

Robust Techniques for Feature-based Image Mosaicing

Thesis submitted in partial fulfilment of the requirements for the degree

of

Master of Technology

in

Electronics and Instrumentation Engineering

by

Vimal Singh Bind

Roll No: 211EC3304



Department of Electronics & Communication Engineering

National Institute of Technology Rourkela

Rourkela, Odisha-769008

May 2013

Robust Techniques for Feature-based Image Mosaicing

Thesis submitted in partial fulfilment of the requirements for the degree

of

Master of Technology

in

Electronics and Instrumentation Engineering

by

Vimal Singh Bind

Roll No: 211EC3304

Under the Supervision of

Prof. Umesh Chandra Pati



Department of Electronics & Communication Engineering

National Institute of Technology Rourkela

Rourkela, Odisha-769008

May 2013

Dedicated
to
My Family and Friends



Department of Electronics & Communication Engineering
National Institute of Technology, Rourkela

CERTIFICATE

This is to certify that the Thesis Report entitled “ **ROBUST TECHNIQUES FOR FEATURE-BASED IMAGE MOSAICING** ” submitted by **Mr. VIMAL SINGH BIND** bearing roll no. **211EC3304** in partial fulfilment of the requirements for the award of Master of Technology in Electronics and Communication Engineering with specialization in “**Electronics and Instrumentation Engineering**” during session 2011-2013 at National Institute of Technology, Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the thesis has not been submitted to any other University / Institute for the award of any Degree or Diploma.

Prof. Umesh Chandra Pati

Place:

Associate Professor

Date:

Dept. of Electronics and Comm. Engineering

National Institute of Technology

Rourkela-769008

Acknowledgement

The real spirit of achieving a goal is through the way of excellence and austere discipline. I would have never succeeded in completing my task without the cooperation, encouragement and help provided to me by various personalities.

I am grateful to all who have contributed towards shaping this thesis. At the outset, I would like to express my sincere thanks to Dr. Umesh Chandra Pati for his advice during my thesis work. As my supervisor, he has constantly encouraged me to remain focused on achieving my goal. His observations and comments helped me to establish the overall direction of the research and to move forward with investigation in depth. He has helped me greatly and been a source of knowledge.

I am also thankful to all the professors of the department for their support and encouragement. I must acknowledge the academic resources that I have got from NIT Rourkela. I would like to thank administrative and technical staff members of the Department who have been kind enough to advise and help in their respective roles.

I am really thankful to my all friends. My sincere thanks to everyone who has provided me with kind words, a welcome ear, new ideas, useful criticism, or their invaluable time, I am truly indebted.

I would conclude with my deepest gratitude to my parents, my classmates and all my loved ones. My full dedication to the work would have not been possible without their blessings and moral support. This thesis is a dedication to them who did not forget to keep me in their hearts when I could not be beside them.

Date:

Roll No: 211EC3304

Place:

Dept. of ECE

NIT, Rourkela

Abstract

Since the last few decades, image mosaicing in real time applications has been a challenging field for image processing experts. It has wide applications in the field of video conferencing, 3D image reconstruction, satellite imaging and several medical as well as computer vision fields. It can also be used for mosaic-based localization, motion detection & tracking, augmented reality, resolution enhancement, generating large FOV etc.

In this research work, feature based image mosaicing technique using image fusion have been proposed. The image mosaicing algorithms can be categorized into two broad horizons. The first is the direct method and the second one is based on image features. The direct methods need an ambient initialization whereas, Feature based methods does not require initialization during registration. The feature-based techniques are primarily followed by the four steps: feature detection, feature matching, transformation model estimation, image resampling and transformation. SIFT and SURF are such algorithms which are based on the feature detection for the accomplishment of image mosaicing, but both the algorithms has their own limitations as well as advantages according to the applications concerned. The proposed method employs this two feature based image mosaicing techniques to generate an output image that works out the limitations of the both in terms of image quality. Again to achieve this task, the image fusion technique has been employed. Image Fusion for multi-sensor images at an altering resolution can be fruitfully implemented by means of wavelet based Multi Resolution Analysis (MRA).Hence in this case Harr based wavelet transform technique has been used for achieving image fusion. The developed robust algorithm takes care of the combined effect of rotation, illumination, noise variation and other minor variation. Initially, the input images are stitched together using the popular stitching algorithms i.e. Scale Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF). To extract the best features from the stitching results, the blending process is done by means of Discrete Wavelet Transform (DWT) using the maximum selection rule for both approximate as well as detail-components. The SIFT provides scale as well as rotational invariance property. The SURF provides better computation speed and illumination invariance. The robustness and quality of the above mosaicing techniques are tested by means of three-dimensional rotational images.

Keywords: Mosaicing, Panorama, Image Fusion, SIFT, SURF, DWT

TABLE OF CONTENTS

Items	Page No.
Acknowledgements	v
Abstract	vi
Table of Contents	vii
List of Figures	ix
List of Abbreviations	x
Chapter 1 Introduction	1
1.1 Overview	2
1.2 Literature review	3
1.3 Objectives	4
1.4 Thesis organization	5
Chapter 2 Harris Corner Detector	6
2.1 General Concept	7
2.2 Flow chart	7
2.3 Feature Extraction	7
2.4 Image Registration	9
2.6 Homography Computation	10
2.7 Image Warping	12
2.8 Image Blending	13
2.9 Results and Discussion	13
2.10 Conclusion	17
Chapter 3 Scale Invariant Features Transform	18
3.1 General Concept	19
3.2 Flow chart	19
3.3 Scale-Space Extrema Detection	20
3.4 Key point Localisation	21
3.5 Orientation Assignment	22
3.6 Key point Descriptor Generation	22

3.7 Results and Discussion	23
3.8 Conclusion	26
Chapter 4 Speeded-Up Robust Features	27
4.1 General Concept	28
4.2 Flow Chart	28
4.3 Interest Point Detection	29
4.3.1 Integral images	29
4.3.2 Hessian matrix	30
4.3.3 Scale space representation	32
4.4 Interest Point Description	33
4.4.1 Orientation assignment	33
4.4.2 Descriptor based on sum of Haar wavelet responses	34
4.5 Results and Discussion	34
4.6 Conclusion	37
Chapter 5 Image Mosaicing using Image Fusion	38
5.1 Introduction	39
5.2 Proposed method	40
5.2.1 SIFT	40
5.2.2 SURF	42
5.2.3 Image fusion	43
5.3 Performance evaluation	45
5.4 Results and Discussion	47
5.5 Conclusion	51
Chapter 6 Conclusion	52
6.1 Conclusion	53
6.2 Suggestions for Future Work	54
References	55
Dissemination	58

LIST OF FIGURES

Figure No.	Page No.
2.1 Flow chart for image mosaicking using Harris corner detector	7
2.2 Left portion of the original scene	14
2.3 Right portion of the original scene	14
2.4 Harris corner points of left portion of the original scene	15
2.5 Harris corner points of right portion of the original scene	15
2.6 Image showing matches between left and right portion of the original scene	16
2.7 Image showing inliers between left and right portion of the original scene	16
2.8 Output image using Harris corner detector	17
3.1 SIFT feature detection algorithm	19
3.2 DoG at varying octave	20
3.3 Neighbouring pixels with which key point is compared	21
3.4 Generation of feature vector	23
3.5 Left portion of the original scene	24
3.6 Right portion of the original scene	24
3.7 Image showing inliers matching between the two portions of the original scene	25
3.8 Output image generated using SIFT algorithm	25
4.1 SURF feature detection algorithm	28
4.2 An integral image	30
4.3 Box filters of order 9x9 (a) x direction, (b) y direction, (c) xy direction	31
4.4 Box filters for two successive scale level (9x9 and 15x15)	32
4.5 Haar wavelet response within 60 degree	33
4.6 Sum of Haar wavelet responses for different intensity pattern	34
4.7 Left portion of the original scene	35
4.8 Right portion of the original scene	35
4.9 SURF points with top 200 matches	36
4.10 Image showing inliers matching between the two portions of the image	36
4.11 Output image using SURF algorithm	37
5.1 Flow chart for proposed technique	40
5.2 Local extremum in DoG scale space	41
5.3 Feature vector generation	42
5.4 Multi-directional box filters	43
5.5 Discrete wavelet decomposition with filter banks	44
5.6 Subjective analysis of DWTs	45
5.7 DWT based Image Fusion Flow chart	45
5.8 Input image-I	48
5.9 Input image-II	48
5.10 SIFT response	49
5.11 SURF response	49
5.12 Fused image	50
5.13 Haar Wavelet decomposition tree at 4th level	50

LIST OF TABLES

Table no.	Page no.
Table 5.1 Performance analysis of SIFT, SURF and Proposed method	47

LIST OF ABBREVIATIONS

SIFT – Scale Invariant Feature Transform

SURF – Speeded Up Robust Feature

PSNR – Peak Signal to Noise Ratio

MI – Mutual Information

SD – Standard Deviation

FOV – Field Of View

FSIM – Feature Simliarity Index Metric

NAE – Normalised Absolute Error

DWT – Discrete Wavelet Transform

DoG – Difference of Gaussian

MRA – Multi Resolution Analysis

LoG – Laplacian of Gaussian

Chapter 1

INTRODUCTION

Overview
Literature Review
Objectives
Organisation of the Thesis

INTRODUCTION

Image mosaicing is the process of combining multiple photographic images with overlapping fields of view to produce a segmented panorama of high-resolution image. It is commonly performed through the use of computer software; most approaches to image stitching require nearly exact overlaps between images and identical exposures to produce seamless results

1.1 Overview

An image mosaic is a synthetic composition generated from a sequence of images and it can be obtained by understanding the geometric relation between the images. The geometric relations are the coordinate system that relates the different image coordinate systems. By applying the appropriate transformation via warping operation and merging the overlapping regions of warped images, it is possible to construct a single image indistinguishable from a single large image of the same object, covering the entire visible area of the scene. The resultant image is the motivation for image mosaicing. Various steps in mosaicing are feature extraction, registration, stitching, warping and blending. Most approaches to image stitching require nearly exact overlap between images and identical exposures to produce seamless results [1].

Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It is alignment of two images geometrically - the reference and sensed images. Data can be multiple photographs, data from different sensors taken at different times, or from different viewpoints. Image registration is a very important step among all image analysis tasks in which the final information is received from the combination of sources like in image mosaicing, image fusion, change detection, and image restoration. Registration is necessary for comparison or integration of the received data from different measurements. Registration methods can be divided into following types: algorithms that use image pixel value directly, e.g., correlation method [2]; algorithm that use frequency domain e.g., FFT-based method [3]; algorithm that use low level features such as edges and corners, e.g., feature based methods [4]; algorithm that uses high level features such as identified objects, or relation between features, e.g., graphical methods.

Homography is the mapping between two spaces which is often used to represent the correspondence between two images of the same scene.

Image warping is the process, where the digital manipulation of an image can be done such that any shapes portrayed have been significantly distorted. It can be used for removing distortion

in the image and also for creative purposes (e.g., morphing [5]). At this step, the two images that are going to create mosaic are warped, by using the geometric transformation.

Image blending is the technique, which modifies the image gray levels in the vicinity of a boundary to obtain a smooth transition between images by removing these seams and creating a blended image by determining how pixel in an overlapping area should be presented.

Mosaicing is a process of joining two or more images in order to create a single wide area image. Some of the application areas are as follows :

1. The mosaic of large air scape and satellite remote sensing image ;
2. Meteorological and environmental monitoring;
3. Sea-bottom survey and geological survey;
4. Combination of medicine and scientific micro-fragment image ;
5. The 3D rebuilding of objects;
6. The building of virtual scene and virtual walkthrough;
7. Video compression, video search browse, and video edit, etc ;
8. The digitized saving of file;
9. Military reconnaissance and taking evidence.

1.2 Literature Review

Algorithms that allow images to be aligned and seamlessly stitched together are among the oldest most widely used in computer vision.

R. B. Inampudi, in the paper [6] proposed a design a high performance software to perform geometrical and radiometrical corrections of two or more images without using any kind of special hardware. His work is on retina and kutub minar image.

S. Peleg et al., in their paper [7] proposed a method to method for generating mosaic image using planar, cylindrical and general manifold. Manifold projection helps in the fast creation of low distorted panorama mosaics under very slight camera motions.

