An Image Registration Algorithm through Perception based Weighted Cost Optimization of Pixel Labelling

B. Sudhir[1], P. Rajesh Kumar[2]
[1]*Electronics & Communications Engg. Dept., Pragati Engineering College, Surampalam, AP, India*
[2]*Electronics & Communications Engg. Dept., Andhra University, Vishakapatnam, AP, India*

Abstract

Semi-automated image registration techniques depend on human eye validation of final image alignment. This results in non-uniform registration based on color discrimination based on the perception of different users. This paper describes a registration method based on perception that is optimized for a cost function weighted on prominence of features perceived within a group of pixels after labelling them. The weights are calculated so as to minimize the penalty of over-emphasizing non-prominent features for registration. The key points are then matched using nearest neighbor distance ratio (NN-DR) followed by motion estimation and a geometric transformation to obtain the final image. The technique has yielded better results and has been proved excellent especially in the image stitching application after registering the similar pixels.

Keywords—*image registration, cost function, prominence weight, graph-cut, nearest neighbor distance ratio*

I. Introduction

Image registration is a very popular and well-studied image processing technique that has found profound applications in medical imaging, satellite imaging and commercial imaging applications. The vital process behind the image registration process is the alignment of a reference and a template image. This can help in easy diagnosis of certain health ailments from medical images of a patient and in geological analyses on satellite images. Image registration has not been used in any literature we have investigated as a pre-cursor for image stitching. Since alignment being the first step in image stitching algorithm and is the exact process done by registration, can we use in the stitching application has been the problem statement during our investigation. This has led us to study the benefits of registration in commercial images as a useful step for image stitching. In this regard we have looked at semi-automated image registration technique through machine validation after human perception.

Human perception used in manual image registration can lead to non-uniform registration and biased approaches [1]. This arises due to the fact that the attention of an observer is usually drawn to certain prominent features present in the image and may lack sometimes the required attention for details required for registration in certain commercial applications. To avoid this, it would be better if we give more weights to the region of pixels where the most key feature points required for the registration is present and give less prominence or penalize the other pixels. This leads to define a cost function on labelling pixels based on how a human must have observed the images in question for the alignment and in optimizing it. An added averaging term is added to straighten out any irregularities in labelling the pixels.

Many literatures are available in video and image synthesis techniques to optimize such a cost function. A local weighted sum of neighboring pixels is used in [2] to alleviate any geometric distortion that might occur while creating a depth map for virtual views. In [3], a maximum a posterior problem is solved using graph cuts for video stitching to align the artefacts properly. Y. Lou and et al have used an objective energy function in their image registration problem in [4]. They have observed that the final geometric transformation in aligning the two images yield an optimal performance when this energy function is minimized. Image registration through establishment of symmetric image interdependence by using local spatial intensity histograms is proposed in [5].

A regularized gradient kernel anisotropic diffusion method is proposed in [6]. The problem of gradient separation from irregularities is formulated as a nonlinearly separable classification problem. The variance of estimated parameters is minimized for characterizing pathological changes in living

ISSN: 2233-7857 IJFGCN

tissues in [7] by an optimal gradient encoding scheme. A local region matching for image alignment is proposed in [8] that will work even if predefined constraints are discounted.

In this paper, an algorithm for image registration through perception based weighted cost optimization of pixel labelling is proposed. The basic idea of image registration is to find the correct pixel pairs between two images which can be modeled as pixel labelling. The cost function in such a task of pixel labelling can be minimized through any optimization algorithm. However the human bias factor in deciding the parameters of the function can alter the results significantly. This is taken care by weighted cost function optimization. The experimental results show excellent result in achieving optimal registration results on commercial images and also have proved to be effective in image stitching application.

II. Theory Of Energy Function Based Method Of Image Registration

In this section, we first explain the theory behind existing normal framework for image alignment through cost optimization of pixel labelling, then a sigmoid function for color discrimination of salient features is described using which the cost optimization is redefined with a weighted parameter to avoid the bias and finally our framework for pixel labelling for image registration is proposed.

