Mobile Video Mosaicing System for Flat and Curved Documents

Tomokazu Sato[†,‡], Akihiko Iketani[‡], Sei Ikeda[†], Masayuki Kanbara[†,‡], Noboru Nakajima[‡], and Naokazu Yokoya[†,‡]

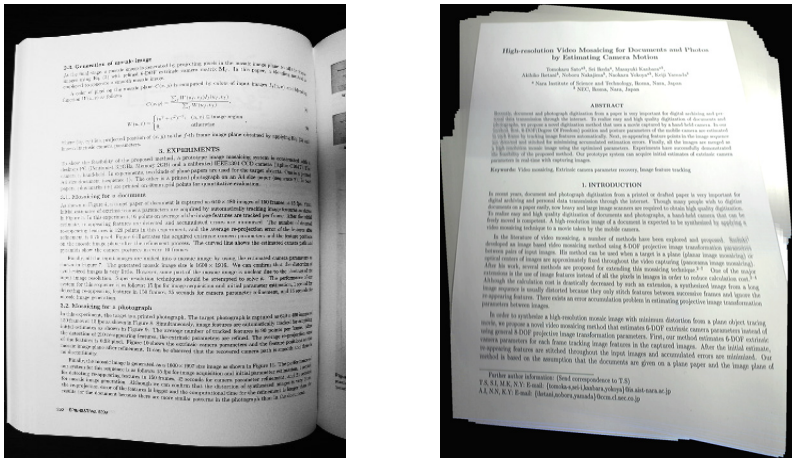
†Nara Institute of Science and Technology, Japan
‡Internet System Research Labs., NEC, Japan

Abstract. This paper presents a real-time video mosaicing system that is one of practical applications of mobile vision. Generally, the video mosaicing techniques give a cheap and low-resolution web-cam device the capability of high resolution capturing of a target with improved imaging quality. To realize video mosaicing on an actual mobile device, in our method, image features are automatically tracked on the input images and 6-DOF camera motion parameters are estimated with a fast and robust structure-from-motion algorithm. A preview of generating a mosaic image is also rendered in real time to support the user. Our system is basically for the flat targets, but the system also has the capability of 3-D video mosaicing in which an unwrapped mosaic image can be generated from a video image sequence of a curved document. In experiments, the video mosaicing system for flat and curved targets is demonstrated to show the feasibility of our approach.

1 Introduction

Document and photograph digitization from a book and printed paper has been very important for digital archiving and electronically transmitting them over the internet. Flat-bed scanners are one of the most commonly used devices for this purpose. These scanners, however, are too large and heavy to be portable. Thus, there has been a strong demand for a high quality digitization of documents using portable imaging devices, such as cameras on cellular phones. One of the problems here is the resolution of the image acquired with these devices. 2M pixel cameras enable full A4 pages to be sampled at about 150 dots per inch (dpi), whereas flat-bed scanners enable sampling at a few thousands dpi. There also exists an image distortion problem when a book is captured by digital cameras, as shown in Figure 1(a).

For the low resolution problem, a video mosaicing technique is one of the most promising solutions. In the video mosaicing method, partial images of the document are captured as a video sequence, and multiple frame images are stitched seamlessly into one large, high resolution image. There are a number of methods for video mosaicing. Conventional methods usually achieve pairwise registration



(a)Book captured by a digital camera. (b)Mosaic image with slanted effect.

Fig. 1. Problems in digitized documents.

between two successive images, and construct a mosaic image by warping all the images to a reference frame (in general, the first frame). Szeliski [1] has developed a method using 8-DOF projective image transformation parameters called homography. In this method, for every pair of consecutive frames, homography which minimizes the sum of squared differences between the two frames is estimated. After his work, various extensions to this method have been proposed [2–9]. One of the major extensions is the use of image features instead of all the pixels in images for reducing computational cost [2–4].

All of these methods, however, align the input images to the reference frame, and thus will generate a mosaic image with slanted effect if the image plane in the reference frame is not parallel to the target document. A mosaic image with this slanted effect is shown in Figure 1(b). Moreover, the homography based methods are just applicable when a target is a plane (planar image mosaicing), or the optical center of camera is approximately fixed throughout the video capturing (panoramic image mosaicing). Thus, if the target is a curved surface, the above assumption no longer holds, and misalignment of images will appear in the resultant image. Although there are some video mosaicing methods which can deal with a curved surface, they require an active camera and a slit light projection device [9]. On the other hand, the usability of the system is also one of the important factors in mobile vision application. Unfortunately, conventional video mosaicing researches have not focused on the usability of the system because interactive video mosaicing systems have not been developed yet.

