

# VIEW AND ILLUMINATION INVARIANT ITERATIVE BASED IMAGE MATCHING

Y. Ramesh<sup>1</sup>, J. Sofia Priya Dharshini<sup>2</sup>

<sup>1</sup>Student, <sup>2</sup>Associate Prof., Department of ECE, RGM CET, Nandyal, Kurnool Dist., Andhra Pradesh, INDIA  
ramesh7792@gmail.com, sofi\_prida@yahoo.co.in

## Abstract

In this paper, the challenges in local-feature-based image matching are variations of view and illumination. Different methods have been recently proposed to address these problems by using invariant feature detectors and distinctive descriptors. However, the matching performance is still unstable and inaccurate, particularly when large variation in view or illumination occurs. In this paper, we propose a view and illumination invariant image matching method. We iteratively estimate the relationship of the relative view and illumination of the images, transform the view of one image to the other, and normalize their illumination for accurate matching. The performance of matching is significantly improved and is not affected by the changes of view and illumination in a valid range. The proposed method would fail when the initial view and illumination method fails, which gives us a new sight to evaluate the traditional detectors. We propose two novel indicators for detector evaluation, namely, valid angle and valid illumination, which reflect the maximum allowable change in view and illumination, respectively.

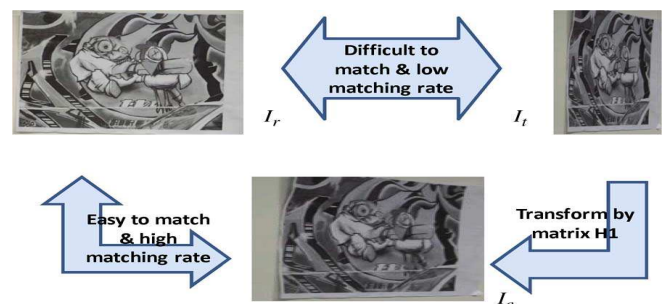
**Keywords-**Feature detector evaluation, image matching, Iterative algorithm.

\*\*\*

## 1. INTRODUCTION

Image matching is a fundamental issue in computer vision. It has been widely used in tracking, image stitching, 3-D reconstruction, simultaneous localization and mapping (SLAM) systems, camera calibration, object classification, recognition, and so on. Image matching aim to find the correspondence between two images of the same scene or objects in different pose, illumination and environment. In this paper, we focus on local feature-based image matching. The challenges of this work reside in stable and invariant feature extraction from varying situations and robust matching. Generally speaking, the framework of a region feature based image matching consists of three steps. *Detecting* stable regions. Interesting points are extracted from images, and the region of interest is the associated circular (or elliptical) region around the interesting point. Generally, researchers use corner (Harris [1], SUSAN [2], CSS [3], etc.) or center of silent region (SIFT [4], SURF [5], DoH [6], HLSIFD [7], etc.) as the interesting point since they are stable and easy to locate and describe. The radius of the region is determined by *a priori* setting (Harris corner) or the region scale (scale invariant features). The total number of features detected is the minimum number of the features extracted from the matched images. *Describing* regions. Color, structure, and texture are widely used to describe images in the recent literature. Descriptors with edge orientation information (SIFT and HOG) are also very popular since they are more robust to scale, blur, and rotation. *Matching* features. Local features from two images are first matched when they are the nearest

pair. A handful of distances can be used in practice, such as L1 distance, L2 distance, histogram intersection distance [8], and earth mover's distance [9]. If the nearest distance is higher than  $k$  times ( $k \in (0, 1)$  empirically) of the second nearest distance, the nearest matching pair will be removed. These are the very initial matching results.



**Fig1** Illustration of the proposed matching algorithm  $I_r$  and  $I_t$  are the images to be matched.  $I_e$  is simulated from  $I_t$  by transformation  $T$ .  $I_r$  is difficult to match with  $I_t$  for the difference of view point and illumination, whereas  $I_e$  is easier to match with  $I_t$  since they are closer in the parameter space.

The three parts of the detect–describe–match (DDM) framework determine the performance of image matching. The first step is the basis of this framework. Unstable and variant features increase the difficulties of the next two steps. Researchers mostly focus on the first step for invariant feature extraction and have proposed many excellent detectors [1],

[4], [5], [7]. However, an important experience of a previous work is that all the aforementioned feature detectors are not strictly invariant to the changes of view and illumination. For larger changes, there would be few invariant features that can be extracted from both images to be matched. This motivates us to think the essential difference of images with different view and illumination. Normally, a question need to be answered: whether an object in two images with different views and illumination looks like the same one, supposing there are two images with a large view change, as shown in Fig. 1. The two top images are the same object in different views. They are so different in appearance that they can be considered as two different objects. We do not attempt to find invariant local feature detectors as in a previous work but focus on a better framework for image matching.