B. Rousso et al., [8] proposed the pipe projection model which is used to define high-quality mosaicking even for the most challenging cases of forward motion and zoom. The model also helps in removing parallax during complex motion.

J. Wang et al., [9] proposed algorithm uses mean value seamless cloning .It uses the visual attention model to extract salient region, and to use the regional registration technology to achieve image matching. The paper automatically and accurately obtains salient regions and also reduces

the complexity of the image registration and thus improves the quality of image mosaicing.

P. Azzari et al., [10] proposed the work with a complete evaluation methodology including data sets, ground-truth information and performance metrics. D. Ghosh et al., [11] proposed the work uses several metrics instead of a single one to validate the reliability of the mosaicing using SIFT algorithm. The performance metrics are all based on simple pixel wise comparison, thus they maintain the computational simplicity.

D. K. Jain et al., [12] proposed a three step automatic image mosaic method. The first step is taking two input images and finding out the Harris corners in both the images, second step is removing out the false corners in both the images and then by using homography, their matched corner pair are found and final output mosaic is obtained.

D.G.Lowe proposed an algorithm called Scale Invariant Feature Transform (SIFT), which is a technique to extract feature points. These features are from local scale-invariant features for object recognition [13], to recognise 3d objects [14] and also for shape indexing using approximate nearest-neighbour search in high-dimensional spaces [15]. He also proposed a method to determine distinctive image features from scale invariant key points [16]. SIFT algorithm is robust to scale & rotational variation. The algorithm provides the feature points which are accurate, stable, reliable, and efficient. In the algorithm he proposed an automatic image mosaicing method which also includes automatic registration.

H.Bay et al., [17] proposed a feature detector Speeded Up Robust Features (SURF), which is believed to have high speed in the feature detection steps: detection, description and matching. This detector is approximately 5 times faster than SIFT in terms of feature extraction. This algorithm is rotational invariant. It is mostly used for real-time application

1.3 Objectives

Vision allows humans to observe and be aware of our surrounding world. The field of view of human eye is restricted to $200 \times 135^\circ$, in order to increase the field of view $360 \times 180^\circ$ panorama is used. Image mosaicing is a technique to create panorama, where images are assembled with some common FOV. It helps in increasing the FOV. It provide us a lot of useful data that are required to extract valuable information not only from a single image but from mosaicing image (i.e. multiple assembled image). Since image mosaicing has been an emerging field which focuses on feature extraction, matching, registration, homography, warping and blending. Moreover, the images not only contain the structure, shape and colour information about the picture, but also

the possible camera motion, calibration and also on movements of objects in the scene.

There are a lot of existing image mosaicing algorithm and still research is going on. Each algorithm is considering a few variation in account while generating the output. So here improvement in the quality of image mosaicing by improving the robust mosaicing algorithms considering the combined effect of rotation, illumination, noise variation and other minor variation. It is precisely this information that is to be produced in this thesis

1.4 Organisation of the Thesis

Including the introductory chapter, the thesis is divided into 6 chapters. The organization of the thesis is presented below.

Chapter 2: Harris Corner Detector

It introduces the Harris corner detector, which is the most popular and widely used corner detection algorithm. It includes basic process of image mosaic, steps involved for the algorithm. And at last, there is analysis and simulation of the detector using two images having some common field of view.

Chapter 3: Scale Invariant Feature Transform

In this chapter, SIFT is discussed, which is an algorithm to extract the feature points. It discusses the basic steps involved in the algorithm. Also the simulation and discussion of result has been done in this section.

Chapter 4: Speeded-Up Robust Features

This chapter describes about the SURF algorithm, which helps in the extraction of interest points. It includes the basic process involved in the algorithm. Moreover, it analyses the simulation results.

Chapter 5: Image Mosaicing using Image Fusion

This is the chapter which describes the, “Robust Technique for Feature-based Image Mosaicing using Image Fusion”, i.e., the title of the thesis. It discusses about the proposed technique which works on 3d rotational images. The results are discussed and shown with the help of comparison table.

Chapter 6: Conclusion

This chapter is concluded with the highlights of the research work along with the suggestions for future research.

Chapter 2

Harris Corner Detector

General Concept
Flow Chart
Feature Extraction
Image Registration
Homography Computation
Image Warping
Image Blending
Results and Discussion
Conclusion

2.1 General Concept

Image Mosaicing algorithm based on random corner method is proposed. An image mosaic is a method of assembling multiple overlapping images of same scene into a larger one [18]. The output of image mosaic will be the union of two input images. In this chapter, three step automatic image mosaic method is used. The first step is taking two input images and finding out the corners in both the images, second step is removing out the false corner in both the images and then by using homography, their matched corner pair are found out and final output mosaic is obtained

2.2 Flow Chart

Fig 2.1 shows the flow chart for image mosaicking using Harris corner detector.

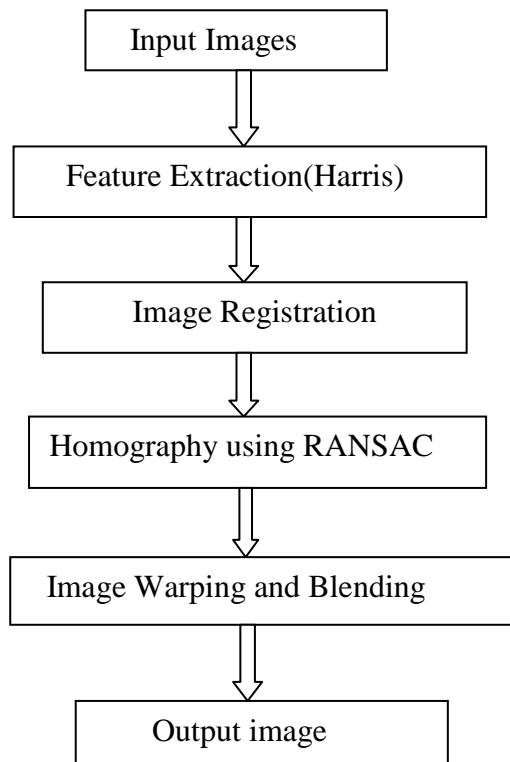


Fig 2.1 Flow chart for image mosaicking using Harris corner detector

2.3 Feature Extraction

Initially, the features were objects manually selected by an expert. Due to the automation of the registration process, two main approaches for feature understanding have been built. The

approach is based on the extracting the salient structures—features—from the images. Significant points (region corners, line intersections) are understood as features here. These feature points should be distinct and spread all over the image, also these should be efficiently detectable in both the images. These are expected to be stable with variation in time to stay at fixed positions during the whole experiment in order to get proper result.

The efficiency of extracted feature points in the two images is ascertained by the invariance and accuracy of the feature detector in the overlapping region. We can also say that the number of common feature points detected from the set of images should be sufficiently high, regardless of the variation in image geometry, radiometric conditions, presence of noise, and other minor variations etc. The effectiveness of the features is given by its definition. On contrary to the area-based methods, the feature-based methods are not directly working on the intensity of image. The features represent higher level information. These properties of feature-based methods make it suitable for situations dealing with illumination changes or multi sensors.

2.3.1 Harris corner detector

This operator was developed by Chris Harris and Mike Stephens in 1988 as a low-level processing step to aid researchers trying to build interpretations of a robot's environment based on image sequences. Specifically, Harris and Stephens were interested in using motion analysis techniques to interpret the environment based on images from a single mobile camera. Like Moravec , they needed a method to match corresponding points in consecutive image frames, but were interested in tracking both corners and edges between frames [19]. Harris and Stephens developed this combined corner and edge detector by addressing the limitations of the Moravec operator. The result is a far more desirable detector in terms of detection and repeatability rate at the cost of requiring significantly more computation time. Despite the high computational demand, this algorithm is widely used.

Considering the gray intensity of pixel (u, v) be $I(x, y)$, the variation of gray pixel (x, y) with a shift of (u, v) can be denoted as

$$E(u, v) = \sum_{x, y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (2.1)$$

With the application of Taylor series expansion

$$E(u, v) = \sum_{x,y} w(x, y) [I_x u + I_y v + O(u^2, v^2)]^2 \quad (2.2)$$

For a small shift of (u, v) we have the following approximation,

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (2.3)$$

Where M is a matrix of 2x2 which has been calculated from the image derivatives:

$$M = \sum_{x,y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (2.4)$$

If λ_1, λ_2 are the eigen values of matrix M, then corner, edge and flat area of the image can be computed from the Eigen value as follows

- Flat area: both λ_1, λ_2 are very small.
- Edge: one of λ_1, λ_2 is smaller and the other is bigger
- Corner: both λ_1, λ_2 are bigger and are nearly equal

2.4 Image Registration

Image registration is not only a most important part of image mosaicing, but also is the foundation of image mosaicing. In accordance with certain image transform parameters generated by similarity measurement, registration of multi-source images collects two or more images which are focused on the same target but produced from different sensors, different perspectives, and different times [20]. This process makes the multi-source images transformed into a same coordinate system and achieves the best match on pixel level. Using two two-dimensional arrays, $I_1(x,y)$ and $I_2(x,y)$, to represent the reference image and the image to be matched, image registration can be stated as follows:

$$I_1(x,y) = g(I_2(f(x,y))) \quad (2.8)$$

f is a transform function for two-dimensional coordinates. It is the key point of image registration and can be described by global transformation model.

There are three common global transformation models:

1. Rotation - Variable ratio - Translation (RST) Model :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = k \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (2.9)$$

2. Affine Transformation Model :

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (2.10)$$

3. Perspective Transformation Model :

The affine transformation function tells that a line in the first image can be mapped to a line in the second image and keeps balance through the affine transformation model.

2.5 Homography Computation

It is mapping between two spaces which is often used to represent the correspondence between two images of the same scene. It is widely used for the project where multiple images are taken from a rotating camera center ultimately warped together to produce a panoramic view

The steps for Homography Detection Algorithm using RANdom Sample Consensus (RANSAC) scheme is

1. Firstly, corners are detected in both images.
2. Variance normalized correlation is applied between corners, and pairs with a sufficiently high correlation score are collected to form a set of candidate matches.
3. Four points are selected from the set of candidate matches, and a homography is computed.
4. Pairs agreeing with the homography are selected. A pair (p, q), is considered to agree with a homography H, if for some threshold: $\text{Dist}(Hp, q) < \epsilon$
5. Steps 3 and 4 are repeated until a sufficient number of pairs are consistent with the computed homography.
6. Using all consistent correspondences, the homography is recomputed by solving step 4.

2.5.1 Random Sample Consensus algorithm:-

RANSAC, is an iterative method to calculate the parameters of a mathematical model from a set of observed data which contains many outliers. This is a non- deterministic algorithm in the way that it produces a reasonable result only with a certain probability, with increase in probability more iteration are allowed. It is assumed that the data consists of “inliers”, i.e., data whose distribution can be explained with the help of parameter model, and “outliers” are data that do not fit the model.

Input of RANSAC algorithm is a set of observed data, a parameterized model which can explain or fit to the observations, along with some confidence parameters. RANSAC achieves its goal by iterative selection of a random subset of the original data. These data are supposed inliers and this supposition is then tested as follows:

- i. A model is fitted according to the supposed inliers
- ii. All other data are then tested to fit the model and if a point fits to the estimated model, it is considered as a hypothetical inlier.
- iii. The estimated model is efficient if sufficiently many points have been selected as hypothetical inliers.
- iv. The model is re estimated using all hypothetical inliers.
- v. Finally, the model is assessed by calculating the error of the inliers with respect to the model.

Given a fitting problem with parameter x , it estimates the parameters considering the following assumption:

- Parameter can be estimated from N data items.
- Available data items are totally M
- The probability of a randomly selected data item being part of a good model is P_g
- The probability that the algorithm exit without finding a good fit if one exists is P_f

Then the algorithm is:

1. N data items at random are selected.
2. Parameter x is estimated.

3. Number of data items which fit the model with parameter vector x are found out within a user defined tolerance. Let it be K .
4. If K is large enough, it is accepted and exit with success.
5. The process is repeated L times
6. The process is failure if it again enters the loop.