In this paper, we solve such problems as mentioned above by using a 3-D reconstruction technique for the image sequence with real-time processing. Our method is constructed of two stages as shown in Figure 2. In the first stage, a user captures a target document using a web-cam attached on a mobile com-

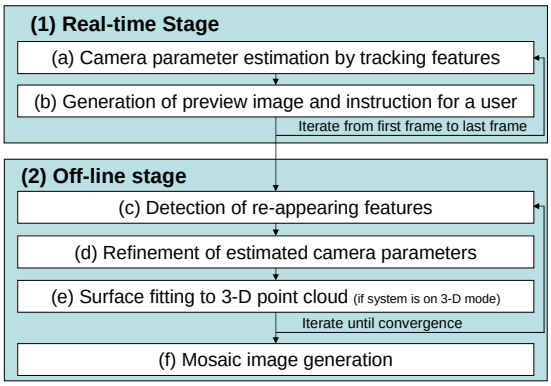


Fig. 2. The flow of the proposed method.

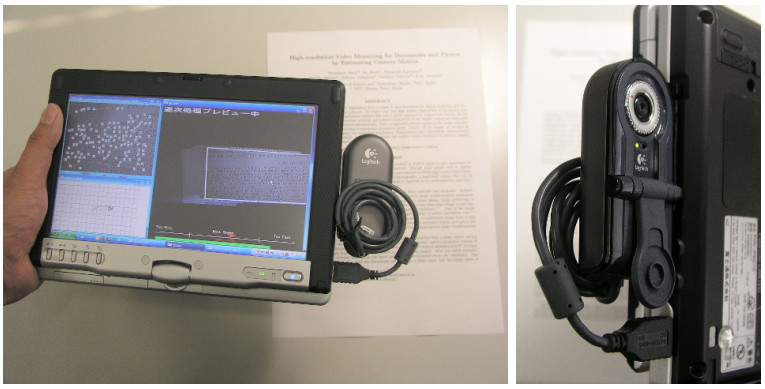


Fig. 3. Mobile computer with web-cam.

puter (Figure 3). In this stage, camera parameters are automatically estimated in real-time and the user is supported interactively through a preview window for displaying a mosaic image on the mobile computer. After this real-time stage, camera parameters are refined, and a high-resolution and distortion-free mosaic image is generated after a few minutes. Contributions of our work can be summarized as follows: (1) slanted effect removal from a mosaic image for the flat target, (2) unwrapped mosaic image generation for the curved document, and (3) interactive user support on the mobile system. The assumptions made in the proposed method are that intrinsic camera parameters are fixed and calibrated in advance to correct lens distortion. For curved documents, the curve of the target must lies along one direction and its curvature changes smoothly along this direction.

The rest of this paper is structured as follows. First, the proposed video mosaicing method is detailed in Section 2. In Section 3, experimental results for flat and curved documents are demonstrated to show the validity and feasibility of the proposed method. Finally, Section 4 gives conclusion and future work.

2 Video Mosaicing for Flat and Curved Document

This section describes a method for generating a high-resolution and distortion-free mosaic image from a video sequence. The flow of the proposed method is given in Figure 2. Although our system has two modes, one for flat target (flat mode) and the other for curved target (curved mode), the basic flow for these two modes are almost the same. In the real-time stage, the system carries out 3-D reconstruction processes frame by frame by tracking image features (a). The preview of generating a mosaic image is rendered in the monitor of the mobile computer and that is updated in every frame to show which part of the document has already been captured (b). After the real-time stage, the off-line stage is automatically started. In this stage, first, re-appearing image features are detected in the stored video sequence (c), and camera parameters and 3-D positions of features estimated in the real-time stage are refined (d). If the system works in the curved mode, surface parameters are also estimated by fitting a 3-D surface to estimated 3-D point cloud (e). After some iteration, finally, a high-resolution and distortion-free mosaic image is generated (f). In the following sections, first, extrinsic camera parameters and an error function used in the proposed method are defined. The stages (A) and (B) are then described in detail.