## 2. VIEW AND ILLUMINATION INVARIANT IMAGE MATCHING

### 2.1. General Definition of Image Matching

Two images of the same object or scene are shown as two points in parameter space  $P$  of the object (scene). Let  $I$  be the original appearance of an object,  $I_r = \mathbf{L}(\mathbf{H}(I))$  be the real appearance of the object shown from an image, where  $\mathbf{L}$  indicates the illumination, and  $\mathbf{H}$  is the object transformation factor from a normal pose. Here, we define the parameter space of a given image  $I$  as  $P_I = \{\mathbf{H}_I, \mathbf{L}_I\}$  (simply written as  $P = \{\mathbf{H}, \mathbf{L}\}$  in the following). Translation  $\mathbf{L}$  and  $\mathbf{H}$  is a point in the parameter space; thus, the observed image is shown as a point in the parameter space, which is expanded by object  $I$ . Therefore, the purpose of image matching is to find transformation  $T$  between the two points in the parameter space  $\{\mathbf{L}_r, \mathbf{H}_r\} \xrightarrow{T} \{\mathbf{L}_t, \mathbf{H}_t\}$  or, in other words,  $T|I_r = T(I_t) \cdot I_r, I_t \in P_I$ . The purpose is to find the coordinate differences between the two points. The norm of this space is difficult to define since illumination factor  $\mathbf{L}$  and transformation  $\mathbf{H}$  are totally independent and cannot be combined together. In this paper, we simply use images with planar objects; therefore,  $\mathbf{H}$  is the homograph transform matrix, and  $\mathbf{L}$  is the histogram matching function that transforms the histogram of one image to a specific one.

### 2.2. Proposed Method

Denote the reference image and test image to be matched as  $I_r$  and  $I_t$ . Suppose that the true pose transformation matrix from  $I_t$  to  $I_r$  is  $\hat{\mathbf{H}}$  and the illumination change function is  $\hat{\mathbf{L}}$ . The relationship between  $I_r$  and  $I_t$  is

$$I_r(\mathbf{X}) = \hat{T}(I_t) = \hat{\mathbf{L}}(\hat{\mathbf{H}}(I_t)) = \hat{\mathbf{L}}(I_t(\hat{\mathbf{H}}\mathbf{X})) \quad (1)$$

Where  $\hat{T}$  is the true transformation between  $I_t$  and  $I_r$ , is the homogeneous co-ordinates, and  $X = (x, y, 1)$ . If there exists approximate estimations about illumination and

transformation, the  $I_t$  could be transformed to an estimated image  $I$ , i.e.,

$$I(\mathbf{X}) = T(I_t) = \mathbf{L}(I_t(\mathbf{H}\mathbf{X})) \quad (2)$$

Where  $\mathbf{H}$  denotes the view point transformation  $\mathbf{L}$  and denotes the illumination transformation. If  $T$  is not a very rough estimation between  $I_r$  and  $I_t$ , the estimated image  $I$  would be more similar to  $I_r$  than  $I_t$  itself. In other words,  $I_r$  is closer to  $I$  than to  $I_t$ . Thus, the matching between  $I$  and  $I_r$  will be easier, as shown in Fig. 1.

In this way, we propose the following iterative image matching process:

$$I_1X = T_1(I_0) = \mathbf{L}_1(I_0(\mathbf{H}_1X^T)) \quad (I_0 = I_t)$$

$$I_iX = T_i(I_{i-1}) = \mathbf{L}_i(I_{i-1}(\mathbf{H}_iX^T)) \quad (i > 1) \quad (3)$$

### 2.3. Algorithm 1: The Proposed Method:

Initial:  $T_0 = \{\mathbf{H}_0, \mathbf{L}_0\} = \{\mathbf{E}, \vec{1}\}, T = T_0, \sigma_H, \sigma_L;$

Iterate

$i = i + 1;$

Estimate  $T_i: \mathbf{H}_i, \mathbf{L}_i;$

$T = T_i \circ T;$

$\mathbf{H} = \mathbf{H}_i * \mathbf{H}.$

Transform  $I_{i-1}$  to  $I_i$  by (3);

Until  $\mathbf{H}_i - \mathbf{E} < \sigma_H, |\mathbf{L}_i - \vec{1}| < \sigma_L$  or  $i > n$ . ( $\mathbf{E}$  is the unit matrix.  $L_i$  is a histogram transformation vector,  $\sigma_H$  and  $\sigma_L$  are convergence thresholds.)

