A Novel Approach for Document Image Mosaicing Using Wavelet Decomposition

P Shivakumara G Hemantha Kumar D S Guru P Nagabhushan

Abstract

There are some situations where it is not possible to capture or scan a large document with given imaging media such as scanner or Xerox machine as a single image in a single exposure because of their inherent limitations. This results in capturing or scanning a large document image into number of split components of a document. Hence, there is a need to mosaic the several split components into a single large document image.

In this work, we present a novel and simple approach for mosaicing the two split document images using wavelet decomposition to generate single and large document image. The proposed method uses the wavelet decomposition to speed up the mosaicing process by means of Multi Resolution Analysis (MRA). The pixel based and Column –Block matching procedures are used here to identify the overlapping region in the split images. The overlapping region is a common region which helps in obtaining mosaiced image from its split images. The proposed methods work based on assumption that the overlapping region is present at the right end of split image1 and at the left end of split image2

Keywords : Wavelet decomposition, Pixel value matching, Column-Block matching, Overlapping region, Document image mosaicing.

0. Introduction

The concept of image mosaicing is a phenomenon that occurs in the vision system of human beings because the human brain mosaics the split images of a large object that are automatically captured through eyes. Each eye functions as a camera lens. But, it is impossible to cover very large area with the help of an eye than a pair of eyes. Keeping this in mind, one can infer that two eyes capture the two split images of a large object but essentially with certain amount of overlap between the split images, which are later mosaiced into a single complete large image by deriving the knowledge from the Overlapping Region (OLR). Similarly, even in the real world the concept of mosaicing is essential because it may not be possible to capture a large document with a given camera or a Xerox machine in a single exposure. It has got to be captured as two or more split images due to the inherent limitations of the capturing medium. In such man made multiple camera exposures to cover a large image, the split images should necessarily contain OLR between the images, so that the stitching of two or more such split images into a single image becomes easier. Therefore, the proposed technique demands small amount of OLR in the split images such that the OLR is present at the right end of the first image and the left end of the second split image respectively.

There is a great demand for developing an algorithm for mosaicing the split images obtained by scanning of the large document part by part in order to restore original and large document image.

The structure of paper is as follows. The proposed methodologies are discussed in section 2 with suitable algorithms and mathematical models. In section 3, the comparative study is given. The experimental results are reported in section 4. Finally, the conclusion is given in section 5.

1. Document Image Mosaicing

Mosaicing is defined as the process of assembling the multiple components that are obtained either by scanning or capturing a large document part by part, in order to restore an original image without any duplication of portions. For example, a Xerox machine handles the documents of sizes A_4 (210mm X 297mm) and A_3 (297mm X 420mm). But the document of sizes A_2 (420mm X 594mm) such as a full newspaper cannot be scanned in a single stroke because of its inherent limitation. Hence, bigger sized documents such as newspapers have got to be split into number of smaller documents of A_4 or A_3 dimensions with little overlap between the split images.

Several researchers addressed the methods for obtaining the large image from its split images. (Schutte and Vossepoe, 1995) described the usage of flat bed scanner to capture large utility map. The method selects the control points in different utility maps to find the displacement required for shifting from one map to the next. These control points are found from pair of edges common to both the maps. However, the process requires human intervention to mask out the region not common to both the split images in image mosaicing.

The researchers (Zappala et al., 1997; Peleg, 1997) have worked on Document Image Mosaicing (DIM). A feature-based approach through estimation of the motion from point correspondence is proposed. They have exploited domain knowledge, instead of using generic corner features, to extract a more organized set of features. The exhaustive search adopted was computationally expensive because of the rotation of an image employed during matching. However, the approaches are limited to only text documents and are prone to failure in case of general documents containing pictures. But in practice, a typical document contains both text and pictures.

An automatic mosaicing process for split document images containing both texts and pictures, based on correlation technique is proposed by (Whichello and Yan, 1997). Here correlation technique was used to find the position of the best match in the split images. However, accuracy is lost at the edges of the images. Moreover, the correlation of two images of practical size is computationally very expensive. In order to find a solution, additional constraints like *a priori* knowledge were introduced. Here, the sequence in which the images were captured and their placement (generally, referred as image sequencing) is known. Template matching procedure was used to search OLRs, present in the split document images. Usually, template-matching procedure is a time consuming method. In addition, this approach assumes that the printed text lies on straight and horizontal baselines, which is not always possible in many of the pragmatic applications.

2. Proposed Methodology

The authors of this paper have proposed (Shivakumara et al, 2001) a new technique to tackle the above-mentioned problems. The proposed technique works for any type of document without considering the nature of the content present in the document to produce a complete large document image without having *a priori* knowledge about the order of image sequence. The proposed technique demands at least one pixel wide (1-2%) OLR in the split images. The OLR is present at the right and the left ends of the first and the second split images respectively. The technique is based on PMA (Pattern Matching Approach). A PMA is employed to determine the OLR in the split images of a large document image. The order of image sequence is obtained by considering the split images in all sixteen possible ways of matching sequence exist between them, as each image is associated with four faces. Sixteen possible ways of matching sequence exist if we consider the perfect square shape images. Out of sixteen possible matches, some of them are accepted as possible mosaiced images based on the overlapping region existing in the split images. Subsequently, the original complete image of the document is obtained by mosaicing the split images without any duplication of portions in the mosaiced image. The presented technique requires 16 O(n2) + 16n comparisons search time for finding the right sequence for two split images under the worst case, where n is the length of String of Column Sums (SCS). However, this method is time consuming method.