Value of L is found by the following formulae:

P_{fail} = Probability of L consecutive failures

= Probability that a given trial is a failure) L

= $(1 - (\text{Probability that a random data item fits the model})^N)^L$

$$P_{fail} = (1 - (P_g)^N)^L \quad (2.11)$$

$$L = \frac{\log(P_{fail})}{\log(1 - (P_g)^N)} \quad (2.12)$$

2.6 Image Warping

Image warping is the process of digitally manipulating an image such that any shapes portrayed in the image have been significantly distorted. Basically we can simply warp all the input images to a plane defined by one of them known as reference image. Warping can also be used for correcting image distortion as well as for creative purposes. The same techniques are equally applicable to video. The two images that will form the mosaic are warped, by using the geometric transformation. While an image can be transformed in various ways, pure warping means that points are mapped to points without changing the colors. It can be mathematically based on any function from the (part of) plane to the plane. If the function is put in the original then it can be reconstructed.

The following list give idea where image warping can be applied.

- Images can be distorted due to simulation of optical deviation.
- Images may be viewed as if they had been projected on to a curved or mirrored surface.
- Images may be divided into polygons and every polygon is distorted.
- Images may be aberrated using morphing.

There are two methods for generation of an image for any type of distortion.

- Forward-mapping: a given mapping from sources to images is directly applied
- Reverse-mapping : for a given mapping from sources to images, the source is found from the image

To determine the type of warping which takes place between consecutive images, optical flow estimation techniques are used.

2.7 Image Blending

The final step is to blend the pixel colours in the overlapped region to avoid the seams. Simplest available form is to use feathering, which uses weighted averaging colour values to blend the overlapping pixels.

Image blending is the technique, which modifies the image gray levels in the vicinity of a boundary to obtain a smooth transition between images by removing these seams and creating a blended image by determining how pixel in an overlapping area should be presented.

2.8 Results and Discussion

The algorithm proposed here has been implemented in Matlab R2010a and has been executed in system with configuration i5 processor, 4Gb RAM, 2 Gb cache memory and 2.8GHz processor. Fig 2.2 and Fig 2.3 are the input images of original scenes taken from www.hdweb4.com. Harris corner technique has been applied on these figure has been obtained as shown in Fig 2.4 and Fig 2.5 respectively. In Fig 2.2, 121 Harris points and in Fig 2.3, 169 Harris points are detected. Fig 2.6 shows the rough matching between the respective Harris corner points that is why some matches are seen between the images which are irrelevant. Those irrelevant points are removed by RANSAC algorithm. The inconsistent points i.e. outliers, which don't fit to the model parameters are removed by it and the algorithm take only consistent points i.e. inliers which fit to the model parameters. Fig 2.7 shows the inlier matching between the left and right portion of the original scene. Robust estimation of homography using RANSAC with 67 inliers in left and right portion are computed. Fig 2.8 shows the final output mosaic image after warping and blending



Fig 2.2: Left portion of the original image



Fig 2.3: Right portion of the original image



Fig 2.4: Harris corner points of left portion of the original image



Fig 2.5: Harris corner points of right portion of the original image



Fig 2.6: Image showing matches between left and right portion of the original image

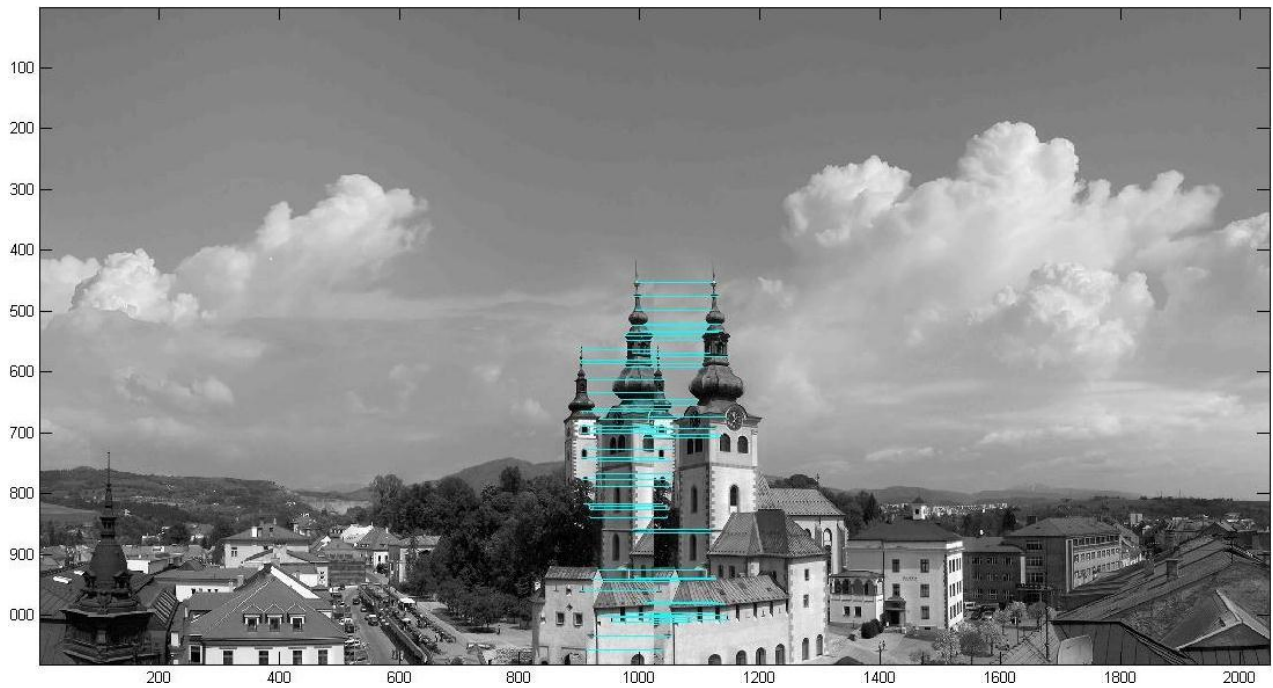


Fig 2.7: Image showing inliers between left and right portion of the original image



Fig 2.8: Output image using Harris corner detector

2.9 Conclusion

Harris corner detector is the most popular and widely used corner detector. It can be combinely used for corner and edge detection but in this chapter it is used for detection of corner points, which are nothing but the interest points. Interest points are used for finding the match between two images having some common FOV. This corner detector is efficient, cheap and rotation invariant but it can't handle the images which suffer scale variation. Also, the detector is taking more time for the computation of corner points.

Chapter 3

Scale Invariant Feature Transform

General Concept

Flow chart

Scale-Space Extrema Detection

Key point Localisation

Orientation Assignment

Key point Descriptor Generation

Results and Discussion

Conclusion

3.1 General Concept

Scale Invariant Feature Transform (SIFT) is an algorithm developed by D.G.Lowe. The algorithm helps in the detection and description of local features of the image. The features detected by the algorithm are accurate, stable; moreover it is invariant towards scale and rotation. Its applications include object recognition, robotic mapping and navigation, image stitching, 3D modelling, gesture recognition, video tracking, individual identification of wildlife and match moving.

3.2 Flow chart

The Flowchart of SIFT algorithm is depicted in Fig.3.1.

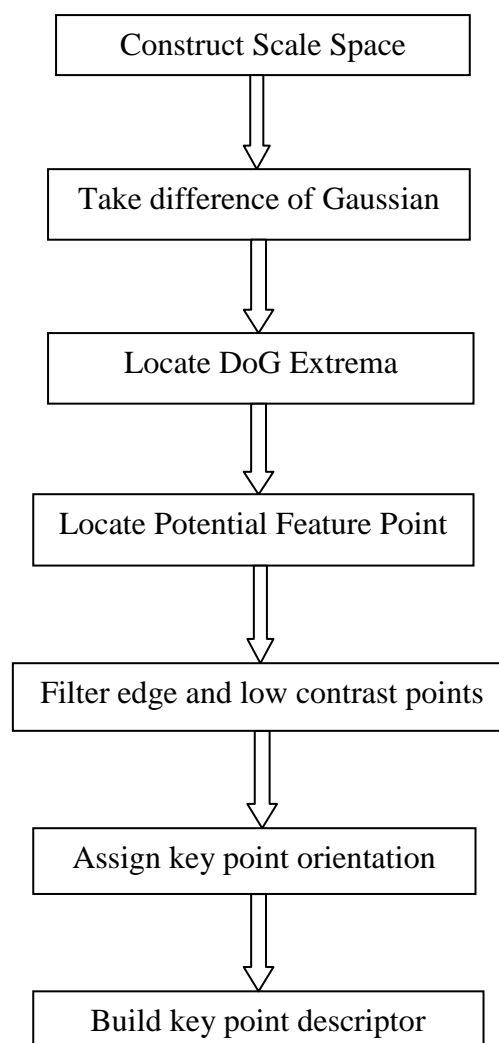


Fig 3.1 SIFT feature detection algorithm

3.3 Scale-Space Extrema Detection

Under this heading the scale space theory is used to determine the key points that is nothing but interest points. For detecting the key points first of all consider an image say $I(x, y)$ and convolve that image with the Gaussian filter, $F(x, y, k\sigma)$ at varying scales. The convolved images having different scales are grouped by octave and a variable ‘k’ is selected in order to get a fixed number of blurred images per octave. The formed images are subtracted at different scales in order to get the difference of Gaussian (DoG) as shown in Fig.3.2. DoG helps in removing the problems arising due to key point localisation. DoG acts as effective tunable band pass filter and extract a range of components which can be used as key points. The DoG can be used as an approximation for the Laplacian operator. The convolution of Gaussian filter to the image is represented in Equation 3.1 and DoG is represented using Equation 3.2.

$$L(x, y, k\sigma) = F(x, y, k\sigma) \otimes I(x, y) \quad (3.1)$$

$$D(x, y, \sigma) = L(x, y, k_i\sigma) - L(x, u, k_j\sigma) \quad (3.2)$$

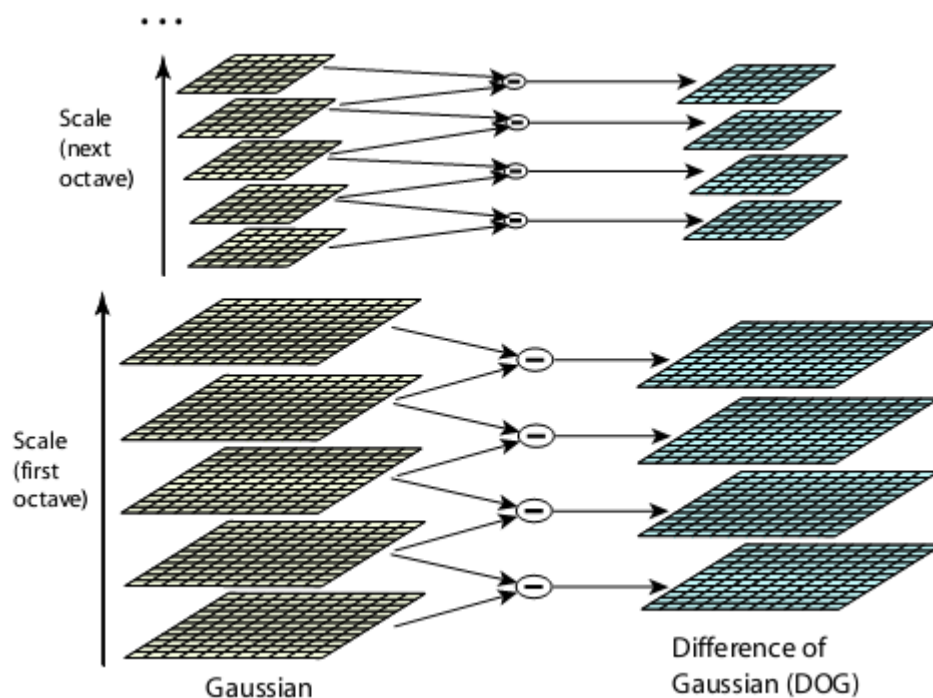


Fig 3.2 DoG at varying octave

3.4 Key point Localisation

Key points are nothing but the local maxima or minima. The key points are selected after taking the DoG. The key point is known as key point when the key point is local maxima or minima. Each pixel from the DoG images are compared with 8 neighbouring pixels having same scale on the same plane, whereas rest eighteen pixels are at the plane lying above and below it with different scales. The twenty-six neighbouring pixels for the candidate key point are shown in Fig 3.3. For determining the accurate position of key point, interpolate the key point with the help of nearby data that was the initial approach. But according to new approach determine the location and scale of the candidate key point using Taylor series expansion. Using Taylor expansion, the extreme points and location are carefully determined. The Taylor expansion of DoG at origin is given by the following equation:

$$D(x) = D + \frac{\partial D}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (3.3)$$

This new approach of interpolation helps in improvising the stability as well as matching.