Return  $\mathbf{T}, \mathbf{H}$

The algorithm is summarized in Algorithm 1. The final estimation of the  $\hat{T}$  is

$$\hat{T} = \dots \circ T_m \circ T_{m-1} \circ \dots \circ T_2 \circ T_1 \quad (4)$$

$$\approx T_n \circ T_{n-1} \circ \dots \circ T_2 \circ T_1 \quad (5)$$

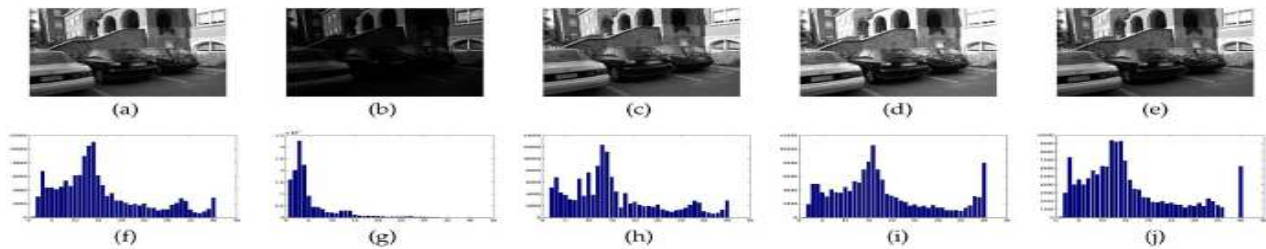
Where “ $\circ$ ” denotes function composition. Our experiments in Section IV-B show the convergence of the iteration with SIFT and the performance with respect to the number of iterations

### 2.4. Estimate the Parameters $\mathbf{H}$ and $\mathbf{L}$

General image-matching methods by local features focus on the first parameter  $\mathbf{H}$  since the concerned issue is the space correspondence between the two images. One of the advantage of the proposed method is that it also estimates the illumination change, which makes matching much better when illumination has changed.

The purpose of general image-matching methods is to find the transformation matrix between the reference image and the test image. These methods are invariant to rotation, scale, and partially affine changes. The  $\mathbf{H}$  can be easily estimated by the general methods without other information. First, we extract features from the matching images and obtain features descriptions (which method is used is not important). Then, we match two features when they are the nearest pair in the feature space. Here,  $L_2$  norm is used to calculate the distance

between to the features The RANSAC algorithm is employed to calculate transformation matrix. The general methods, i.e., HarAff, HesAff, SURF, SIFT, and HLSIFD, all can be used as the feature extraction method. We call them I-HarAff, I-HesAff, ISURF, ISIFT, and IHLSIFD (“I” indicates “Iterative”), respectively. Moreover, image matching is usually used in video sequences.



**Fig.2** Illustration of the histogram transformation (a) The original image. (b) Darker image. (c) Transformed image from (b) according to the histogram of (a). (d) Brighter image. (e) Transformed image from (d) according to the histogram of (a). (f)–(j) The corresponding histograms of (a)–(e).

We assume that the difference between two consecutive frames is not large, and the object or the camera smoothly moves. Thus, the  $t$ th frame’s transformation  $\mathbf{H}_t$  can be approximated by the previous results. Different detectors and descriptors have been developed to extract illumination invariant local features. The gradient direction histogram is normalized to form the descriptors. There is usually a tradeoff between the distinction and the invariance. If we do not normalize the descriptors, they will be sensitive to illumination changes but more distinctive. Computing detectors and descriptors also cost much time. Conversely, the detector will be more efficient if we do not require the detector to be invariant to illumination change. We want to keep both illumination invariant and descriptor distinctive in our method. Thus, it is necessary to estimate the illumination change between the two images.

other image is  $h_2$ . We calculate the cumulative functions of  $h_1$  and  $h_2 - F_1$  and  $F_2$ . Finally, the translation function is

$$\mathbf{L} = \mathbf{F}_2^{-1}\mathbf{F}_1. \tag{6}$$

Since the cumulative function of gray histogram is always monotonically increasing, inverse function  $\mathbf{F}^{-1}$  always exists. We transform the histogram of the test image according to the histogram of the reference image to normalize the illumination between the pair, as shown in Fig. 2, and the whole procedure is illustrated in Fig. 3. To sum up, we estimate transformation matrix  $\mathbf{H}$  between the matching pairs by feature detector, estimate the illumination relationship, and change one of the images according to the color histogram of the other to map the pose and illumination of the object in one image to the other.

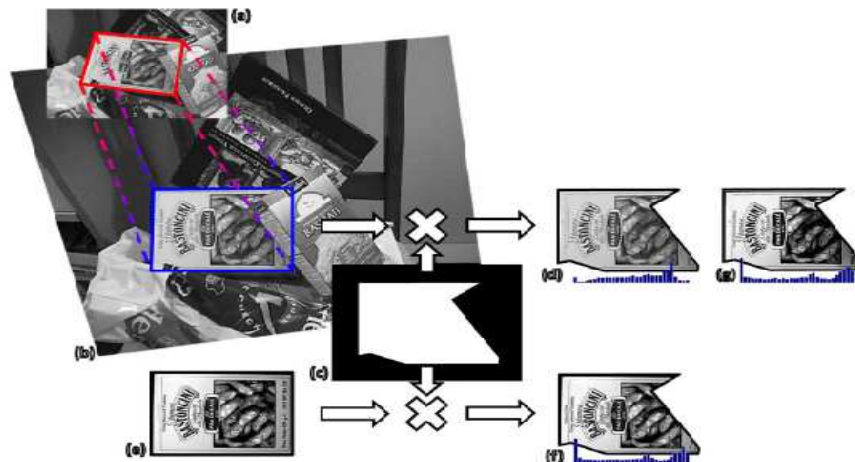
### 2.5. Relationship between the Iterative Algorithm and ASIFT