We propose two methods one is simple and pixel based and Column-Block matching procedure to identify the overlapping region in the split images to produce a single large document image. The two methods work based on wavelet decomposition. The proposed method is based on assumption that the overlapping region is present at the right end of split image1 and at the left end of split image2. The following sub sections explain the wavelet decomposition in detail, simple algorithm for mosaicing and column-block matching procedure for mosaicing.

2.1 Wavelet Decomposition

The wavelet transform of a 2D image f (x, y) is defined as the correlation between the image and a family of wavelet function { $ø_{s,t}$ (x, y)}: W_f (s , t; x, y) = f (x, y) * $ø_{s,t}$ (x, y) .Wavelets are generated from a mother wavelet function as follows: Ø s,t (x, y) = 1/s ø (x – t_x/s, y – t_y/s) Where s is the scale parameter, and (t_x, t_y) the translation parameters in the x axis and y axis. in most practical applications, one never explicitly calculates the mother wavelet. The pyramid-structured wavelet decomposition operation (mallt, 1989) produces four subimages f_{LL} (x, y), f_{LH} (x, y), f_{HL} (x, y) and f_{HH} (x, y) for one level of decomposition. f_{LL} (x, y) is a smooth sub image, which represents the coarse approximation of the image. f_{LH} (x, y), f_{HL} (x, y) and f_{HH} (x, y) are detail subimages, which represent the horizontal, vertical, and diagonal directions of the image, respectively. The 2D pyramid algorithm can iterate on the smooth subimage f_{LL} (x, y) to obtain four coefficient matrices in the decomposition level. Fig.1 depicts one stage in multi-resolution pyramid decomposition of an image (Tsai and Chiang, 2002).

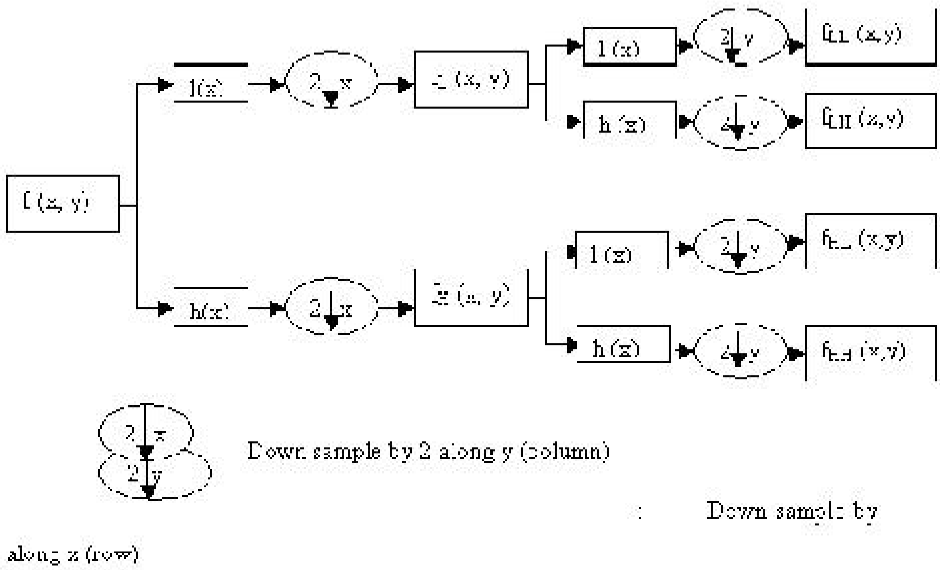


Fig.1 One stage in multi-resolution image decomposition

The reduction factor of an image size is given by 4j, where j is the number of decomposition levels. For level-2 decomposition, the size of an original image can be reduced by a factor of 16. This results in great computational saving in the matching process. In practice, the effective size of the smallest subimages in the decomposition should be used as a stopping criterion for determining the maximum number of decomposition levels. If the decomposed subimage has an over-down sampling size, the locations and wavelet coefficient values of object features may change dramatically from one sample to sample, and generate a false match accordingly. The experimental results on a variety of test images showed that the smallest size of decomposed template subimage should be larger than 20X20 pixels. The matching process can be performed either on the decomposed smooth subimage or on the decomposed detail subimages at a lower multi-resolution level. In this study, we consider decomposed smooth subimage for mosaicing purpose.

2.2 Simple Pixel based Algorithm for Mosaicing

This section presents a simple approach to generate mosaiced image from its split images containing wavelet coefficients at one level. The wavelet coefficients are obtained by the above section (section 2.1). Let S_1 and S_2 be the split image1 and split image2. The method compares the values of pixels of First column (F_c) of S1 with First column of S_2. If match occurs then the pointer (i) pointing column of S_1 moves to next pixel value in the same column. Similarly, in S_2 also the pointer (j) moves to next pixel value of same column. If it doesn't match then the pointer moves to next column mean while the pointer of S_2 comes back to first column. This procedure repeated till all the columns match continuously since once overlapping starts it ends at the end of image in S_1. If whole column match (CM) in S_1 and S_2 then both the pointers go to next column. The number of matching columns decides the overlapping region in the split image. Finally, the algorithm terminates when the pointer of S_1 reaches n where n is the number of column in S_1 without overlapping region. The algorithm also terminates if pointer of S_1 reaches n with overlapping region.

