MOSAICING OF TORN DOCUMENT IMAGES

Nagaraj B. Patil[1], V. M Viswanatha[2] & Sanjay Pande M. B[3]

[1]Dept. of Computer Science and Engg., Government Engg. College Raichur Karnataka, India
[2]Dept. of Electronics and Communication And Engg., Sln College of Engg. Raichur, India
[3]Dept. of Computer Science & Engg., Gsssiet for Women, Mysore, India
[1] and [2] are research scholars of Singhania University, Rajastan, India
Email: [1]Nagarajbpatil@yahoo.co.In , [2]vmviswanath@yahoo.co.in ,[3]rakroop99@gmail.com

ABSTRACT

Recovery of ripped-up documents is a problem that often arises in archival study and investigation science. The growing interest in featureless image mosaicing is because of its demand as it is still in incubation stage. The use of mosaicing torn document images is that it has reliable performance, fast document recovery capability and on-screen visual verification. Further, Most document reconstruction problems can be solved in two steps first, finding an initial set of matching fragment pairs, then resolving the ambiguity among these fragments to reconstruct the original document. The principle technique used here is the featureless image mosaicing technique. The proposed architecture for mosaicing of torn document images involves various stages. The first stage includes the scanning of the two fragments of the ripped up document such that the non-uniform sides face each other. The fragments include both uniform and non-uniform boundaries. We have also considered the non-uniform boundary values since, they had to be separated from the uniform boundary values. In present work, we have planned to design an algorithm to re-construct or mosaic a ripped-up document. We have corroborated the methodology by considering two fragments of the same torn/ripped up document.

Key words: *Torn Image, Featureless Image Mosaicing, Stitching Images.*

1. INTRODUCTION

An *image* is a matrix whose row and column indices determine the location of a point; the corresponding elements are called picture elements, pixels. Images are two-dimensional, such as a photograph, screen display. They may be *captured* by optical devices such as cameras, mirrors, lenses, telescope, microscopes etc. and natural objects or phenomena like human eye or water surface. *Image processing* is a form of signal processing for which the input is an image, such as photographs or frames of video; the output of this processing can be either an image or a set of characteristics or parameters related to the image. *Image Mosaicing* is the process of reconstructing or stitching a single, continuous image from a set of separate or overlapped sub-images. The various methods adopted for image mosaicing can be broadly classified into direct methods and feature based methods. Direct methods are found to be useful for mosaicing large overlapping regions, small translations and rotations. Feature based methods can usually handle small overlapping regions and in general tend to be more accurate but computationally intensive. Some of the basic problems in image mosaicing are Global alignment, Local adjustment, Automatic selection, Image blending, Auto exposure compensation.

A *Document* is a bonded physical or digital representation of information designed with the capacity to communicate. In prototypical usage, a *Document* is understood as a paper artifact containing information in the form of ink marks. Increasingly documents are also understood as digital artifacts. *Document mosaicing* is one of the most crucial steps for automatically reconstructing ripped-up documents is to find a globally consistent solution from the ambiguous candidate matches. The candidate disambiguation problem is formulated in a relaxation scheme in which the definition of compatibility between neighboring matches is proposed, and global consistency is defined as the global criterion. Initially, global match confidences are assigned to each of the candidate matches. After that, the overall local relationships among neighboring matches are evaluated by computing their global consistency. Then, these confidences are iteratively updated using the gradient projection method to maximize the criterion. This leads to a globally consistent solution and, thus, provides a sound document reconstruction. Recovery of ripped-up documents is a problem that often arises in archival study and investigation sciences such as forensic, crime detection. Documents may be ripped up due to aging, by hand or shredded by a machine. In any of these cases, the automatic or semiautomatic reconstruction of the original document would alleviate the manual effort, which is complex and time-consuming.