A. The framework for Image Alignment

Let I_T and I_R be a pair of images to be aligned under image registration process. Assume they have a small area of overlapping pixels and let $\mathcal{P}$ be this set of overlapping pixels. If $\mathcal{L} = \{0,1\}$ is a set of labels for each pixels from the two images from the pair then a normal cost function would mean some energy spent in assigning a label to each pixel from each image. More formally if we assign "0" as the label to all the pixels from I_T and "1" as the label to that of I_R, then a label $l_p \in \mathcal{L}$ should be assigned to each pixel $\wp \in \mathcal{P}$. The problem of cost optimization in image alignment then can be modeled as a mapping from $\mathcal{P}$ to $\mathcal{L}$ such that it minimizes the below energy function:

$$E(l_p) = \sum_{\wp \in \mathcal{P}} C_p(l_p) + \sum_{(\wp,q) \in \mathcal{N}} A_{p,q}(l_p, l_q) \qquad 1$$

Where $\mathcal{N} \subset \mathcal{P} \times \mathcal{P}$ is a pixel system in the neighborhood of feature points. The term $C_p(l_p)$ represents the cost of assigning a label to a pixel and $A_{p,q}(l_p, l_q)$ is an averaging parameter that helps in smoothening any irregularities and represents the cost of assigning a pair of labels to a pair of pixels in the pixel system in the neighborhood of our feature points that are to be used for alignment. The cost of labelling is defined as:

$$\begin{cases} C_p(1) = 0, C_p(0) = \mu, if\ p \in \partial I_T \cap \partial \mathcal{P} \\ C_p(0) = 0, C_p(1) = \mu, if\ p \in \partial I_R \cap \partial \mathcal{P} \\ \quad C_p(0) = C_p(1) = 0, otherwise \end{cases} \qquad 2$$

Where μ is a parameter of penalty for any mislabeling and thus avoid it. $\partial I_T \cap \partial \mathcal{P}$ and $\partial I_R \cap \partial \mathcal{P}$ represents the common pixel border in the target and reference images respectively with the pixel under consideration for labelling. The average term is modeled as:

$$A_{p,q}(l_p, l_q) = \frac{1}{2}\left|l_p - l_q\right|\left(I_*(p) + I_*(q)\right) \qquad 3$$

$$I_*(\cdot) = \|I_T(\cdot) - I_R(\cdot)\|_2 \qquad 4$$

Where $I_*(\cdot)$ is the Euclidean color difference that is measured by finding the difference of distances within a color space as shown in Figure 1. The cost function can be minimized by any optimization technique available in the literature. We have chosen graph-cut optimization which is a popular energy minimization algorithm used in many image processing applications. From the cost equation it is to be noted that the modeling of this function plays the most vital role in the framework for alignment in image registration since we are looking for the least cost in labelling pixels from a pair of images correctly.

ISSN: 2233-7857 IJFGCN

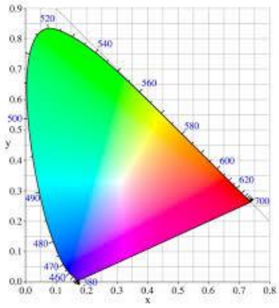


Figure 1:Euclidean color space

B. A perception based method of image alignment

Studies have shown that there always exists a pixel system of feature points that has more energy but is invisible during the process of registration than a system with less energy. This leads to false alignment due to biased perception or lack of attention and overlooking. Hence it is absolutely vital in such applications that the most invisible pixel system of feature points should possess the least energy. For this a perception consistent cost function should be modeled.

1) Sigmoid function for color difference:

Some salient features become invisible despite more cost due to lesser penalty parameter μ by the Euclidean color difference function. A sigmoid-metric color difference could have captured this fault and distinguish the features more vividly and avoid the local misalignment. Sigmoid function being a quality metric that can measure the visibility of color difference makes itself an excellent choice far better than Euclidean color difference in capturing these discriminations among the salient features. By the virtue of definition of the sigmoid function the cost of invisible terms will be approximated to zero and the cost of visible terms will approximated to one as defined below:

$$sigmoid(x) = \sigma(x) = \frac{1}{1+e^{-4k(x-\tau)}} \qquad 5$$

Where τ is the color discrimination threshold and k is a parameter which will be discussed next.

For a pair of images under consideration for registration, τ is a value which divides their pixels into correctly aligned and misaligned area according to their color difference which is exactly like choosing a threshold to differentiate the foreground and background of a grey color or binary image. Ostu's algorithm is used in this regard to determine the threshold value suitably with maximum variance between the classes of pixels or colors. The parameter k denotes the rate of change of color discrimination sensitivity around the suitable threshold value. If ϵ is the width of the bins of the histogram in Ostu's algorithm, then $k = \frac{1}{\epsilon}$. This choice of k value has proved to have very good practical performance in many literatures.