2.1 Definition of Extrinsic Camera Parameter and Error Function

In this section, the coordinate system and the error function used for 3-D reconstruction is defined. In the proposed method, the coordinate system is defined such that an arbitrary point $\mathbf{S}_p = (x_p, y_p, z_p)$ in the world coordinate system is projected to the coordinate $\mathbf{x}_{fp} = (u_{fp}, v_{fp})$ in the f-th image plane. Defining 6-DOF extrinsic camera parameters of the f-th image as a 3×4 matrix $\mathbf{M}_f$, the relationship between 3-D coordinate $\mathbf{S}_p$ and 2-D coordinate $\mathbf{x}_{fp}$ is expressed as follows:

$$(au_{fp}, av_{fp}, a)^T = \mathbf{M}_f (x_p, y_p, z_p, 1)^T, \tag{1}$$

where a is a parameter. In the above expression, $\mathbf{x}_{fp}$ is regarded as a coordinate on the ideal camera with the focal length of 1 and without radial distortion induced by the lens. In practice, however, $\mathbf{x}_{fp}$ is the projected position of $\hat{\mathbf{x}}_{fp} = (\hat{u}_{fp}, \hat{v}_{fp})$ in the real image, which is given by transferring using known intrinsic camera parameters including focus, aspect, optical center and distortion parameters. In the rest of this paper, this transformation from $\hat{\mathbf{x}}_{fp}$ to $\mathbf{x}_{fp}$ is omitted for simplicity.

Next, the error function used for 3-D reconstruction is described. In general, the projected position $\mathbf{x}_{fp}$ of $\mathbf{S}_p$ to the f-th image frame does not coincide with

the actually detected position $\mathbf{x}' = (u'_{fp}, v'_{fp})$, due to errors in feature detection, extrinsic camera parameter and 3-D feature position estimation. In this paper, the squared error E_{fp} is defined as an error function for the feature p in the f-th frame as follows:

$$E_{fp} = |\mathbf{x}_{fp} - \mathbf{x}'_{fp}|^2. \tag{2}$$

2.2 Real-time Stage for Image Acquisition

In the real-time stage, extrinsic camera parameters are estimated in real-time to show some information that helps users in image capturing. First, user should very roughly set the image plane of camera parallel to the target paper. Note that this setting is not laborious for users because it is done only for the first frame. This setting is used to calculate initial values for the real-time camera parameter estimation. After this setting, the user starts image capturing with free camera motion.

As shown earlier in Figure 2, the real-time stage is constructed of two iterative processes for each frame. First, the extrinsic camera parameter is estimated by tracking features (a). A preview image of generating a mosaic image and instruction for controlling camera motion speed are then shown in the display on the mobile computer and updated every frame (b). The following describes each process of the real-time stage.

Step (a): Camera parameter estimation by tracking features.
An iterative process to estimate the extrinsic camera parameters and 3-D position $\mathbf{S}_p$ of each feature point is described. This process is basically an extension of the structure from motion method in [10].

In the first frame, assuming that the image plane in the first frame is approximately parallel to the target, rotation and translation components in $\mathbf{M}_f$ are set to an identity matrix and 0, respectively. For each feature point detected in the first frame, its 3-D position $\mathbf{S}_p$ is set to $(u_{1p}, v_{1p}, 1)$, based on the same assumption. Note that these are only initial values, which will be corrected in the refinement process (Figure 2(d)).

In the succeeding frames ($f > 1$), $\mathbf{M}_f$ is determined by iterating the following steps until the last frame.

Feature point tracking: All the image features are tracked from the previous frame to the current frame by using a standard template matching with Harris corner detector [11]. The RANSAC approach [12] is also employed for eliminating outliers.

Extrinsic camera parameter estimation: Extrinsic camera parameter $\mathbf{M}_f$ is estimated using the tracked position (u'_{fp}, v'_{fp}) and its corresponding 3-D position $\mathbf{S}_p = (x_p, y_p, z_p)$. Here, extrinsic camera parameters are obtained by minimizing $\sum_p E_{fp}$, the sum of the error function defined in Eq. (2), by using the Levenberg-Marquadt method. For 3-D position $\mathbf{S}_p$ of the feature point p, the estimated result in the previous iteration is used.

