IMAGE MOSAICING

A Project Work Submitted in Partial Fulfillment

of the requirements for the Degree of

BACHELOR OF TECHNOLOGY

in

ELECTRONICS & COMMUNICATION ENGINEERING

by

PARNASHI HAZARIKA (*Gau/C/10/L-114*)
SNEHA KUMARI (*Gau/C/10/L-113*)
KHANTHAI MALA BASUMATARY (*Gau/C/10/L-121*)
&
DEBADYUTI ROY (*Gau/C/10/03*)

Under the supervision of
Mr. Haradhan Chel
Asst. Professor, Dept. of ECE

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

केन्द्रीय प्रौद्योगिकी संस्थान कोकराझार
CENTRAL INSTITUTE OF TECHNOLOGY KOKRAJHAR
(A Centrally Funded Institute under Ministry of HRD, Govt. of India)
BODOLAND TERRITORIAL AREAS DISTRICTS :: KOKRAJHAR :: ASSAM :: 783370
Website: www.cit.kokrajhar.in , www.cit.ac.in

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING

केन्द्रीय प्रौद्योगिकी संस्थान कोकराझार

CENTRAL INSTITUTE OF TECHNOLOGY, KOKRAJHAR

(An Autonomous Institute under MHRD)

Kokrajhar – 783370, BTAD, Assam, India

CERTIFICATE OF APPROVAL

This is to certify that the work embodied in this project entitled as ***"Image Mosaicing"*** submitted by **Parnashi Hazarika (*Gau/C/10/L-114*), Sneha Kumari (*Gau/C/10/L-113*), Khanthai Mala Basumatary (*Gau/C/10/L-121*)** and **Debadyuti Roy (*Gau/C/10/03*)** to the Department of Electronics & Communication Engineering, is carried out under my direct supervisions and guidance.

The project work has been prepared as per the regulations of Central Institute of Technology and I strongly recommend that this project work be accepted in partial fulfillment of the requirement for the degree of B. Tech.

Supervisor

Mr. Haradhan Chel
Assistant Professor, Dept. of ECE

Countersigned by

Mr. Arindum Mukharjee
HoD (i/c), Department of ECE

DEPARTMENT OF ELECTRONICS & COMMUNICATION ENGINEERING
केन्द्रीय प्रौद्योगिकी संस्थान कोकराझार
CENTRAL INSTITUTE OF TECHNOLOGY, KOKRAJHAR
(An Autonomous Institute under MHRD)
Kokrajhar – 783370, BTAD, Assam, India

CERTIFICATE BY THE BOARD OF EXAMINERS

This is to certify that the project work entitled as "***Image Mosaicing***" submitted by **Parnashi Hazarika (*Gau/C/10/L-114*)**, **Sneha Kumari (*Gau/C/10/L-113*)**, **Khanthai Mala Basumatary (*Gau/C/10/L-121*)** and **Debadyuti Roy (*Gau/C/10/03*)** to the Department of Electronics & Communication Engineering of Central Institute of Technology, Kokrajhar has been examined and evaluated.

The project work has been prepared as per the regulations of Central Institute of Technology and qualifies to be accepted in partial fulfillment of the requirement for the degree of B. Tech.

Project Co-ordinator

Board of Examiners

____

____

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompanied the successful completion of any task would be incomplete without mentioning about the people whose constant guidance and encouragement made it possible. We take pleasure in presenting before you, our project, which is the result of studied blend of both research and knowledge.

First of all, we give our sincere thanks to **The Principal, *Central Institute of Technology,*** for providing us all the facilities that were required for the completion of the project.

We would also like to have the privilege of thanking **Mr. Arindum Mukharjee,** ***HoD (i/c), Dept. of ECE*** for giving us the opportunity to work on this project.

We express our earnest gratitude to our Project Guide **Mr. Haradhan Chel,** ***Asst. Professor, Dept. of ECE,*** for his constant support, encouragement and guidance throughout the session. We are grateful for his cooperation and valuable suggestions.

Date:

Parnashi Hazarika
University Roll: ***Gau/C/10/L-114***
University Registration No: ***59453 (2009-10)***

Sneha Kumari
University Roll: ***Gau/C/10/L-113***
University Registration No: ***59452 (2009-10)***

Khanthai Mala Basumatary
University Roll: ***Gau/C/10/L-121***
University Registration No: ***59641 (2009-10)***

Debadyuti Roy
University Roll: ***Gau/C/10/03***
University Registration No: ***57438 (2009-10)***

ABSTRACT

There are situations where it is not possible to capture large documents with the given imaging media such as cameras, scanners or copying machines in a single stretch because of their inherent limitations. This results in capture of a large document in terms of split components into the original and put together the document image. Through Image Mosaicing, we can have an image of a large area with more detail and better quality.

In this paper, we present a simple approach to mosaic two separate images of the same object based on pixel value matching. The method compares the values of pixels in the columns of split images to identify the common overlapping region in then, which helps in mosaicing of split images of a large document.

CONTENTS

INTRODUCTION

Many a time, it may not be possible to capture image of a large document in a single exposure as most image capturing media work with documents of definite size. In such cases, the document has to be scanned part by part producing split images. Thus document image analysis and processing require mosaicing of the split images to obtain a complete final image of the document. Hence document image mosaicing is the process of merging split images that are obtained by scanning different parts of single large document image with some sort of overlapping region or to produce a single and complete image of the document. In this paper, we consider only two split images to obtain the final fused image.

OBJECTIVE AND MOTIVATION

The objective of our project entitled as "***Image mosaicing***" is to form a large image by composing different sections of the image captured separately with some common portions between the sections with the help of MATLAB. To do so, we have studied the sequential methodology of the process. So our target is to mosaic more than two sections of a large image by opting different optimization methods. For that, we will first try to perform the image mosaicing for the same image taken twice with positional changes and with the same sensor. After that, we will try to implement the same process for more than two images by applying various methods, so that we can get an efficient alignment of the image.