The proposed iterative method is similar to ASIFT in ASIFT, the features are not invariant to affine change, but they cover the whole affine space, as shown in the middle block in Fig. 4. Every simulation of the reference image is one pose of the image in the affine space. Therefore, parts of the simulations of the reference image and the test image should have similar poses in the affine space theoretically. The simulations of the reference image and the test image are independently constructed. No mutual information is used in the simulations

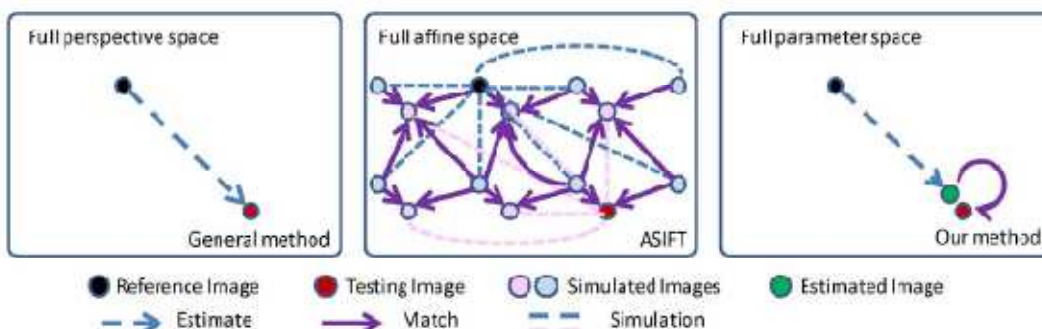
**Table 1:** Comparison of Asift and our Method

	ASIFT	Our method
Simulation to ref image	Yes	No
Simulaton to test image	Yes	Yes
Number of simulations	Many	Few
Number fo features	$10^4 \sim 10^5$	$10^3$
Pose simulation	Yes	Yes
Illumination simulation	No	Yes
NCM	High	High
RS	Vey low	High
Affine invariancy	Full	Partial
Computational cost	High	low
Rea-time	No	Yes

Simulating in a high density in the affine space, may supposed image poses are constructed, and then, they are matched in a general way. The number of matches increases with the number of the simulations. ASIFT indeed increases the invariability of the image-matching method. However, it does not care what the transformation matrix between the reference and test images is, by trying many possible transformations and combining the matches. Thus, ASIFT can be regarded as a sampling method around the original points in parameter space, whose properties are shown in the left column of Table I. Essentially, our method also constructs “simulation.” We simulate the image not only in the pose but also in illumination, as shown in the right part of Fig. 4. In addition, we transform one simulation per iteration, and in most tasks, two iterations are enough. We will give an experiment to illustrate this in Section IV-B. Benefiting from few simulations, the computational cost of our method is very low, compared with ASIFT, which simulates much more images than our method.



**Fig. 3** Procedure of illumination estimation (a) The test image. Warp (a) by the estimated transformation matrix to generate (b). (c) Mask with the matched regions labeled as 1, and the unmatched regions labeled as 0. (d) The inner product of (b) and (c). (e) The reference image. (f) The inner product of (c) and (e). (g) Illumination simulated image from (d) according to the histogram of (f).



**Fig. 4** Relationship among the general framework, ASIFT, and the proposed method. (Left block) The general framework, (middle block) ASIFT, and (right block) ours. The general DDM framework directly estimates the transformation between two images. It is simple but coarse. ASIFT simulates many poses of the two images to cover the affine space, whereas our method estimates the transformed pose first and then accurately matches in the projective space.

A coarse-to-fine scheme can reduce the computational time of ASIFT to three times of the SIFT, whereas our method only costs two times. One drawback of the proposed method is that it does not increase the invariability of the original method. When the initial method fails in matching images, the proposed method also fails. One promising method to overcome this shortage is to combine the proposed method with the ASIFT, which improves both the invariability and the accuracy. Furthermore, the histogram matching may amplify noise that seems to affect the performance. A few more key points would be extracted after the histogram matching, but they would not affect the performance too much. We will show this in Section IV-C. Experimental results show that the performance of the proposed framework reaches a comparable level, compared with ASIFT with much fewer features totally detected, as shown in Table III. Therefore, the RS of our method is much higher than that of ASIFT. The computational cost of our method is much lower than that of ASIFT because much fewer features are required.

Above all, there are some common properties between iterative SIFT (ISIFT) and ASIFT. Instead of directly matching the original images, both methods find good simulations of the original pairs. ASIFT samples the imaginary images in the whole affine space, whereas our method directly estimates in the whole parameter space. We should point out here that these comparisons and the experiments shown in the following section are all under the situation that the original method, i.e., SIFT, still works. When it fails, the proposed method also fails, whereas the ASIFT can still obtain a valid result.

