An Affine Invariant Iterative Image Matching Approach for Matching Images with Different Views and Illumination

D.Rajasekhar1, T.Jayachandra Prasad2, K.Soundararajan3
1Ph.D. Scholar, Department of Electronics and Communication Engineering, JNTUA Anantapur, Ananthapuramu, Andhra Pradesh,India-515002.
2Professor and Principal, RGM CET, Nandyal, Andhra Pradesh,India-518501.
3Professor and Dean of R&D, TKREC, Hyderabad, Telangana,India-500097.
1dkethanaraj@gmail.com
2jp.talari@gmail.com
3soundararajan_jntucea@yahoo.com

Abstract—Image matching is a critical and challenging task in many mission critical applications. Especially with the developments in domain of Computer Vision, the need for better image matching algorithms are ever increasing. The critical nature of many applications demands that image matching algorithms be robust and fast. This paper presents the design of one such algorithm which can deliver an accurate performance in the presence of different lighting conditions and difference in the view of the image. The proposed iterative approach employs affine transforms to identify the descriptors initially which are further grouped and classified using the Bayes approach. The results have been validated and presented. The results delivered by the proposed approach has been compared and found to be better than the matching done by ASIFT. The result indicates the high degree of accuracy of the proposed image matching algorithm and its reduced computational complexity makes it extremely suitable for real time application.

Keywords: Image Matching, Computer Vision, Affine, Bayes, ASIFT.

I. INTRODUCTION

One of the most important aspects in digital photogrammetry is image matching. It finds its application in automated modeling and mapping [1] where it is used for identifying corresponding pixels in two or multiple images [2, 3]. Most of the existing image matching algorithms have been created over the past 50 years; these include many empirical techniques as well. Even though wide variety of approaches can be found in the literature, there is no universal method that is generic for most of the image matching applications. It’s important that different attributes of the images under study influence the type of image matching approach that has to be adopted. The matching algorithm has to take in to account the following important features in an image; a) Scale: this concept arises when at least two elements of the set of images have different scales, b) Occlusions: this happens when interference happens in the 2D projected image plane for two objects that are spatially separated in the 3D world. For example in the case of object recognition occlusion is a problem which has to be specifically tackled, it necessitates the use of more robust algorithms, c) Orientation: this occurs when image views are rotated at a particular angle in relation to each other. For most of the algorithms to perform a reliable match a maximum orientation of 30° is a typical maximum value, d) Object to be matched: this defines the characteristics of the object to be matched, if the object is a planar, textured or an edgy object, e) Clutter: this implies the conditions of the image background. When the image has a cluttered background it is difficult for the algorithm to identify the boundaries of the object of interest and f) Illumination: variation in illumination also presents a critical problem for accurate feature matching. Current image matching algorithms may deliver reasonably well in regard to some of the image conditions described. But the fact is in general, it is very difficult to accomplish total invariance to these parameters. A good deal of research is towards imparting these algorithms with features that will make them invariant to the common matching problems. It should be also kept in mind that it will be very difficult to accommodate all the necessary elements in a single algorithm without enhancing the computational complexity.

Initial approaches in image matching can be classified under two typical categories of direct and feature based approaches. Feature based methods work by extracting edges, corners and other features. Local information like correlation is limitedly used by these methods to identify specific image patches [4, 5]. In the case of direct methods images are iteratively aligned by considering all the pixel values [6, 7]. A variant of these approaches referred as invariant approaches form invariant descriptors for image matching and indexing. They compute these descriptors by using large amounts of local image data around salient features.
The pioneering work in the area of image matching was developed by Schmid and Mohr [8] who demonstrated the application of Gaussian derivatives in developing a rotationally invariant descriptor around a Harris corner. Lowe [9] further extended this work by incorporating scale invariance and published his work as Scale Invariant Feature Transform (SIFT) algorithm. He identified distinctive invariant features in an image and used them to identify robust and reliable matching. Reliable image matching and distinctive invariant features are two important attributes of image matching. A number of researchers have come up with a variety of approaches for identifying image descriptors that are invariant under affine transformation [10, 11 and 12]. These approaches Harris corner detector, maximally stable region and stable local phase structures [13]. Literature also presents different methods that have been evolved for evaluating these approaches. The evaluation has typically been carried in regard to interest point repeatability [14] and descriptor performance [15].

