of the same scene or object observed under different viewing conditions. Spatial-Frequency Technique describes the frequency content of an image. The Fourier transform decomposes the image content into the basis functions. However, in this representation, the spatial relations between points are not explicit and the basis functions are infinite; therefore, it is difficult to adapt to a local approach. In Differential Descriptors a set of image derivatives computed up to a given order, approximates a point in neighborhood [6]. The properties of local derivatives are investigated and derived differential invariants, combine the components to obtain rotation invariance. These filters differ from the Gaussian derivatives by a linear coordinate which change in filter response domain.

Our proposed algorithm uses distribution-based descriptor. During the process of image registration, the major problem called as corner clustering occurs which results in increasing false matches as matching the feature point. This is overcome by using this descriptor which is nothing but Scale-invariant feature transform (SIFT) technique [6]. The size of this descriptor is reduced with Principle Component Analysis (PCA) technique. The number of correct matches and correspondences is determined with the overlap error. The overlap error measures how well the regions correspond under a transformation.

The third step involves the template matching algorithm. Feature-based method is usually applied when the structural information matches than the intensity information. They cannot handle image distortions. Harris corner edge detector has rotation invariance, but it is sensitive to scale changes. Area-based approaches are preferable when there are many prominent details and characteristic information such as color size shape etc[2],[3],[12]. Templates and input images must have either statistical dependence or intensity similarity. If they have intensities similarity then correlation methods can be used. In geometric points of view, only small amounts of shifts and rotations are allowed. The proposed method is the hybrid of feature and area based method. This method is focusing best Gaussian filtering at highest 80% noise density and gives maximum cross correlation coefficient for different features such as contrast and brightness.

II. PROPOSED ALGORITHM

Following figure shows the block diagram of the proposed system.

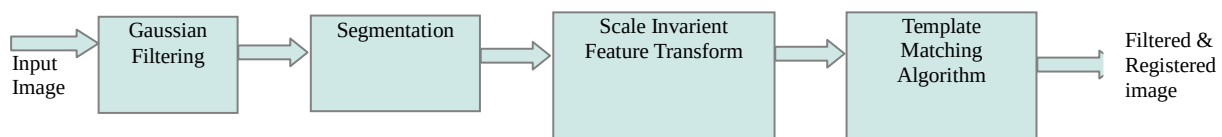


Figure 1: Block Diagram of Image Filtering and Registration by using Hybrid Area and Feature Based Method

It is classified by two algorithms, first Gaussian filtering and second image registration by using segmentation, SIFT and template matching algorithm. In first algorithm, at 80% noise density, the image is filtered and in second stage, maximum cross correlation coefficient is obtained for filtered image by using template matching algorithm [11].

III. STAGE 1: IMAGE FILTERING BY GAUSSIAN FILTER

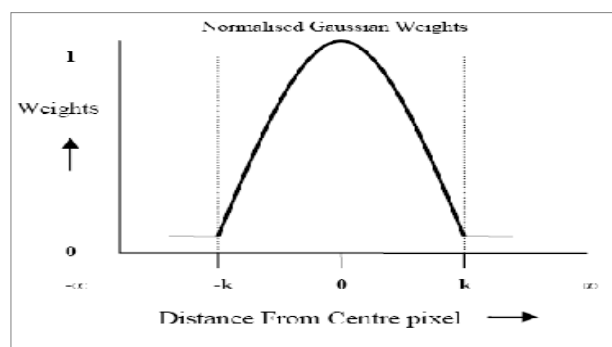


Figure 2: A Normalized Gaussian Weighting Function

Here we consider the two dimensional Normalized Gaussian weighting model over the neighborhood of the corrupted pixel under consideration as shown in Figure 2 and given in equation (1). In the first stage, the distribution of impulses is stimulated in terms of their position and quantity using an impulse detector. The second stage involves restoring the corrupted pixels as a weighted sum of their neighboring uncorrupted pixels. The algorithm of Gaussian filter is explained as follows:

Stage I:

- 1) Get the noisy image $S(p, q)$ of initially 3×3 window size.
- 2) Classify noisy image into 255 and 0. If it is 255 or 0, then flag matrix $f(i, j) = 1$, otherwise $f(i, j) = 0$
- 3) Find the minimum, maximum and median of noisy image i.e. S_{Min} , S_{Max} , S_{Med} and noisy image is $S(i, j)$.
- 4) For window size selection, if $S_{Min} < S(p, q) \& S(p, q) < S_{Max}$ Output $(p, q) = S(p, q)$; Else increase the window size

Stage II:

- 1) The filtering of noisy pixel is carried out recursively. Find the Gaussian weights by using Equation (1).

$$W_{i,j} = (1/2\pi) \{ e^{-|i-p|^2/2} \times e^{-|j-q|^2/2} \} \quad (1)$$

- 2) Area under Gaussian curve is unity, so weights are reduced to zero wherever there is a noisy pixel, so we get (2)

$$W_{i,j} = (1/2\pi) \{ (1-f_{i,j}) \times e^{-|i-p|^2/2} \times e^{-|j-q|^2/2} \} \quad (2)$$

- 3) All weights are scaled by scaling factor 'a' in order to maintain the unit area as shown in (3)

$$W_{i,j} = a \times [(1/2\pi) \{ (f_{i,j}) \times e^{-|i-p|^2/2} \times e^{-|j-q|^2/2} \}] \quad (3)$$

Note that scaling factor will differ for each noisy pixel. Scaling factor can be calculated by only considering the non noisy pixel. Take the summation of non noisy weights and to maintain unity area, scaling factor will be calculated by dividing $1/(\text{sum of current weights})$.

- 4) Replace the noisy pixels as a weighted sum of their neighboring uncorrupted pixels.

IV. STAGE 2: IMAGE REGISTRATION

It is classified into following three stages.

A. Segmentation

The proposed system uses Segmentation in color space by K-means Clustering Algorithm. Each image point is mapped to a point in a color space. It is many to one mapping. The points in color space are grouped together in clusters. The clusters are then mapped to regions in the image which is represented in HSV color space [1], [4].

The steps of segmentation are as follows.

1. The properties of the HSV color space are analyzed and studied.
2. Next the variations in hue saturation and intensity values of an image pixel from viewpoint of visual perception is done.
3. Then the pixel features are extracted choosing the hue or intensity as dominant property which is based on saturation value of pixel.
4. Then the segmentation of image by grouping pixels of similar features using K-means clustering algorithm is done.
5. Given an image of N pixels, the goal is to partition the image into K clusters. The data features give rise to vector space and finds natural clustering by K-means clustering algorithm.
6. Clusters provide a grouping of pixels. This is dependent on pixel values in the image but not on the location of the image unless the location is a specified property.
7. The hue and intensity values belong to the same number space, hence are clustered separately, so that the gray value and the color pixels are not considered in the same cluster.
8. Thus in short K-means clustering steps are as follows.
 - a. Initialization of Parameter:- Let the image be of N pixels. Let us denote $V(x_i)$ as the property vector which is associated with the pixel x_i .
 - b. Then values of potential property vectors are initialized.
 - c. Choose the values of each element of the property vector. This is done randomly from the set of all possible values for that element [4].
 - d. Hard Assignment of Pixels to the Clusters:- Each pixel x_i of the property vector is given to the cluster having closest mean. The points are all clustered around the centroid as shown in (4).

$$C_{i,j}^{(i)} = \operatorname{argmin} \| x_{i,j}^{(i)} - \mu_j \| \quad (4)$$

It has seen that after the process of segmentation, memory required for the storage is reduced. The throughput is increased. The computational efficiency is also increased.

