International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 07, July-2015

Mosaicing of Text Contents from Consecutive Frames in Pedestal Shot Videos

Nagabhushan. P
Department of Studies in Computer Science
University of Mysore
Mysore, India

Vimuktha Evangeleen Jathanna
Department of Studies in Computer Science
University of Mysore
Mysore, India

***Abstract*— Mosaicing of frame contents from video sequences has gained growing interest in the vision community. The applications of video mosaics include panoramic creation for forming one single image that increases visual field of view or mosaicing of textual contents for improving the quality of texts by combining multiple frames for the purpose of text localization, extraction and recognition. In this paper we propose an approach to mosaic the text contents from a pedestal shot video. The paper utilizes the SIFT – match algorithm to find the matches between frames as it is invariant to scale, rotation and geometric distortions like blurring / resampling of local image orientation planes. In order to improve the time efficiency we are proposing horizontal strip based matching technique. The stitching process includes estimation of homography using RANSAC and blending them with a transformation function.**

Keywords*— *Horizontal strips, Pedestal Shot Video, Pedding Shot, SIFT, RANSAC, Mosaicing, Homography

I. INTRODUCTION

The present day smartphone, PDA, IPAD contain integrated digital cameras that have the ability of capturing high resolution images. The trend of using these cameras has also facilitated to capture wide variety of images that forms a huge dataset for research projects and commercial applications. Especially the videos captured through these cameras pose different challenges while processing them. The major issue in processing these videos includes redundancy in spatial content, camera motion, blurring, illuminations and distortions. Also, as the videos are captured using hand held devices, they pose challenges in capturing methods. E.g Some document videos imitate the reading pattern of human beings, for these type we need o capture the document source using track movement i.e. left to right scan. In some special cases where the capturing is dependent on the language of use we adapt reverse scan or right to left scan. In some cases where we need to capture multiple columns in vertical patches as in case of journals or bills or time tables we use pedestal shot or vertical top to bottom movement.

The paper describes the approach of mosaicing document videos containing document source captured in vertical motion also called Pedding Motion or Boom Scan [17]. Mosaicing of document videos refers to the process of stitching successive video frames containing more than 50% of overlapping textual information into one large high resolution composite.

Also, we propose a method to mosaic the successive frame information with previous frame information by tracking only the spatial transition in the rows by decomposing the frames into horizontal image strips and matching them using SIFT – Scale Invariant Feature Transform descriptors. We have utilized the available homography estimation algorithm using RANSAC. The estimated homography was then used to transform the image for stitching using simple translation function.

The remainder of this paper is structured as follows: in section II we have carried out a brief survey on the available literature on mosaicing. In section III we present the motivation for developing the horizontal strip based mosaicing approach. In section IV presents the proposed design, detailed development of the algorithm for the horizontal strip based mosaicing. Section V describes the experimental data set, implementation, results and complexity analysis of the approach is discussed. In section VI the conclusion drawn from this paper is described.

II. LITERATURE SURVEY

There are several ongoing researches on camera based document mosaicing and video mosaicing. Few of the methods depicted in the literature are discussed below.

Lowe et al’s SIFT -Scale Invariant Feature Transform was used in [1] to reconstruct panorama of images. The SIFT features were extracted from video frames and were matched using k - nearest neighbourhood. The homographies were estimated between the matched pairs using RANSAC and was verified by probabilistic model. Then each of the connected components derived from the graph search method was subjected to bundle adjustment with joint camera parameters and was subjected to multi band blending to provide panoramic view which was invariant of scaling, rotation and geometric distortions

Nagabhushan P et al [2] proposed a vertical strip based mosaicing technique based on SIFT for track movement videos. The reference frame were matched initially with the other adjacent frames and then the vertical strips were created. The false matches from SIFT were fitted using RANSAC and a simple transformation blending function was proposed.

Liang et al [3] proposed a mosaicing technique for camera captured images. Overlapping regions within a small area with perspective distortions were used for image registration.

The seamless blending was obtained by sharpness based text component selection. Mosaicing was done based on feature based alignment.

Hemanth Kumar et.al [4] proposed a novel approach for mosaicing split images based on simple pixel correspondence and Euclidian distance

Hemanth Kumar et.al in [5] proposed a technique to mosaic the two split images of a large document based on matching the sum of values of pixels of window in the split images. The method compared the sum of values of pixels of window in split images to identify overlapping region in the split images.

Lian et al. in [6] proposed a mosaicing method for camera captured document images. the method stitches document images captured from arbitrary angles using a digital camera. In this method, perspective distortions of document images are removed based on vanishing points estimated from text line direction and vertical character stroke direction. Then feature points of fronto-parallel document images are extracted and matched using PCA-SIFT. [11]

Tomohiro et.al. in [7] proposed a mosaicing method for camera-captured document images. by calculating the corresponding feature points using an image retrieval method since it was invariant to perspective distortion. Feature points were matched without compensating perspective distortion.. Document images were aligned using a perspective transformation parameter estimated from the correspondences.

Ashwini P et al[8]proposed mosaicing algorithm based on SIFT and corner detection algorithms to enhance mosaicing quality.

Srinath P et al [10] proposed a novel approach for Braille document mosaicing based on stitch the two split pieces of Braille document.

Isgr´o et in [12] proposed a feature based image mosaicing method. In this method, the feature points were extracted from one of two images to be stitched, and then the corresponding points in the other image were calculated. An Euclidean transformation parameter were estimated from the corresponding points. The Euclidean transformation parameter was used to stitch the images. However, the method is unable to deal with scaling and perspective distortion because the Euclidean transformation includes only translation and rotation

Zappala et al [13] and Peleg & Gee [14] proposed methods on document image mosaicing by estimating the motion through point matching. They considered the features based on the domain through exhaustive search procedure to extract best matches for mosaicing. The method works for 50% overlapping regions.