This paper presents one such approach for image matching which is not prone to view and illumination variations. The proposed iterative approach first defines the affine invariant data points in each image with the help of ASIFT. Descriptors based on gradient magnitude and orientation is assigned to be followed by the calculation of maximum disparity range. The area around all the assigned key points within the maximum disparity range is then selected iteratively. A disparity category having similar pairs of regions is identified and initial probabilities are assigned using Bayes theory. The probabilities are further updated with the help of certain updating rules. Once all the iterations are over the final classes are identified and displayed. The proposed approach is validated by performing experimental validation and empirical analysis of images with different tilts and illumination. Both absolute and transitional tilts are considered. The transitional and absolute tilts were considered for tilts varying from 10 degree to 80 degrees. In order to verify the illumination invariance the brightness of the images were varied. The results are compared with that of ASIFT based image matching. The results demonstrates the suitability of the proposed approach in delivering image matching that is not prone to view and illumination changes, the results also suggests that the performance is close to ASIFT based matching.

II. RELATED WORK

SIFT is one of the pioneering methods used for extracting the points of interest in an image. SIFT operates by extracting points of interests in addition to extracting those features that are present around these points of interest. It is one of the most reliable approaches for matching different viewpoint of a scene or object in an image. SIFT approach is robust in that it is not only invariant to image orientation but also to image scale. It’s hence capable of providing matching even in the presence of affine distortions. The performance of SIFT has also been found to be better in regard to 3D viewpoint, variation in illumination and presence of noise. SIFT operates by extracting features from reference images and storing them in a database. Image matching is then performed by comparing feature of the image with those stored in the data base. Euclidean distance is employed for finding candidate matches. Ke and Sukthankar [16] presented an “improved” version of SIFT descriptor by employing Principal Component Analysis (PCA) to identify the local features instead of the SIFT smoothed weighted histograms. The PCA-SIFT enhanced speed of implementation of SIFT matching process by an order of magnitude. Even though it could speed up SIFT it failed in delivering results that are as distinctive as that of the SIFT. Bay. H developed SURF [17],SURF stands for Speeded-Up Robust Features and its objective is to improve the strength of the leading existing feature detectors and descriptors (i.e. SIFT and PCA-SIFT). Unlike PCA-SIFT, SURF speeded up the SIFT’s detection process without damaging the quality of the detected points.

Over the decade many significant developments in image matching have been published. Zhang [18] provided a dense pattern of mass points for Digital Surface Model (DSM) generation using Geometrically-Constrained Cross-Correlation (GC3) algorithm. Since SIFT is only capable of detecting blob-like feature points and result in relatively limited matching results [19], many researchers have used it to achieve reliable seed points. Long [20] and Silveira [21] presented a hybrid matching approach that combines a feature-based algorithm and an area-based algorithm. Wu [22] and Zhu [23] proposed a triangulation-based hierarchical image matching method that employs few seed points to generate the initial Delaunay triangulations. The interest points are then matched under the triangle constraint and the epipolar constraint.

Xiong [24, 25] developed a robust interest point matching algorithm that first detects the super points with the greatest interest strength. Then, a control network is constructed, and each interest point is assigned a unique position and angle as its descriptor. IAlruzouq and Habib [26] used Modified Iterated Hough Transform (MIHT) to estimate the corresponding linear features between stereo pairs. Colerhodes et al. [27] introduced a registration algorithm that combined a simple yet powerful search strategy based on a stochastic gradient with two similarity measures, correlation and mutual information. Eugenio and Marques [28] developed two contour-matching techniques for satellite imagery, based on a general affine transformation, which modeled directly the corrections in the image domain without an explicit identification of the distortion sources.
Semi-Global Matching (SGM) [29, 30] successfully combines the concepts of global and local stereo methods for accurate, pixel-wise matching at low runtime; the core algorithm considers pairs of images with known intrinsic and extrinsic orientation, and the method have been implemented for rectified and unrectified images [31]. Humenberger et al. [32] introduced a stereo matching approach consisting of a combination of census-based correlation, SGM disparity optimization, as well as segmentation-based plane fitting for enhancements on textureless and occluded areas.

III. PROPOSED APPROACH

The primary objective of the proposed work is to design and develop a novel image matching algorithm which is not prone to view and illumination variations. This section describes in detail various steps involved in the proposed image matching algorithm.

Step 1: Key-points selection in both the images with the help of ASIFT.

Step 2: For every keypoint from both the images a descriptor is computed.

Step 3: An area is selected around every right keypoint node, considering possible maximum disparity range.