Further, Since document fragments often have similar contour segments, invalid matching occurs frequently that leads to intensively backtracking. To overcome this difficulty, we make use of a global approach to disambiguate the candidate matches. The rationale of approach is based on the fact that, although some incorrect matches may score higher than the correct one locally, the global score of the correct configuration will probably be much higher than that of an invalid one. The relationships among all candidate matches are exploited to search for a globally consistent solution for reconstructing the original document, which might make a method more efficient and robust local searching strategies. The reconstruction of ripped-up documents is somewhat similar to the problem of automatic assembly of jigsaw puzzles. In a jigsaw puzzle, it is assumed that there exists four corner points for the canonical jigsaw puzzle piece and the piece boundary curve can be separated into four edges at the four corner points. In such case of document fragments, however, there are no restrictions on the shapes of the fragments and corners are usually hard to identify. The act of ripping often produces irregularities in the fragment contours, which makes it difficult to get a perfect curve matching. Moreover, no prior knowledge about the original content exists for the document reconstruction (in many cases). The document reconstruction problem can also be regarded as a special case of the fragmented object reassembly problem occurring in archaeology, art restoration, forensics, and other disciplines. Since matching between two fragments usually occurs over only a fraction of their corresponding contours, partial curve matching is must. Many methods have been proposed to solve this problem, which can be roughly divided into two categories based on whether the fragment contours are sampled uniformly or not. One is *string matching*- based methods, in which fragment contours are represented by uniformly sampled points. The other is *feature-matching*-based methods, where fragment contours are usually represented by critical points or Polygonal approximations. To completely reconstruct the original object, a global technique is needed to eliminate the ambiguity resulting from local curve matching. However, there is surprisingly no much work is done on solving this global reconstruction problem.

2. LITERATURE SURVEY

Mosaicing is one of the techniques of Image processing which is useful for tiling digital images which generally is blending together of several arbitrarily shaped images to form one large radio-metrically balanced image with boundaries between the original images are not visional. It can be a special case of geometric correction where registration takes place in the existing image. The concept of Ground Control Points being collected from input image to be transformed or else if both images already have compatible geo-referencing, then an appropriate translation and scaling can be applied instead of polynomial transformation. An attempt to develop a package for mosaicing multiple images is presented by *Ramesh.*

Stitching blends the aligned images seamlessly. Image stitching is a technique to merge a sequence of images into one blended picture [1] Integration of plenty of small images to create one large Image. This can be even defined as the process of perspective warping of images [11]. *Kamesh et.al.,* proposes an image stitching technology wherein there is no restriction to the space between the two stitched images. A larger white space between the two images will yield the same result as that by two images with very less distance between the stitched images. Image alignment procedure establishes geometric correspondences among the images. The automatic construction of Image stitching is an active research area of research in the fields of photogrammetry, computer vision, and computer graphics [2].

Constructing a panoramic image from digital still images was proposed by *Naoki et.al.,* The method is fast and robust enough to process non-planar scenes with free camera motion which includes two techniques. One conventional method is a cylindrical panorama that covers a horizontal view for creating virtual environments. Using Image mosaicing technique to create vast scene images and panoramic images as been presented in *Gao and* Jia*;* In this method firstly, feature points are extracted using *Harris corner detector* to perform correlation. Then, filtering for matching points is presented by choosing pairs of correlation feature points with clustering algorithm before purifying them with RANSAC algorithm. But, directly use of RANSAC algorithm will result in the great waste of calculation because a lot of interest points are mismatched. To get a perfect stitched picture we can use blending images, there are some de-ghosting methods based on graph cuts and pyramid blending [6]. This paper proposes an efficient method using optimal route and luminance center-weighting algorithm.

An automatic mosaicing process for document images as presented in *Adrian and Hong.* It introduces a concept wherein if it is impossible to capture the entire image in one scan with the available equipment, a montage can be made from separately scanned pieces. Image placement and overlap is used to reject incorrect solutions. Fourier transforms can be used to speed up the registration of images. Although this is useful for template matching, the process requires human intervention to mask out regions not common to both images when used for image mosaicing. The concern is with assembling mosaics obtained using the same device, under similar conditions but with large shifts of the field of view across the document. The imaging capabilities of modern mobile phones have created an opportunity for more flexible document image acquisition.