From the literature survey it is evident that mosaicing technique is useful in creating huge image composites and panoramas and is applied in various document , image and video mosaicing applications. In this paper we are extending the mosaicing technique with horizontal strip based approach for pedestal shot videos. The motivation for developing this approach is discussed in the next section.

III. BACKGROUND

In general when we capture a document video we capture the text content by scanning the source either left to right called track movement / pan movement, or top to bottom called tilt / pedding shot / pedestal shot. E.g. when the individual words in a line of a document needs to be scanned we prefer track movement but while capturing whole paragraphs or columns present in journal, magazine or bill slip we prefer pedding or tilt movement.

Pedestal shot means moving the camera vertically up or down with respect to the subject. A pedestal scan in practice is different from camera tilt. In case of a pedding movement the whole camera is moved but not just the angle of view. But in case of tilt, the camera will be in the same position but tilts the angle of view up and down. When a document is captured using pedestal or pedding movement it results in many content wise similar frames resulting in high redundancy in the spatial content. The purpose of capturing the document video is to extract and recognize the text information present in
them. The cost of localizing and recognizing the text information from each individual frame becomes computationally expensive. Thus it is advantageous to mosaic the frame content by finding only the spatial transition occurred in the successive frame with already exiting present frame content. i.e. in case of vertical camera movement according to observation information is added at the bottom of the image and subtraction of information is at the top of the image is successive frames. Figure 1 depicts the camera movement using pedestal shot.

Fig.1. Camera movement in pedestal shot and difference between pedestal shot and tilt

IV. PROPOSED METHODOLOGY

The horizontal strip based mosaicing procedure undergoes eight steps to mosaic the contents from the consecutive frames. The detailed description of the steps is discussed in the following sections. Figure 2 depicts the steps involved in mosaicing.

A. Video Acquisition Procedure

To acquire video in pedestal shot mode we shall consider two methods. In the first method a mobile phone with a camera or digital camera is held in the hand and the document in placed parallel to the camera at a suitable distance. The hand is then moved vertically, either top to bottom or bottom to top covering the entire content of the document. It is noted that in this method of acquisition the video suffers lot of handshakes, blur due to hand movement. This can be overcome in the second method

In the second method the same mobile phone or digital camera is fixed on a tripod stand. The tripod's vertical rod is moved along with the mobile camera vertically up or down and the document is captured. This exhibits a smooth movement and the blur, handshakes are reduced.

B. Fragmenting video into frames

Once the video is acquired from the acquisition device it is stored in the computer for further processing. To mosaic the stored video it is necessary to fragment the video into frames. The number of frames is dependent on the frame trigger rate of the acquisition device. The general frame rate is 25 fps to 26 fps but we have set the frame rate to 1 fps. This avoids lot of redundancy. When the video is fragmented, the result will be as follows

$$\text{Video} \rightarrow \{ f_1, f_2, f_{2\ldots} f_n \} \quad (1)$$

where, f1..fn represents the frame i.e individual static images with three dimension representing spatial and time domain

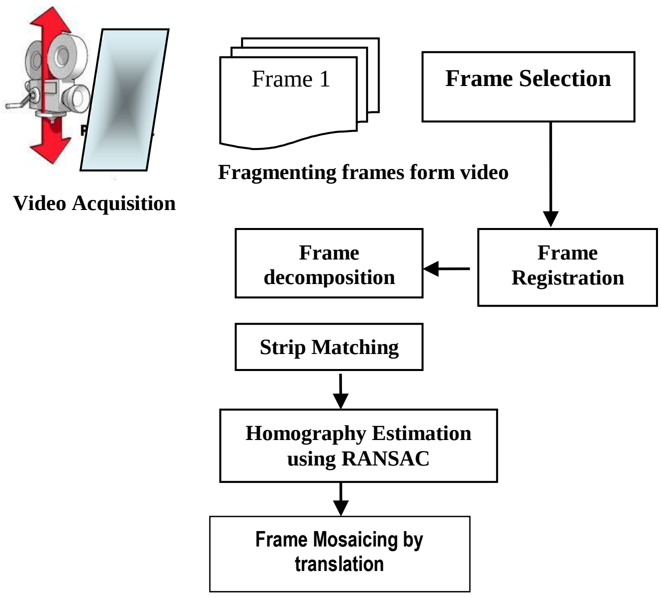


Fig.2. Block Diagram of proposed method

C. Frame Selection:

After fragmentation, the frames are subjected to selection procedure. The video captured using mobile phone either by hand or by tripod results in several artifacts like linear blur, defocusing or noise due to the hand or tripod movement. As a result, it is necessary to find the best quality frame before stitching the individual frames. We have considered using the no reference perceptual blur metric as stated by R. Ferzli in [15]. The metric is used to find the sharpest intensity frame present among the frames fragmented and returns value either 0 or 1. . Here it is noted that 0 refers to the sharpest frame and 1 refers to the blur frame.

We have also used the algorithm stated in [2]. A 3x3 median filter is applied to remove noise. This is effective at preserving edges by removing very fine noise

D. Frame Registration:

The frame registration involves the process of setting the reference frame. After selecting the best frame, the first in the list of frames is set as the reference frame. The rest of the frames are decomposed into strips, are matched and stitched to this reference frame.

E. Decomposition of frames into horizontal strips

After selecting the reference frame, the next step is to divide the consecutive frames into horizontal strips. Most of the mosaicing approaches combine full images without performing the strip decomposition. The vertical strip decomposition method is stated in [2] by Nagabhushan P et.al. A similar approach is suggested here in this paper. But instead of using vertical strip decomposition we are dividing the frames into horizontal strips as the video is captured in pedestal shot mode.