Estimation of 3-D feature position: For every feature point p in the current frame, its 3-D position $\mathbf{S}_p = (x_p, y_p, z_p)$ is refined by minimizing the error

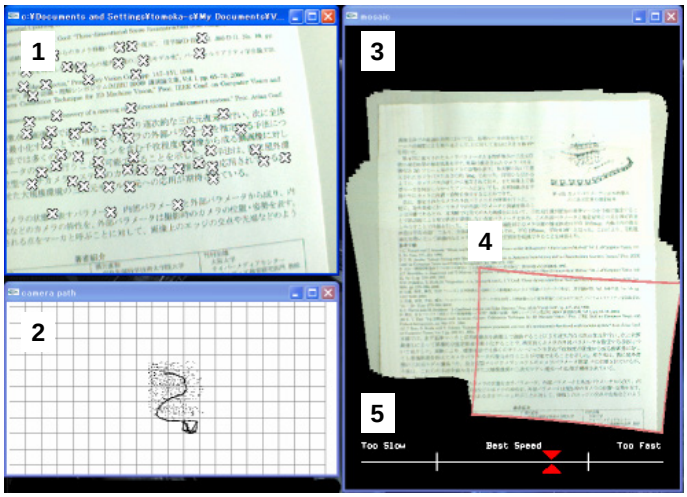


Fig. 4. Displayed information for user in real-time stage. 1: input image and tracked feature points. 2: estimated camera path and posture. 3: preview of generating a mosaic image. 4: capturing image area on mosaic image. 5: instruction for speed of camera motion.

function $\sum_{i=1}^{f} E_{ip}$. If the system works in the flat mode, z value of the $\mathbf{S}_p$ is fixed to a constant to increase the accuracy of estimated parameters.

Addition and deletion of feature points: In order to obtain accurate estimates of camera parameters, good features should be selected. The set of features to be tracked is updated by evaluating the reliability of features [10].

By iterating the above steps, extrinsic camera parameters $\mathbf{M}_f$ and 3-D feature positions $\mathbf{S}_p$ are estimated.

Step (b): Generation of preview image and instruction for user.
Figure 4 shows displayed information for a user in our system. Usually, the user watches the right side window in the real-time stage. In necessary, the user can check the input images and estimated camera path on the left side windows. The preview image helps the user to confirm a lacking part of capturing the document.

For each frame, the captured image is resized to a lower resolution and stored to texture memory. Every stored texture is warped to the mosaic image plane by texture mapping using Eq. (1). For the texture of the current frame, the boundary of the current frame image is colored in the preview so that the user can easily grasp which part of the target is currently captured.

On the other hand, the instruction shown in the bottom of the right window in Figure 4 helps the user to control the speed of camera motion. In our system, there is optimal speed for camera motion which gives both good quality of mosaicing result and little computational cost. Faster camera speed gives a worse

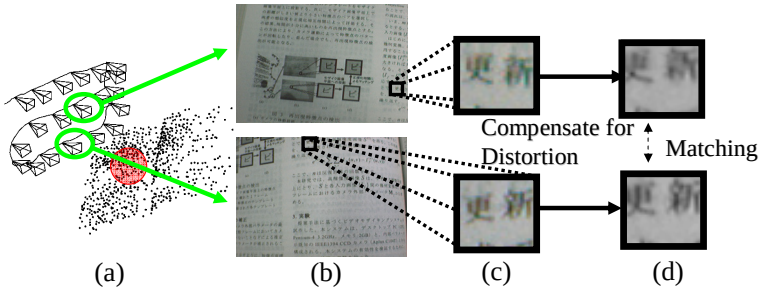


Fig. 5. Detection of re-appearing features. (a) camera path, posture and 3-D feature positions, (b) temporarily distinct frames in the input video, (c) templates of the same feature in different images, (d) templates without perspective distortion.

mosaicing result because an error of estimated camera parameters increases due to tracking errors and shortage of tracking span of each feature. In contrast, too slow motion will consume computational resources and memory in vain because the number of images is increased to capture the whole part of the document. Thus, the speed of video camera is shown by an arrow mark in the speed gauge, and the speed gauge is divided into "too slow", "best speed", and "too fast" indications. To obtain a good result in video mosaicing, the user should control camera speed so that the arrow mark stays in the "best speed" range.

2.3 Off-line stage for Parameter Refinement and Target Shape Estimation

This section describes the process to globally optimize the estimated extrinsic camera parameters and 3-D feature positions, and to approximate the shape of the target by fitting a curved surface. First, as shown in Figure 2(B), feature points which reappear in temporally distinct frames are detected (c). Using these reappearing features, 3-D reconstruction result is globally optimized (d). Then, by fitting a curved surface to 3-D feature points, the target shape is estimated (e). After a few iterations of steps (c) to (e), an unwrapped mosaic image of the target is generated using the target shape and estimated extrinsic camera parameters (f). Details on steps (c) to (f) are described below.