In *Jari et.al* a full view of a document page such as a poster or a map which is usually too large to capture in detail by taking a single snapshot is taken up. The resolution can be increased by zooming in or moving closer and then taking part of the images. A high resolution image is finally achieved by stitching captured images together. The quality of the result is mainly affected by two important tasks namely Image Registration and Image Blending. A number of methods have been proposed for the special case of document image mosaicing. Most of these are developed for flatbed scanners where image stitching is easier because geometric transformation between images can be modeled using a 2-D similarity model and lighting conditions are fixed. The captured images are matched by applying image mosaicing with the help of the mouse movement. The other methods used are over-the-desk camera to take images of a document that is moved on the desk.

Liang et al [7] proposed a method for camera captured document images. It allows registration of images with an overlapping area as small as 10%with significant perspective distortion. They also apply sharpness based text component selection in order to obtain seamless blending. In this paper, a method for document image scanning with mobile phones is presented. During online scanning, motion estimation with low-resolution images is used to control the interaction between the user and the device. Images with coarse alignment information are downloaded to a desktop computer, where mosaicing of images can be constructed automatically using a feature-based alignment method.

Mosaicing is a common and popular method of effectively increasing the field of view of a camera, by allowing several views of a scene to be combined into a single view. *C*onstruction of accurate panoramic mosaics from multiple images with a rotating camera, or alternatively of a planar scene is presented in *Philip and Allan.* The novelty of the approach lies in *(i)* the transfer of photogrammetric bundle adjustment techniques to mosaicing; *(ii)* a new representation of image line measurements enabling the use of lines in camera self calibration, including computation of the radial and other non-linear distortion; and *(iii)* the application of the Variable State Dimension Filter (VSDF). To use the technique one must assume that either the camera is rotating in a stationary scene, or that the scene is approximately planar. In these cases there is no parallax induced by the motion, and so no truly 3D effects are involved, simplifying the task of combining views. Existing methods are based on either image-to-image warping [6,8], or corner feature location and matching [9].

There are also commercial mosaicing systems such as the RealViz "Stitcher" software. In all cases the aim is to recover the *homographies* that map images to each other, and hence allow the images to be transformed and combined in a single coordinate system. When seamless, high resolution mosaics are required it makes sense to consider the techniques that are used in full 3D structure-from-motion and photogrammetry, which have been developed over many years for high-accuracy metrology.

The mosaic representation associates a transformation matrix with each input image, rather than explicitly projecting all of the images onto a common surface**.** *Heung and Richard*. In particular, to construct a full view panorama in there paper, a *rotational mosaic* representation that associates a rotation matrix with each input image. A *patch-based alignment* algorithm is developed to quickly align two images. In order to reduce accumulated registration errors, globally aligning (*block adjustment*) the whole sequence of images will result in an optimally registered image mosaic. To compensate for small amounts of motion *parallax* introduced by translations of the camera and other unmodeled distortions, a local alignment (*deghosting*) technique is developed which warps each image based on the results of pair-wise local image registrations. By combining both global and local alignment, the quality of image mosaics can be significantly improved, thereby enabling the creation of full view panoramic mosaics with hand-held cameras.

Prabhakara and Mahidhar have presented three methods in mosaicing a sequence of multiple still images with some amount of overlapping between every two successive images. Firstly, featureless registration method which handles rotation and translation between the images using phase correlation even under blur and noise. Secondly an efficient method of stitching registered images which removes the redundancy of pasting pixels in the overlapped regions between the images with the help of a binary canvas. The third is a minimal blending approach to remove the seams in the mosaic and preserve the quality close to reality [10].

Image mosaicing method based on feature points and color transfer is proposed in *Ruwei and Xiaowei* Use Scale-invariant feature transform algorithm which search regions being limited around the overlapping regions, and probabilistic screening to eliminate the false matching points to improve the accuracy of matching. Further, lαβ color model is used to improve and adjust the brightness of the images to solve the problem of different illumination. Graph cut and multiresolution fusion techniques are used to cope with ghosting.