B. Scale-Invariant Feature Transform (SIFT) Technique

The most powerful approach to obtain the local descriptors is obtained by SIFT technique. The SIFT approach transforms image data into scale-invariant co-ordinates relative to local features. It has four major stages such as Scale-Space Extrema Detection, Key-point Localization, Orientation Assignment and Key-point Descriptor [6]. This approach is named the Scale Invariant Feature Transform (SIFT), because it transforms image data into scale-invariant coordinates which has relation with local features. It describes image features which has many properties. These properties makes features suitable for matching differing images of an scene or objects. The features are invariant to image rotation and scaling, and they also invariant to some amount to change in illumination. They are well localized in both the frequency and spatial domains. This reduces the probability of disruption by noise, occlusion and clutter. It is described using following stages

C. Scale Space Extrema Detection

It detects key-points using a cascade filtering approach [6]. The first step is to search scales and locations which can be repeatably assigned under different views and times of the given object. For obtaining locations that are not variant to changes in scale of the image data, we can search for stable image features at all possible scales, and use a continuous function of scale known as scale space. The scale space kernel is the Gaussian function. Therefore, the scale space of an image is defined as a function, that is produced from the convolution of a variable-scale Gaussian, with original input image as shown in equation (5) and (6) .

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (5)$$

$$G(x, y, \sigma) = (1/2\pi \sigma^2) * e^{-(x^2+y^2)/2\sigma^2} \quad (6)$$

Next Step is to compute the difference of Gaussian as shown in equation (7). It efficiently detects stable key-point locations in the scale space[6].

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (7)$$

Now once the key-points are obtained next step is to obtain stable key-points. So we do Local Extremum Detection. To detect the local maxima and minima of $D(x, y, \sigma)$ comparison of each sample point with its eight neighbors in the same image and nine of its neighbors in the scale below and above is done. The sample point is selected only if it is larger or smaller than all of these neighbors. The frequency of sampling is determined in the image and scale domains so that it helps in detecting the extrema. From these extrema obtained not all the key points will be selected but Low Contrast Key-points will be detected and rejected. The function value at the extremum $D(x)$ as shown in equation (8) is used to reject unstable extrema with minimum contrast [6].

$$D(x) = D + ((1/2)(\partial D^T / \partial x))_x \quad (8)$$

In equation (8) if the value of LHS is less than equal to 0.09 then that key points are rejected.

Next we have to eliminate edge responses. The difference- of-Gaussian function has a very strong response along the edges. This is because the location along the edge is poorly determined. Therefore it is unstable to small amounts of noise [6]. A poorly defined peak in the difference of Gaussian function has a large principle curvature at the edge. It has small curvature in the perpendicular direction which is found from a 2×2 Hessian matrix which is computed across the scale and location of the key-point and is known as H . The derivatives are calculated as we can take differences of neighboring sample points. Now the eigenvalues of H are directly proportional to the principle curvatures. Here we are interested with the ratio of eigenvalues. Let us assume that α is the eigenvalue with the largest magnitude, and β is the eigenvalue with lowest magnitude [9]. Then we compute the sum of the eigenvalue from the trace of matrix and product from the determinant of the obtained matrix. The ratio thus obtained is r and if r is greater than equal to 10 then we can eliminate that key points.

D. Localization of Accurate Key point

Now a key point candidate has been found by taking a comparison of a pixel to its neighbors. Then we can perform a detailed fit of the point to the nearby data for scale, ratio of principle curvatures and locations [6]. By this information we can reject points that have low contrast or the points which are poorly localized near an edge. This approach uses Taylor's series.

E. Assignment of Orientation

We can assign a consistent orientation to each key point and this assignment is based on local image properties. The key point descriptor is represented based on this orientation and hence it is able to achieve invariance to image scaling and rotation [6]. The scale of the key point is used to obtain the Gaussian smoothed image with the closest scale. The computations are performed in a scale-non-variant manner. At this scale for each image sample the gradient magnitude $m_1(x, y)$ and gradient orientation $\theta_1(x, y)$ is to be precomputed by using the pixel differences as shown in equation (9) and (10) taken from [6] .

$$m(x, y) = ((L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2)^{1/2} \quad (9)$$

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (10)$$

Here each sample which is added to the histogram is weighted by its gradient magnitude. Peaks in the

orientation histogram will give the dominant directions of each gradients. The maximum peak in the histogram is observed, and then any other local peak that is within eighty percent of the maximum peak is obtained and is used to create a key-point with that orientation. The locations with multiple peaks of equal magnitude, takes multiple key-points at the same location and same scale but with different orientations.