The averaging term from equation 3 can now be rewritten as

$$\tilde{A}_{p,q}(l_p, l_q) = \frac{1}{2}|l_p - l_q|\left(I_{\dagger}(p) + I_{\dagger}(q)\right) \qquad 6$$

$$I_{\dagger}(\cdot) = \sigma(I_*(\cdot)) \qquad 7$$

The sigmoid function has proved to yield better results than Euclidean metric used for measuring color difference in that it effectively helps in avoiding the feature points to cross the misaligned area and thus helps in choosing the required feature points for registration.

2) Feature Prominence weights:

From a human's perspective on looking at the alignment and misalignment we observe that there are a lot of non-uniform perception of images. This is more profound in manual image registration techniques where the registration is done based on the human perception of the key features in the image. This is because human eyes have an inclination to salient objects in the images and thus pay more

ISSN: 2233-7857 IJFGCN

attention to those features and may miss out on certain key features essential for registration. This leads us to believe that features in the prominent regions should be more pronounced than in the non-prominent regions. This can be achieved by defining a weight parameter which will give more weights to artifacts in feature points rich region, called prominence weights. This can be mathematically described as below:

$$W_{p,q} = \begin{cases} 0, if\ p|q \in \partial_{\#}\mathcal{P} \\ 1 + \frac{\omega(p)+\omega(q)}{2}, otherwise \end{cases} \quad 8$$

Where $\omega(\cdot)$ is the average pixel prominence in a system of pixels $\mathcal{P}$. The prominence weights thus computed are normalized in the range of [1,2] to avoid penalizing the weight parameter going out of bounds. The cost function in equation 1 can now be rewritten with the prominence weight added as below

$$\tilde{E}(l) = \sum_{\mathcal{p}\in\mathcal{P}} C_p(l_p) + \sum_{(\mathcal{p},\mathcal{q})\in\mathcal{N}} W_{p,q} \cdot \tilde{A}_{p,q}\left(l_p, l_q\right) \quad 9$$

Now based on the average prominence of artifacts in the pixel region under consideration we assign to the pixels in $\mathcal{P}$ based on perception, the penalty on averaging parameter will be varied which will in effect determine the cost function. Thus those features whose cost need to be optimized will be more vivid in our analysis in discriminating the color difference more accurately.

C. Feaure matching and motion estimation

After the cost function has been optimized we would have sampled the pixel groups that contain our feature points from the reference and the template images. The next step would be to do the matching of these labelled pixels for validation. A pixel by pixel matching can be done by combining Euclidean distance with NN-DR thresholding [9]. The Euclidean distance between the points under consideration can be calculated as

$$d_{k,l}^2 = \| I_T - I_R \|_2^2 \quad 10$$

The NN-DR thresholding is done to take out only the best matching vectors for further calculation and thus avoid any mismatches. Before final registration a measure of alignment deviation would help us in implementing the geometric transformation correctly. This is done by estimating the motion between the matched vectors. We have used optical flow to measure the motion and represent it as motion vectors to act as input for the final geometric transformation.

The apparent velocities of pixel movements in the pair of images used for alignment is approximated by finding the differential by local Taylor series. The brightness constraint used is shown below

$$I(x, y, t) = I(x + \Delta x, y + \Delta y, t + \Delta t) \quad 11$$

The higher order terms in the series are truncated linearly and divided by the differential time interval resulting in the velocity of the pixel relative to the reference image. Lucas-Kanade method is used to determine the optical flow by assuming the velocity of the pixel to be constant within the neighborhood and uses least squares to solve the problem of motion estimation.

Finally, the images can be aligned using an affine geometric transformation by establishing a point by point matching. The affine transformation is defined as

$$\vec{y} = f(\vec{x}) = A\vec{x} + \vec{b} \quad 12$$

where A is a matrix that represents the linear map, $\vec{b}$is the translation vector, f is the affine map on vector $\vec{x}$. The vector $\vec{y}$ then gives the new pixel values of the registered image.