Although some authors have proposed a post-processing step to reduce the registration errors in these situations, there have not been attempts to compute the optimal solution, *i.e.*, the registrations leading to the panorama that best matches the entire set of partial views. *Bernardo a*nd *Pedro i*n there paper have discussed the problem of recovering, in an automatic way, a panoramic image, or a mosaic, from a set of uncelebrated partial views is been answered, *e.g.*, a set of video frames there model / algorithm would compute an efficient way to estimate the Maximum Likelihood of all the unknowns involved that is the parameters describing the alignment of all the images and the panorama itself was taken up.

One of the most crucial steps for automatically reconstructing ripped-up documents is to find a globally consistent solution from the ambiguous candidate matches. However, little work has been done so far to solve this problem in a

general computational framework without using application-specific features. *Liangjia et.al.,* presented a global approach for reconstructing ripped-up documents by first finding candidate matches from document fragments using curve matching and then disambiguating these candidates through a relaxation process to reconstruct the original document. The candidate disambiguation problem is formulated in a relaxation scheme in which the definition of compatibility between neighboring matches can be proposed which leads to a globally consistent solution, thus, providing a sound document reconstruction.
Image mosaicing is an effective means of constructing a single panoramic image from a series of snapshots taken in different viewing angles. However, in the case of congested traffic scenes with a cluttered environment including vehicles or pedestrians, there are severe difficulties in aligning a pair of snapshots. In such cases, some objects would be taken only in one of the image pair, thereby resulting in failure in stitching the pair of images. The three types of techniques for performing an image mosaicing: Homography estimation for determining geometrical relationships between the image pair, Expectation-Maximization algorithm for removing inconsistent overlapping region, and Graph Cuts for seamless stitching and background estimation is presented in *Takeaki and Wataru*.
Panoramic images have a broader field of view than standard images[3, 6]. Among these methods, panoramas can be created by mosaicing a series of images from one camera. The construction of such mosaics requires appropriate image registration to put them in a common coordinate system. Cylindrical panoramic mosaics can be created by aligning and stitching images from a series, captured by a camera rotating around its optical center. Existing methods compute a global transformation for the whole image starting from pixels in the overlapping region. This global transformation, due to local distortions often results in ghost problems in the overlapping region which have to be handled with some deghosting algorithm. *Nadege and Marie* proposed to directly use optical flow in order to find a pixel-wise registration in the overlapping region, reducing ghost problems.
Thus from the survey papers, we can confirm that most document reconstruction problems can be solved in two steps, that is, first, finding an initial set of matching fragment pairs as candidate matches, then resolving the ambiguity among these candidate matches to reconstruct the original document, which is the key approach. Further, document fragments can be represented by their contours and candidate matches obtained by local curve matching. Since local curve matching can produce many ambiguous matches, a search technique can be required to resolve these ambiguities for global reconstruction. The best first strategy is to select and merge the best matching pair to form a new fragment and this process is repeated until there is only one fragment left. However, such local search strategies may be inefficient in real. A global approach for reconstructing ripped-up documents is to first find candidate matches from document fragments using curve matching and then disambiguating these candidates through a relaxation process to reconstruct the original document.

3. IMPLEMENTATION

The proposed approach for the mosaicing of the torn document image has five stages which involves both hardware and software components. In first stage the ripped fragments of the document are scanned, scanned images are fed into the algorithm as input in the second stage. The third stage involves matching the fragments. In the fourth stage the matched fragments are translated and displayed in the fifth stage.