Step 4: All the keypoints are found in the area selected in step 3, around a right keypoint node in the right image.

Step 5: The procedure of step 3 and 4 is iteratively performed for all the right keypoints and the area around each right keypoint is selected considering the approximate maximum disparity range.

Step 6: The procedure in steps 3 and 4 is repeated for all left keypoints from left image iteratively. But here area around the left keypoint is a fixed sample area of size 16 x 16.

Step 7: The left keypoint node \( b_l \) is paired with every right keypoint node \( c_r \) and the pair is called as category pair.

Step 8: For every category pair, Euclidean distance between the descriptors of the keypoints is calculated.

Step 9: Weight is assigned to every right keypoint \( b_l \), in the selected area. The weight is inversely proportional to the Euclidean distance between the corresponding descriptors.

Step 10: For every category \( c_r \) the weight is calculated as

\[
w_l(c) = \frac{1}{k \times s_l(c) + 1}, c \neq \bar{c}
\]

(1)

Step 11: For every category set \( c, \bar{c} \) is undefined disparity category.

Considering weight \( w_l(\bar{c}) \) for \( \bar{c} \) which is undefined i.e. keypoint \( b_l(x, y) \) from left image does not correspond to any keypoint in the right selection area of right image. The weights cannot be used as probability estimates as \( w_l(c) \) is undefined and weights will not sum up to 1.

Step 12: Considering the keypoint matching as a classification problem, \( b_l \) is classified to one of the category \( c_r \). Initial probability for undefined category is given as

\[
p_l(c) = 1 - \max_{c \neq \bar{c}} (w_l(c))
\]

(2)

Step 13: Application of Bayes rule

\[
p_l(c) = p_l(c|l) \times (1 - \bar{p}_l(\bar{c})) \quad , c \neq \bar{c}
\]

(3)

\( p_l(c|l) : \) Conditional probability that \( b_l \) has category \( c \) as matching, given that \( b_l \) is matchable

\( (1 - \bar{p}_l(\bar{c})) : \) prior probability that \( b_l \) is matchable

Step 14: Estimating \( p_l(c|l) \) as below

\[
p_l(c|l) = \frac{w_l(c)}{\sum_{c \neq \bar{c}} w_l(c)}
\]

(4)

Step 15: Initial probabilities are assigned to every category \( c_r \) from right selection by equation (2), (3) and (4).

Initial probabilities which depend only on the similarity of neighborhood of candidate matching points can be improved using consistency property. The probability updating rule should have the property that the new probability \( p_{l,k+1}(c) \) should tend to increase when descriptors with highly probable category consistent with \( c \) are found nearby the keypoint region. Categories are considered consistent if they represent nearly the same disparity i.e.

\[|d(c_l) - d(c_m)| < \text{Threshold}\]
The threshold has to be decided empirically by inspecting the images and the results through trial and error method. For computing new probability $p_{ij}^{k+1}(c)$ for all $c_i$ in category set $c$, likelihood estimation is done. The degree to which the $c_j$ of $c$ strengthen $p_{ij}(c)$ should be related to estimated likelihood.

$$q_{ij}^k(c) = \sum_{m=1}^{L} p(c_m)$$

$|d(c_i) - d(c_m)| < \text{Threshold}$

$q_{ij}^k(c)$: Estimated likelihood considering the neighborhood of $c_j$, $L$: Number of neighbours in the category set.

Step 16: The category probability is updated using

$$p_{ij}^{k+1}(c) = \frac{p_{ij}(c)q_{ij}^k(c)}{\sum_{c_{m=1}}^{L} p(c_m)q_{ij}^k(c)}$$

where denominator acts as normalizing factor

$$p_i^{k+1}(c) = \sum_{j=1}^{N} a_j p_{ij}^{k+1}(c)$$

Here category probability is updated iteratively. The values in the $k^{th}$ iteration are used to calculate values in $k + 1^{th}$ iteration. $a_j$ are the weights associated with contribution of different neighbors of $b_i$. $N$ is the number of neighborhood points of $b_i$.