The growing field of image mosaicing and its increasing future prospects on various fields has encouraged us to take over this project. Image mosaicing is a research based project and a very interesting field. There are a lot of works already have been done with this field as well as several researches are going on and we are really keen to enrich our knowledge in this field and to focus in its implementations in various technology in future days.

APPLICATIONS AND SCOPE:

1) **Formation of High-Resolution image:** Resolution of an image can be improved through image mosaicing. A low resolution image can be converted into a high resolution image by mosaicing several snaps of the same object since in every capture, the sensor senses different points of the same object. So in every image, there are some information that remains hidden and those details can be found out with the help of Image mosaicing. This application of the process is broadly applied in the Medical Science where Super-Resolution images are required for getting the accurate output of any diagnosis. For example, in Ultra-Sonography, we need the finer detail of the image for diagnosis. So it can be obtained by taking more than ten images of the same object and on them, image mosaicing is performed.

2) **Navigation Systems:** It is widely used in the navigation systems to determine the location of a particular place with the help of the maps of the small parts of the area. It is preferably included in the satellite navigation system.

3) **3-D Modelling:** It is very helpful for creating a 3-Dimensional model from the 2-Dimensional images of any object. By capturing the image of the object from different angles and by applying some transformation, we can get the 3-D model of that particular object through the fusion of the images.

4) **Security Systems:** Image mosaicing helps to identify any person by capturing different parts of the body. It also helps to track the changes of the location of any object by scanning its features.

OVERVIEW OF THE RELATED WORK:

There are three basic steps involved in the process of Image Mosaicing. They are briefly discussed below:

1) **Image Capturing:** Images of different sections of any large are or building are captured using a camera. The sensor used for capturing and the time of taking the image may be same or different. But it is necessary to keep some common portion between each consecutive pair of the images.

2) **Image Registration:** It is the alignment of the images considering the changes occurred in the images taken from different viewpoints or at different times or both. In this process, the pixels with almost same values are found out from the given images and aligns them so that when the processed image is fused over the reference image, it completely merges with the reference with negligible error, giving more details. Through Image Registration, we can form a high resolution image from several low resolution images.

   The three primary components required for the registration of an image are:

   a) **The Reference Image,**
   b) **Transformation Model and**
   c) **Optimization Method**

3) **Image Fusion:** Once the images are aligned through Image Registration, the processed image is fused over the reference image. After superimposing the image, if the images are aligned properly, then the quality of the image will be improved and we will get a better resolution image.

The basic steps of image mosaicing can be described using a schematic block diagram:

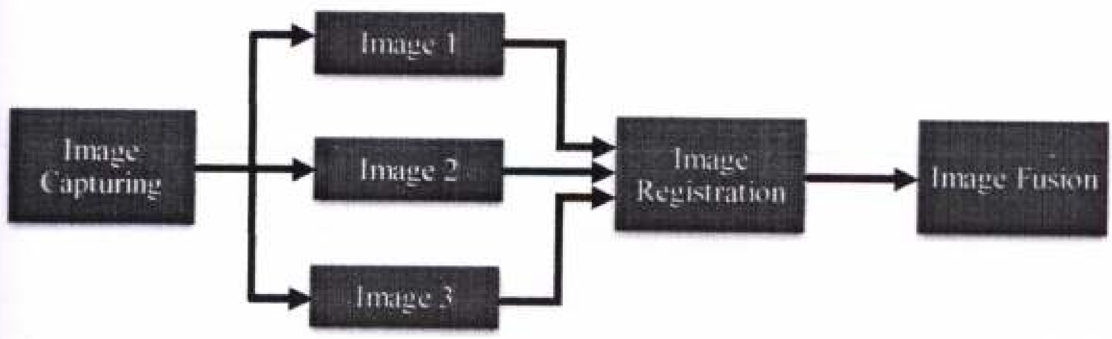


Image Capturing:

The different sections of a large area are captured in such a way that some common portion remains in each consecutive pair of the images.

Sample captured images:

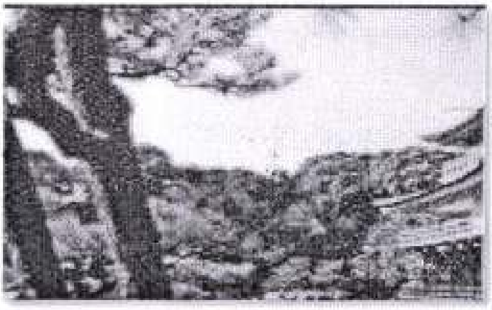


Image 1

Image 2

Image Registration:

Image Registration is primarily classified into two types:-

1) **Rigid Registration and**
2) **Non-Rigid Registration**

1) Rigid Registration:

In this type of registration, the images are aligned using some geometric transformations. Here, basic type of transformations are considered which are generally occurred during the capturing of an image.

There are certain things that may happen as follows during capturing of an image:

a) **Translation:** The sections may have been captured from different distances which will cause in the difference of the size of the objects between the images. In this case, we have to consider a Scaling parameter that will bring the objects inside the image to an equivalent size.
b) **Shifting:** Since the portions of the area are captured from different viewpoints, so shifting of the pixel coordinate will be arose. Hence, for the alignment of the objects, we will have to consider some shifting parameter that will reduce the difference between the pixel values.

The shifting equation is given as:

$$P' = SP$$

where, $\mathbf{P'}$ → Shifted or processed image

$\mathbf{P}$ → Original or Reference Image

$\mathbf{S}$ → Shifting matrix

The shifting matrix is given as:

$$S = \begin{pmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{pmatrix}$$

where, t_x , t_y → shifting factors along x and y-coordinate.

c) **Rotation:** Some rotation may occur during capturing of an image due to the holding positions of the capturing device, hence we have to consider some rotation parameter also to align the images.

The rotation equation is given as:

$$P' = RP$$

where, $\mathbf{P'}$ → Rotated or Processed image

$\mathbf{P}$ → Original or reference Image

$\mathbf{R}$ → Rotation matrix

The shifting matrix is given as:

$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

d) **Perspective changes:** During capturing of an image, due to the distance of the viewpoint, some perspective changes may occur in the images such as upward inclination or downward inclination. So we have derive some mathematical model considering some virtual axes of rotation so that we can achieve the required alignment with minimum error possible.
e) **Affine transformation:** Due to movement of the sensor and the changes in the reflections coming from the sources, some type of stretching or radial changes of the captured image may occur. To reduce this type of complexity of the registration, we have to use some geometric transformations.