Strip decomposition is used because of its ability to align the image correctly, easy to determine image manifolds [4], and is computationally inexpensive as we do not match all the features of two images but match only the part of the image strip where there is spatial transition in the content.

Now to perform strip decomposition we need to observe the transition in contents of the frames with respect to time t. It is seen in [2] that in track movement the information changes mainly occurred in first few columns in the left and last few columns in the right. But in pedestal shot videos it is observed that the due to the gradual top-down vertical camera movement the information present in the previous frame is subtracted from the top of the current frame and new information is added at the bottom. And in case of bottom-top vertical camera movement it vice versa.

In the frame decomposition stage we are dividing the reference frame and its consecutive frames into three horizontal strips creating three sub images from that of the original and is stored in a buffer memory. We have created three sub images since atleast 30% of the match is required for attaining best match results using SIFT. Figure 3 depicts the decomposition of a frames

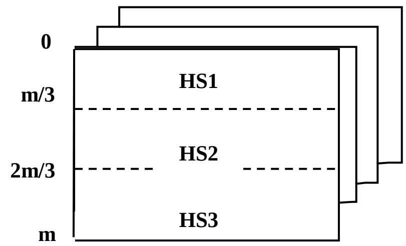
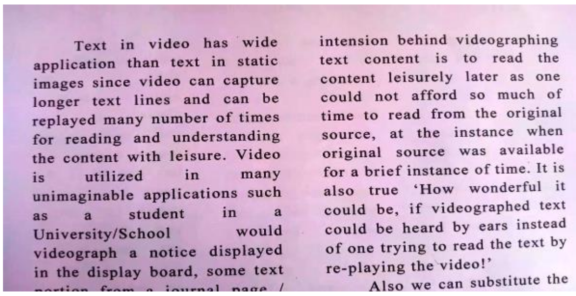


Fig 3. Depicts the decomposition of frames into horizontal strips

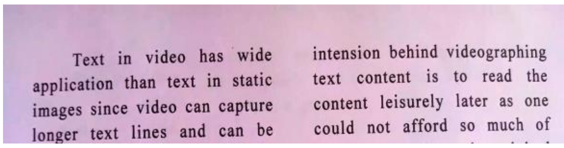
In the figure above, the horizontal strip HS1, HS2 and HS3 forms sub images of size (M/3, N). For creating the horizontal strips programmatically we have set the row indexes for HS1 starting from 0 to M /3 and column index as N. For HS2 row index starts from M/3+1 to (2xM) / 3 and column index N. And For HS3 the row index starts from 2xM/3 + 1 to M column remaining as N. The algorithm for strip decomposition is given below. Figure 4 depicts the frame and figure 5 depicts the sub images.

//Algorithm for strip decomposition
Step 1: Start
Step 2: [row, col] = size(Image)
Step 3: for i = 1 to row / 3
Step 4 : for j = 1 to n
Step 5 HS1(i, j) = Image(i, j)
End
End
Step 6: for i = (row / 3) + 1 to (2 x row) / 3
Step 7: for j = 1 to n
Step 8: HS2(i, j) = Image(i, j)
End End
Step 9 for i = (2 x row) + 1 / 3 to row
Step 10: for j = 1 to n
Step 11: HS 3(i, j) = Image(i, j)
End End
Step 12: Return (HS3)
End


Text in video has wide application than text in static images since video can capture longer text lines and can be replayed many number of times for reading and understanding the content with leisure. Video is utilized in many unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text

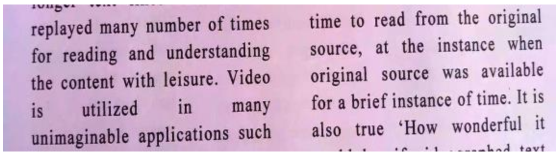
intension behind videographing text content is to read the content leisurely later as one could not afford so much of time to read from the original source, at the instance when original source was available for a brief instance of time. It is also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the

Fig. 4 Original Frame 1


Text in video has wide application than text in static images since video can capture longer text lines and can be

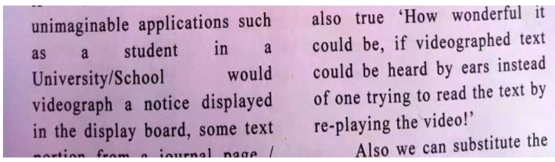
intension behind videographing text content is to read the content leisurely later as one could not afford so much of

Fig. 5(a) Horizontal strip HS1


replayed many number of times for reading and understanding the content with leisure. Video is utilized in many unimaginable applications such

time to read from the original source, at the instance when original source was available for a brief instance of time. It is also true 'How wonderful it

Fig. 5(b) Horizontal strip HS2


unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text

also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the

Fig. 5(c) Horizontal strip HS3
Fig. 5 Decomposition of frame 1 into horizontal strip HS1, HS2 and HS3

F. Horizontal Strip Matching

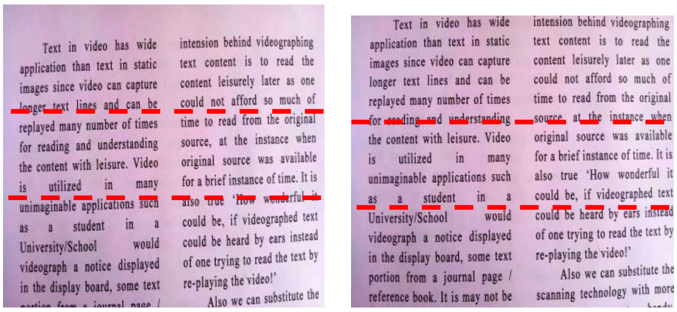
After decomposing the frames into horizontal strips, the next step is to match the strips with the reference frame and mosaicing it to the reference frame. Before matching the horizontal strip we shall understand the information transition in pedestal shot video. Let us assume that the camera is moved vertically either top to bottom or bottom to top . According to experimental observations the transition in the information change in the consecutive frames can be visualized as

a) Case 1: Top to bottom camera scan
   1. New information is moved in to the current frame at the bottom. i.e in horizontal strip HS3.
   2. Previous frame information is moved out of the current at the top i.e. horizontal strip HS1.