Algorithm: Simple

Input: Split image1 (S_1) and Split image2 (S_2)
Output: Mosaiced image
Method:

```
    Step1: For each Column (C) of S1 and S2
                For each pixel value of C of S1 and S2
                If (Pi = Pj) in Ci of S1 and S2 (Where Pi is the pixel values of column of S1 and
                Pj is the pixel values of column of  S2)
                        i = i + 1 and j = j + 1 (i is pointing to Column values of S1 and j
                         is pointing to column values of S2)
                        if (Pi = Pj) and (i = Ec of S1) (here Ec is represents the end of column in )
                                CM = 1 (if the whole column matches in S1 and S2)
                        Else exit from the for loop
                else exit from for loop
                Step2: If (CM = 1)
                        Cs1 = Cs1 + 1 and Cs2 = Cs2 +1
                        Else  Cs1 = Cs1 + 1 in S1 and Cs2 in S2 comes back to Fc
                ( Cs1 is the pointer pointing to Column of S1 and Cs2 is the pointer pointing to
                column of S2)
    For end
    Step3:      If (CM = 1) and (Cs1 = n)  then OLR = 1
                Else OLR = 0
                If (OLR = 1) Mosaic the split images
                Else algorithm terminates with out overlapping region
    Method ends
Algorithm simple ends
```

2.3 Algorithm Column-Block Matching Procedure for Mosaicing

This section presents an algorithm Column-Block matching procedure for mosaicing of two split images containing the wavelet coefficients to produce single large document image. Let S_1 and S_2 be the given two split images containing local FT coefficients obtained by the above algorithm (section 2.1) The algorithm begins by matching the pixel values of F_c (First column) of S_1 with F_c of S_2. IF match is found then it goes to next pixel values of corresponding columns of S_1 and S_2. After finding whole column match (CM) the algorithm considers rest of the portion in the split images as a block from next to CM to end of S_1. Similarly in S_2 also. Next the method computes total sum of the values of pixels in both the blocks of S_1 and S_2. If sums are match then that portion is considered as actual overlapping region in the split images. If the pixel values in the column or sums do not match then the pointer C_p pointing to S_1 moves to next column mean while the pointer C_p pointing to S_2 comes back to F_c. This is because of assumption that the overlapping region is present at the ends of the split images. That means the overlapping region in S_1 begins at middle column and in S_2 the overlapping region begins from first column of S_2. The algorithm terminates when C_p of S_1 reaches n where n is the end of column of S_1 without overlapping region. The algorithm also terminates if the overlapping region is found in the split images.

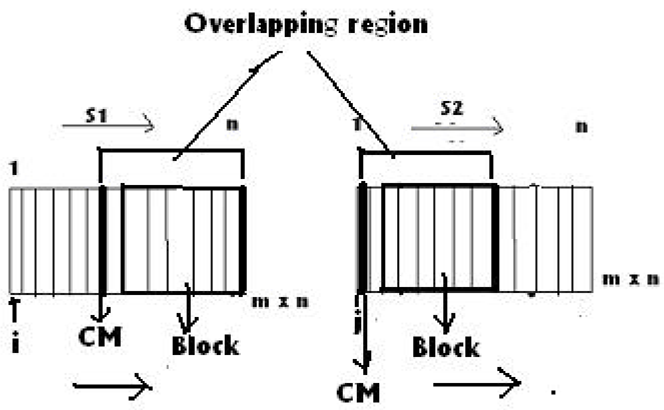


Fig. 1. The method to finds overlapping region in the split images

In Fig. 1, the Column Match (CM) denotes the matching column in the split images and Blocks denotes rest of the overlapping region in the split images. i and j are the pointers pointing to split image 1(S_1) and split image 2 (S_2). The actual overlapping region is represented by both CM and Blocks of split images.

Algorithm for CB
Input: S_1 and S_2 containing local FT coefficients
Output: Mosaiced image
Method:

Step1: For each Column (C) of S_1 and S_2
  For each pixel value of C of S_1 and S_2
    If ($P_i = P_j$) in C_i of S_1 and S_2 (Where P_i is the pixel values of column of S_1 and P_j is the pixel values of column of S_2)
      i = i + 1 and j = j + 1 (i is pointing to Column values of S_1 and j is pointing to column values of S_2)
      if ($P_i = P_j$) and (i = E_c of S_1) (here E_c is represents the end of column in)
        CM = 1(if the whole column matches in S_1 and S_2)

Else exit from the for loop
else exit from for loop
for end
Step2: If (CM = 1) then B_1 = N – CM + 1th = W in S_1
B_2 = CM + 1th to W in S_2 (B_1 represents the Block of S_1,
N is number of column in the S_1, W is the width of Block of S_1 and B_2 represents the Block of S_2)
Else $C_p = C_p + 1$ in S_1 and C_p in S_2 comes back to F_c (C_p is the pointer pointing to Column of S_1 and S_2)
For end
Step3: For B_1 of S_1

$Sum_1 = \sum_{p=1}^{N} \sum_{q=1}^{M} B_{pq}$ (Where p and q are the pointers of Block and N is the number of rows in Block and M is the number of column in the Block)
For B_2 of S_2

$Sum_2 = \sum_{p=1}^{N} \sum_{q=1}^{M} B_{pq}$

Step4: If (Sum_1 = Sum_2) then OLR = 1 (overlapping region is found)
Else OLR = 0
If (OLR = 1) Mosaic the split images
Else if (i = n) algorithm terminates with overlapping region is not found
Method ends
Algorithm ends

3. Comparative Study

In this section, we present comparative study of two methods by considering time for obtaining wavelet coefficients, time for matching area is to be found and number of comparisons with respect to levels for particular data set.