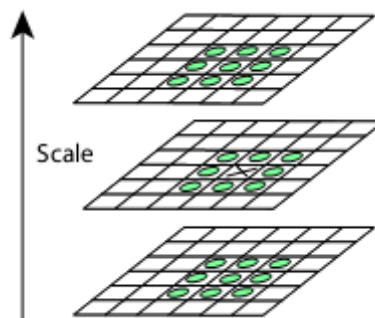


Fig 3.3 Neighbouring pixels with which key point is compared

During the process, the candidate key points with low contrast are removed using the second order Taylor expansion method at particular offset, if the value is more than 0.03 then it is considered otherwise it is discarded. Also the DoG function have high responses along edges, even though the candidate keypoints are prone to small intensity of noise. So, for increasing the stability, remove the key points that are not having properly determined locations and are having strong edge responses.

3.5 Orientation Assignment

Here, every key point is having one or more orientations depending on the local image gradient directions. This step helps in achieving invariance to rotation because the key point descriptor can be presented relative to this orientation and thus invariance to image rotation achieved.

First of all Gaussian-smoothed image $L(x, y, \sigma)$ having scale σ is considered so that all the computations are scale-invariant. For an image $L(x, y)$ with scale σ , the gradient magnitude, $m(x, y)$, and orientation, $\theta(x, y)$, are calculated using pixel differences:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3.4)$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (3.5)$$

The computation of magnitude and direction for the gradient are done at every pixel in the neighbouring region around the candidate key point in the Gaussian-blurred image L . An orientation histogram of 36 bins is created, with each bin having 10 degrees. Each sample of the neighbouring window is added to the histogram bin which is weighted by its gradient magnitude also by a Gaussian-weighted circular window with 1.5 time σ to the scale of the candidate key point. The peaks of the histogram correspond to dominant orientations. Once the histogram is completely filled, the orientations which corresponds to the highest peak and local peaks which are within 80% of the highest peaks assigned are considered as the key point. For multiple orientations, an additional key point is created which have the same location and scale as that of the original key point for each additional orientation.

3.6 Key point Descriptor Generation

This step insures invariance to image location, scale and rotation. Now compute a descriptor vector for every key point such that the descriptor is highly distinctive and partially invariant to the variations such as accuracy, illumination, 3D viewpoint, stability etc. This step is performed on the image scale which is closest in scale to the scale of the key point.

Firstly a set of orientation histograms are generated on 4x4 pixel neighbourhoods with 8 bins each. These magnitude and orientation values are computed from histograms of samples in a 16 x 16 region around the key point such that 4 x 4 sub regions of the original neighbourhood region are formed . The diagram showing generation of feature vector is shown in Fig 3.4. The

magnitudes are then weighted by a Gaussian function with 1.5σ of the descriptor window. The descriptor now becomes a vector from all the values of the histograms. Since there are $4 \times 4 = 16$ histograms each having 8 bins, so the vector has in total 128 elements. This vector is normalized to unit length to enhance invariance due to affine changes in illumination. For reducing the effects of non-linear illumination, set a threshold of 0.2 and once again the vector is normalized.

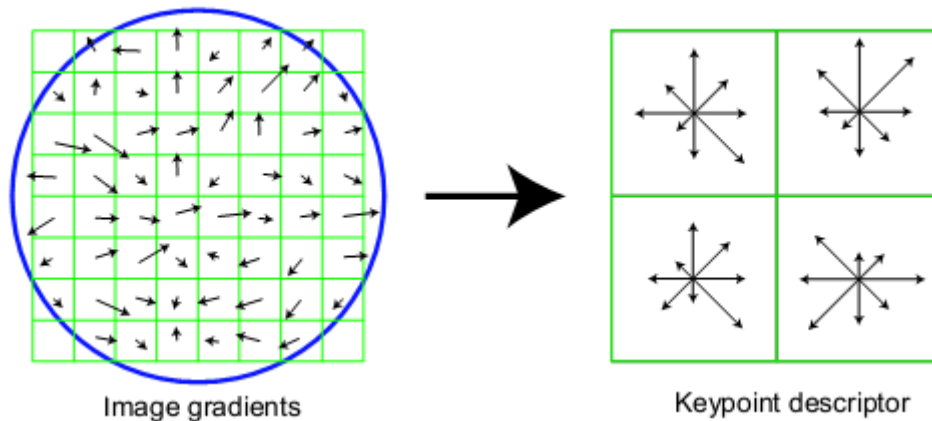


Fig 3.4 Generation of feature vector

Even though the dimension of the descriptor (128) is high, descriptors having lower dimension than this are not performing well across the matching range and thus the computational cost remains low because of the approximate Best Bin First (BBF) method used to find the nearest-neighbour. It is also seen that feature matching accuracy is above 50% for variation in viewpoint up to 50 degrees. Thus SIFT descriptors prove invariant to small affine changes

3.8 Results and Discussion

The algorithm proposed here has been implemented in Matlab R2010a and has been executed in system with configuration i5 processor, 4Gb RAM, 2 Gb cache memory and 2.8GHz processor. Fig 3.5 and Fig 3.6 are the input images of original scene. SIFT algorithm has been applied on these figure. Some matches are seen between the images which are irrelevant. Those irrelevant points are removed by RANSAC algorithm. The inconsistent points i.e. outliers, which don't fit to the model parameters are removed by it and the algorithm take only consistent points i.e. inliers which fit to the model parameters. Fig 3.7 shows the inlier matching between Fig 3.5 and Fig 3.6 portion of the original scene. Robust estimation of homography using RANSAC with 83 inliers in left and right portion are computed. Fig 2.8 shows the final output image after warping and blending



Fig 3.5: Left portion of the original scene



Fig 3.6: Right portion of the original scene

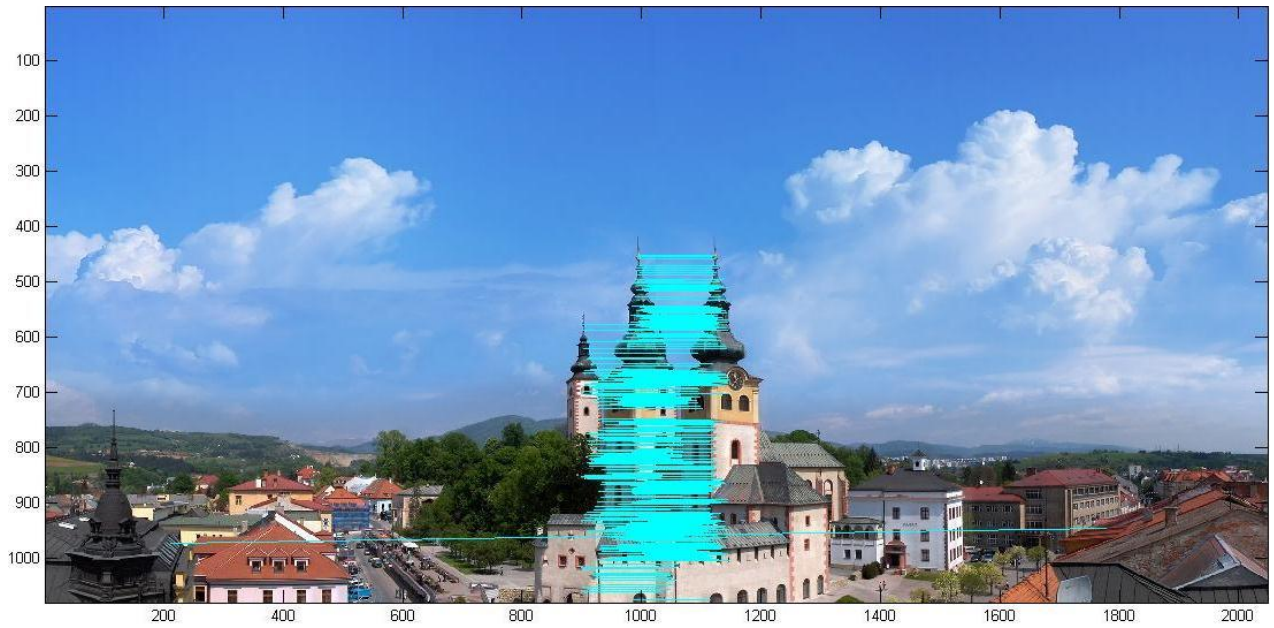


Fig 3.7: Image showing inliers matching between the two portions of the original scene



Fig 3.8: Output image generated using SIFT algorithm

3.9 Conclusion

SIFT algorithm brought a revolution in the field of image mosaicing, with its coming the concept of automatic image stitching came into existence. It extracts highly distinctive features from the set of images. It helps in the correct object identification with low probability of mismatch. This algorithm is robust to change in illumination, noise and minor changes in viewpoint. It is robust towards the scale and rotation variation of the image.

Chapter 4

Speeded-Up Robust Features

General Concept
Flow chart
Interest Point Detection
Interest Point Description
Results and Discussion
Conclusion

4.1 General Concept

SURF (Speeded Up Robust Features) is a robust local feature detector, first presented by Herbert Bay et al. in 2006, that can be used in computer vision tasks like object recognition or 3D reconstruction. It is partly inspired by the SIFT descriptor. The standard version of SURF is several times faster than SIFT and more robust against different image transformations than SIFT. SURF is based on sums of 2D Haar wavelet responses and efficiently use the integral images.

It uses an integer approximation to the determinant of Hessian blob detector, which can be computed extremely quickly with an integral image (3 integer operations). For features, it uses the sum of the Haar wavelet response around the point of interest. Again, these can be computed with the aid of the integral image.

4.2 Flow chart

The flow chart of SURF algorithm is portrayed in the Fig.4.1.

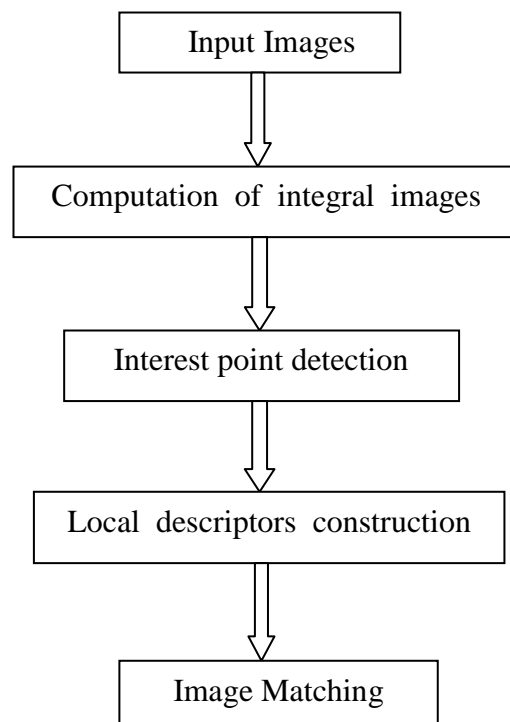


Fig 4.1 SURF feature detection algorithm

4.3 Interest Point Detection

The most popularly used detector is the Harris corner detector but because of its variance towards scale leads to the improvement of other interest point detectors. Several scale invariant detectors have been proposed For detecting the interest point using a SURF approximation of basic Hessian matrix is used. For the feature detection step, local maxima from the Hessian determinant matrix is applied to the scale-space and are computed to select feature point candidates. These candidates are then tested if the response is above a defined threshold. Both scale and location of the candidate points are then refined using an iterative procedure to satisfy a quadratic function. In short a threshold value selected, if the interest point is greater than that value, then it is compared with its twenty- six neighbouring pixels. If extreme point is found greater than the neighboring value, then that point is known as a feature point. Normally, a few hundreds of feature points are detected in a digital image with a size of 1 Mega-pixels. The more detailed description of interest point detection can be understood under the heading of integral images and hessian matrix

4.3.1 Integral Images

Let us consider a digital image defined over pixel grid . In the following steps, only consider the gray value of the images (range 0 to 255), which is a simple process to be robust to colour variations (as a white balance correction). The integral image of the considered image, 'u' is defined as :

$$U(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} u(i, j) \quad (4.1)$$

Integrate images help in faster processing of the box type convolution filter. An example of integral image is shown in Fig 4.2. Convolve the considered image 'u' with a 2-D rectangular function. The pre-computation of integral images permit to convolve with the box type filter in three operations and four memory accesses. Since computation time does not depend upon the size of the box, because it performs only the addition operation, so it is better to use bigger filter sizes.