b) Case 2: Bottom to top camera scan
   1. New information is moved in to the current frame at the top i.e. in horizontal strip HS1
   2. Previous frame information is moved out of the current at the top i.e. horizontal strip HS2

Hence depending on the user defined camera direction we can mosaic either HS3 in case of top to bottom scan or Hs1 in case of bottom to top scan.

The example for information transition in vertical top to bottom document scan is shown in figure 7 and 8.

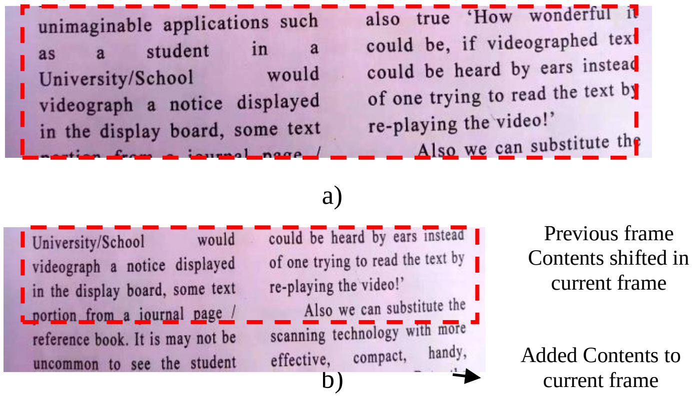

Text in video has wide application than text in static images since video can capture longer text lines and can be replayed many number of times for reading and understanding the content with leisure. Video is utilized in many unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text

intension behind videographing text content is to read the content leisurely later as one could not afford so much of time to read from the original source, at the instance when original source was available for a brief instance of time. It is also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the

Text in video has wide application than text in static images since video can capture longer text lines and can be replayed many number of times for reading and understanding the content with leisure. Video is utilized in many unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text portion from a journal page / reference book. It is may not be

intension behind videographing text content is to read the content leisurely later as one could not afford so much of time to read from the original source, at the instance when original source was available for a brief instance of time. It is also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the scanning technology with more

Figure 6 a) Decomposed Reference frame b) Decomposed Frame 2



unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text

also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the

University/School would videograph a notice displayed in the display board, some text portion from a journal page / reference book. It is may not be uncommon to see the student

could be heard by ears instead of one trying to read the text by re-playing the video!' Also we can substitute the scanning technology with more effective, compact, handy,

Fig. 7 a) Reference frame HS3 b) Frame 2's HS3

Figure 6 a) shows the decomposed reference frame and figure 6 b) shows the decomposed next consecutive frame, frame 2. In figure 7 the horizontal strip HS3 contains the reference frame information shown in the red dotted box. At time instance t it is clearly evident that the reference frame HS3 contents has shifted to the top and certain amount of information is added at the bottom in the frame 2's HS3 which is shown in the red dotted box of figure 7 b).

Once the information transition is identified we shall match the HS3 of reference frame and HS3 of the consecutive frames using SIFT – match feature descriptors that is defined in [1]. The matching is done to all the rest of the frames and are mosaiced later using homography estimation and blending function

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 07, July-2015

The algorithm for the above described procedure is given below and figure 9 shows the result of SIFT match

// Algorithm for vertical top down pedestal shot video
Step 1: Start
Step 2: Set reference frame as fr = 1
Step 3: for I = 2 number of frames
Step 4: u = image_strip(fr1)
Step 5: v = image_strip(I)
Step 6: Match_Score = SIFT_match(u, v)
Step 7 : if (Match_Score > theshold)
Step 8 : IH =SIFT_Match(fr, v)
Step 9: Store outliers in table
Step 10 : Compute homography
Step 11: Call Blending Function
Step 10: End

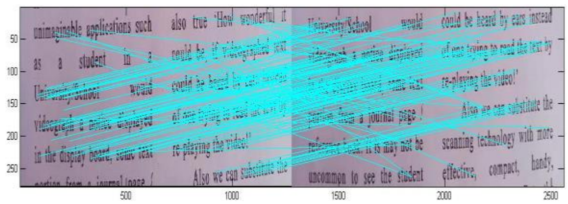


Fig. 9. SIFT_Match Result

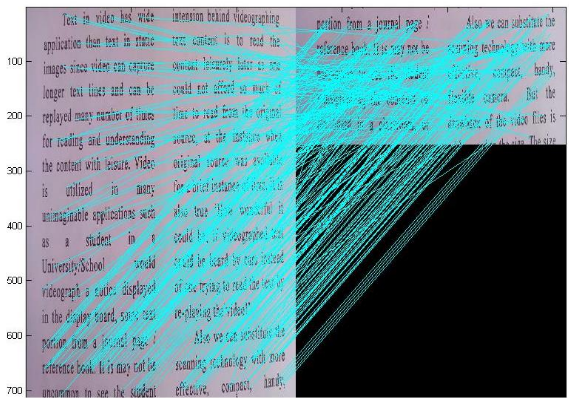


Fig. 10 SIFT_Match for mosaicing

G. Homography Estimation Using RANSAC

Homography maps the points in two images with one to one correspondence. Mathematically homography refers to projective linear transformation. Among computer vision community it is most commonly called as linear transformation between two image planes. According to [1] a 2D homography is defined by 3x3 matrix represented as H which corresponds to pixel P of horizontal strip HS3 of reference frame 1 to pixel P^1 of horizontals strip HS3 in consecutive frames i +1. It can be estimated using 4 or more corresponding points obtained by SIFT descriptors using the equation

$$HP = WP^1 \qquad (2)$$

where W is the scale parameter.