IV. TESTING DATA

The data set used for testing comprises of images with different degrees of tilt (both absolute and transitional tilt). These images help to validate the suitability of the proposed approach in performing image matching under different viewing conditions. Similarly, images with different degree of illumination are also considered for analysis and validation. Fig.1 illustrates images having an absolute tilt of specified degrees. It has a total of 9 images including the reference image with zero tilt. Fig. 2 illustrates the image dataset with images having different degrees of transitional tilt with $t=4$.
Illumination invariance property of the proposed approach is demonstrated by testing the algorithm for images with different degrees of illumination. Fig.3 illustrates 2 images from that dataset, having different degrees of illumination.
V. RESULTS AND DISCUSSION

As described in the previous section, different images with different viewpoints characterized by different degrees of absolute tilt and transitional tilt have been considered. Similarly images with different illuminations have also been considered to demonstrate the algorithm being not prone to view and illumination changes. One of the important and the foremost steps in the algorithm is to identify the key points. The following figures illustrate the keypoints / descriptors identified for certain images chosen from the data set. Number of keypoints identified eventually influences the number of matches identified and hence play very key role in the image matching algorithm. Fig. 4 illustrates different keypoints identified in the frontal image with absolute tilt of zero degrees, while Fig.5 depicts the descriptors identified in the image having an absolute tilt of 60 degrees compared to the frontal image.

![Fig. 4. Descriptors identified for the frontal image with 0 degrees absolute tilt.](image1)

![Fig. 5. Descriptors identified for the image with 60 degrees absolute tilt.](image2)

The descriptor for the image with transitional tilt of 4 is given in the Fig.6 while the descriptors for two different images with different degrees of illumination are given in Fig.7 and Fig.8
It can be inferred from the Fig.6 and Fig.8 representing transitional tilt and lesser illumination respectively, the number of descriptors identified are comparatively less when compared to the other images. This serve as
pointer towards the challenges in identifying matching points in images having a great degree of view change and illumination.

The results of image matching using the proposed approach, between the frontal image and the image with a 10 degree absolute tilt is depicted in the Fig. 9

Fig.9. Results of image matching between the frontal image and an image with absolute tilt of 10 degrees

A total of 827 matches were delivered by the proposed method, in order to visualize the accuracy of the proposed algorithm for the same set of images only 30 data points are plotted to show the accuracy of the matching. This is visualized with the help of Fig.10

Fig.10. Best 30 matches between the frontal image and an image with absolute tilt of 10 degrees

A total of 415 matches were delivered by the proposed method, when frontal image is matched with an image with an absolute tilt of 80 degrees, it can be seen from the Fig.11 there are very few mismatches even for a tilt of 80 degrees

Fig.11. Results of image matching between the frontal image and an image with absolute tilt of 10 degrees

The result of image matching when two images with a transitional tilt of $t=4$ at 0 degrees and a transitional tilt of $t=4$ at 80 degrees is illustrated in the Fig.12. A total of 298 matches were identified by the proposed algorithm.
Fig. 12. Results of image matching between an image with transitional tilt \( t=4 \) @ 0 degrees and transitional tilt \( t=4 \) @ 80 degrees.

Fig. 13 illustrates the luminance invariance property of the proposed algorithm in delivering matches when images having two different illuminations are matched. One image is the brightest image available in the dataset considered while the other one is the darkest image in that particular data set. A total of 1071 points were matched. For the same image set to visualize the accuracy of matching, the top 30 matched points are depicted using Fig. 14.

It can be inferred through visual inspection of Fig. 14 that the proposed approach delivers a matching accuracy of a very high degree. The points in two different images are accurately matched even with varying degree of illumination.
### TABLE I. Performance of ASIFT and Proposed method for images with absolute tilt

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Image Pair</th>
<th>Number of Matches Delivered by ASIFT Approach</th>
<th>Number of Matches Delivered by Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Frontal image, image with Tilt =10 degrees</td>
<td>749</td>
<td>827</td>
</tr>
<tr>
<td>2</td>
<td>Frontal image, image with Tilt =20 degrees</td>
<td>802</td>
<td>813</td>
</tr>
<tr>
<td>3</td>
<td>Frontal image, image with Tilt =30 degrees</td>
<td>635</td>
<td>771</td>
</tr>
<tr>
<td>4</td>
<td>Frontal image, image with Tilt =40 degrees</td>
<td>486</td>
<td>701</td>
</tr>
<tr>
<td>5</td>
<td>Frontal image, image with Tilt =50 degrees</td>
<td>367</td>
<td>608</td>
</tr>
<tr>
<td>6</td>
<td>Frontal image, image with Tilt =60 degrees</td>
<td>493</td>
<td>549</td>
</tr>
<tr>
<td>7</td>
<td>Frontal image, image with Tilt =70 degrees</td>
<td>225</td>
<td>575</td>
</tr>
<tr>
<td>8</td>
<td>Frontal image, image with Tilt =80 degrees</td>
<td>164</td>
<td>415</td>
</tr>
</tbody>
</table>