Rigid Registration is again classified into two sub categories:

a) **Manual** and
b) **Automatic**

a) Manual Rigid Registration:

In this type of registration, the image is observed and few similar points from both the images are chosen. Based on the changes of the pixel coordinate in the second image with the reference image, some predicted geometric transformations are applied on the image and the images are fused over one another. It is called Manual Registration because the similar points are to be selected manually with the help of observation by human-eye.

Drawbacks of Manual Registration:

1) This method does not give efficient result or perfect alignment of the images because the selected points are not perfectly similar with the other image. As it is determined by human-eye, though the points look similar, the pixel values may differ with a non-negligible amount. Hence giving a wrong prediction of the geometric transformation.

2) This method cannot be implemented on a device because it needs human contribution to perform the transformation. It is completely dependent on human observation.

b) Automatic Rigid Registration:

In this type of registration, some predefined methods are used to determine the transformation to be applied on the image. It detects some sort of calculations and comparisons and accordingly applying a transformation on the image.

There are many types of automatic registration for determining the parameters of transformation, but three common types are:

a) **Feature Based,**
b) **Mean Square Error Based and**
c) **Correlation Based**

b.1) Feature Based Image Registration:

This method detects some certain features of the image by dividing it in a grid and compares the features with the reference image. According to the comparison results, the geometric transformation is determined and then applied over the whole image.

One limitation of this process is that, a sufficient number of features must be calculated to determine the parameters of transformation. If it fails to detect sufficient features, the whole transformation gets failed.

Another limitation is that, it is less efficient for higher resolution images. The more the number of pixels, the more number of blocks will be there and the variation of features will be more, resulting in slower process of registration.

b.2) Mean Square Error Based Image Registration:

In this type of registration, the alignment is done by computing pixel to pixel difference and hence calculating the MSE (Mean Square Error). Using some optimization method, the parameters of transformation are determined where minimum MSE is found and then that transformation is applied over the whole image.

The MSE is calculated as:

$$MSE = \sqrt{\frac{\sum_{y=0}^{N-1}\sum_{x=0}^{M-1}\{\hat{f}(x,y)-f(x,y)\}2}{MN}}$$

where,
$\hat{f}(x,y)$ → pixel intensity of the processed image at (x, y)
$f(x,y)$ → pixel intensity of the reference image at (x, y)
M, N → Dimensions of the image

b.3) Correlation Based Image Registration:

In this process, the alignment is done by calculating the cross-correlation between the two images. The pixels where the correlation is minimum that is determined as the matched pixels and the transformation according to the changes of that particular coordinate is applied over the whole image.

The Cross-Correlation of two signals f(t) and g(t) is given as:

$$f(t) * g(t) = \int_{-\infty}^{\infty} f^*(t)g(t+\tau)d\tau$$

where,
f(t) → 1^{st} signal
g(t) → 2^{nd} signal
τ → delay or shifting factor

In this project, we are capturing a single image from the same viewpoint with a minimum of perspective changes. And we are considering only two parameters viz. Shifting and Rotation. After the proper alignment of the two images through image registration, the two images are fused over one another. Since we have considered only two parameters of changes, hence the result is not that efficient but comparatively we get a good result and the image quality is increased.

PROPOSED METHODOLOGY:

There are few popular optimization methods that are widely used for image registration. Some of them are 1st order and others are 2nd order methods. The 1st order methods are comparatively slower than the 2nd order methods. Some of the methods are:

1) **Gradient Descent Method,**
2) **Conjugate Descent Method,**
3) **Particle Swarm Optimization Method.**

In this project, we have opted the ***MSE*** (Mean Square Error) based Image registration using ***Gradient Descent Optimization Method.***

Gradient Descent Method:

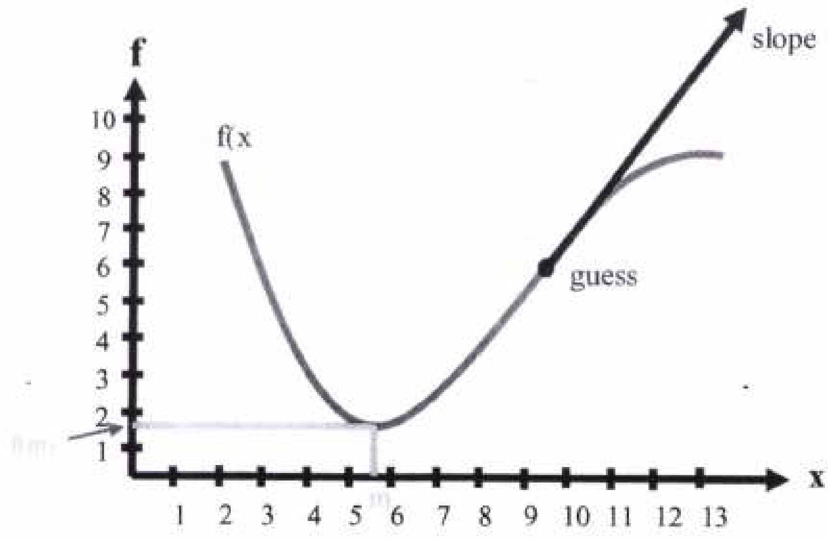
It is a 1st order method in which the parameters of transformation are determined by computing the pixel to pixel difference. The gradient on each point is calculated and from them the values of the transformation factors are calculated based on MMSE (Minimum Mean Square Error).