Step (c): Detection of Reappearing Features.
Due to the camera motion, most of image features come into the image, move across toward the end of the image, and disappear. Some features, however, reappear in the image, as shown in Figure 5. In this step, these reappearing features are detected, and distinct tracks belonging to the same reappearing feature are linked to form a single long track. This will give tighter constraints among camera parameters in temporally distinct frames, and thus makes it possible to suppress cumulative errors in the global optimization step described later.

Reappearing features are detected by examining the similarity of the patterns among features belonging to distinct tracks. The problem here is that even if two patterns belong to the same feature on the target, they can have different appearance due to perspective distortion. To remove this effect, first, templates of all the features are projected to the fitted surface (described later). Next, feature pairs whose distance in 3-D space is less than a given threshold are selected and tested with the normalized cross correlation function. If the correlation is higher than a certain threshold, the feature pair is regarded as reappearing features. Note that in the flat mode, the shape of the target is simply assumed as a plane. When the system works in the curved mode, this step is skipped at an initial iteration of steps (c) to (e) because surface parameters have not been estimated at the first iteration.

Step (d): Global Optimization of 3D Reconstruction.
Since 3-D reconstruction process described in Section 2.2 is an iterative process, its result is subject to cumulative errors. In this method, by introducing bundle-adjustment framework [13], the extrinsic camera parameters and 3-D feature positions are globally optimized so as to minimize the sum of re-projection errors, which is given as follows:

$$E_{all} = \sum_{f} \sum_{p} E_{fp}. \tag{3}$$

As for re-appearing features, all the tracks belonging to the same reappearing feature are linked, and treated as one single track. This enables the extrinsic camera parameters and 3-D feature positions to be optimized, maintaining the consistency among temporally distinct frames.

Step (e): Target Shape Estimation by Surface Fitting.
In this step, assuming the curve of the target lies along one direction, the target shape is estimated using 3-D point clouds optimized in the previous step (d). Note that this stage is skipped if the system works in the flat mode. First, as shown in Figure 6, the principal direction of curvature is computed from the 3-D point clouds. Next, 3-D position of each feature point is projected to a plane perpendicular to the direction of minimum principal curvatures. Finally, a polynomial equation of variable order is fitted to the projected 2-D coordinates, and the target shape is estimated.

Let us consider for each 3-D point $\mathbf{S}_p$ a point cloud $\mathbf{R}_p$ which consists of feature points lying within a certain distance from $\mathbf{S}_p$. First, the directions of maximum and minimum curvatures are computed for each $\mathbf{R}_p$ using local quadratic surface fit. For a target whose curve lies along one direction, as assumed in this paper, the minimum principal curvature must be 0, and its direction must be the same for all the feature points. In practice, however, there exists some fluctuation in the directions of minimum curvature, due to the estimation errors. Thus, a voting method is applied to eliminate outliers and the dominant direction $\mathbf{V}_{min} = (v_{mx}, v_{my}, v_{mz})$ of minimum principal curvatures for the whole target is determined.

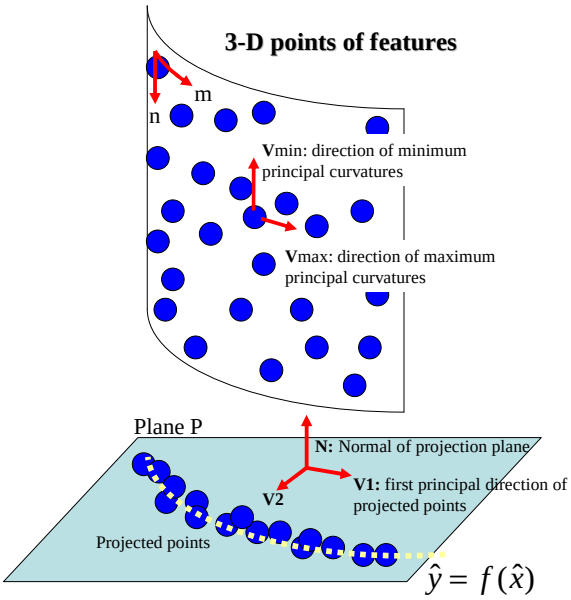


Fig. 6. Target shape estimation by polynomial surface fitting.