Here we are estimating homography using RANSAC. The main reason behind using RANSAC is for removing the falsely detected points called outliers detected by SIFT match. RANSAC in general matches 4 random pairs of matched points by SIFT algorithm and uses it to compute homography. Later on it checks the remaining key points obtained by SIFT with obtained consistent homography points. As RANSAC is an iterative algorithm the number of iterations to find consistent points depends on the amount of outliers generated by SIFT. The algorithm for RANSAC is given below

// Algorithm for RANSAC Estimation
1. Choose number of samples N
2. Choose 4 random potential matches
3. Compute H using normalized DLT
4. Project points from x to x' for each potentially matching pair:
5. Count points with projected distance < t – E.g t = 3 pixels
6. Repeat steps 2-5 N times – Choose H with most inlier

H. Image Blending

After fitting the homography the image is wrapped using a simple translation function available by a solving a set of affine transformation using 1*st* order polynomial

$$x_0 = a_0 + a_1 x + a_2 y$$
$$y_0 = b_0 + b_1 y + b_2 y$$

Then the text contents of the adjacent frames are stitched at the bottom of the reference frame. The mosaiced image after wrapping is as shown below

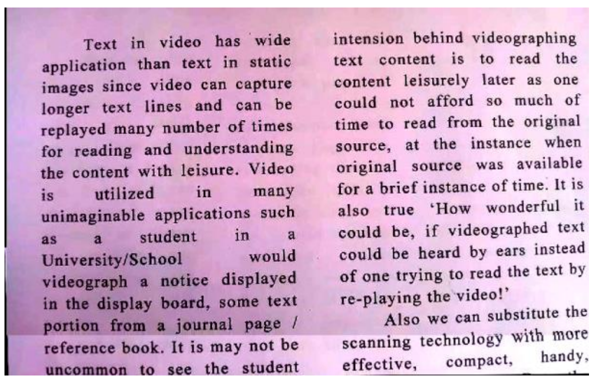

Text in video has wide application than text in static images since video can capture longer text lines and can be replayed many number of times for reading and understanding the content with leisure. Video is utilized in many unimaginable applications such as a student in a University/School would videograph a notice displayed in the display board, some text portion from a journal page / reference book. It is may not be uncommon to see the student intension behind videographing text content is to read the content leisurely later as one could not afford so much of time to read from the original source, at the instance when original source was available for a brief instance of time. It is also true 'How wonderful it could be, if videographed text could be heard by ears instead of one trying to read the text by re-playing the video!'
Also we can substitute the scanning technology with more effective, compact, handy,

Fig 11. Mosaiced Image

V. EXPERIMENTAL ANALYSIS

A. Video Data Set

Experiments were conducted to capture pedestal shot videos. The videos were captured using mobile camera with 5.0 mega pixel. Table 1 depicts the type of textually rich content video frames that are captured using mobile phon. We considered documents with homogenous background (white) with black foreground and few colored samples. The documents were from printed books, journal pages and printed pages. The videos were captured with gradual top to bottom movement at a distance of 15 inches from the source placed parallel to the camera.

The experiments were conducted in ordinary lighting conditions without using mobile camera flash. The results of mosaicing are discussed in the next section.

B. Implementation

We have implemented the entire mosaicing algorithm using Matlab 2014(a). The result obtained by the algorithm is depicted in Table 2. Mean square error metric can be used to assess the quality of the mosaiced image [2], [16]. As these text rich document videos are mosaiced for the purpose of recognition of text the overall quality of the mosaiced image depends on the OCR recognition rate of text present in the mosaiced image. Also, the result of the approach is evaluated based on the readability of mosaiced image by human perception.

We have analyzed the time complexity of the proposed horizontal strip based mosaicing. The complexity analysis is explained in the next section.

C. Complexity

Most of the processing time is taken to match the SIFT features of reference frame. Hence this step has a complexity of O (N M D) comparisons, where N is the number of features in consecutive frame, M is the number of features in the reference frame and D is the dimension given by row x col. As we are reducing the dimension row by row/3 rows we are saving the time consumed for matching and fitting the data. The graph for comparison of two frames for with and without strip decomposition for time complexity analysis is shown in figure 12.

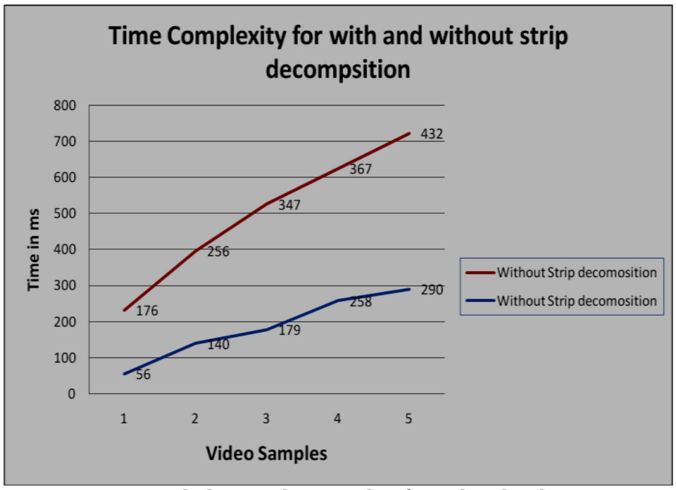


Figure 12 .Graph showing the time taken for with and without strip decomposition

VI. CONCLUSION

The horizontal strip based method for mosaicing the text contents present in the consecutive frames is proposed in this paper. The use of SIFT descriptors for matching works on different types of documents since it is invariant of scale, rotation and geometric distortions. The experimental studies reveal that the mosaicing time is comparatively reduced in case of this approach. We can use the same approach for bottom to top scan. Future works includes adapting similar procedure for reverse and hybrid scan patterns in capturing the document video.