The Algorithm for the Method is as follows:

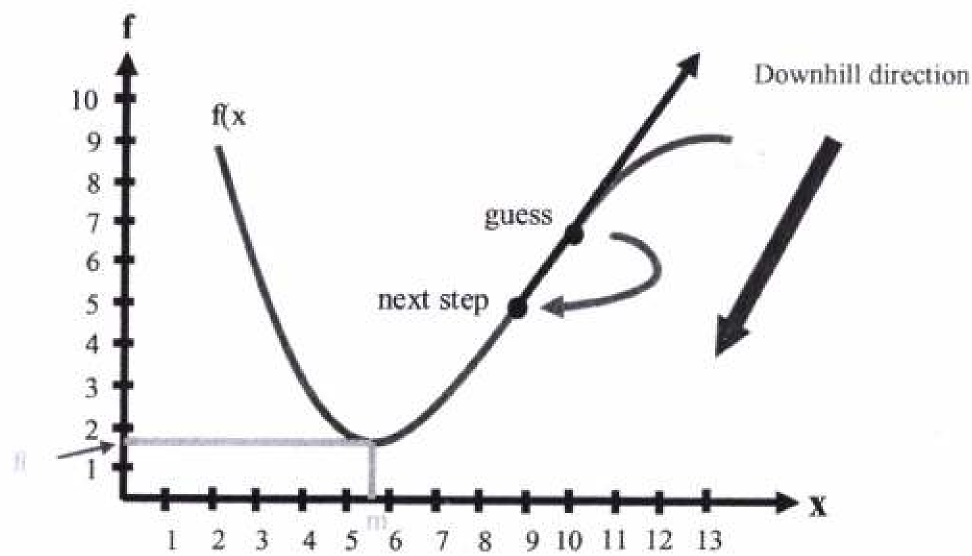
1) **Starting with a chosen guess point x € domain f**
2) **Repeat:**
   a) Determine a descent direction: $\Delta x := \nabla f(x)$
   b) Line search: Choose a step size h via exact or backtracking line search.
   c) Update: $x := x + t\Delta x$
3) **Until the stopping criterion is satisfied.**

The stopping criterion is usually of the form $\| \nabla \mathbf{f(x)} \|_2 < e$
where, $e \rightarrow$ a very small chosen value generally in the range of $\mathbf{10^{-5}}$

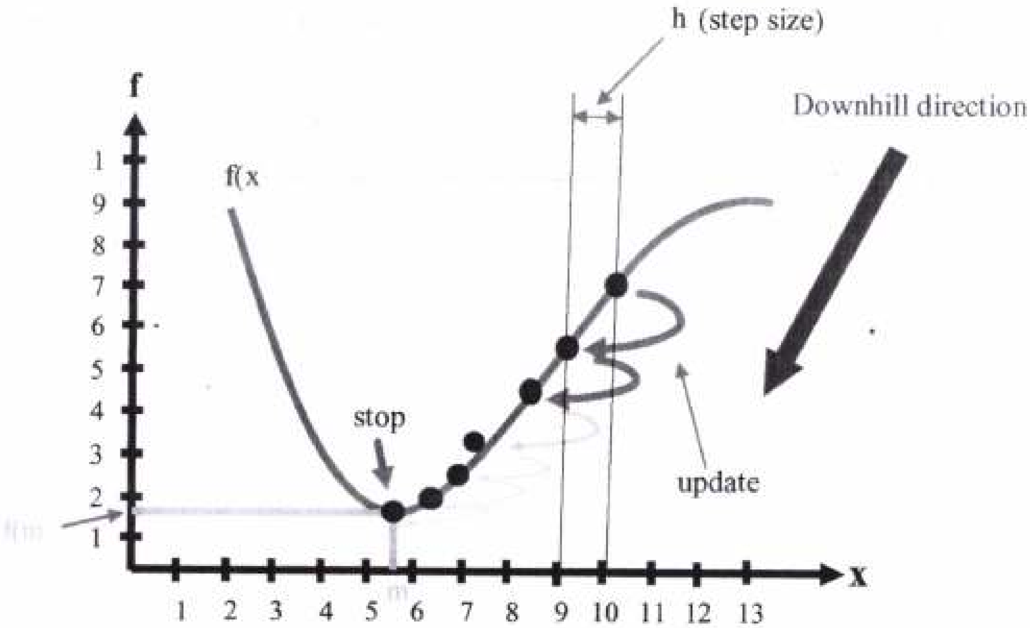
(1) Taking an initial guess value:



(2) Determining the Descent Direction:



(3) Choose a step size h and update by $x := x + t\,\Delta x$

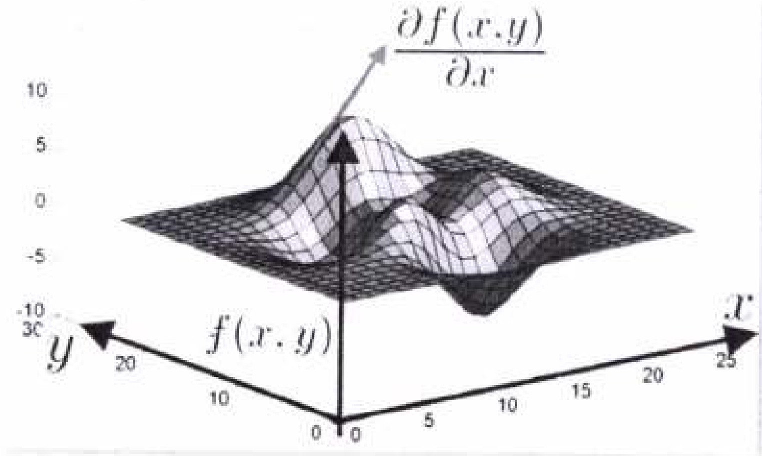


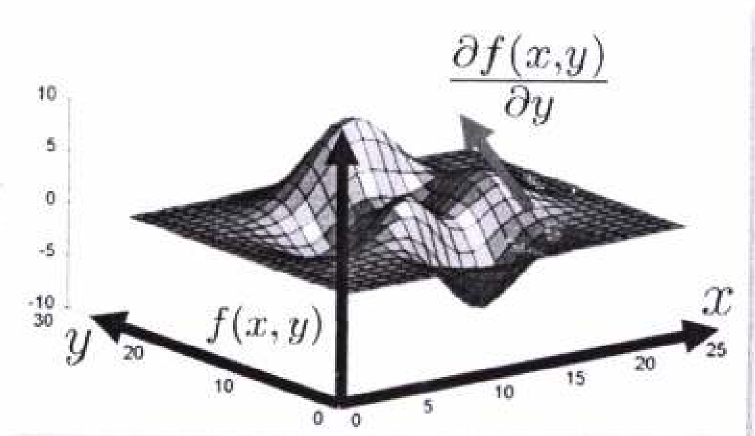
On each point, the gradient is calculated and according to that the mean error is calculated. When f(m) is minimum i.e., global minima is reached, the process is stopped and the found parameters are applied to the whole image.