### 3. EXPERIMENTAL RESULTS

#### 3.1. Database

In the first experiment, we want to show the performance of the proposed method. We capture two images with changes both in illumination and view. This experiment is not used for comparison, but it only shows the effectiveness of the proposed method. To evaluate the performance of the proposed image-matching framework, we do experiments on the database provided by Mikolajczyk.1 This database contains eight groups of images with challenging transformations. Parts of them are shown in Fig. 5. We compare the proposed method with ASIFT and the usual DDM framework with the state-of-the-art detectors: SURF, SIFT. In addition, two evaluations on the detectors through our strategy are proposed. One of them tests the adaptive capacity on the view change, and the other tests the capacity on the illumination change. To finish the two evaluations, we build two databases. One of them contains 88 frames with view changes from  $0^\circ$  to  $87^\circ$ . The other one contains 55 frames with light exposure changes from 40 to 14 (0.1 EV). The two databases contain continuous transformation frames. Thus, we can evaluate the view invariant ability of the

detectors at a 1 interval and the illumination change invariant ability at a step of 0.1 EV. Such databases seldom appear in the open literature, and they will be currently available on the Internet.

#### 3.2. Convergence

As we mentioned in Section III-B, the number of iteration is an important parameter. A question that should be answered is whether more iterations bring better performance. Experiments show that, under the proposed framework, our method converges very fast. Fig. 6 shows an experiment on matching two images. The reference image is captured from a frontal view, and the test image is captured from a view angle of  $60^\circ$ , as shown in Fig. 6(a) and (b), respectively. Here, SIFT is used as the base detector. The RS and NCM of our method and the DDM framework with SIFT are drawn for comparison, as shown in Fig. 6(c) and (d).

**Table 2:** Performance of sift, isift with only pose estimation, isift both pose and illumination estimation

	SIFT	ISIFT(H)	ISIFT(H&L)
Total detected	436	388	2021
Total matches	50	64	169
NCM	39	57	153
RS(%)	8.95	14.7	7.57
MP(%)	78.0	89.1	90.5

The results show that more iteration does not necessarily increase the performance significantly, whereas it increases the computation time linearly. When,  $n=2$  the performance significantly increases. The NCM increases more than 300 matches from only 12 to 365, and the RS increases from 12.1% to 37.1%. However, as  $n$  further increases the performance little, the NCM only moves around 360, and the RS moves around 37%. Thus, two iterations are enough in general situations, and we use  $n=2$  in the following experiments. Moreover, all the features in this experiment and the following experiments are described by a SIFT descriptor, except SURF, which is described by a SURF descriptor. Following the general evaluation, three criteria are often used as feature evaluator.

- 1) NCMs are the number of total correct match pairs.
- 2) RS is the ratio between the NCM and the minimum of total number of features detected from the image pair  $RS = \frac{NCM}{\min(TOTAL)}$ .
- 3) Matching precision (MP) is the ratio between the NCM and the number of matches  $MP = \frac{NCM}{Matches}$ .

#### 3.3. Performance

In this experiment, a brief view of the performance of the proposed method is given. We use SIFT as the base detector in this experiment (ISIFT). Two images with both view and illumination changes are matched here.

We first match the two images by SIFT, and then, we only simulate the pose of the left image in our strategy. Finally, we simulate both pose and illumination. The matching results are shown in Fig. 7 and Table II. View and illumination changes both degrade the performance of the general method. SIFT could achieve 8.95% RS with 39 correct matches. ISIFT, with the pose estimation only, could achieve 14.7% RS with 57 correct matches. When we estimate the pose and illumination changes, the number of total detected features rapidly increases, and the NCM increase to 153. Because histogram matching amplifies noise in simulation, many fake features are detected, and the RS is reduced to 7.57%. This experiment is only a brief view of our strategy, and more experiments will

be presented in the following. We estimate the global illumination change between the matching pair to increase the NCMs. The illumination change is usually continuous in the image. Thus, revising the illumination of part of the image could benefit to other regions. Our algorithm does not increase the invariance of the original detector, but it increases the accuracy, stability, and reliability of the matching results. When SIFT fails, our method also fails. However, when SIFT works, but not robust, the proposed method will play an important role. More matches could not increase the invariance, but it can increase the accuracy of alignment when the matching by SIFT is inaccurate.



Fig.5. Four groups of images that we used for comparison [33] Each group contains one or two transformations with six images, and only parts of them are shown here. (a) Boats (scale and rotation). (b) Graf (view). (c) Wall (view). (d) Leuven (illumination)