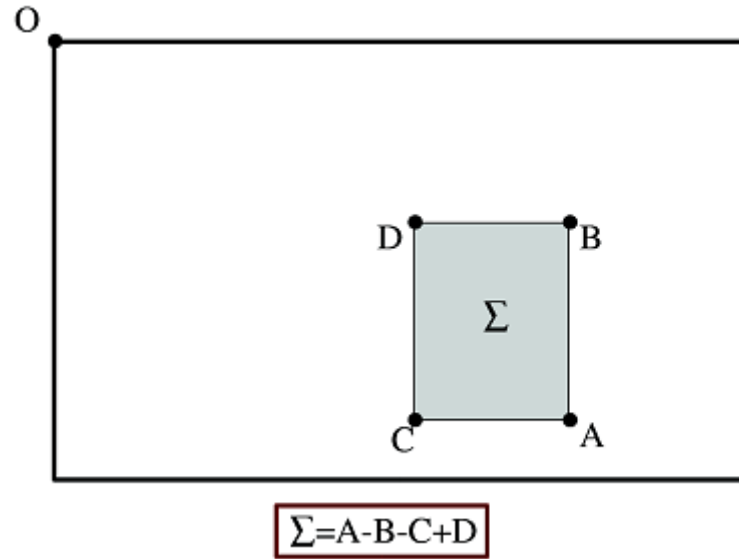


Fig 4.2 An integral image

4.3.2 Hessian Matrix

SURF uses the Hessian matrix because of its lesser computation time and good accuracy. Here the reliability is on the Hessian matrix for determining both the scale and the location of the interest point. This algorithm is also based on scale space theory. Without down sampling, it generates a stack in order to restore the same resolution. Here, the local maxima are estimated using Hessian matrix (H). Given a point $x = (x, y)^T$ in an image I, the Hessian matrix H (x, σ) in x at scale σ is defined as follows:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (4.2)$$

Where, $L_{xx}(x, \sigma)$ represents the convolution of middle point X with the Gaussian filter $\frac{\partial^2 g(\sigma)}{\partial x^2}$

To enhance the computing speed, the box filter approximation is taken instead of Gaussian filter. Gaussian filters are optimized for scale-space analysis. In general, however, the Gaussian filter needs to be discretised and cropped, also with Gaussian filters aliasing occurs when the resulting images are (further sampled) sub-sampled. Since Gaussian filters are not ideal in any case, and observing Lowe's success with approximations of LoG, an approximation can be done even further by utilizing box filters. These evaluations of approximate second order Gaussian

derivatives, are very fast using integral images, and are independent of size. The multi-directional box filters of order 9x9 are shown in Fig.4.3.

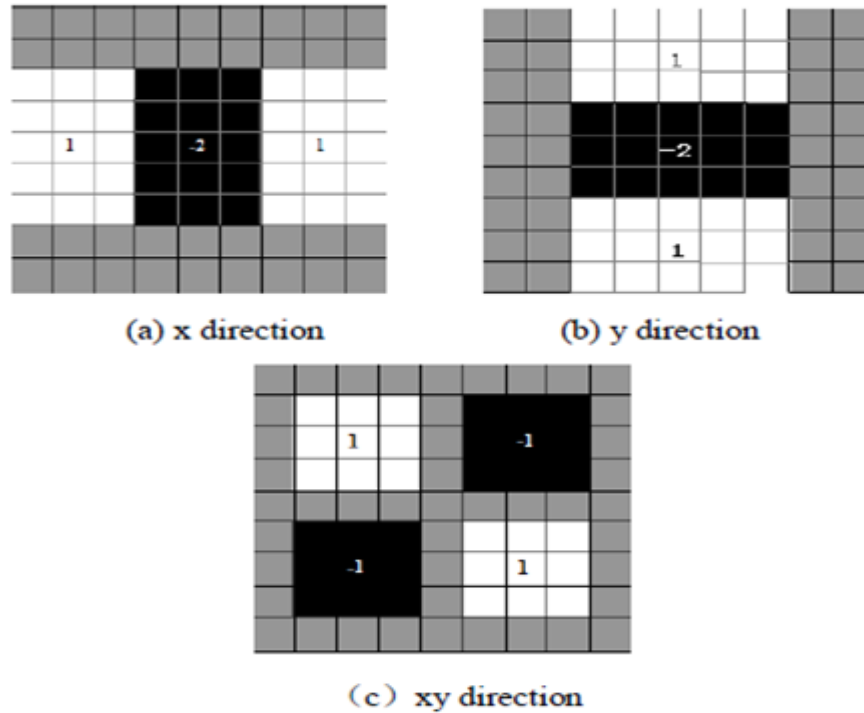


Fig 4.3 Box filters of order 9x9 (a) x direction, (b) y direction, (c) xy direction

The determinant of Hessian matrix, ΔH can be reduced to

$$\Delta H = D_{xx}D_{yy} - (wD_{xy})^2 \quad (4.3)$$

The response for each spot can be determined by assigning $\omega = 0.9$ as per the Frobenius norm. Also the responses of Gaussian filter are normalised with respect to the size of mask. A threshold is set for non-maxima suppression to detect the extreme points. The stable feature points are selected by comparing it with the neighboring values followed by the interpolating it in scale space. Implementations of scale space are usually done with the help of image pyramids. Sub sampling of image are done so as to achieve higher pyramid level. It is the advantage of box type filters and integral images that iterative application of same filter to the previously filtered layer output. So, the analysis of scale space is done by up-scaling the size of filter rather than repeatedly reducing the image size.

The output of the 9×9 filter is considered at the initial scale layer, for which consider a scale $s = 1.2$. For large scales, the step between consecutive filter sizes should also scale accordingly.

Thus, the filter size will be doubled for all new octave. The ratios of filter layout and Gaussian derivatives scale remain constant.

So for localising the interest points in the desired image and over scales, apply a non- maximum suppression of $3 \times 3 \times 3$ in neighbourhood. The maxima of the Hessian matrix's determinant are then interpolated with scale and image space. Scale space interpolation is very crucial for the case, where the difference in scale between the first layers of every octave is relatively large.

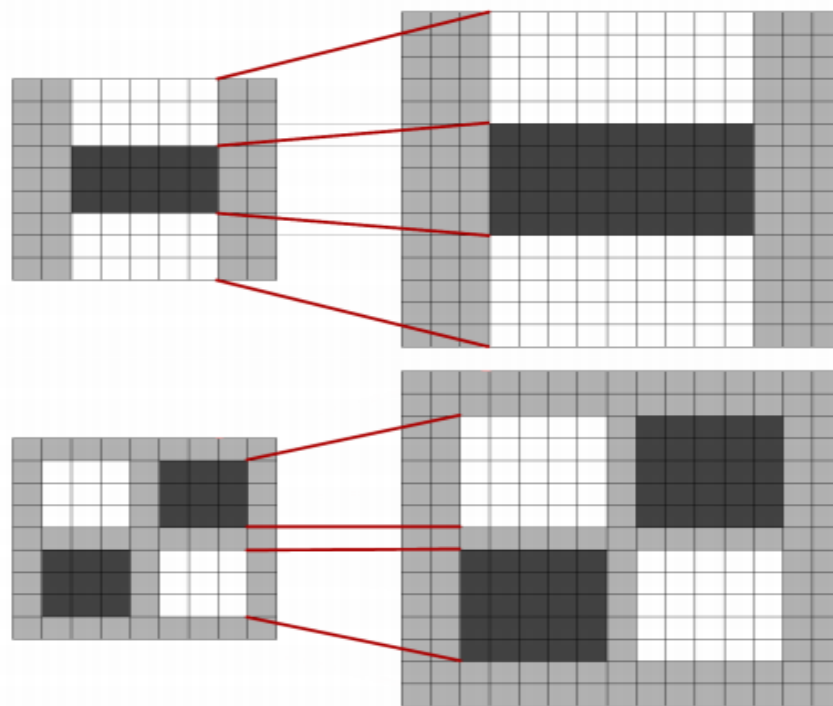


Fig 4.4 Box filters for two successive scale level (9x9 and 15x15)

4.3.2 Scale space representation

Similar to SIFT, SURF approach is also based on a regular sampling of the scale parameter σ . The division of scale is done into octaves: a new octave represents to the doubling of the kernel size. Again each octave is further divided into different levels (or intervals), where each level corresponds to the smallest increase in the size of the involved discrete filter.

In SURF, let o be the octave index and I be the interval index, and then the scale sampling is found with the following relation:

$$\sigma = \frac{1.2}{3} (2^o \times i + 1) = \frac{1.2}{3} l \quad (4.4)$$

where $3l$ is the bandwidth for the corresponding box filter.

4.4 Interest Point Description

The proposed SURF descriptor uses the similar properties as SIFT, with the complexity reduced down even further. The first step consists of forming a circular region around the feature point with a reproducible orientation. Then, form a square where the alignment is selected for the orientation, in order to extract the SURF descriptor from it. These two steps are elaborately explained in the following sections.

4.4.1 Orientation assignment

For interest points to be rotation invariant, identify a reproducible orientation. For that, first calculate the Haar-wavelet responses in x and y direction, within a circular neighbourhood with a radius of $6s$ around the interest point. The circle around the interest point should be at the scale with which the interest point was detected. Moreover the s is chosen to be sampling step and is scale dependent. Since, for large scales the big size of wavelets is used.

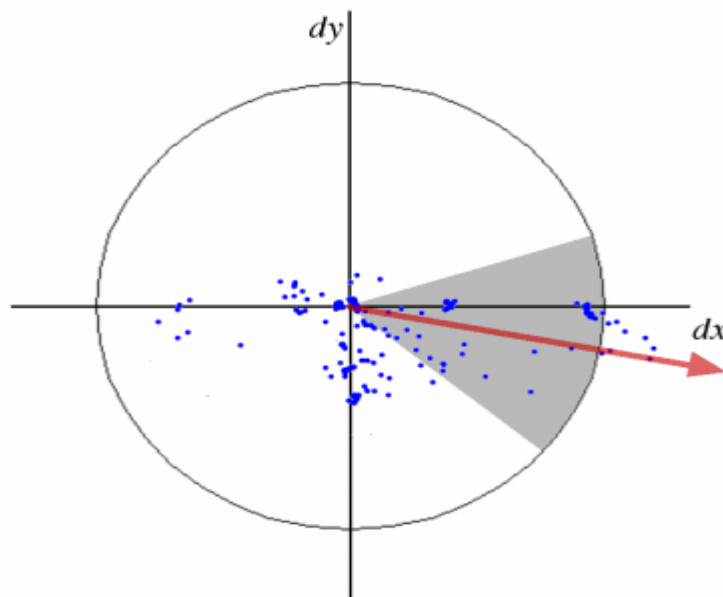


Fig 4.5 Haar wavelet response within 60 degree

Therefore, for fast filtering again use integral images. Only six operations are required to compute x or y direction response at any scale. The dominant orientation direction is estimated by computing the sum of all the responses that are within a sliding direction window covering a range of 60 degree. The horizontal and vertical responses inside the window are added up. The summed up responses then form a new vector. The longest vector thus formed is the orientation to the interest point. The sliding window size is chosen experimentally

4.4.2 Sum of Haar wavelet responses

For the descriptor extraction, first construct a square region which will be centered around the feature point, and is oriented along the orientation chosen in the previous section.

The size of the selected window is $20s$. The region is divided regularly into 4×4 smaller square sub-regions, which helps in storing spatial information. For every sub-region, compute simple features at 5×5 evenly spaced sample points.