Next, 3-D position $\mathbf{S}_p$ for each feature point is projected to a plane whose normal vector $\mathbf{N}$ coincides with $\mathbf{V}_{min}$; i.e. $P(x, y, z) = v_{mx}x + v_{my}y + v_{mz}z = 0$. The projected 2-D coordinate $(\bar{x}_p, \bar{y}_p)$ of $\mathbf{S}_p$ is given as follows:

$$\begin{pmatrix} \bar{x}_p \\ \bar{y}_p \end{pmatrix} = \begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \end{pmatrix} \mathbf{S}_p, \tag{4}$$

where $\mathbf{V}_1$ is a unit vector parallel to the principle axis of inertia of the projected 2-D coordinates $(\bar{x}, \bar{y})$, and $\mathbf{V}_2$ is a unit vector which is perpendicular to $\mathbf{V}_1$ and $\mathbf{V}_{min}$; i.e. $\mathbf{V}_2 = \mathbf{V}_1 \times \mathbf{N}$.

Finally, the target shape parameter $(a_0, a_1, \cdots, a_m)$ is obtained by fitting the following variable-order polynomial equation to the projected 2-D coordinates $(\bar{x}, \bar{y})$.

$$\bar{y} = f(\bar{x}) = \sum_{i=0}^{m} a_i \bar{x}^i. \tag{5}$$

Here, the optimal order m in the above equation is automatically determined by using geometric AIC [14].

In case where the target is composed of multiple curved surfaces, e.g. a thick bound book shown in Figure 1(a), the target is first divided with a line where the normal vector of the locally fitted quadratic surface varies discontinuously, and then the shape parameter is computed for each part of the target. The estimated shape is used for generating an unwrapped mosaic image in the next

process, as well as for removing the perspective distortion in the reappearing feature detection process (Figure 2(c)).

Step (f): Mosaic Image Generation.
Finally, a mosaic image is generated by using extrinsic camera parameters and surface shape parameters. Let us consider a 2-D coordinate (m, n) on the unwrapped mosaic image, as shown in Figure 6. The pixel value at (m, n) on the unwrapped mosaic image is given by computing the average of the pixel values at all the corresponding coordinates (u_f, v_f) in the input image.

The relation between (m, n) and its corresponding 3-D coordinate $(\bar{x}, f(\bar{x}), \bar{z})$ on the fitted surface is given as follows:

$$(m, n) = (\int_0^{\bar{x}} \sqrt{1 + \{\frac{d}{dx} f(x)\}^2} dx, \bar{z}). \tag{6}$$

The coordinate $(\bar{x}, f(\bar{x}), \bar{z})$ is transferred to the corresponding 2-D coordinate (u_f, v_f) on the f-th image plane by the following equation.

$$\begin{pmatrix} au_f \\ av_f \\ a \end{pmatrix} = \mathbf{M}_f \begin{pmatrix} \mathbf{V}_1 \\ \mathbf{V}_2 \\ \mathbf{N} \end{pmatrix} \begin{pmatrix} \bar{x} \\ f(\bar{x}) \\ \bar{z} \end{pmatrix}. \tag{7}$$

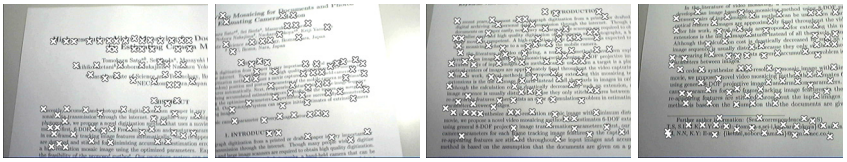
Assumption of the flat target shape (the flat mode) simplifies these equations. In this paper, description for the flat target is omitted.

3 Experiments

We have developed a prototype video mosaicing system which consists of a mobile PC (Fujitsu lifebook P1510D, Pentium-M 1.2GHz, Memory 1GB) and an USB Web-cam (Logitech QuickCam for Notebooks Pro). The appearance of the prototype system has been shown earlier in Figure 3. In our system, a user starts image capturing by touching the monitor of mobile computer. The real-time stage for image capturing is automatically finished when a video buffer becomes full. Experiments have been carried out for flat and curved documents. In these experiments, the intrinsic camera parameters are calibrated in advance using Tsai's method [15], and are fixed throughout image capturing. Note that in the current version of the prototype system, curved mode is not implemented on the mobile system. Thus, the latter experiment for curved surface is carried out by using a desktop PC (Pentium 4 Xeon dual 3.2 GHz, Memory 2GB), and the initial camera parameter estimation is processed in offline after image capturing using an IEEE 1394 type web-cam (Aplux C104T).