REFERENCES

[1] Brown, M. Lowe, David G. (2003). "Recognizing Panoramas". Proceedings of the ninth IEEE International Conference on Computer Vision. 2.pp. 121825.doi:10.1109/ICCV.2003.1238630

[2] Nagabhushan P, Vimuktha Evangeleen Jathanna. "Mosaicing of Text Contents From Adjacent Video Frames". International Journal of Machine Intelligence. Vol 3, Issue 4, 2011.

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

[4] Hemanth Kumar, P.Shivkumar, D.S.Guru, P.Nagabushan (2004) Document Image Mosaicing: A novel approach. Sadhana Vol 29, Part 3, June 2004, pp-329=341, India

[5] Hemanth Kumar, P.Shivkumar, D.S.Guru, P.Nagabushan (2004) Sliding window based approach for document image mosaicing. Elsevier Image and Vision Computing Volume 24, Issue 1, 1 January 2006, Pages 94-100

[6] Lian, J., DeMenthon, D., and Doermann, D., "Camera-based document image mosaicing," in InternationalConference on Pattern Recognition, (2006).

[7] Tomohiro Nakai, Koichi Kise and Masakazu Iwamura," Camera-based document image mosaicing using LLAH", 7th International Workshop DAS 2006

[8] Ashwini P, Jayalakshmi H, "Image Mosaicing Using SIFT and corner Detection Algorithm", IJATER, Vol 4, Issue 2, (2014).

[9] *Harshal Patil* ,"Image Mosaicing Approach and Evaluation Methodology", *IJPRET, Volume (8): 2013*

[10] Srinath S and C.N Ravikumar, "Braille document Image Mosaicing : A Novel Approach", IJCA, Vol 103; Issue 6, 2014

[11] R. Ferzli and L. J. Karam, "A No-Reference Objective Image Sharpness Metric Based on Just-Noticeable Blur and Probability Summation", ICIP'07, pp.445-448, 2007

[12] Isgr´o, F. and Pilu, M., "A fast and robust image registration method based on an early consensus paradigm,"Pattern Recognition Letters 25(8), 943–954 (2004).

[13] Zappala A R, Gee A H, Taylor M J 1997 Document mosaicing. Int. Proc. British Machine Vision

[14] Peleg S, Gee A H 1997 Virtual cameras using image mosaicing. Haifa Research Laboratory, Hebrew University, Jerusalem

[15] S. Battiato, G. Di Blasi, G.M. Farinella, G. Gallo ,"A Survey of Digital Mosaic Techniques**"**, Eurographics Italian Chapter Conference (2006)

[16] Pietro Azzari, Luigi Di Stefano and Stefano Mattoccia, "An Evaluation Methodology for Image Mosaicing Algorithms " in in Proceedings of ACIVS 2008

[17] Pedestal Shot or Boom Movement. Retrieved from www.mediacollege.com/video/shots/pedestal.html