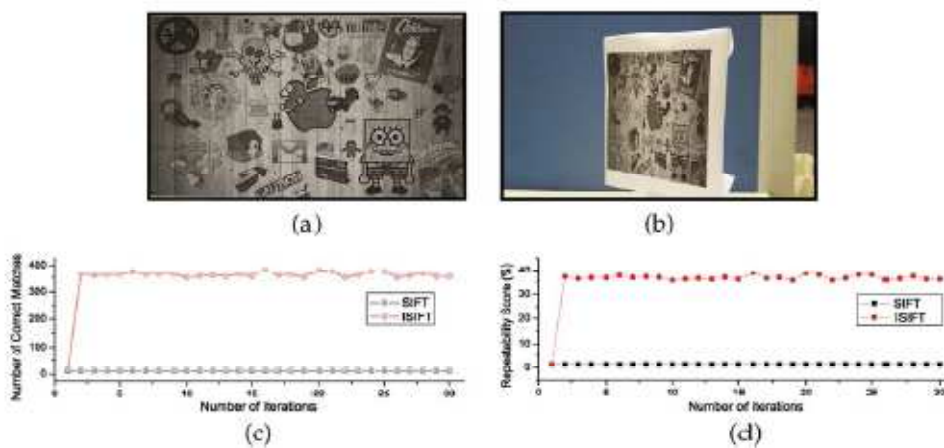
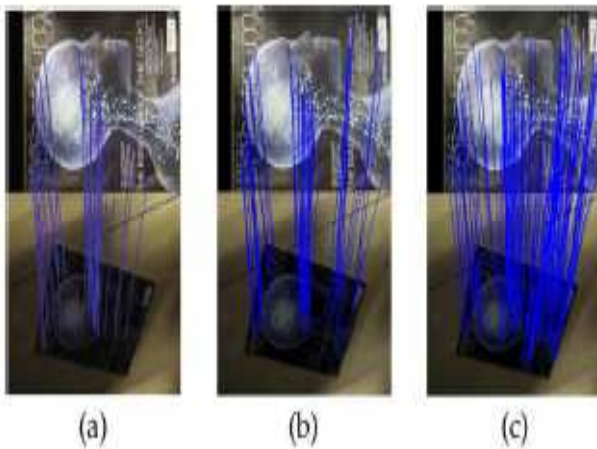


Fig.6. Experiments of convergence. (a) The reference image. (b) The test image. (c)–(d) The NCM and RS of ISIFT compared with SIFT.



**Fig.7.** Matching results of SIFT and ISIFT. (a) Matching result of SIFT. (b) Matching result of ISIFT with only pose simulation ( $\mathbf{H}$ ). (c) Result of ISIFT with both pose and illumination ( $\mathbf{H}$  and  $\mathbf{L}$ ) simulation.

In other words, the advantage of the proposed method is that the performance does not degrade with the increase in the pose change or transition tilt, which is addressed in the valid range. Additionally, the local key point location will be more accurate than that of the original detected point. To corroborate this point of view, we show an extra experiment in the following. The first row in Fig. 8 is the matching results of SIFT, and the second row is the results of ISIFT. Both the matches and the alignment residual error are shown. From this experiment, we can find that our algorithm can obtain less error than SIFT, and the NCM affects the accuracy of matching very much.

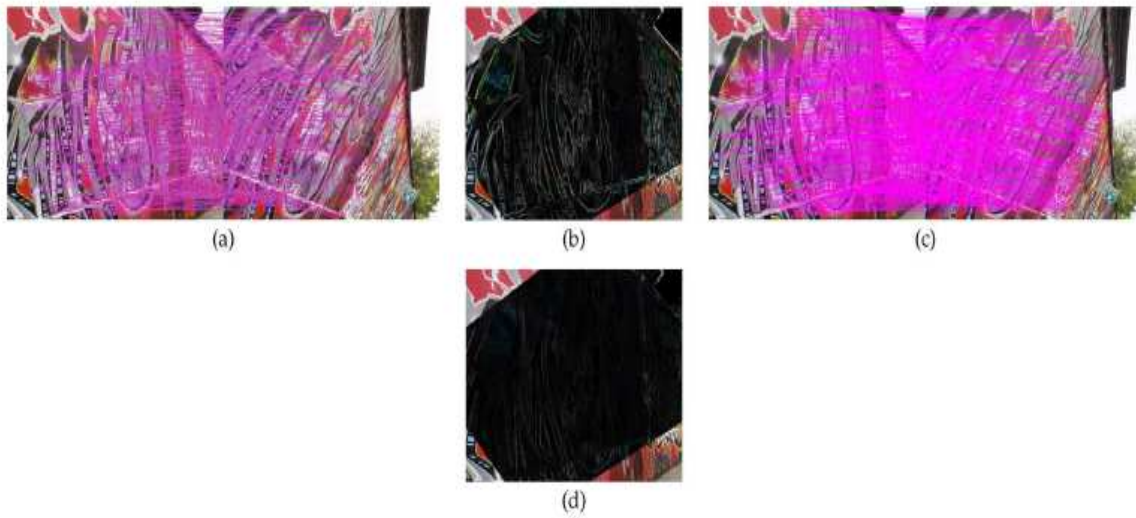
### 3.4. Comparison

We choose the database provided by Mikolajczyk and compare them with SURF, SIFT and ASIFT. Four pairs of images with scale, view, and illumination change are tested, as shown in Fig. 9. The images on top are the reference image, and those at the bottom are the test image. Table 3 is a comparison of this experiment in terms of NCM, RS, and MP. Our method estimates the pose and illumination of the matching pairs and simulates the reference image. Therefore, the simulated image is closer to the original image, which contains most information of the original image, shortening the distance of the matching pairs in the parameter space. First, the NCM of the ISIFT is much higher than that of the traditional methods. They obtain 584 matches, whereas SURF obtains 9 matches, and SIFT obtains 46 matches in the Graf (affine change situation; second row in Fig. 9). SURF and SIFT obtain 793 and 2837 features, respectively. Thus, the RS of ISIFT increases to 36.4%, whereas that of SURF and SIFT is only 1.14% and 1.62%. This implies that the efficacy of IIM framework is much better than the traditional DDM