E. Key-point Descriptor

All the operations mentioned above have assigned an image scale, location and orientation to each key-point [5]. The next step is computation of a image descriptor for the local image region. It has to be highly distinctive and also invariant to change in illumination or 3D viewpoint. One method is to sample the intensities of the local image across the key point by using proper scale, and then match these using a normalized correlation measure. First the image orientations and gradient magnitude are sampled around the location of the key-point . It is done by using the key point scale to select the Gaussian blur level [5], [6]. To achieve orientation invariance, the coordinates of the descriptors and the gradient orientations are rotated in relation to key point orientation.

G. Template matching Algorithm

The description of this algorithm is as follows.

a. To Obtain Matching Candidates

Template Matching is a technique for finding areas of an image that match to template of an image. Here we normalize the output of the matching procedure. First it loads an image[12]. Then we perform template matching. Then we localize the location with highest matching probability. Basically it is process automatically selects lots of detailed and unique templates from one image and locates the templates in another image of the same scene[12]. In this process the centroids of corresponding templates are used as corresponding control points and the best correspondences that minimizes the error criteria are used to determine the scaling, rotational and translational parameters needed to register the images. If the centers of two image volumes correspond to each other and these image data have a small rotational difference, then we can say that the average distance between the corresponding voxels in the image will be larger than the average distance between corresponding voxels in the two sub volumes of these images when the centers of the sub volumes will coincide[12]. So it makes possible to find template correspondences accurately even when the images to be registered have some small rotational differences. Using a set of corresponding control points obtained from the centroids of the corresponding templates in the image data set, it will be possible to determine the translational, rotational and scaling parameters that can bring the image volume into registration[5].

b. Template Selection

To achieve highly accurate matches, the templates selected in the target image should represent highly detailed and unique regions[12]. For template to be highly detailed, it should contain a large number of high gradient edges. This is measured by computing the sum of gradient magnitudes in the template.

c. Template Matching

It is the process to find the locations of the sub-image and are called as templates inside the image. Once we find the number of corresponding templates, then their centers are used as corresponding control points to determine the registration parameters. Template matching involves comparing a given template with windows of the similar size in an image data and identifying the windows that is most similar to the template[12].

d. Affine Transformation

The affine transformation class represents a 2D affine transform that performs a linear mapping from 2D co-ordinates to other 2D co-ordinates that preserves the straightness and paralleness of the lines[]. Affine transformations can be constructed using the sequence of translation, flips, scales, rotations, and sheres [5],[7].

V. EVALUATION PARAMETER

One way to characterize uniqueness is to compute the correlation of a template with windows of the same size in its neighborhood [7], [12]. A sharp peak among the correlation coefficient at the template's position is evidence that the template is unique. Here the value of the co-relation coefficient is minimum. The less sharp this peak, the more is the co-relation coefficient value and hence we can say that there is more similarity between the template and the windows in its neighborhood. So when a template that is not locally unique is selected for matching, it may match rather well with many windows in its neighborhood, making distance between the correct matches and an incorrect one difficult[7], [12]. The algorithm finds the co-relation between the original image and the sensed image [8]. It applies the transformation parameters on the sensed images so that the maximum co-relation between original image and the sensed image are achieved. If $P(m,n)$ is the reference image and $Q(m,n)$ is the sensed image then the co-relation coefficient (r) between P and Q is found by equation (11)

$$r = \frac{\sum_m \sum_n (P_{mn} - P_o)(Q_{mn} - Q_o)}{(\sum_m \sum_n (P_{mn} - P_o)^2 \sum_m \sum_n (Q_{mn} - Q_o)^2)^{1/2}} \quad (11)$$

VI. RESULTS

Some standard images such as Lena, Barbara, Pepper and House of 512×512 are used. Results are found using these standard images by using Visual Studio 2008 with Open CV library. The following figure shows original image, filtered image, segmented image, scale-invariant feature transformed image and registered image. The image matches with itself with different angle of rotation, scaling and also for brightness and contrast. Our result shows maximum cross-correlation coefficient at 45 rotation at 80% noise density.