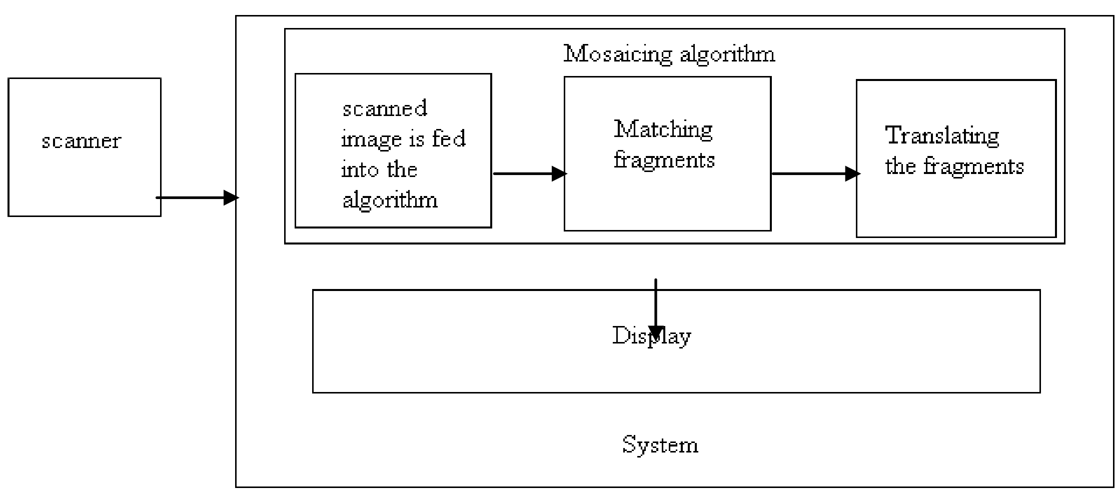


Figure 1: Proposed architecture for Mosaicing of torn document images.

4. DATA COLLECTION

The process of preparing and collecting data is done during data collection. Here we have used the scanner, HP scanjet 2400 with the following specification: Scan type: Flatbed; Optical scanning resolution: Up to 1200 dpi; Bit depth: 48-bit; Gray scale levels: 256; Scan size flatbed (maximum): 21.59 x 29.72 cm; Power consumption: 17 watts maximum.

The approach taken is presented in the form of algorithm. The fragments are scanned with the help of a scanner and taken as input to the algorithm. Initially the algorithm detects the boundary of each fragment and marks them explicitly. Later these boundary values are plotted with the help of a graph. The fragment contains both uniform and non-uniform boundaries. As we consider only the non-uniform boundary values, they must be separated from the uniform boundary values. Next the non-uniform boundary values have to be recursively matched. Once the boundary values are matched, the fragments are translated and the reconstructed document is displayed. If the boundary values are not matched, then the matching process is continued.

5. IMPLEMENTATION ALGORITHM

Step 1: Boundary Detection:

Stage 1.1.Read the image using imread function.

Stage 1.2 Convert the image into binary image

(Pixels with the value 0 are displayed as black; pixels with the value 1 are displayed as white).

Step 2: Plotting Boundary Values:

Stage 2.1 Choose a appropriate scale to plot boundary values on to a graph.

Stage 2.2 The boundary values are plotted using the function; plot(Y) plots the columns of Y versus their index if Y is a real number.

Stage 2.3 If Y is complex, plot(Y) is equivalent to plot(real(Y),imag(Y)).

(The scale considered for each plot is based on the width and length of the document fragments from the top-left corner of the scanned image).

Note: The extraction of the plotted boundary values and bifurcating these values as uniform boundary and non-uniform boundary values is very difficult because while scanning the fragments, even if placed with at most care on the scanner bed since, skew problem occurs.

Stage 3: Bounding box:

Stage 3.1 The scanned input image is first converted into a binary image using:

Stage 3.2 IM2BW( ); Produces binary images from indexed, intensity, or RGB images.

(The output binary image BW has values of 0 (black) for all pixels in the input image with luminance less than LEVEL and 1 (white) for all other pixels).

Stage 4: Crop image:

Stage 4.1 Working area that is selected within the bounding box

Stage 4.2 IMCROP ( ); imcrop crops an image to a specified rectangle.

(It displays the input image and waits for you to specify the crop rectangle with the mouse).

Stage 5: Boundary detection:

Stage 5.1 Detect the boundaries of the fragments.

Stage 5.2 BWBOUNDARIES(BW); It traces the exterior boundaries of objects, as well as boundaries of holes inside these objects, in the binary image BW.

(BW must be a binary image where nonzero pixels belong to an object and 0 pixels constitute the backgrounds).