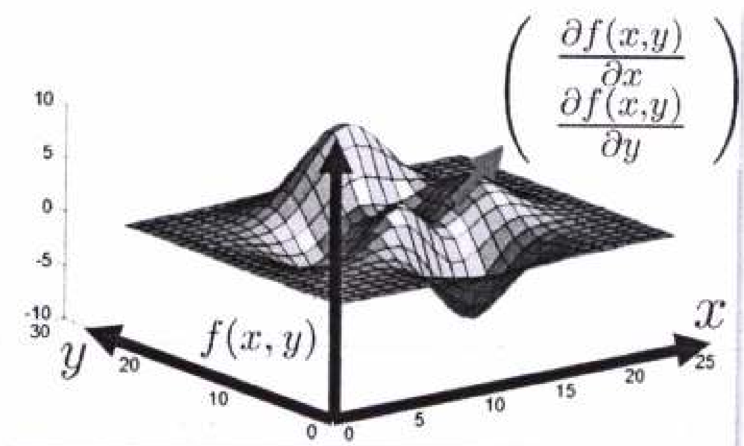
The mathematical approach is given as:

guess = $\mathbf{x}$
direction = $-f'(x)$
step = $\mathbf{h} > 0$
$\mathbf{x} = \mathbf{x} - \mathbf{h} f'(x)$
$f'(x) \sim 0$

Illustration of gradient in 2-D:

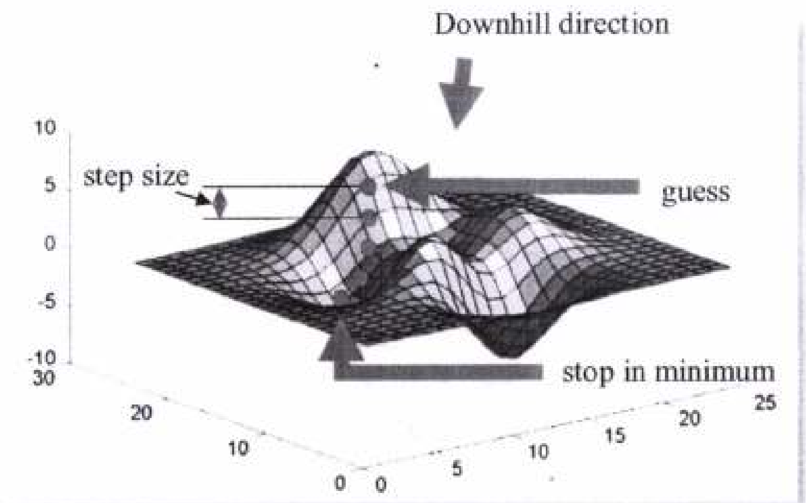






The gradient in 2-D is defined as:

$$\nabla f(x,y) = \begin{pmatrix} \dfrac{\partial f(x,y)}{\partial x} \\ \dfrac{\partial f(x,y)}{\partial y} \end{pmatrix}$$



2-D Mathematical Approach is given as:

guess = (x, y)

direction = $-f'(x, y)$

step = $\mathbf{h} > \mathbf{0}$

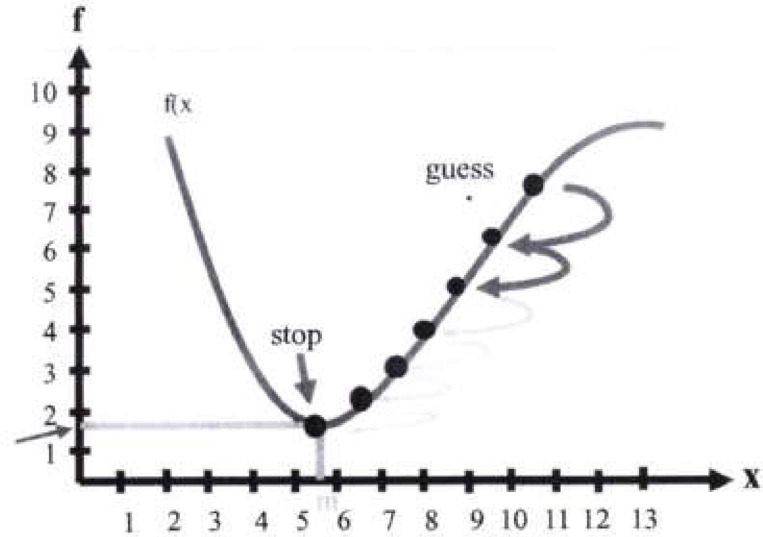
$\mathbf{x} = \mathbf{x} - \mathbf{h}\, \nabla f'(x)$

$\mathbf{y} = \mathbf{y} - \mathbf{h}\, \nabla f'(y)$

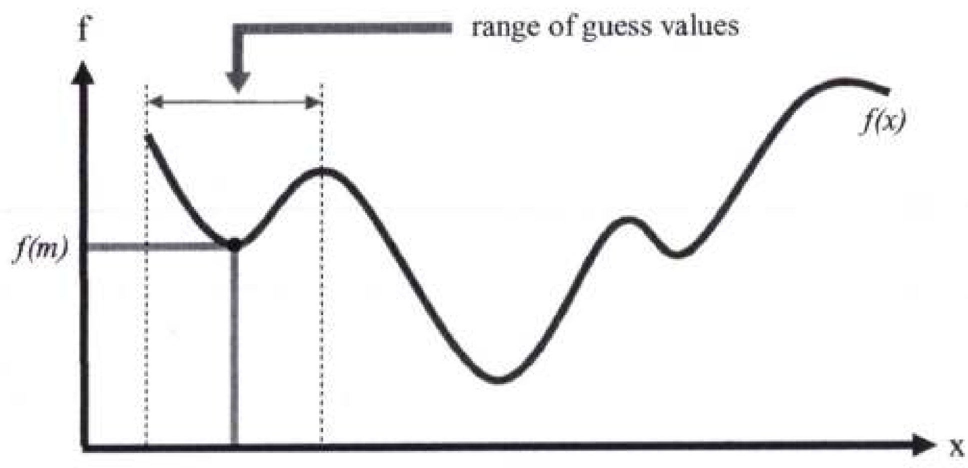
$\nabla f'(x, y) \sim \mathbf{0}$

Limitations of Gradient Descent Optimization Method:

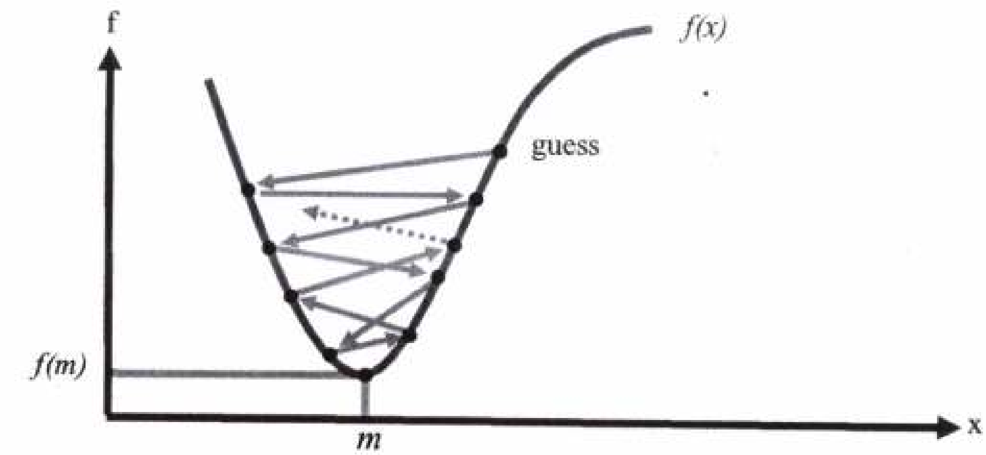
1) This method is very slow. Since it calculates the MSE (Mean Square Error) by computing the difference between the pixels of the two images, so more the number of the pixels i.e., higher the resolution of the image or larger the dimension of the image, slower the process.



2) This method may not give efficient result due to the wrong selection of guess value. Since we are calculating the initial guess value within a range of values, so the range of values may not contain the global minima; instead it may give us the local minima as the global minima and hence according to that guess value, the minimum error will be calculated and the result will be not that much efficient or may be completely wrong.

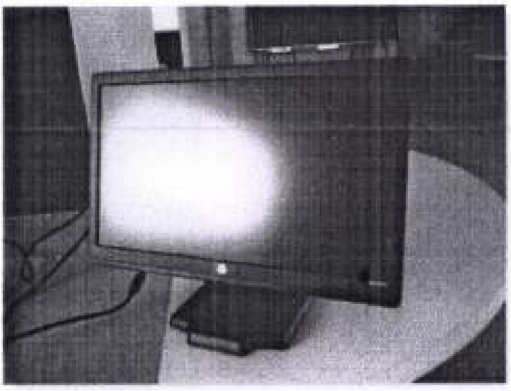


3) A wrong selection of the step size may result into Ping-Pong Effect. The step size is a very important factor in calculating the next value of the parameter. Since in Gradient Descent Method, the step size is fixed, so if it is chosen wrongly, it can update the value to a range out of the requirement and it may never reach to the global minima, instead it will remain fluctuating within the curve. This effect is called Ping-Pong Effect.



EXPERIMENTAL RESULTS:

An image is captured twice from the same capturing device with a minimum perspective change. The captured images are shown below:

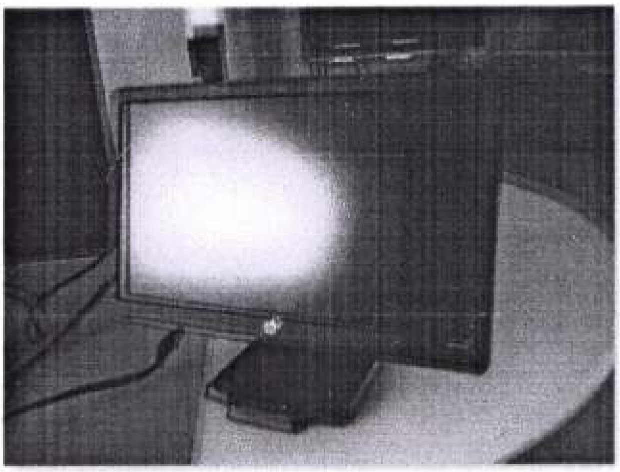


Though the two images look quite similar, but they are not perfectly same. Some change in the coordinates of the pixels are there.

The images are converted into gray image as we are doing our project on gray images only considering gray values ranging from 0 to 255.

Dimensions of the two images are same and since we are omitting the scaling parameter, so the two images are captured from the same distance from different viewpoints.

After fusion of the two images without performing image registration, we got the fused image like the below:



The two images are not aligned and hence two different images are distinguishable resulting like a motion blur occurrence of the image.