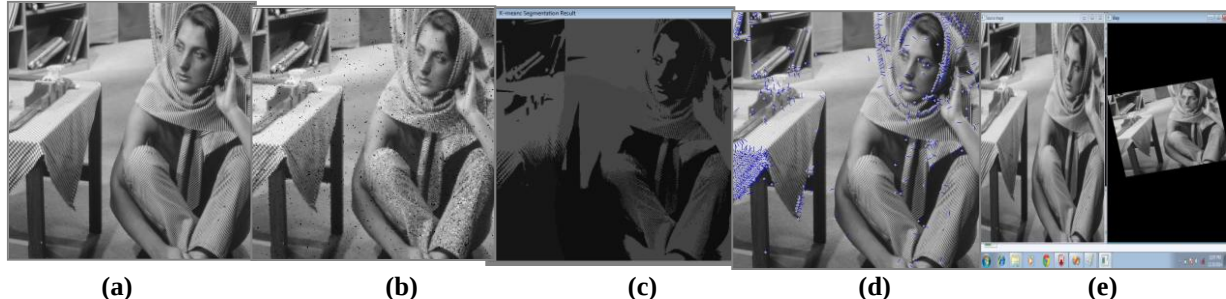


Figure 3: (a) Original Barbara.jpg (512x512 size) (b) Filtered image at 80% noise density (c) Segmented image (d) SIFT descriptor image (e) Registered Image

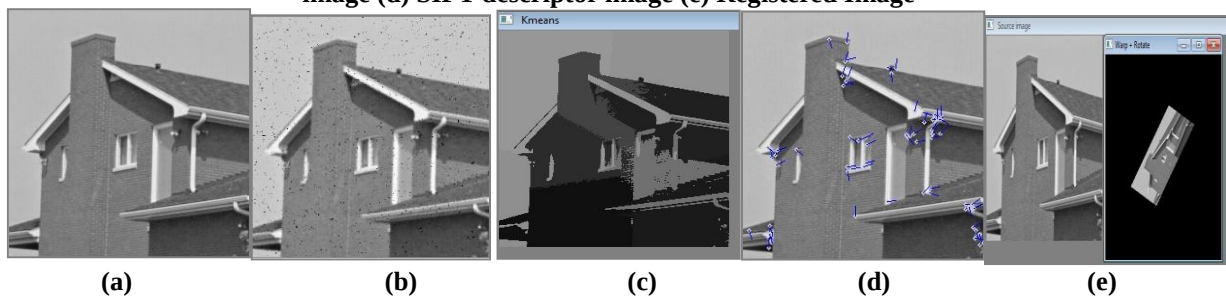


Figure 4: (a) Original House.jpg (512x512 size) (b) Filtered image at 80% noise density (c) Segmented image (d) SIFT descriptor image (e) Registered Image



Figure 5: (a) Original Lena.jpg (512x512 size) (b) Filtered image at 80% noise density (c) Segmented image (d) SIFT descriptor image (e) Registered Image

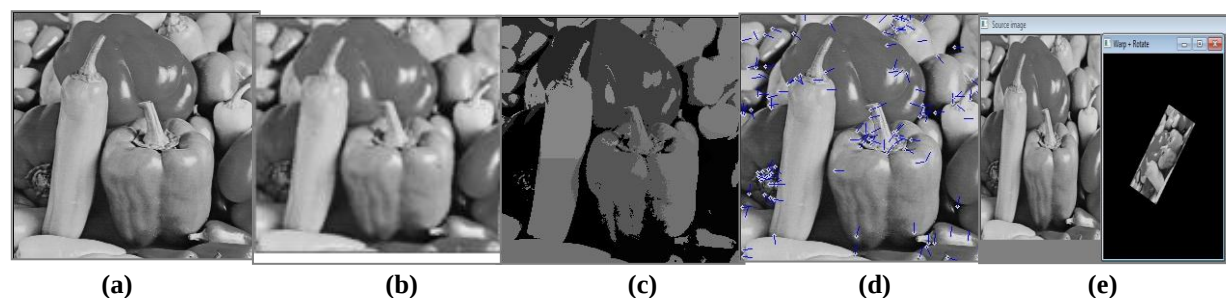


Figure 6: (a) Original Pepper.jpg (512x512 size) (b) Filtered image at 80% noise density (c) Segmented image (d) SIFT descriptor image (e) Registered Image