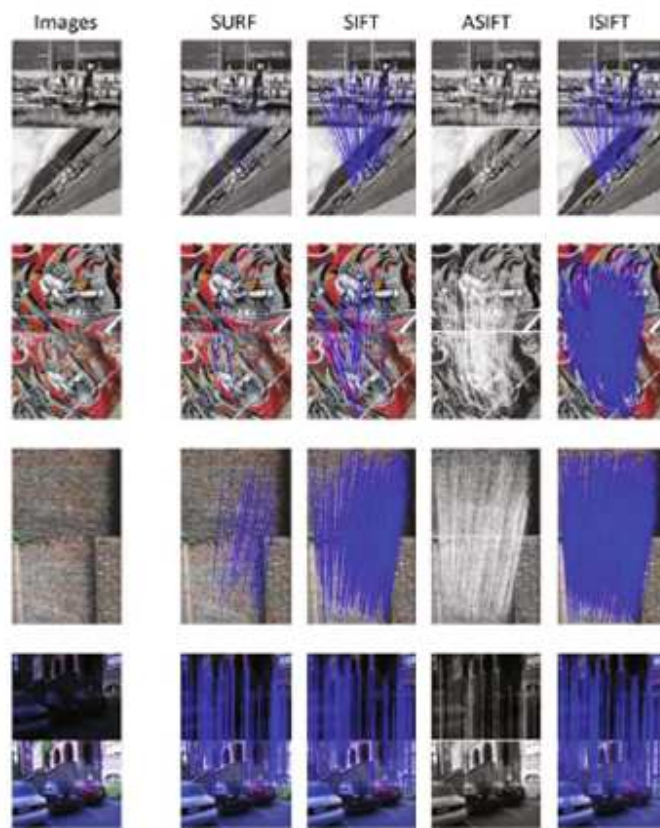
framework. We increase about 32 times and 22 times RS in this view-change experiment. With the significant increasing performance, we can make the matching more stable and reliable. Similarly, more correspondences are found in other experiments, particularly under affine and illumination change situations. Our method does not significantly increase NCM under only scale change comparing to SIFT, SURF, and HLSIFD since they are theoretically scale invariant. The RS and MP also significantly increase.

However, in extreme situations when SIFT fails in the first matching, our algorithm also fails. The proposed method can increase the stability, reliability, and accuracy of the original detector, but it cannot increase the invariance. A solution is integrating the proposed method into ASIFT as the second layer to refine the original matching results. We will show an experiment in Section IV-E. ASIFT also obtains 105, 465, 556, and 157 matches from Boat, Graf, Wall, and Leuven matching images (61, 46, 409, and 259 matches are found by SIFT, respectively). However, these matches are calculated from 29 985, 45 151, 64 908, and 22 562 extracted features. Indeed, ASIFT increases the NCM, but they need to extract much more features from the images, which cost much time in computation. More detail results are summarized in Table III.

In this paper, we try to link our method with the general optimization theory. Essentially, the target of image matching is finding the correspondence. We want to find the transformation function between the matching pair, which can minimize the matching error. Thus, we optimize the view difference and then optimize the illumination. With the two-step optimization, our method can find more accurate transformation function. Different from ASIFT, the proposed method does not increase the invariance of the original detector, but it increases the stability and reliability.



**Fig.8.** Matching error of the SIFT and the proposed method. (a) The matches of SIFT. (b) The residual error of SIFT. (c) The matches of ISIFT. (d) The residual error of ISIFT.



**Fig. 9** Matching results of four groups of images. (Test images from top to bottom) Boat, Graf, Wall, and Leuven. The results of the correct matches are drawn in blue or white lines.