### TABLE II. Performance of ASIFT and Proposed method for images with Transitional tilt

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Image Pair</th>
<th>Number of Matches Delivered by ASIFT Approach</th>
<th>Number of Matches Delivered by Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 10 degrees</td>
<td>259</td>
<td>298</td>
</tr>
<tr>
<td>2</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 20 degrees</td>
<td>235</td>
<td>289</td>
</tr>
<tr>
<td>3</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 30 degrees</td>
<td>142</td>
<td>298</td>
</tr>
<tr>
<td>4</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 40 degrees</td>
<td>122</td>
<td>275</td>
</tr>
<tr>
<td>5</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 50 degrees</td>
<td>22</td>
<td>280</td>
</tr>
<tr>
<td>6</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 60 degrees</td>
<td>51</td>
<td>298</td>
</tr>
<tr>
<td>7</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 70 degrees</td>
<td>19</td>
<td>298</td>
</tr>
<tr>
<td>8</td>
<td>Transitional Image t=4, 0 degrees, image t=4, 80 degrees</td>
<td>27</td>
<td>298</td>
</tr>
</tbody>
</table>
TABLE III. Performance of ASIFT and Proposed method for images with different illumination

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Image Pair</th>
<th>Number of Matches Delivered by ASIFT Approach</th>
<th>Number of Matches Delivered by Proposed Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Image1, Image2</td>
<td>349</td>
<td>1417</td>
</tr>
<tr>
<td>2</td>
<td>Image1, Image3</td>
<td>480</td>
<td>1278</td>
</tr>
<tr>
<td>3</td>
<td>Image1, Image4</td>
<td>338</td>
<td>1188</td>
</tr>
<tr>
<td>4</td>
<td>Image1, Image5</td>
<td>315</td>
<td>1113</td>
</tr>
<tr>
<td>5</td>
<td>Image1, Image6</td>
<td>310</td>
<td>1071</td>
</tr>
</tbody>
</table>

Table (1), Table (2) and Table (3) depict the comparison in the performance; image matching as done by ASIFT based approach and the proposed one. It can be clearly observed from the results that the proposed method delivers a better performance across all the three scenarios when compared with ASIFT. It is interesting to note that the performance delivered in terms of number of matches identified by the proposed approach is extremely high in the case of images with transitional tilt.

VI. CONCLUSION

A Novel framework for view and illumination invariant image matching approach has been designed and successfully presented in this paper. The performance of the proposed algorithm was evaluated with images having different degrees of tilt and illumination. It can be clearly inferred from the results that the proposed approach has outperformed the ASIFT in terms of identifying the number of image matches. The accuracy of the proposed approach can also be visualized from the results; where in few best matching points have been displayed for better visualization. It can be observed from those depictions that the proposed approach has a high degree of accuracy in matching images under different view and illumination. It was also observed during the experimentation, that the proposed approach is faster than ASIFT making it suitable for real-time applications.

REFERENCES


AUTHOR PROFILE

D.Rajasekhar: He is Research Scholar at JNTUA, Ananthapuramu, Andhra Pradesh, India. His interest includes Image Processing and Communications.

Dr. T. Jayachandra Prasad: He is working as a Principal and Professor in the Department of Electronics and Communication Engineering at RGM CET, Nandyal-518501, Kurnool (DT), A.P. INDIA. His interest includes Digital Signal and Image Processing.

Dr. K. Soundararajan: He born at Tirupathi and graduated from S.V.U college of Engineering, Tirupathi in the year 1976. He got his post graduation from JNTU College of Engineering Kakinada in the year 1979 and obtained his Ph.D. from University of Roorkee, U.P, India in the year 1992. He produced 130 papers at both International and National journals and conferences. He got best technical paper award in consecutive years from Institution of Engineers India. His area of interest includes Micro Computers, Image Processing and Operating System. He got best teacher award from Andhra Pradesh Govt. in the year 2006 and also got best teacher award from lead India 2020. He worked as Professor at JNTU College of Engineering, Anantapur and former Rector of JNTUA and also Officer on Special Duty to Vice Chancellor JNTUA. Presently he is Professor and Dean (R&D) at TKR Engineering College, Hyderabad, Telangana, India-500097.