VII. CONCLUSION

Thus we can conclude that our image is filtered at 80% noise density and registered with maximum cross-correlation coefficient. The problem of corner clustering is overcome by using SIFT. Minimum memory is required by using clustering in color space for segmentation. At 80% noise density we get properly registered image. The comparison of various images is shown in the table 1.

Table1: Comparison of three types of registration methods at 80% noise density for different standard images

Images	Area-based (Affine Transform)	Feature-based (Discrete Cosine Transform)	Proposed Hybrid Algorithm
Lena	0.896499	0.879982	0.956845
Pepper	0.920	0.8895	0.96
House	0.8956	0.9021	0.9128
Barbara	0.95	0.91926	0.9568

From the table it is concluded that maximum cross-correlation coefficient is obtained for Pepper image by proposed hybrid algorithm.

ACKNOWLEDGEMENT

Both the authors are thankful to Usha Mittal Institute of Technology for the facilities provided.

REFERENCES

- [1] Shamik Sural, Gang Qian and Sakti Pramanik, "Segmentation and Histogram Generation using the HSV Color Space for Image Retrieval", IEEE, pp. 589-592 2002.
- [2] Ruhina B. Karani and Dr. Tanuja K. Sarode, "Image Registration using Discrete Cosine Transform and Normalized Cross-correlation", Int Conf Workshop on Recent Trends in Technology, pp. 28-34 2012.
- [3] Ming-Sui Lee and Meiyin Shen and Akio Yoneyama and C-CJay Kuo, "DCT-Domain Image Registration Techniques for compressed Video", Integrated Media Systems Centre and department of Electrical Engineering 2012.
- [4] Hernani Goncalves and Luis Corte, "Automatic Image Registration through Image Segmentation and SIFT," IEEE Transactions on Geoscience and Remote Sensing, Vol. 49 no 7, pp. 2589-2600 July 2011.
- [5] Jimmy Singla and Raman Goyal, "A systematic way of Affine Transformation using Image Registration", Int Jou of Information Technology and Knowledge Management, Vol. 5 no 2, pp. 239-243 November 2004.
- [6] D.G.Lowe, "Distinctive Image Features from Scale Invariant Keypoints", Int Jou Computer Vis, Vol. 60 no 2, pp. 91-110 November 2004.
- [7] Mehfuza Holia and Dr. V.K.Thakur, "Image Registration for Recovering Affine Transformation using Nelder Mead Simplex Method for Optimization", Int Jou of Image Processing, Vol. 3 no 5, pp. 218-228 November 2004.
- [8] Medha V. Wyawahare and Dr. Pradeep M. Patil and Hemant K. Abhyankar, "Image Registration Techniques, An Overview", In Proc. International Image Processing and Pattern Recognition, pp.11-25 September 2009.
- [9] Rafael Gonzalez Richard Woods, "Digital Image Processing", Pearson Publications.
- [10] N. C. Gallagher Jr and G. W. Wise, "A theoretical analysis of the properties of median filters", IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-29, pp. 1136-1141, Dec. 1981.
- [11] Czerwinski, R.N.; Jones, D.L.; O'Brien, W.D., Jr., "Ultrasound speckle reduction by directional median filtering", International Conference on Image Processing, 1995. Proceedings.
- [12] Lijun Ding, Ardeshtir Goshtasby and Martin Satter, "Volume Image Registration by Template Matching", Department of Computer Science and Engineering, Wright State University, Dayton, OH45435.