D. Algorithm of perception based weighted cost optimization of pixel labelling for image registration

The proposed algorithm for image registration through perception based weighted cost optimization of pixel labelling is summarized in the below table

ISSN: 2233-7857 IJFGCN

Table 1

Algorithm 1 Image Registration Algorithm through perception based weighted cost optimization of pixel labelling

Input: Template image $\boldsymbol{I}_T \in \mathbb{C}^{m \times n}$ and reference image $\boldsymbol{I}_R \in \mathbb{C}^{m \times n}$;

Output: Image $\boldsymbol{I}_{final} \in \mathbb{C}^{m \times n}$ after coregistration;

1: Compute $I_*(\mathcal{P})$ from Eq. 4

2: Evaluate τ from Eq. 5 using Otsu's algorithm.

3: Calculate $I_\dagger(\mathcal{P})$ from Eq. 7 and $\tilde{A}_{p,q}(l_p, l_q)$ from Eq. 6

4: Compute $\omega(\mathcal{P})$ based on perception of salient features from the image and $W_{p,q}$ from Eq. 8

5: Compute $C_p(l_p)$ in Eq. 2

6: Optimize $\tilde{E}(l)$ in Eq. 9 using graph-optimization technique

7: Match the features using NN-DR

8: Estimate the motion vectors of the matched features

9: Use geometric transformation for image registration to obtain $\boldsymbol{I}_{final}$

III. Experimental Results

The algorithm explained in the previous sections has been experimented on a set of rgb images. The set of images has been chosen such that the matching key feature points in some are very few and easy to perceive and in other examples are more in numbers and might easily be missed by a human eye. Such a choice of images would really help in proving the algorithm we have built to be effective while experimenting on real world examples. We have also split the same images in most cases to prove the fact that our algorithm by matching the key points can very well be used in stitching the images together. If experimented for positive results, we believe this would be the first time an image registration algorithm can be extended to image stitching.

ISSN: 2233-7857 IJFGCN

Figure 2: Test-1 input image1

Figure 3: Test-1 input image2

Figure 4: Image Registration

Figure 5: Weighted Cost image

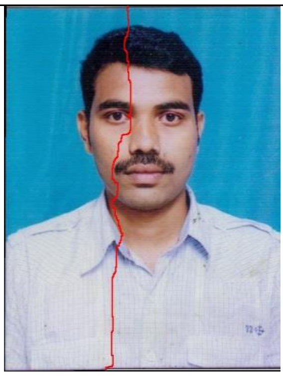


Figure 6: Cost Optimized image

Figure 7: Reconstructed output image

Figure 2 - Figure 37 shows the experimental results we have done on 6 different set of sample rgb color images. Table 2-Table 4 show the different parameters we have used in the series of tests we have done on these set of images. The first 2 sets of images in each example represent the reference and the target image for registration. To demonstrate the potential of our registration algorithm toward image stitching we have taken the same image and broken it down into two set of images with overlapping areas that contain matching feature points for registration. Thus this overlapping regions serve as the set of pixels to be matched. The first step would be to identify these feature points or set of pixels to be labelled. The registration based on cost optimization follows as have been shown in Figure 4-Figure 6. Upon registration using the matching pixels the original image to be constructed after splitting could be easily done as the pixels would have matched. This has been excellently and successfully done by our algorithm as shown in the final reconstructed image in Figure 7.