Table 1. Comparative study of Simple and Column-Block method with respect number of comparisons and its time in second

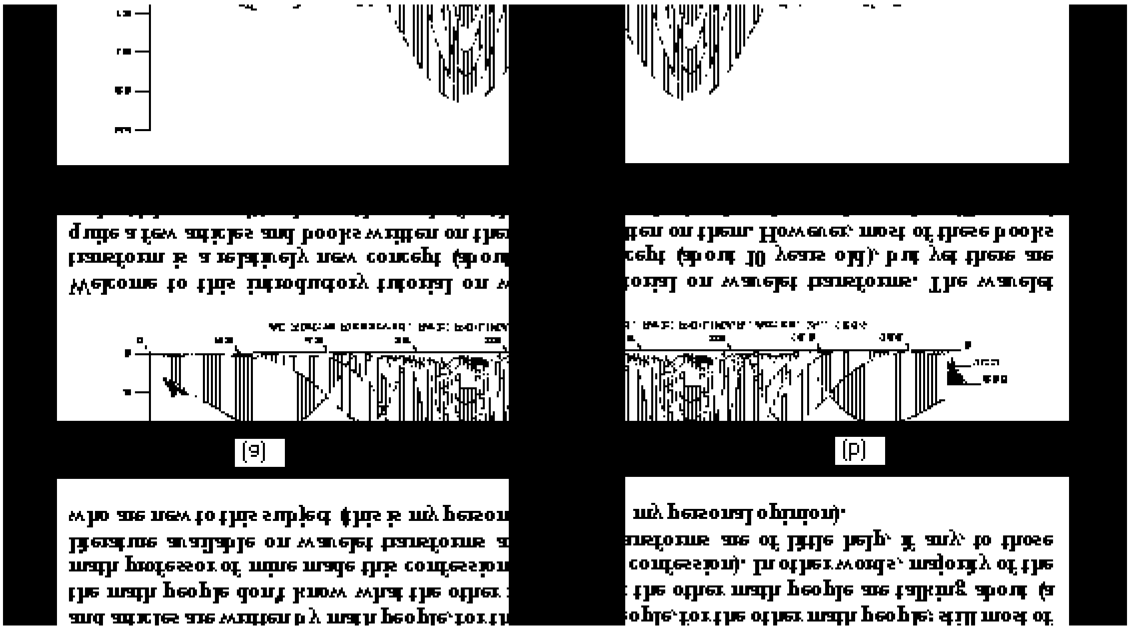
Levels	Column-Block method			Simple (pixel based)		
	TFW	TC	No.C	TFW	TC	No.C
128X128(I)	0.49 sec	0.05 sec	27	0.55 sec	0.05 sec	87
64X64(II)	0.60 sec	0.05 sec	13	0.87 sec	0.05 sec	43

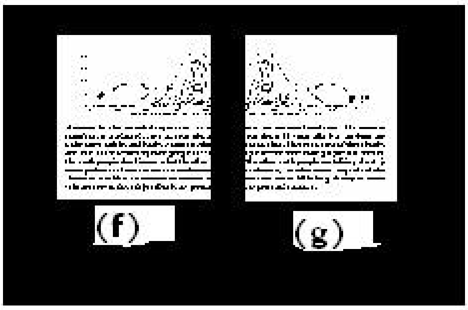
From the above table, it is observed that the column-block matching procedure takes very less number of comparisons compared to simple method in both the levels. In the above table TFW denotes time for obtaining wavelet coefficients at first level, TC means time for comparisons i.e time required for getting a matching area in the split images and No.C means number of comparisons required to identify the actual overlapping region in the split images. With this we conclude that the column-block matching procedure is better than simple method in all the way. This is because in column-block method after whole column matches it requires one match to decide overlapping region whereas in simple method every values in the columns should be compare to decide overlapping region.

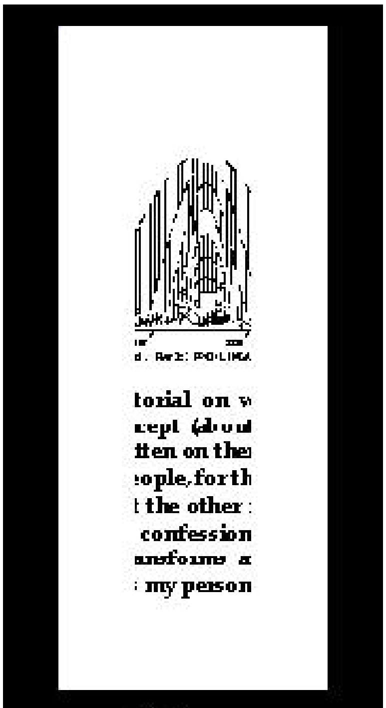
4. Experimental Results

In this section, we present experimental results based on proposed methodology. We have conducted several experimental results out of them a few of are presented here. For both the methods we get same results. Therefore, we have given only one data set for both the algorithms. The experimental results showed that the proposed method work for any type of documents. In the following Examples Figs. (a) and (b) are the input images, Figs. (f) and (g) are the results of wavelet decomposition at one level, Fig. (c) is the overlapped images and Fig. (e) is the mosaiced image in wavelet domain and Fig. (d) is the reconstructed and original mosaiced image. For sample we have given wavelet decomposition effect for only one data set. In all examples, the proposed methods use the reduce sized split images. After finding control points of overlapping region the methods reconstruct the original image (output image) using inverse wavelet transform.