Let dx be the Haar wavelet response in horizontal and dy be the Haar wavelet response in vertical direction. Now, the wavelet responses dy and dx are added up over each sub-region to create a first set of entries for the feature vector. For the polarity change in intensity, extract the sum of the absolute values of the responses, $|dx|$ and $|dy|$. Hence, every sub-region has a four-dimensional descriptor vector 's', which is given as $s = (dx, dy, |dx|, |dy|)$. The sum of Haar wavelet responses for different intensity pattern is shown in Fig 4.6. This results in a 64 dimensional descriptor vector for all 4×4 sub-regions. Invariance to contrast and illumination is achieved by normalising the descriptor.

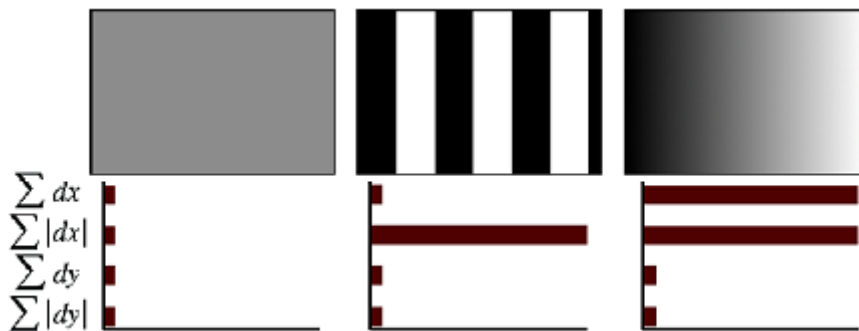


Fig 4.6 Sum of Haar wavelet responses for different intensity pattern

4.5 Results and Discussion

The algorithm proposed here has been implemented in Matlab R2010a and has been executed in system with configuration i5 processor, 4Gb RAM, 2 Gb cache memory and 2.8GHz processor. Fig 4.7 and Fig 4.8 are the left and right portion of the input images of original scene. SURF algorithm has been applied on these figures and shown in Fig 4.9. The Fig 4.9 shows top 200 matched points out of 1902 SURF points. Some matches are seen between the images which are irrelevant. Those irrelevant points are removed by RANSAC algorithm. The inconsistent points

i.e. outliers, which don't fit to the model parameters are removed by it and the algorithm take only consistent points i.e. inliers which fit to the model parameters. Fig 4.10 shows the inlier matching between Fig 3.5 and Fig 3.6 portion of the original scene. Robust estimation of homography using RANSAC with 23 inliers in left and right portion are computed. Fig 4.11 shows the final output image after warping and blending.



Fig 4.7 Left portion of the original scene



Fig 4.8: Right portion of the original scene

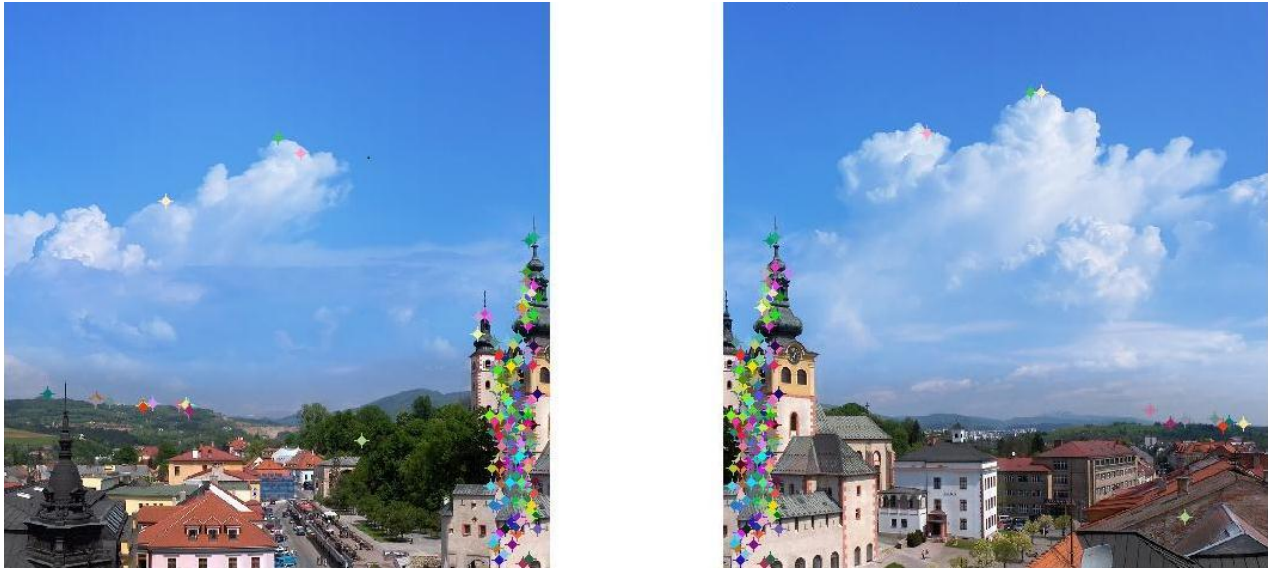


Fig 4.9 SURF points with top 200 matches



Fig 4.10: Image showing inliers matching between the two portions of the image



Fig 4.11: Output image using SURF algorithm

4.6 Conclusion

Speeded Up Robust Features is the most advanced algorithm for image mosaicking. It is a robust local feature detector. This algorithm use 2D Haar wavelet response and make an efficient use of integral images. This algorithm works effectively in presence of noise and other minor variation. Also this algorithm is scale and illumination invariant. This algorithm can be used for the real time application as because the standard version of SURF is several times faster than SIFT.

Chapter 5

Image Mosaicing using Image Fusion

Introduction
Proposed method
Performance evaluation
Results & Discussion
Conclusion

5.1 Introduction

Virtual Environment and panoramic imaging have been an emerging field of research with the improved brain-computer interfacing to deal with real-time applications. Image mosaicing plays a vital role in developing the panoramic view. The complementary information on individual image scenes in spatial and temporal domain can be combined to produce unsegmented panorama using images of smaller field of view. A number of images mosaicing algorithms have been proposed to generate a seamless, wide view image to interpret real world more clearly. In this chapter, a robust technique for panoramic image mosaicing by means of image fusion. The proposed technique consisting of two efficient stitching algorithms i.e. SIFT and SURF. The SIFT algorithm performs better for images with scale and rotational variance. These properties compensate the requirement of SURF. Again, the SURF is known for its illumination invariance and better computational speed. The response of both is blended together using the optimum image fusion rule. Here, the fusion process takes place using Haar Discrete Wavelet Transform (DWT).

Image stitching algorithms can be categorized into two broad horizons. The first is the direct method and the second one is based on image features . The direct methods need an ambient initialization, whereas, feature based methods do not require initialization during registration [23]. The feature-based techniques primarily consist of the four steps: feature detection, feature matching, transformation model estimation, image resampling and transformation . In 2004, David G. Lowe proposed an algorithm known as SIFT, which is considered as a feature-based method invariant to scale, rotation as well as affine transformation [15,16]. In 2006, Herbert Bay developed SURF [17]. It is mostly used for real-time applications. Again, the image fusion job for multi-sensor images at an altering resolution can be fruitfully implemented by means of wavelet based Multi Resolution Analysis (MRA). The review paper by S. Krishnamoorthy et al. [26], Haar wavelet transform fusion technique has been appraised as the salient method especially for subjective analysis.

This chapter is organized in the following manner. In Section 5.2, the proposed image mosaicing technique is depicted. The performance evaluation are discussed in Section 5.3. The simulation results are vividly discussed in the Section-5.4. Finally, the paper is concluded with the highlights of the proposed technique in the Section-5.5

5.2 Proposed method

The proposed method mainly consists of the following major steps:

- i. Concerned image acquisition
- ii. Image stitching algorithms
- iii. Image fusion

The flow chart for the proposed methodology is shown below

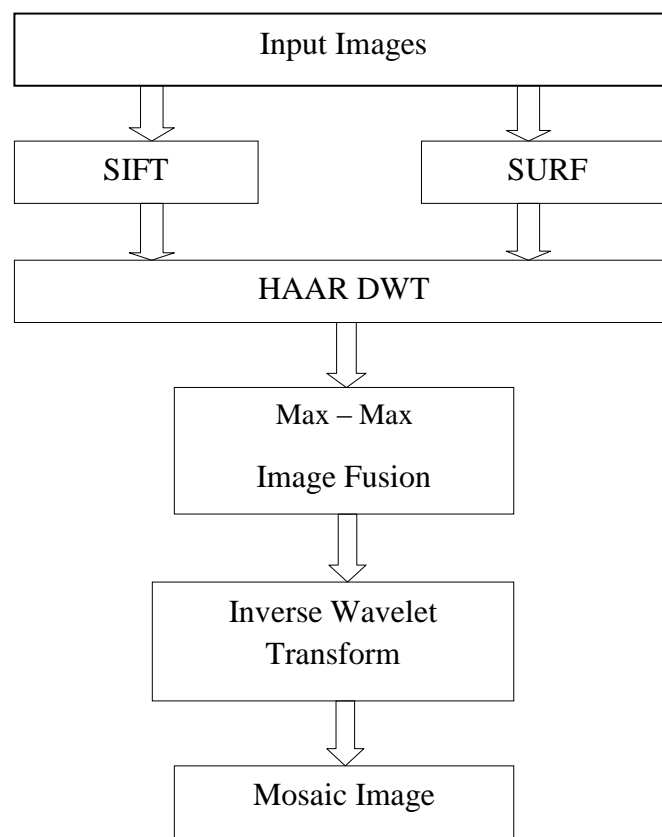


Fig.5.1 Flow chart for proposed technique

5.2.1 Scale Invariant Feature Transform

Lowe proposed a scale invariant feature transform algorithm in the year 1999. It has the unique features, such as rotation, affine transformation, scale invariance and noise immunity. SIFT algorithm is based on feature spotting in scale space. The four major steps of this algorithm are:

- i. Scale space detection, preliminary confirm the key points, location and the scale as shown in Fig.5.2 The middle point is compared with its 26 neighborhood points (8 points on the

same plane and 18 points on planes lying above and below that point) to detect utmost points. The local maxima or minima is considered as key points for further processing.

- ii. Using Taylor expansion, the extreme points and location are carefully determined using the following equation:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (5.1)$$

- iii. By the help of key point neighborhoods, the gradient $m(x, y)$ and the direction are estimated for an image $L(x, y)$. The gradient and direction can be formulated as:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (5.2)$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (5.3)$$

Taking the gradient value and characteristic into consideration, each sample point is added to the histogram. The direction of the feature points is estimated from the maximum peak values from the histogram.

- iv. Feature vectors are generated, which is shown in Fig.5.3. The arrow in each cell stands for gradient direction along with the amplitude of pixels. The seed point can be formed by aligning the unidirectional gradients followed by the normalization.