Stage 6: Plotting points:

Stage 6.1 Two distinct points are manually plotted on each fragment.

Stage 6.2 GINPUT( ); It enables us to select points from the figure using the mouse for cursor positioning.

(Select n points from the current axes and returns the x- and y-coordinates in the column vectors x and y, respectively).

Stage 7: Translation:

Stage 7.1 Based on the reference points plotted fragments are translated for reconstruction.

Stage 7.2 Geometric image rearrangement is used here in order to translate the *fragments.*

Stage 7.3 effectively resize, examine image content and removing "less important" regions.

Stage 7.4 The pixel R(u, v) will be derived from the input pixel I(u + tx, v + ty).

(The optimal shift-map is defined as a graph labeling, where the nodes are the pixels of the output image, and each output pixel can be labeled by a shift (*tx*, *ty*). When an output pixel (*u*, *v*) should originate from location (*x*, *y*) in the input image, the appropriate shift gets zero energy while all other shifts get a very high energy).

Note: Object rearrangement is specified in two parts using the data term. One part forces pixels to appear in a new location using. The second part marks these pixels for removal from their original location using to avoid their duplication in the output.

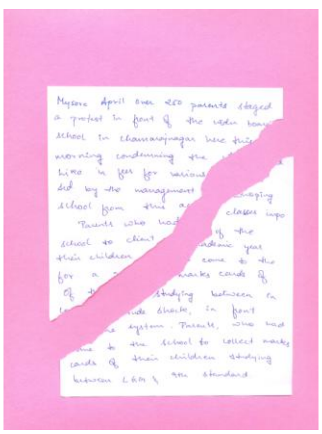
In image rearrangement pixels can be relocated by a large displacement, creating a possible computational complexity and by allowing many possible shifts as labels there can be a exponential explosion thus in order to reduce this complexity, the set of allowed shifts for each pixel will be included as local shifts around its original location.

Stage 7.5 Z = imadd (X,Y) adds each element in array X with corresponding element in array Y and returns the sum in corresponding element of the output array Z.

(X and Y are real, nonsparse numeric arrays with the same size and class, or Y is a scalar double. Z has the same size and class as X, unless X is logical, in which case Z is double).

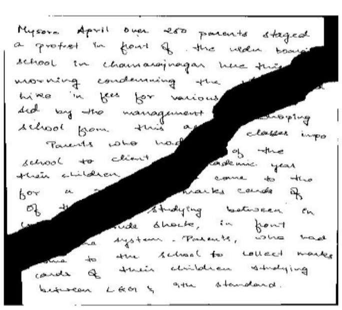
Note: If X and Y are integer arrays, elements in the output that exceed the range of the integer type are truncated, and fractional values are rounded using: *ROUND ( )* Y = *round* (X) rounds the elements of X to the nearest integers.

6. EXPERIMENTATION

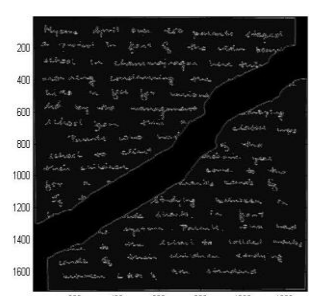


Stage 1: Input image

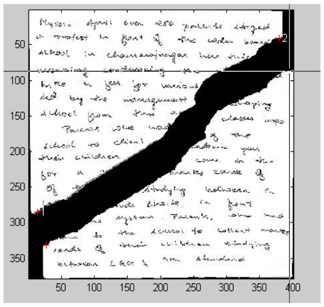
Stage 2: Bounding box applied



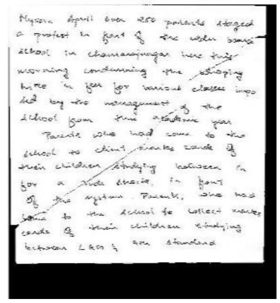
Stage 3: Cropped image



Stage 4: Boundary detected and plotted



Stage 5: Plotting points



Stage 6: Mosaiced torn image

7. DISCUSSION AND CONCLUSION

Document is an evidence of an event. They play a vital role in the day to day activities of all organizations. Agile document management techniques involves managing the documents manually by creating storage space which consumes more space and also increased cost of maintenance due to various factors like insect attack, ageing and careless manual handling caused damage. In order to retrieve the information lost due to the damage, the technique of mosaicing is employed where in much of work done so far in this field involves joining of the fragments using the overlapping method based on scanning each fragment of the ripped document individually.