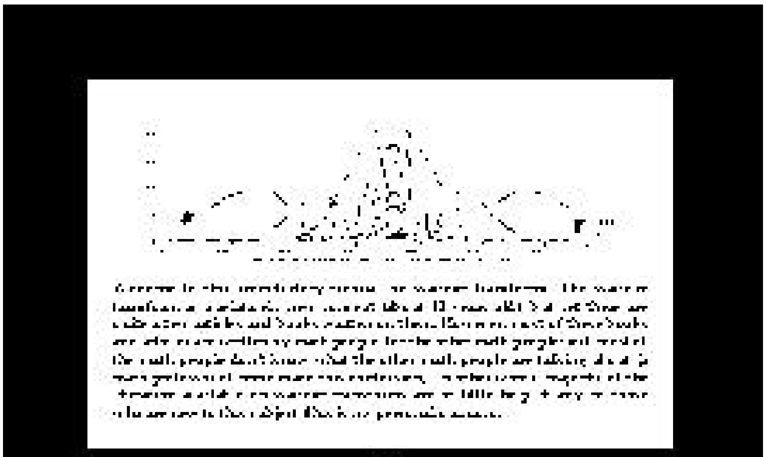
Example1: Here, the split images contain text with graph



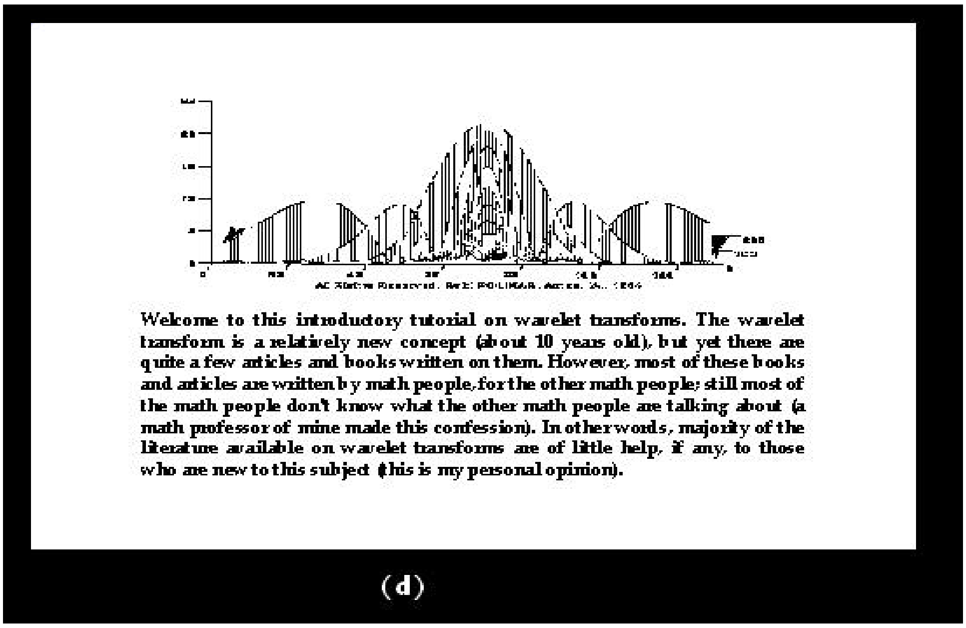



torial on v
cept (abou
tten on the
eople, forth
t the other
confession
ansforme a
my person

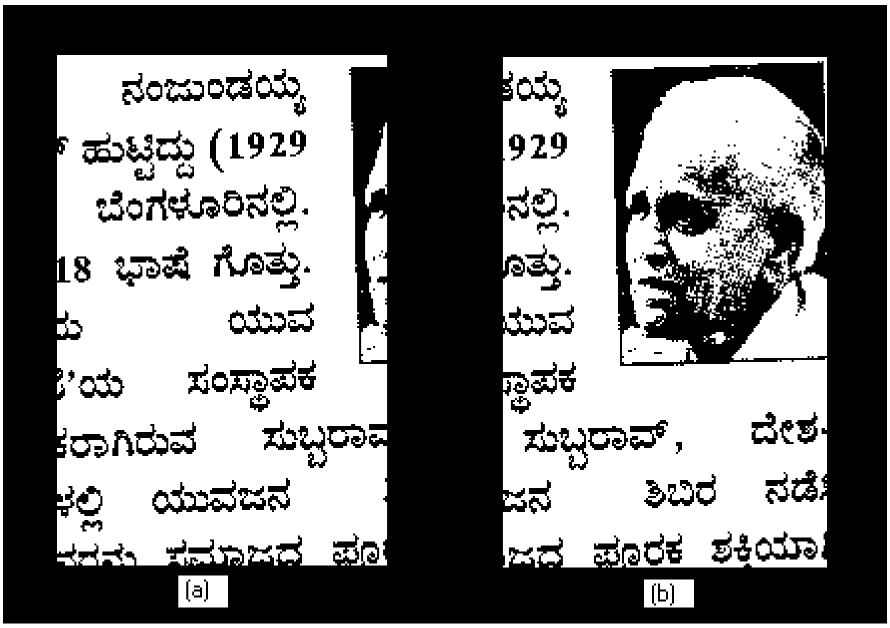
(c)