After performing the image registration using Gradient Descent Optimization Method, we have obtained a comparatively better image than the above one. The resultant image is shown below:

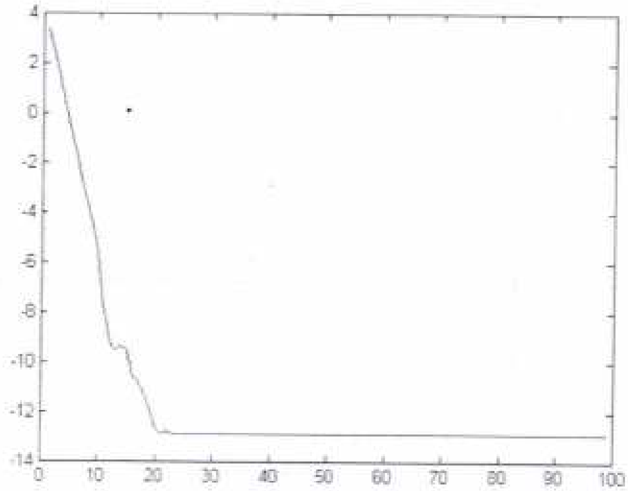
Since in this project, we are considering only two parameters i.e., Shifting and Rotation, so there are still some differences in both the images. But comparatively, we got a better resolution image.

The 2[nd] image contains some upward inclination and it is distinguishable in the edges of the table on which the monitor is placed as well as in edges of the blackboard present in the background.

The changes of the parameters of shifting and rotation in each iteration are plotted in a graph. These error values are calculated automatically during the shifting and rotation operation with the initial guess values.

The variation of x-axis **shifting** parameter with each iteration is shown in graph below:

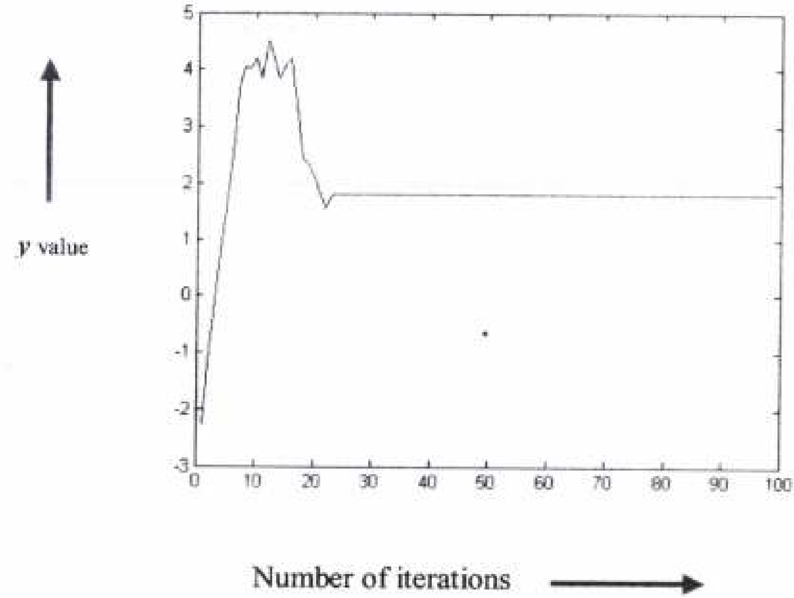
x value



$x = -13$

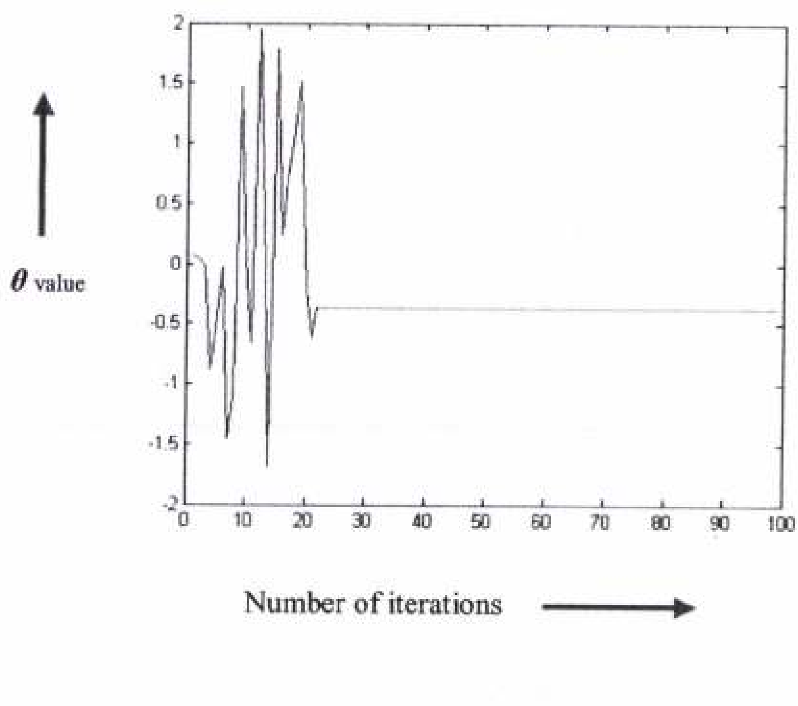
Number of iterations

The variation of ***y***-axis **shifting** parameter with each iteration is shown below:



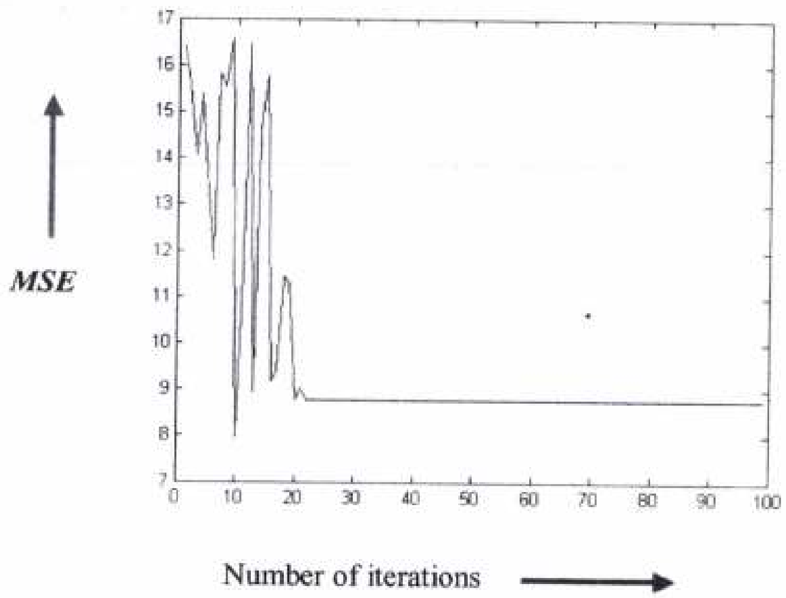
y = 2

The variation of the rotation parameter ***θ*** with every iteration has been shown below:



θ = -0.6389 rad

The variation of the ***MSE*** w.r.t. each iteration is shown below:



***MSE* = 8.712**

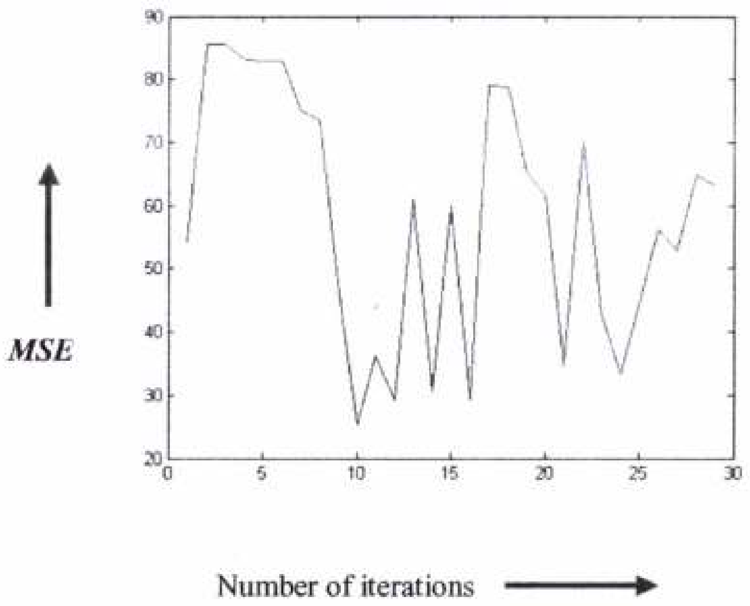
Drawbacks Observed:

1) **Time consumption:** To perform only 100 iterations, this method took around 15 minutes. But in practical applications, we have to perform at least 10,000 iterations which will be nearly impossible to implement with this method.

2) **Inefficiency:** The method does not provide the same alignment of the images in every iteration. Therefore, we have to perform the iterations for too many times to obtain the appropriate output.

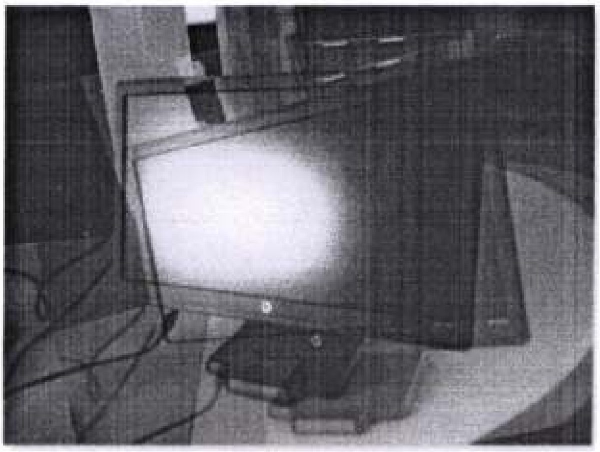
3) **Ping-Pong effect:** By increasing the step size, we observed the Ping-Pong effect during processing. The error starts to fluctuate within a specific range i.e., it does not reach the minima point, does not even decrease with the iterations.

   The error plot with the Ping-Pong effect is shown below:



MSE = 63.2563

The alignment of the images with the Ping-Pong effect is shown below:



CONCLUSION:

With the help of Image Mosaicing, we converted a low resolution image into a better resolution image with finer detail. The remaining perspective changes can be minimized by considering the parameters of perspective views. Image mosaicing is very helpful in image documentation i.e., taking the images section wise from closer viewpoint and then combining them to form the original image so that we can have a large coverage with closer look.

During this process of optimization, there are some problems that we have faced regarding the optimization method that obligated us from getting the efficient result. The problems that we have faced are actually the drawbacks of the proposed methodology. They are briefly discussed below:

FUTURE WORK:

1) We will try to overcome the problems faced in this method by using the following methods:

    a) By using some multiple ranges of values for guess value selection so that we can compare the values and get the best solution.
    b) By considering the rest of the parameters such as Scaling, Perspective changes, Affine transformation etc.
    c) By using 2^{nd} order methods like Conjugate Gradient Method and Particle Swarm Optimization Method.

2) We will try to mosaic more than two different sections of a large image using a 2^{nd} order optimization methods.

Benefits of using 2^{nd} order methods:

1) 2^{nd} order methods are faster than 1^{st} order methods since it calculates the parameters of transformation by computing the 2^{nd} order derivative of each value.
2) There is no confusion of local minima or global minima.
3) In only one single step, we can get the properly aligned image.
4) The step size can be made variable which is a great advantage of using 2^{nd} order methods. The step size can be varied automatically according to the gradient of the curve.
5) Very efficient for mosaicing more than two images.

Drawbacks of 2^{nd} order methods:

1) Calculation complexity is higher. The factor that will calculate the parameters of transformation is very difficult to generate and it takes too high density of variables.
2) Complexity increases proportionally with increase in the dimension or the number of pixels of the image.

REFERENCES:

Books:

1) **Fundamentals of Digital Image Processing,** by *Anil K. Jain*
2) **Digital Image Processing,** by *Gonzalez & Woods*

Journals:

1) **Gradient Descent Algorithm for Image Mosaicing,** by *Rahul Garg & Rohit Khandekar, IBM T. J. Research Centrė, New York.*

2) **Iterative Methods in Non quadratic Optimization,** by *Christopher Mehl, Margo Martinez, Elizabeth Pflum, and Robert Vanden Bosch.*

3) **Adaptive Stochastic Gradient Descent Optimization for Image Registration,** by *Stefan Klein.*

4) **Document on Image Mosaicing,** by *G Hemantha Kumar, D. S. Guru & P. Nagabhushan, Dept. of Studies in Computer Science, University of Mysore.*

THE END