In the present work of Mosaicing Torn Document Images, we have made an effort of stitching two fragments. The fragments here are scanned together in a single frame. During the process, the working area is reduced by applying a bounding box around the fragments. In the next stage, boundaries of these fragments are detected. In order to join the two fragments, two distinct points are plotted on each fragment manually and the fragments are joined based on these reference points.

The major problem, that we faced while working with the initial design was that, the extraction of the plotted boundary values and bifurcating these values as uniform boundary and non-uniform boundary values was very difficult because while scanning the fragments, even though they were placed with utmost care on the scanner bed, skew problem occurred and it was highly impossible to remove this skew manually. Removal of skew from a scanned document is still in research and developmental stage. Hence continuing with this procedure to build a mosaicing algorithm was quite tedious. So we took up another design to solve the problem. In the next design taken-up, first the input image was converted to a binary image. A bounding box was then applied only to the region of interest. Thus selected region was then cropped, eliminating the unwanted values/pixels from the source image thus identifying the boundaries of the fragments. Further, two points from each fragment were selected along the edges of the non-uniform sides which acted as the reference points for the stitching or mosaicing of the two fragments. The final stage was the stitching of the two fragments based on the values of the reference points. The re-constructed document image was displayed for visual verification. The major advantage of our work is that, the algorithm can be applied on all kinds of document irrespective of its content.

8. ACKNOWLEDGMENT

The authors like to thank Singhania University, Rajasthan and Vidya Vikas Institute of Technology, Mysore.

9. REFERENCES

[1]. Digital Image Processing, *Gonzalez R.C and Wintz P*, New York, Academic, 1977.

[2]. "Seamless Image stitching in the gradient domain", *Anat Levin, Assaf Zomet, Shmuel Peleg, and Yair Weiss*, in 8th European Conference On Computer Vision, May 2004, vol. 4, pp. 377-389.

[3]. "The Method for Mosaicing the Low Overlap Image Sequence", *Ruwei Luo, Xiaowei Chen*.2008 International Conference on Audio, Language and Image Processing, Shanghai,China,2008.

[4]. "Eliminating ghosting and exposure artifacts in image mosaics", *Matthew Uyttendaele, Ashley Eden, Richard Szeliski1*. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Hawaii, 2001, pp. 509-516.

[5]. An Improved Algorithm for Image Mosaic. *Fang Xianyong, Pan Zhigeng, Xu Dan*. Journal of computer-aided design & computer graphics, 2003, pp. 1362-1366.

[6]. "Panoramic image mosaicing", *Shum, H.Y. and Szeliski, R.*, Technical Report MSR-TR-97 23, Microsoft Research, 1997.

[7]. Camera-based document image mosaicing. *J. Liang, D. DeMenthon, and D. Doermann*. In *Proc. Int. Conf. on Pattern Recognition*, 2006, pp. 476–479.

[8]. True multi-image alignment and its application to mosaicing and lens distortion correction. *H. S. Sawnhey and R. Kumar*. In Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 1997, pp. 450–456.

[9]. Automated mosaicing with super-resolution zoom. *D. Capel and A. Zisserman*. In *Proc. Of the IEEE Conf. on Computer Vision and Pattern Recognition*, June 1998, pp. 885–891.

[10]. A survey of image registration techniques. *Lisa G. Brown.* ACM Computing Surveys, December 1992, pp. 325–376.

[11]. http://www.library.unsw.edu.au/~thesis/adt-ADFA/uploads/approved/adt-ADFA20051007.144609/public/03chapter2.pdf ,18th october 2007.