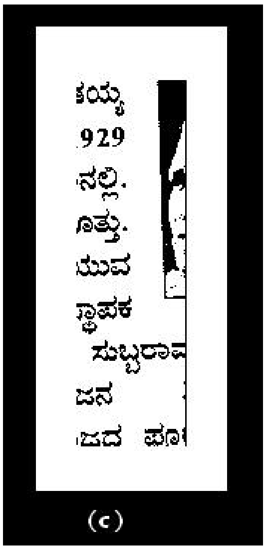


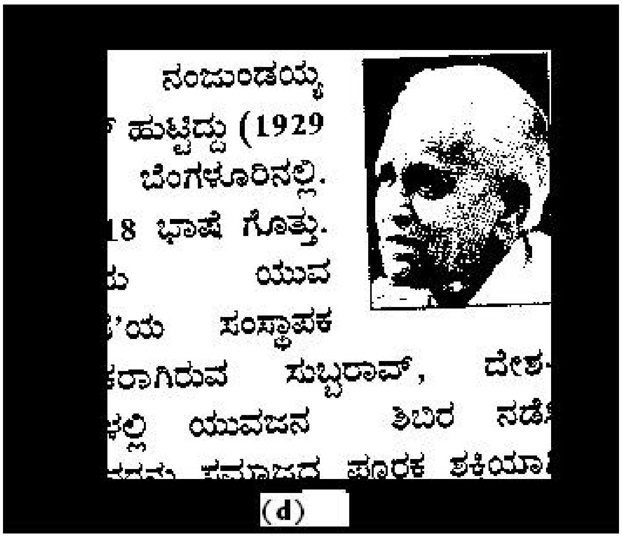
(e)



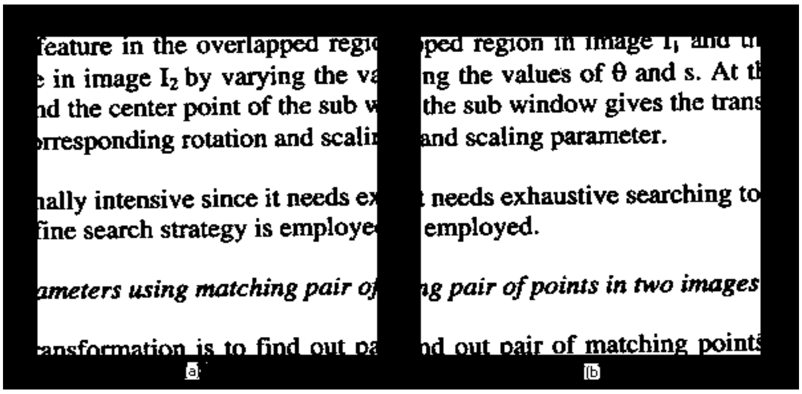
Example 2: Here, the split images contain Kannada with picture