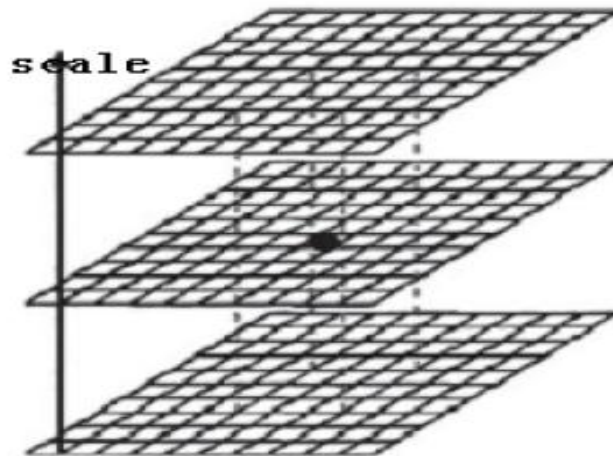


Fig. 5.2 Local extremum in DoG scale space

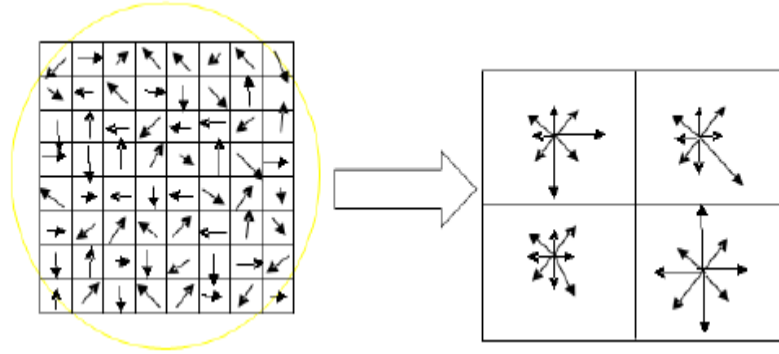


Fig. 5.3 Feature vector generation

5.2.2 Speeded Up Robust Features

A Speeded-Up Robust Features based algorithm [17] developed by Herbert Bay in 2006. It became popular for its computing speed. This algorithm is also based on scale space theory. Without down sampling, it generates a stack in order to restore the same resolution. Here, the local maxima are estimated using Hessian matrix (H). The Hessian matrix of an image at any point $X = (x, y)^T$ is

$$H(X, \sigma) = \begin{pmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{pmatrix} \quad (5.4)$$

Where, $L_{xx}(X, \sigma)$ represents the convolution of middle point X with the Gaussian filter $\frac{\partial^2 g(\sigma)}{\partial x^2}$

To enhance the computing speed, the box filter approximation is taken instead of Gaussian filter. The multi-directional box filters are shown in Fig 5.4. The determinant of Hessian matrix, ΔH can be reduced to

$$\Delta H = D_{xx}D_{yy} - (wD_{xy})^2 \quad (5.5)$$

The response for each spot can be determined by assigning $\omega = 0.9$ [17]. A threshold is set for non-maxima suppression to detect the extreme points. The stable feature points are chosen by comparing with the neighbouring values as follows

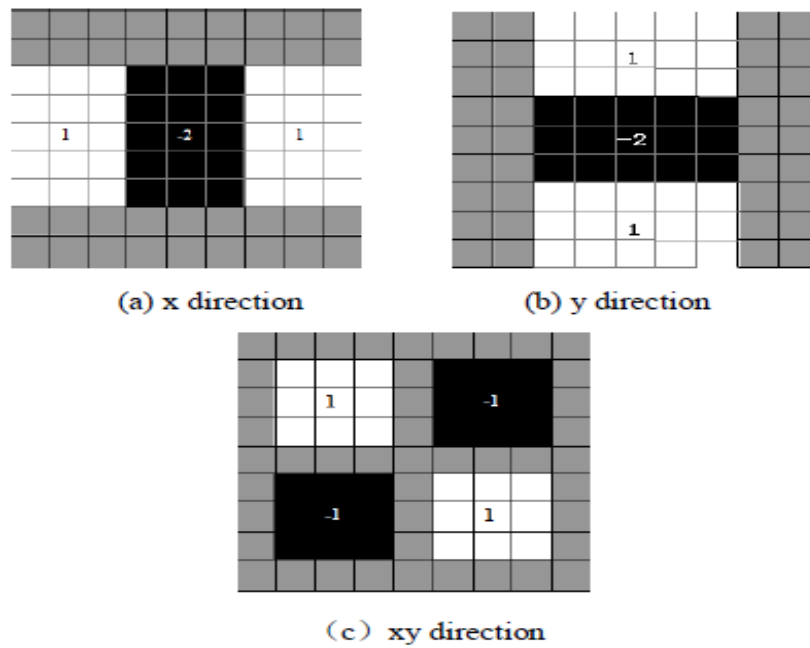


Fig.5.4 Multi-directional box filters

by the interpolation operation in scale space. Gaussian weighting coefficients are merged with Haar wavelet responses to extract the interest points. The Haar wavelet responses in vertical direction (dy) and in horizontal direction (dx) are summed up along with the absolute value of the response as:

$$V_{sub} = (\sum dx, \sum dy, \sum |dx|, \sum |dy|) \quad (5.6)$$

The normalized description vector helps to combat with illumination variance

5.2.3 Image Fusion

Image fusion is the process in which two or more images are blended together to form an image holding all the common as well as complementary information from each of the original images. The fusion process also produces a higher spatial resolution image free from all volatile blurring effects. Pixel level image fusion techniques are mostly stirred by blurring effect and usually time consuming due to a large number of computations. So, in this paper, it is opted for wavelet base multi resolution analysis technique mitigating all issues due to pixel level fusion. The original image is passed through high pass and low pass filters so as to get the detail and approximate components. Again, the down sampling operation takes place followed by the next filtering stage to generate the low-low (LL), low-high (LH), high-low (HL), high-high (HH) image sub band

components. Here, the Haar-wavelet implemented for the decomposition of better subjective analysis . The discrete wavelet based decomposition process flow is shown in Fig.5.5

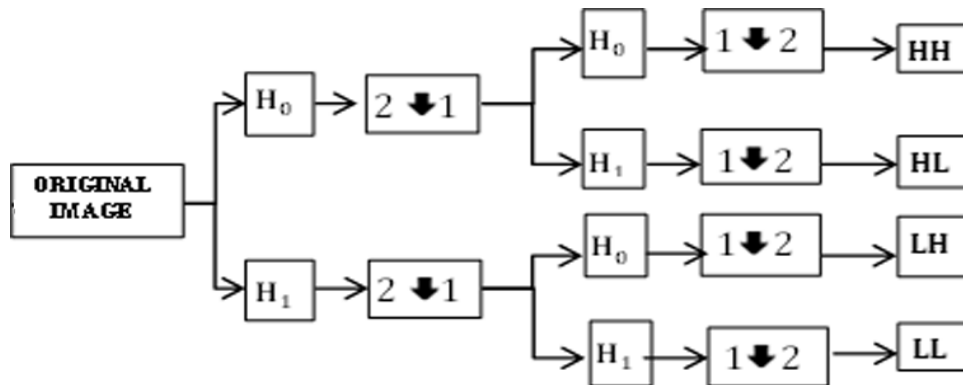


Fig.5.5 Discrete wavelet decomposition with filter banks

The unique features of Haar wavelet transform are its excellent processing speed, simplicity, memory management as well as reversibility. The subjective performance analysis of some of the popular wavelets is shown in Fig.5.6. The Haar mother wavelet function can be presented as:

$$\psi(t) = \begin{cases} 1 & 0 \leq t \leq 0.5 \\ -1 & 0.5 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad (5.7)$$

The scaling function $\phi(t)$ is given by

$$\phi(t) = \begin{cases} 1 & 0 \leq t \leq 0.5 \\ 0 & \text{otherwise} \end{cases} \quad (5.8)$$

The lower order Haar matrix is

$$H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (5.9)$$

The decomposed coefficients can be integrated using inverse discrete wavelets transform (IDWT). The fusion rule for this process is Maximum selection scheme to extract only the dominant sub-band components. The generalized discrete wavelet based image fusion process flow is depicted in Fig.5.7. The fusion of responses from scale invariant feature transform algorithm and speeded-up robust features based algorithm regenerates an panoramic image having the best features of

both the algorithms. The resultant image is robust towards rotation, noise as well as illumination invariant photography

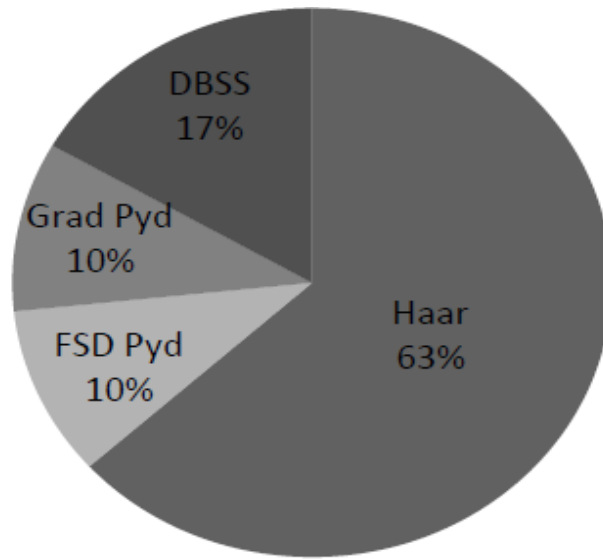


Fig.5.6 Subjective analysis of DWTs [26]

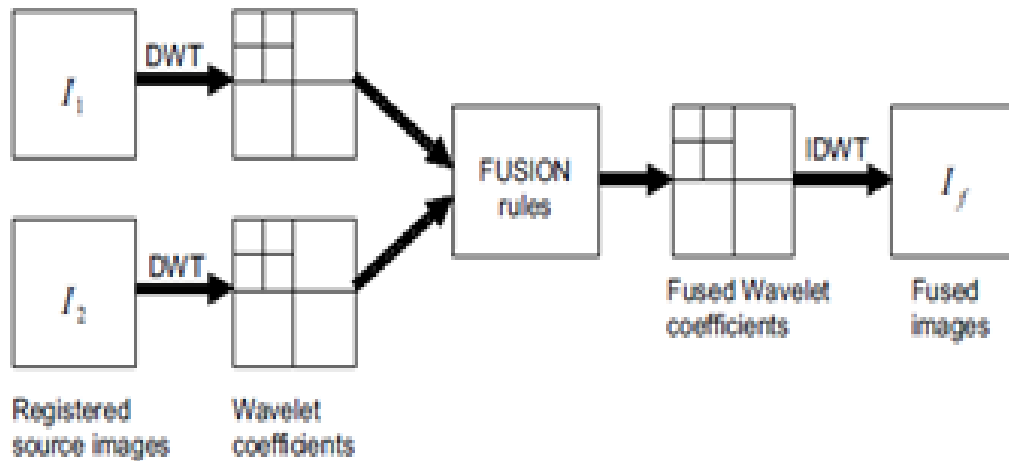


Fig.5.7 DWT based Image Fusion Flow chart

5.3 Performance evaluation

Both objective as well as subjective performance evaluation has been a crucial part of image quality evaluation process. Here, the simulation resultant image quality is verified in terms of

PSNR, Feature Similarity Index (FSIM), Mutual Information (MI), Normalized Absolute Error (NAE) and Standard Deviation (SD).

i. PSNR as Quality Measure

The peak signal-to-noise ratio (PSNR), in decibels is calculated between the reference and processed image. The more the PSNR, the better the quality of the reconstructed image. PSNR can be calculated by using the following equation:

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \quad (5.10)$$

Where, R is the maximum range in the input image data type.

ii. FSIM as Quality Measure

For Combined similarity

$$S_L(x) = [S_{PC}(x)][S_G(x)] \quad (5.11)$$

$$PC_M(x) = \max(PC_1(x), PC_2(x)) \quad (5.12)$$

$$FSIM = \frac{\sum_{x \in \Omega} S_L(x) \cdot PC_M(x)}{\sum_{x \in \Omega} PC_M(x)} \quad (5.13)$$

iii. Mutual Information (MI)

It measures the asymmetry between the two desired image as well as the fluctuation from its mean value. MI for two images M (i, j) and N (i, j) can be expressed as

$$MI = H(M) + H(N) - H(M,N) \quad (5.14)$$

Where.H(M) is the entropy of image M(i, j), H(N) is the entropy of image N(i, j) and H(M, N) is the joint entropy of image M(i, j) and N(i, j).

iv. Enhancement performance measure (EME)

It is a quantitative method to measure the image enhancement. In terms of entropy, it can be defined with the help of an image, which is divided into $k_1 \times k_2$ blocks $w_{k,l}(i, j)$

$$EME = \min_{\phi \in \{\phi\}} (EME(\phi)) \quad (5.15)$$

$$EME = \min_{\phi \in \{\phi\}} \left(\frac{1}{k_1 k_2} \sum \sum 20 \log \frac{I_{\max;k,l}^w(\phi) - I_{\min;k,l}^w(\phi)}{I_{\max;k,l}^w(\phi) + I_{\min;k,l}^w(\phi)} \right) \quad (5.16)$$

v. Normalized Absolute Error (NAE)

The Normalized Absolute Error can be used for image quality metric and formulated as

$$NAE = \frac{\sum_{i=1}^m \sum_{j=1}^n (|A_{ij} - B_{ij}|)}{\sum_{i=1}^m \sum_{j=1}^n (A_{ij})} \quad (5.17)$$