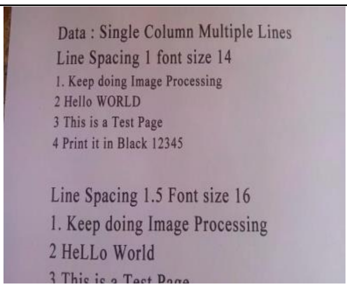
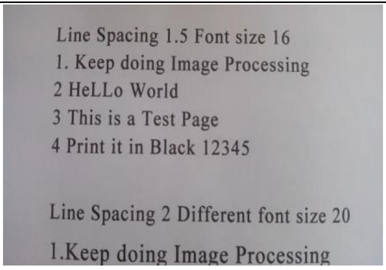
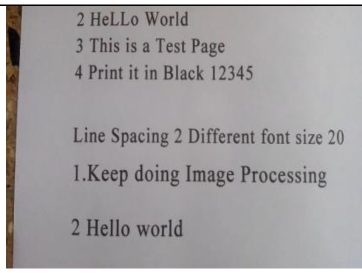
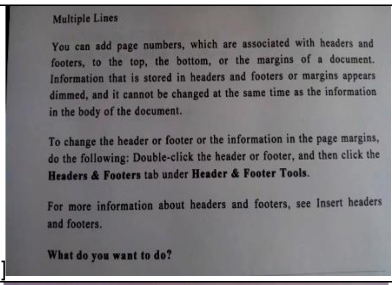
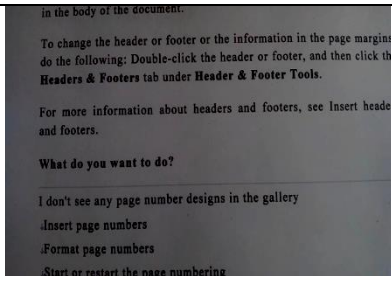
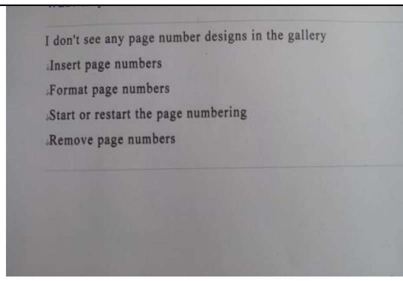
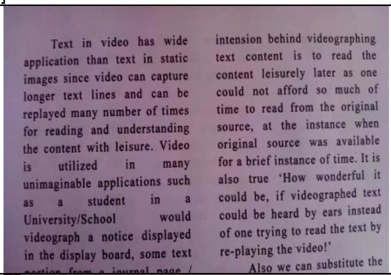
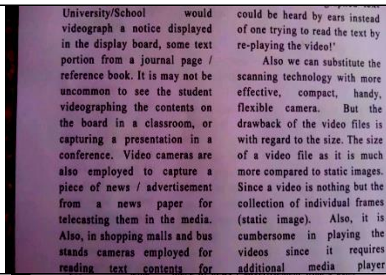
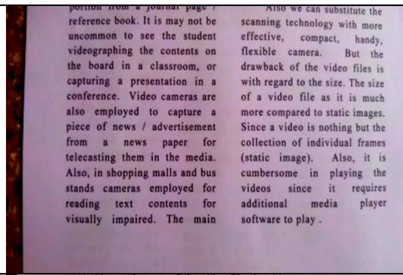
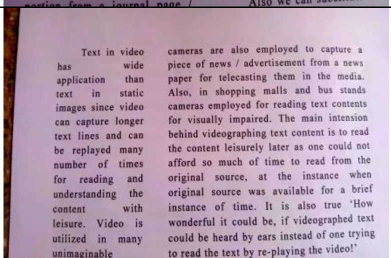
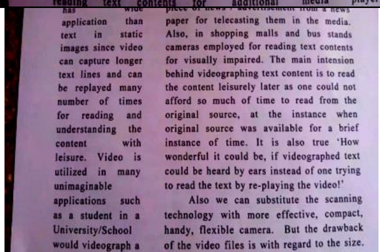
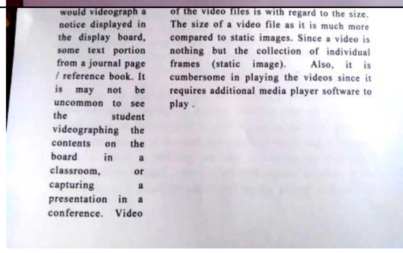
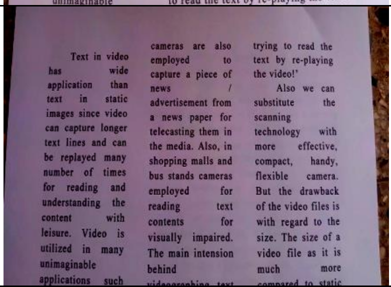
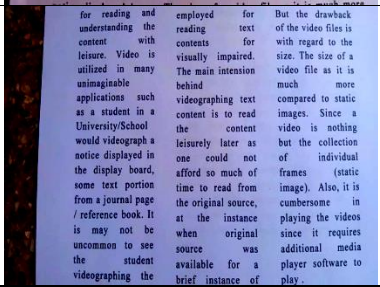
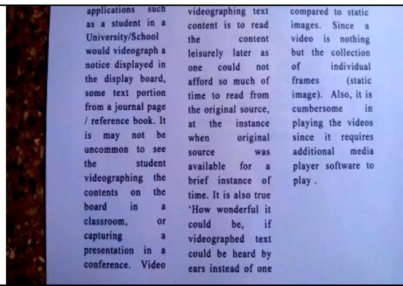
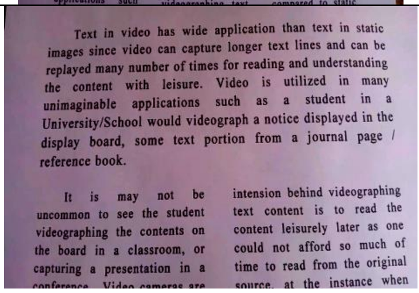
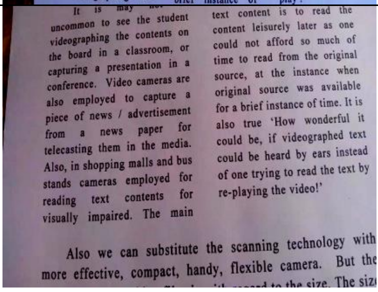
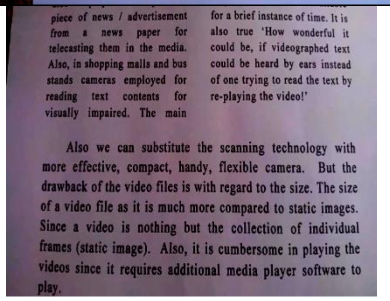
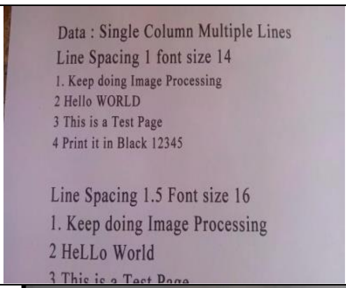
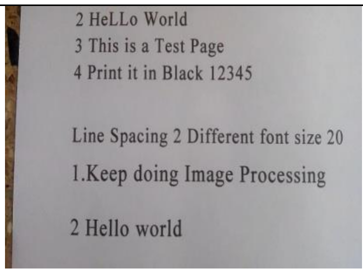
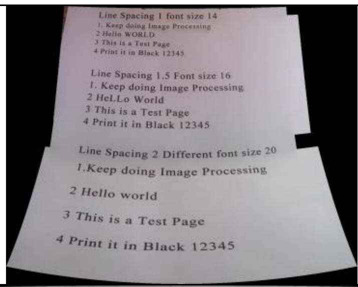
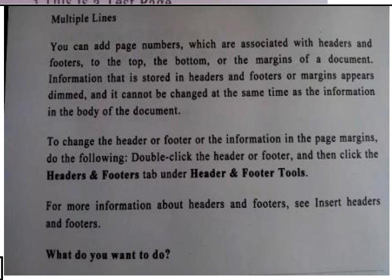
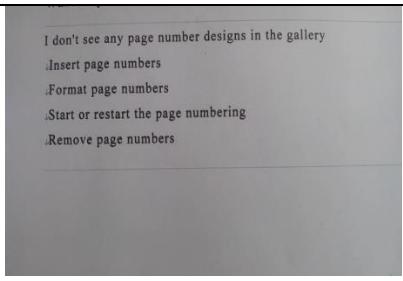
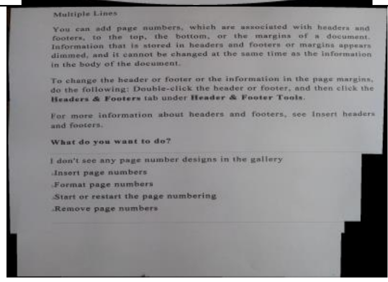
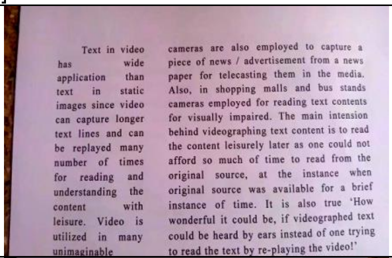
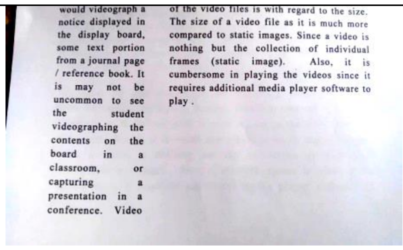
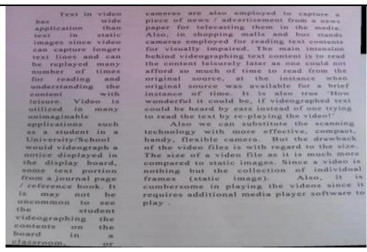
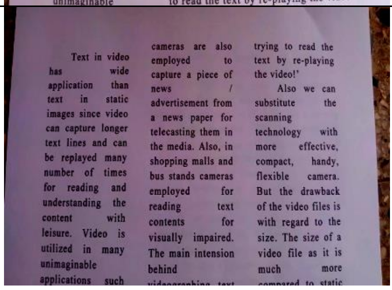
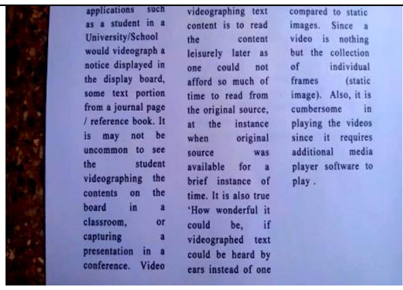
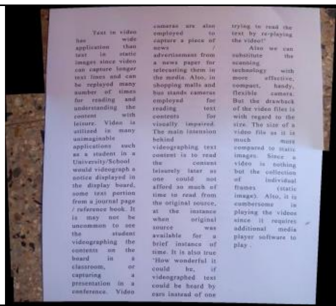
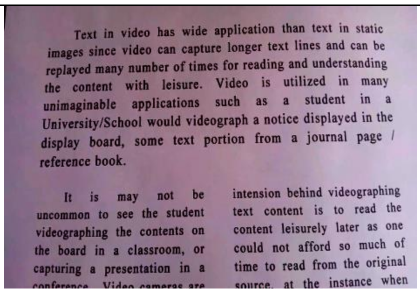
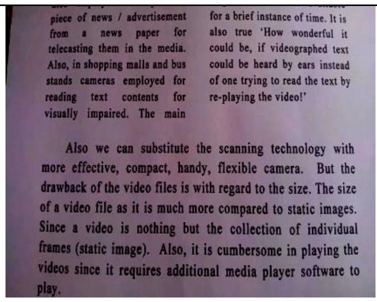
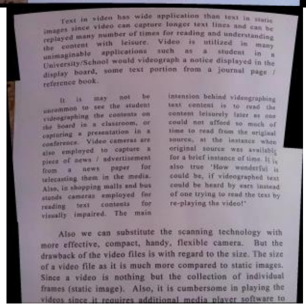
Video Type	Frame 1	Frame 2	Frame n
Single Column Video Sequence			
Single Column			
Double Column			
Double Column			
Triple Column			
Hybrid Column format			