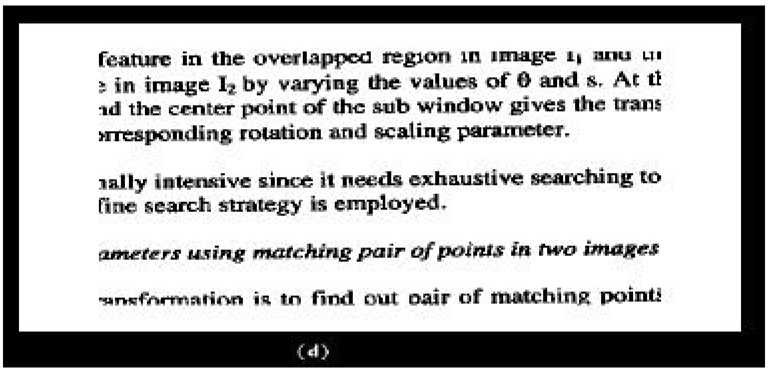






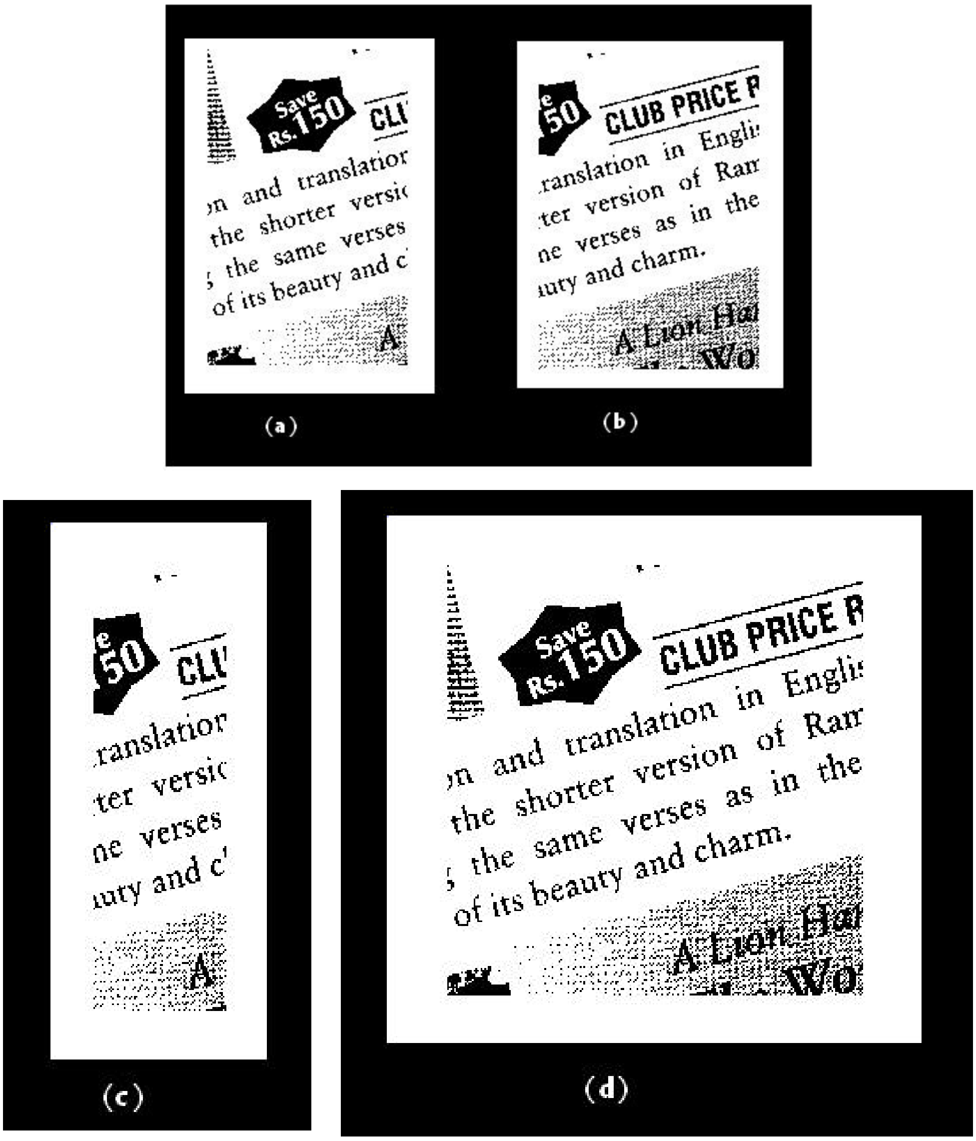
Example 3: Here, the split images contain only English text



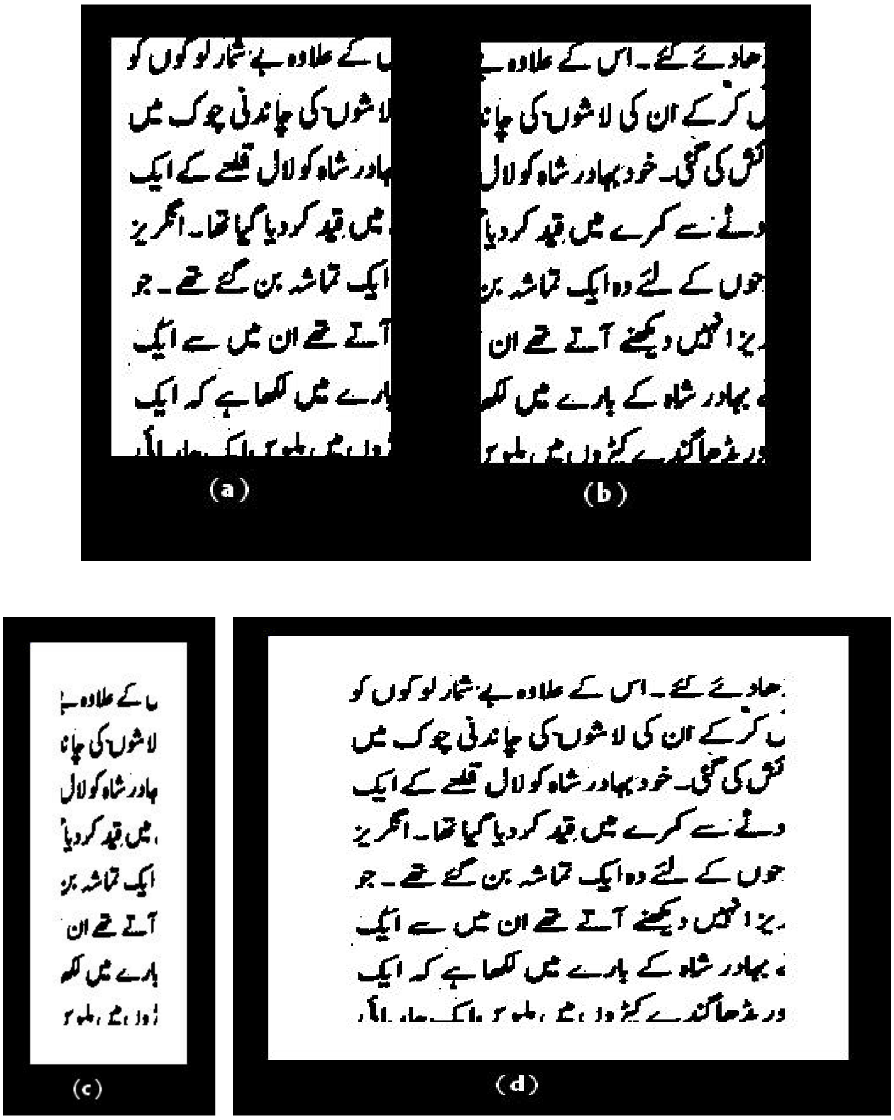


Example 4: Here, the split images contain Malayalam language

Example 5: Here, the split images contain English text with pictures



Example 6: Here, the split images contain the Urdu language



6. Conclusion

The wavelet decomposition based methods are presented in this paper. The comparative study of two methods is also given. The proposed methods take very less computational burden since the methods involve the multi resolution analysis property of wavelet. The experimental results and comparative study showed that the column-block matching procedure is better method in solving real world applications compared to simple method. The proposed methods assume that the overlapping region is present at the right end of the split image 1 and at the left end of the split image 2. The proposed methods fail when the split images are rotated, scaled and skewed differently. In addition, the methods fail to mosaic unequal sized split images.