Where, A- perfect image and B- fused image to be assessed. The image fusion result is appreciable with respect to subjective analysis. The objective evaluation of the fused image is depicted in Table.5.1

Table 5.1: Performance analysis of SIFT, SURF and Proposed method

Algorithm/ Parameters	SIFT	SURF	Proposed
PSNR (dB)	41.693	41.962	42.415
FSIM	0.706	0.714	0.739
MI	1.209	1.264	1.465
EME	8.561	6.332	9.457
NAE	0.147	0.143	0.132
SD	56.846	56.542	57.283

5.4 Results & Discussion

In the proposed technique, the two test images are acquired by means of a camera DSC-WS70, maximum aperture of 2.75, focal length of 14 mm and exposure time of 0.02 sec. The unique feature of the images is that, these are three dimensional rotational images. Here, the two captured images at rotational angle of 10°. The input images are shown in Fig.5.8 and Fig.5.9. The images also have some illumination variation. The ground truth image has been generated using Autostitch software. The images are processed through scale invariant feature transform and speeded-up robust features algorithms separately in a parallel process. The response of SIFT and

SUFT algorithms is shown in Fig.5.10 and Fig.5.11 respectively. The panoramic images generated from these algorithms are passed through the blending process using the Haar discrete wavelet transform. The wavelet decomposition tree is presented in Fig.5.13. Here, implement the maximum-approximate and maximum-detail fusion rule to generate a panoramic image of high contrast, robust towards noise as well as illumination variation. The fused panoramic image is depicted in Fig.5.12. Here the panoramic image generated by the fusion process compensates the complementary features and boosts up the common features of individual stitching images. SURF algorithm has the distinctive property of illumination invariance along with good scale and rotation invariance property, whereas, SIFT is the most effective algorithm for scale and rotate image stitching [14]. But, it cannot cope up with illumination variation. Therefore, the resultant image proves superior as compared to the SIFT as well as SURF algorithms in terms of PSNR, Mutual Information (MI), Normalized Absolute Error (NAE), Feature Similarity Index Metric (FSIM), Standard Deviation (SD) and Measure of Enhancement (EME).



Fig.5.8 Input image-I



Fig.5.9 Input image-II



Fig.5.10 SIFT response



Fig.5.11 SURF response



Fig.5.12 Fused image

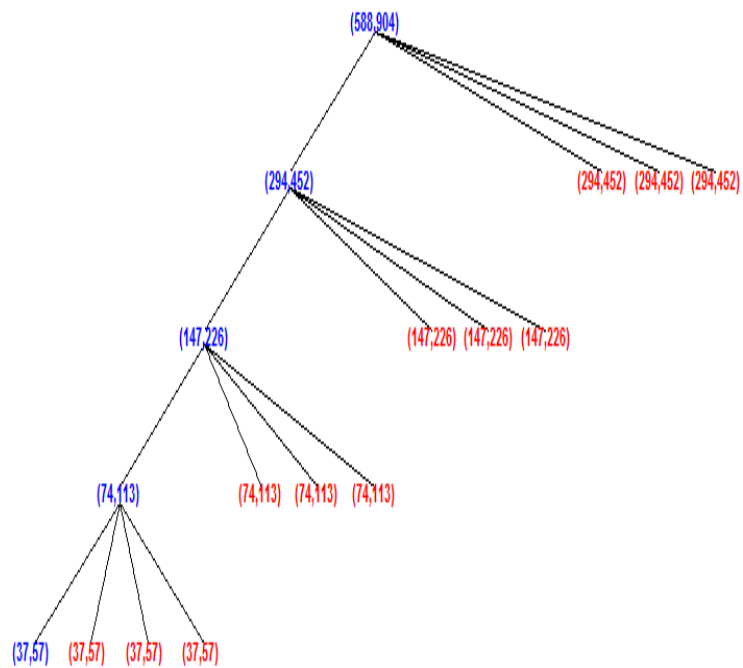


Fig.5.13 Haar Wavelet decomposition tree at 4th level

5.5 Conclusions

In this chapter, a novel panoramic image mosaicking technique for three dimensional, rotational images with illumination variation has been proposed. The input images are passed through two robust stitching algorithms i.e. SIFT and SURF. The Scale Invariant Feature Transform is invariant towards scale and rotational variation. It is also robust towards noisy environment. Speeded-Up Robust Features algorithm is has very similar properties as SIFT. However, it has the properties of illumination invariance and good computational speed. Therefore, the fusion result of these two efficient algorithms gives rise to a panoramic image, which carries all the properties of both. The performance evaluation of the proposed technique is done in terms of PSNR, FSIM, MI, EME, NAE and SD. The proposed method shows superior results as compared to both SIFT and SURF.

Chapter 6

Conclusion

Conclusion

Suggestions for Future Work

This chapter focuses on the advantages and limitations of all the algorithms used for image mosaicing. The scopes of future research work in this domain are also discussed.

6.1 Conclusion

Harris corner detector is the most popular and widely used corner detector. It can be combinely used for corner and edge detection but in this chapter it is used for detection of the corner points, which are nothing but the interest points. Interest points are used for finding the match between two images having some common field of view. This corner detector is efficient, cheap and rotation invariant but it can't handle the images which suffer scale variation. Also the detector is taking more time for the computation of corner points. The SIFT algorithm overcame this problem since it is robust to noise, minor changes in viewpoint and also towards the scale and rotation variation of the image. It extracts highly distinctive features of the set of images. It helps in the correct object identification with low probability of mismatch. But the SURF algorithm was found to perform much better for the illumination variance and also it took lesser time for execution. Hence to supplement the complimentary features of the SIFT and SURF, a new Feature based image mosaicing technique using image fusion has been proposed and implemented based on a robust algorithm which takes care of the combined effect of rotation, illumination, noise variation and other minor variation. The input images are stitched together using the popular stitching algorithms i.e. SIFT and SURF. To extract the best features from the stitching results, the blending process is done by means of Discrete Wavelet Transform using the maximum selection rule for both approximate as well as detail-components. The robustness and quality of the above mosaicing techniques are tested by means of three-dimensional rotational images. The panoramic image generated by the fusion process compensates the complementary features and boosts up the common features of individual stitching images. SURF algorithm has the distinctive property of illumination invariance along with good scale and rotation invariance property, whereas, SIFT is the most effective algorithm for scale and rotate image stitching. But, it cannot cope up with illumination variation. Therefore, the resultant image proves superior as compared to the SIFT as well as SURF algorithms in terms of PSNR, MI, NAE, FSIM, SD and EME.

6.2 Future Work

The given work was based the implementation and the quantative evaluation of the three above mentioned image mosaicing techniques. Also an approach and methodology have been proposed to enhance the performance of the above methods to produce the best quality images by the fusion of the complimentary features specific to these algorithms. The test input images used for the present work were the planar 3-D images; it can be extended for the cylindrical and spherical images as well. Individual algorithms can be further improved at the algorithmic level rather that at the level of implementation to incorporate the added and complimentary features of the other algorithms, and hence the execution time as well the procedural length can be greatly reduced.

References

- [1].G. Ward, "Hiding seams in high dynamic range panoramas". Proc. of the 3rd ACM International symposium on applied perception in graphics and visualization, pp.153-155, 2006.
- [2].D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital registration", IEEE Trans, Comput, vol. C-21, pp. 179-186, 1972.
- [3].C. D. Kuglin and D. C. Hines, "The phase correlation image alignment method", Proc. of IEEE Int. Conf. on Cybernet Society, New York, pp. 163-165,1975.
- [4]. L. G. Brown, "A survey of image registration techniques", ACM Computing Surveys, vol.24, pp.325-376, 1992.
- [5].G. Wolberg, "Digital Image Warping", Proc. of IEEE Computer Society Press, California, pp.169-172, 1990.
- [6].R. B. Inampudi, "Image mosaicing", Proc. of International Conference on Geoscience and Remote Sensing Symposium, Seattle, pp.2363-2365, 1998.
- [7].S. Peleg and J. Herman, "Panoramic mosaics by manifold projection" Proc. of IEEE Computer Society conference on Computer Vision and Pattern Recognition, San Juan, pp.338-343, 1997.
- [8].B. Rousso, S. Peleg, I. Finci and R. Acha," Universal mosaicing using pipe projection", Proc. of Sixth International Conference on Computer Vision, pp.945-950,1998.
- [9].J. Wang and Y. Li, "Image mosaicing algorithm based on salient region and MVSC", Proc. of International Conference on Multimedia and Signal Processing, pp. 207-211, 2011.
- [10]. D. Ghosh, S. Park, N. Kaabouch and W. Semke," Quantum evaluation of image mosaicing in multiple scene categories ", Proc. of IEEE Conference on Electro/Information Technology, pp.1-6, 2012.
- [11]. D.K. Jain, G. Saxena, V.K. Singh, "Image mosaicing using corner technique", Proc. of International Conference on Communication System and Network Technologies, pp.79-84, 2012.

- [12]. L. Juan et. al., "A comparison of SIFT, PCA-SIFT, and SURF", *International Journal of Image Processing*, vol. 3, pp.143-152, 2009.
- [13]. D. G Lowe, "Object recognition from local scale-invariant features", *International Conference on Computer Vision*, vol. 2, pp.1150– 1157,1999.
- [14]. D. G. Lowe, "Local feature view clustering for 3D object recognition", *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, pp.682-688, 2001.
- [15]. J. Beis, and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces", *Conference on Computer Vision and Pattern Recognition*, Puerto Rico, pp. 1000–1006, 1997.
- [16]. D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", *International Journal of Computer Vision*, vol. 60, pp. 91- 110, 2004.
- [17]. H. Bay, et al., "Speeded up Robust Features", *Proc. of the 9th European Conf. on Computer Vision*, Cambridge, U.K., pp. 404-417, 2006
- [18]. R. Szeliski and S. Ken, "Direct methods for visual scene reconstruction", *IEEE Workshop on Representations of Visual Scenes*, Cambridge, pp. 26–33, 1995.
- [19]. D. Marr and E. Hildreth, "Theory of Edge Detection," *Proc. of Royal Soc. London B*, vol. 207, pp. 187-217, 1980.
- [20]. I. Zoghiami, O. Faugeras, and R. Deriche, "Using geometric corners to build a 2D mosaic from a set of images", *Proc. of the International Conference on Computer Vision and Pattern Recognition*, Puerto Rico, pp.340-349, 1997.
- [21]. P. McLauchlan and A. Jaenicke, "Image mosaicing using sequential bundle adjustment", *Image and Vision Computing*, vol. 20, pp.751–759, 2002.
- [22]. C. Harris, "Geometry from visual motion" A. Blake and A. Yuille, editors, *Active Vision*, MIT Press, pp. 263–284, 1992.
- [23]. B. Zitova and J. Flusser, "Image registration methods: A survey", *Image and Vision Computing*, vol.21, pp.977-1000, 2003.

- [24]. K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, pp.1615-1630, 2005.
- [25]. G. Pajares and J. Cruz, "A wavelet-based image fusion tutorial", Elsevier Pattern Recognition. vol. 37, pp. 1855 – 1872, 2004.
- [26]. S. Krishnamoorthy and K. Soman, "Implementation and Comparative Study of Image Fusion algorithms", International Journal of Computer Application, vol. 9, pp. 25-35, 2010.
- [27]. A. P. Witkin, "Scale-space filtering", Int. Joint Conference on Artificial Intelligence, Karlsruhe, West Germany, pp.1019-1022, 1983.
- [28]. Yu Wang and Y. Wang, "The image matching algorithm based on SIFT and Wavelet transform", Journal of Beijing Institute of Technology, vol. 5, pp. 341-352, 2009.
- [29]. L. Huihui, G. Lei and L. Hang, "Research on Image Fusion Using Wavelet Transform Based on Gradient Selective Rule", Computer Engineering and Applications, vol.12, pp.76-78, 2005.
- [30]. R. C. Gonzalez and R. E. Woods, "Digital Image Processing", 3rd ed., Pearson Education publication, 2012.

Dissemination

[1] Vimal Singh Bind, Priya Ranjan Mudul and Umesh Chandra Pati, “A Robust Technique for Feature-based Image Mosaicing using Image Fusion”, *International Journal of Advanced Computer Research (IJACR)*, Vol.3, Issue-8, pp.263-268, March 2013, India