**Table 3:** Comparison of the algorithms on view change pairs

Methods		SURF	SIFT	ASIFT(HR)	ISIFT
Boat / Scale	Total	722	7986	29985	615
	Matches	43	125	-	94
	NCM	8	61	105	79
	RS(%)	1.25	0.764	0.35	12.8
	MP(%)	18.6	48.8	-	84
Graf / Affine	Total	793	2837	45151	1605
	Matches	34	210	-	586
	NCM	9	46	465	584
	RS(%)	1.14	1.62	1.03	36.4
	MP(%)	26.5	21.9	-	99.7
Wall / Affine	Total	1730	7094	64908	5358
	Matches	78	452	-	834
	NCM	40	409	556	833
	RS(%)	2.31	5.77	0.857	15.5
	MP(%)	51.3	90.5	-	99.9
Leuven / Illumination	Total	647	999	22562	1159
	Matches	172	289	-	379
	NCM	161	259	157	344
	RS(%)	24.9	25.9	6.96	29.7
	MP(%)	93.6	89.6	-	90.8

### 3.5. Real-Time Image Matching

An important application of image matching is object detection and poses estimation in video frame. Suppose that the camera smoothly moves and the reference image can be matched with the first frame, the estimation of the transformation matrix from the reference image to certain frame in video can be initialized from the matching of the previous frame. In addition, we match the first frame with the reference image directly by local-feature-based image-matching method. We directly use SIFT here. The RS and NCM of our method and SIFT are shown in Fig. 14, and parts of the matching results of ISIFT and SIFT are shown in Fig. 13. The RS of our method (ISIFT is used here) stays around 30%, and NCM is always higher than 100 pairs in this experiment.

The RS of SIFT is running around 7%. Only a small part of features are useful for the correspondence calculation. The NCM of SIFT is about 70 matches, which is lower than that of the proposed method. The mean of the RS and NCM of the ISIFT and SIFT is, respectively 29.6%, 137, 5.7%, and 66. Our method accurately calculates matches all through the video frames, even in large view changes such as frames 750 to 900. To sum up, ISIFT is very accurate and stable in real applications.

We develop a real-time image-matching system to show the efficiency. The proposed method could cope with a wide range of view and illumination changes with stable matches, as shown in Fig. 15. We compare the real performance of

SURF and SIFT by using them as our basic detector. ISURF is faster than ISIFT; however, it is not as stable as ISIFT. The system is implemented on a computer with two dual-core 2.8-GHz central processing unit, and the processed image size is 640 X 480. The matching could be finished in 80 ms, with parallel coding in a algorithmic level.

### CONCLUSIONS

In this paper, we have proposed a novel image-matching algorithm based on an iterative framework and two new indicators for local feature detector, namely, the VA and the VI. The proposed framework iteratively estimates the relative pose and illumination relationship between the matching pair and simulates one of them to the other to degrade the challenge of matching images in the valid region (VA and VI). Our algorithm can significantly increase the number of matching pairs, RS, and matching accuracy when the transformation is not beyond the valid region. The proposed method would fail when the initial estimation fails, which is relative to the ability of the detector. We have proposed two indicators, i.e., the VA and the VI, according to this phenomenon to evaluate the detectors, which reflect the maximal available change in view and illumination, respectively. Extensive experimental results show that our method improves the traditional detectors, even in large variations, and the new indicators are distinctive.

## REFERENCES

- [1]. C. Harris and M. Stephens, "A combined corner and edge detection," in Proc. 4th Alvey Vis. Conf., 1988, pp. 147–151.
- [2]. S. M. Smith and J. M. Brady, "Susan—A new approach to low level image processing," Int. J. Comput. Vis., vol. 23, no. 1, pp. 45–78, May 1997.
- [3]. F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale space," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 12, pp. 1376–1381, Dec. 1998.
- [4]. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [5]. H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded-up robust features (SURF)," Comput. Vis. Image Understand., vol. 110, no. 3, pp. 346–359, Jun. 2008.
- [6]. T. Lindeberg, Scale-Space Theory in Computer Vision. Norwell, MA: Kluwer, 1994.
- [7]. Y. Yu, K. Huang, and T. Tan, "A Harris-like scale invariant feature detector," in Proc. Asian Conf. Comput. Vis., 2009, pp. 586–595.
- [8]. A. Barla, F. Odono, and A. Verri, "Histogram intersection kernel for image classification," in Proc. Int. Conf. Image Process., 2003, vol. 3, p. III–513–16 [Online]. Available: <http://dx.doi.org/10.1109/ICIP.2003.1247294>, vol.2
- [9]. Y. Rubner, C. Tomasi, and L. J. Guibas, A Metric for Distributions With Applications to Image Databases. Washington, DC: IEEE Comput. Soc., 1998, p. 59.

## BIOGRAPHIES



**Y. Ramesh** received B.Tech degree from GATES Engineering College in the year 2010 and currently pursuing M.Tech in Digital Systems and Computer Electronics at RGM College of Engineering and Technology, Nandyal, Kurnool(dist), Andhra Pradesh. His areas of interest are Digital Image Processing, Data Communication and the application of Digital systems.



**J. Sofia priya dharshini** has received B.Tech degree (Electronics & Communication Engineering) in the year 2005, M.Tech (Digital Systems and Computer Electronics) in 2009 and also pursuing Ph.D. (Wireless Communications and Networking) from JNTU Anantapur, A.P, India. Presently she is working as

Associate Professor in the dept. of ECE, RGM CET, Nandyal. She has published two papers in International Journals and nine papers in National and International Conferences. Her area of interest includes Wireless communications and Networks, Mobile Computing and Video and Image processing.