7. Acknowledgement

The authors acknowledge the support extended by Dr. D.S Ramakrishna, Principal, APS College of Engineering, Somanahalli, Bangalore – 82 and Mr. Bhavani Shankar Hiremath

8. References

1. Shivakumara P, Guru D. S, Hemantha Kumar G and Nagabhushan P, *Document Image Mosaicing: A* Novel Technique Based on Pattern Matching Approach. Proceedings of the National Conference on Recent Trends in Advanced Computing (NCRTAC-2001), Tamil Nadu, Feb 9-10, 2001, pp 01-08.
2. Shivakumara Guru D. S, Hemantha Kumar G and Nagabhushan P, Pattern Matching Approach based Image Sequencing useful for Document Image Mosaicing. Proceedings of the National Conference on Document Analysis and Recognition (NCDAR-2001), Mandya, Karnataka, July 13-14, 2001.
3. Shivakumara Guru D. S, Hemantha Kumar G and Nagabhushan P, Mosaicing of Color Documents: A Technique based on Pattern Matching Approach. Proceedings of National Conference on NCCIT, Kilakarai, Tamilnadu, 24th and 25th September, 2001, pp 69-74.
4. Shivakumara Guru D. S, Hemantha Kumar G and Nagabhushan P, Mosaicing of Scrolled Split Images Based on Pattern Matching Approach. Proceedings of Third National Conference on Recent Trends in Advanced Computing (NCRTAC – 2002), Tamil Nadu, Feb 13-15, 200
5. A drian Philip Whichello and Hong Yan Document Image Mosaicing, Imaging Science and Engineering laboratory, Department of Electrical Engineering, University of Sydney, NSW 2006,1997.
6. Shmnel Peleg, Andrew Gee, Haifa Research Laboratory, Virtual Cameras usingImage mosaicing, Herbrew University, October 1997
7. Zappala A.R, Gee A.H and Taylor M.J Document Mosaicing. In proceedings of the British Machine Vision Conference, volume2, pages 600-609, Colchester, 1997
8. Tsai and Chiang, Rotation-invariant pattern matching using wavelet decomposition, Pattern Recognition Letters, 23, pp 191-201, 2002.

About Authors

Shivakumara P, obtained B.Sc, M.Sc. and M.Sc. Technology by research in Computer Science from the University of Mysore in the year 1996, 1999 and 2001 and his B.Ed. degree from Bangalore University in the year 1997. He has submitted his Ph.D dissertation to the University of Mysore for the award of Doctoral Degree. He has participated in 13 National/ International Conferences/Workshops. He has published around 60 papers in various journal/conferences on Document image Mosaicing, Skew Detection, Character Recognition and Automatic Face Recognition. His research focuses on Pattern Recognition and Image Processing in general, Document Analysis and Document Image Mosaicing in particular. Currently he is working as Assistant Professor in APS College of Engineering, Bangalore.

G. Hemantha, obtained M.Sc and P.h. D in computer Science from the University of Mysore in the year 1988 and 1998. He has published around 80 research papers in various journals/conference proceedings. Currently he is guiding 6 P.h.D candidates and 4 MSc Technology by research candidates. He has delivered talk at many Universities in India and abroad. His research focuses on Pattern Recognition and Image Processing, Document Analysis, Speech Processing, Computer Network and Simulation. He is currently the Reader and Chairman, Department of Studies in Computer Science, University of Mysore.

D. S. Guru, obtained B.Sc. degree, and M.Sc degree in Computer Science from the University of Mysore in 1991 and 1993 respectively. The same University awarded him the Ph.D degree for his work in the field of object recognition in the year 2000. Currently he is the senior faculty member in the Department of Computer Science in University of Mysore. He holds rank positions both at B.Sc and M.Sc levels. He has been identified as reviewer for journals and conference proceedings. He is currently guiding 6 Ph.D and 4 M.Sc Technology (by research) scholars. He has authored about 60 research papers in international/national journals and conference proceedings. His areas of research include Image Processing, Pattern Recognition, Advanced Software Engineering, Data mining, Image retrieval, Object Recognition, Digital Image Processing.

P. Nagabhushan, F. I. E obtained B.E from University of Mysore, M.Tech from Birla Institute of technology and then Ph.D from University of Mysore in the years 1980, 1983 and 1988. He is currently the professor, Department of Studies in Computer Science, University of Mysore. He was the dean of Faculty of Science and Technology, University of Mysore and is working as an advisor to many other Universities in India and abroad. He has authored about 230 research papers. Currently, he is guiding 7 Ph.D and 4 M.Sc Technology (by research) candidates. He has been a visiting Professor at a few abroad Universities. His areas of research include Pattern Recognition, Image Processing, Dimensionality Reduction, Data Mining, Document Analysis, Simulation.