ISSN: 2233-7857 IJFGCN

Figure 8: Test-2 input image1

Figure 9: Test-2 input image2

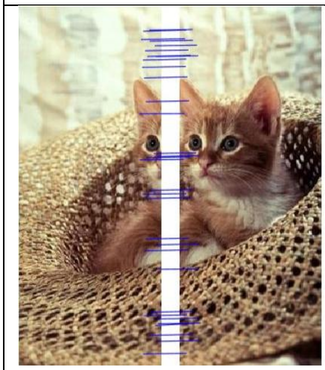


Figure 10: Image Registration

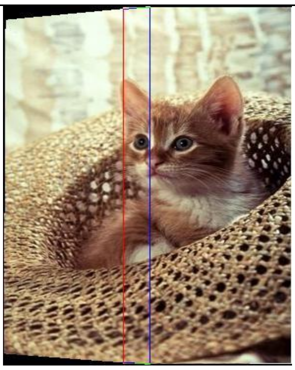


Figure 11: Weighted Cost image

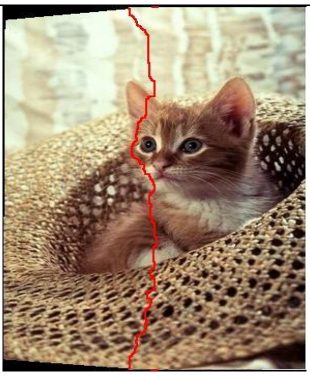


Figure 12: Cost Optimized image

Figure 13: Reconstructed output image

Figure 14: Test-3 input image1

Figure 15: Test-3 input image2

ISSN: 2233-7857 IJFGCN

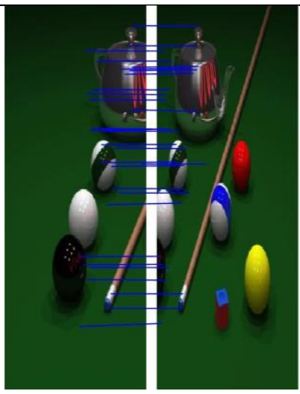


Figure 16: Image Registration

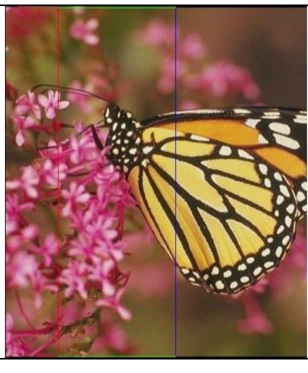
Figure 17: Weighted Cost image

Figure 18: Cost Optimized image

Figure 19: Reconstructed output image

Figure 20: Test-3 input image1

Figure 21: Test-3 input image2



ISSN: 2233-7857 IJFGCN

International Journal of Future Generation Communication and Networking
Vol. 13, No. 4, (2020), pp. 1887–1899

Figure 22: Image Registration

Figure 23: Weighted Cost image

Figure 24: Cost Optimized image

Figure 25: Reconstructed output image

Figure 26: Test-3 input image1

Figure 27: Test-3 input image2

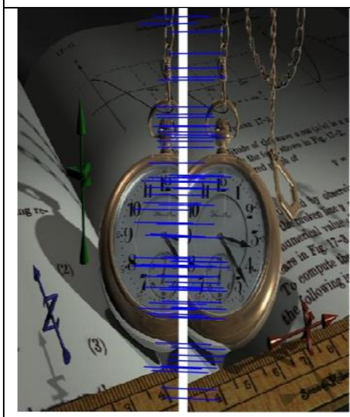


Figure 28: Image Registration

Figure 29: Weighted Cost image

ISSN: 2233-7857 IJFGCN

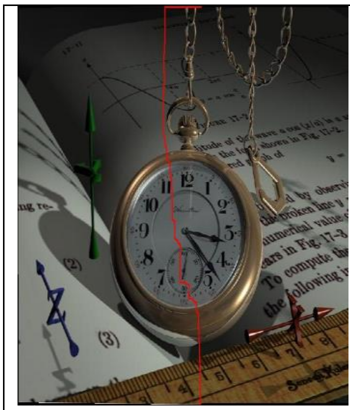


Figure 30: Cost Optimized image

Figure 31: Reconstructed output image

The experimental results have been shown in the other examples using five more different set of examples. In each of the examples we have chosen, we have classified them as of varying number of feature points. For example, the set of images in Figure 14-Figure 15 the number of key points to be identified, selected and matched on a simple perception would prove is very minimal. Whereas the examples in Figure 20-Figure 21, Figure 26-Figure 27 and Figure 32-Figure 33 have far more feature points to be matched for registration. The cost involved and calculation involved also correspondingly increases in each of these examples.

Figure 32: Test-3 input image1

Figure 33: Test-3 input image2

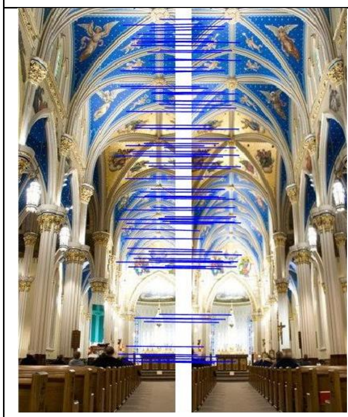


Figure 34: Image Registration

Figure 35: Weighted Cost image

ISSN: 2233-7857 IJFGCN

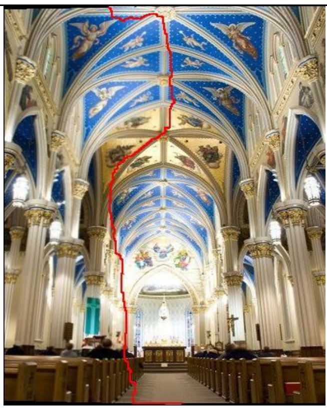
 *Figure 36: Cost Optimized image*	 *Figure 37: Reconstructed output image*