Table 1 Types of documents captured using pedestal shot

Video Type	Frame 1	Frame n	Mosaiced Images using Horizontal strip Match
Single Column Video Sequence			
Single Column			
Double Column			
Triple Column			
Hybrid Column format			

Data : Single Column Multiple Lines
Line Spacing 1 font size 14
1. Keep doing Image Processing
2 Hello WORLD
3 This is a Test Page
4 Print it in Black 12345
Line Spacing 1.5 Font size 16
1. Keep doing Image Processing
2 HeLLo World

2 HeLLo World
3 This is a Test Page
4 Print it in Black 12345
Line Spacing 2 Different font size 20
1.Keep doing Image Processing
2 Hello world

Multiple Lines
You can add page numbers, which are associated with headers and footers, to the top, the bottom, or the margins of a document. Information that is stored in headers and footers or margins appears dimmed, and it cannot be changed at the same time as the information in the body of the document.
To change the header or footer or the information in the page margins, do the following: Double-click the header or footer, and then click the **Headers & Footers** tab under **Header & Footer Tools**.
For more information about headers and footers, see Insert headers and footers.
What do you want to do?

I don't see any page number designs in the gallery
.Insert page numbers
.Format page numbers
.Start or restart the page numbering
.Remove page numbers

Table 2 Results obtained from Horizontal strip based mosaicing