Table 2

Ostu's Parameters		
Test Samples	logistic function(alpha)	Maximum energy
Test-1	0.18	0.000845
Test-2	0.06	0.000781
Test-3	0.3	0.00053
Test-4	0.06	0.000546
Test-5	0.06	0.0013
Test-6	0.06	0.0012
Test-7	0.12	0.0014
Test-8	0.12	0.0027

Table 3

Registration Metrics

	Information Ratio (IR) count	Lower bound on Information Ratio (LIR) count
Test-1	133.1766	75.7107
Test-2	162.9871	89.1469
Test-3	102.2865	59.8249
Test-4	144.1135	81.9415
Test-5	144.3372	82.2035
Test-6	167.7553	90.9503
Test-7	148.5646	83.6912
Test-8	159.1013	88

ISSN: 2233-7857 IJFGCN

Table 4

Registration Metrics

	Mutual Information Ratio (MIR) count	Lower bound on Mutual Information Ratio (LMIR) count
Test-1	41.0239	140.5672
Test-2	42.2008	171.3213
Test-3	28.102	108.238
Test-4	21.3899	159.8084
Test-5	18.2205	157.5916
Test-6	34.6113	176.7551
Test-7	39.6777	158.6011
Test-8	34.0585	168.3226

IV. Conclusion

An image registration algorithm through perception based weighted cost optimization of pixel labelling has been proposed in this paper. The algorithm has been extended in the application of image stitching. Experiments show that our algorithm has given excellent registration results on different set of sample images with varying number of feature points to be matched. The validation of pixel matching has been done through a weighted cost optimization technique. The experiments has also shown successfully that the algorithm as promised can be very well used in the image stitching application. In future, we plan to improve upon cost optimization through other energy functions for pixel labelling and matching for image alignment and generalize the model.

References

[1] T. Zuk and M. Atkins, "A comparison of manual and automatic methods for registering scans of the head," *IEEE Transactions on Medical Imaging,* vol. 15, no. 5, 1996 .

[2] L. Zhou, Z. Wang, Y. Fu, J. Chen, R. Xiang, R. Zhong and S. Wang, "Depth image in-loop filter via graph cut," in *International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 2016.

[3] J. Su, H. Cheng, L. Yang and A. Luo, "Bayesian view synthesis for video stitching," in *International Conference on Multimedia & Expo Workshops (ICMEW)*, Seattle, USA, 2016.

[4] Y. Luo and A. C. S. Chung, "Nonrigid Image Registration With Crystal Dislocation Energy," *IEEE Transactions on Image Processing,* vol. 22, no. 1, pp. 229-243, 2013.

[5] A. Wang, F. Qu, J. Zhang and H. Li, "The Non-Rigid Medical Image Registration Based on the Histogram Decomposition and Improved Energy Function," in *International Conference on Intelligent Networks and Intelligent Systems,* Wuhan, China, 2008.

[6] A. H. Shabani, J. S. Zelek and D. A. Clausi, "Regularized Gradient Kernel Anisotropic Diffusion for Better Image Filtering," in *Conference on Computer and Robot Vision*, 2012.

[7] M. Alipoor, I. Y.-H. Gu, S. E. Maier, G. Starck, A. Mehnert and F. Kahl, "Optimal Gradient Encoding Schemes for Diffusion Tensor and Kurtosis Imaging," *IEEE Transactions on Computational Imaging,* vol. 2, no. 3, 2016.

[8] Z. Lou and T. Gevers, "Image Alignment by Piecewise Planar Region Matching," *IEEE Transactions on Multimedia*, vol. 16, no. 7, 2014.

[9] A. Yong and Z. Hong, "SIFT matching method based on K nearest neighbor support feature points," in *International Conference on Signal and Image Processing (ICSIP),* Beijing, 2017.

ISSN: 2233-7857 IJFGCN

A. K. Mirabadi and S. Rini, "The Information & Mutual Information Ratio for Counting Image Features and Their Matches," 2020 Iran Workshop on Communication and Information Theory (IWCIT), Tehran, Iran, 2020, pp. 1-6, doi: 10.1109/IWCIT50667.2020.9163458.

ISSN: 2233-7857 IJFGCN