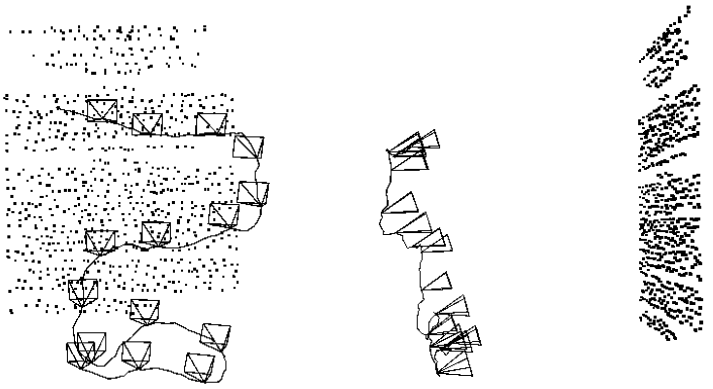
3.1 Flat target

As shown in Figure 7, a flat target document is captured as 150 frame images of 640×480 pixels at 6 fps with initial camera parameter estimation. Image features



1st frame 50th frame 100th frame 150th frame

Fig. 7. Sampled frames of input images and tracked features (flat target).



(a) frontal view (b) side view

Fig. 8. Estimated extrinsic camera parameters and 3-D positions of features (flat target).

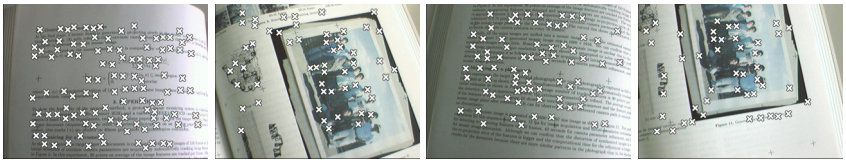
tracked in the real-time stage are depicted with cross marks. Note that none of input image planes are parallel to the target document. Figure 8 illustrates estimated extrinsic camera parameters and feature positions on the mosaic image plane. The curved line shows the estimated camera path and pyramids show the camera postures in every 10 frames. The generated mosaic image is shown in Figure 9. The size of the image is 1600×1936. We can confirm that the slanted effect is correctly removed in the final mosaic image by comparing the result with the homography based result shown in Figure 1(b) where the same input images and the same feature trajectories with this experiment were used. Although the target paper is not perfectly flat, there are no apparent distortions in the final mosaic image. The performance of the prototype system for this sequence is as follows: 6 fps for image acquisition and initial parameter estimation with preview of mosaicing, 17 seconds for camera parameter refinement, and 39 seconds for generating the final mosaic image.

Fig. 9. Generated mosaic images (flat target).

3.2 Curved target

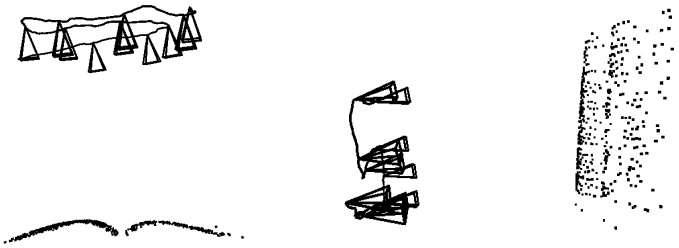
An example of curved documents used in experiments is shown in Figure 10. The target is composed of 2 pages of curved surfaces: one page with texts and the other with pictures and figures. The target is captured with a web-cam as a video of 200 frames at 7.5fps, and is used as an input to our system. Sampled frames of the captured video are shown in Figure 11. Tracked feature points are depicted with cross marks. 3-D reconstruction result obtained by the proposed method is shown in Figure 12. The curved line shows the camera path, pyramids show the camera postures in every 10 frames, and the point cloud shows 3-D positions of feature points. As can be seen, the point cloud coincides with the shape of the thick bound book. The shape of the target is estimated after 3 time iterations of reappearing feature detection and surface fitting process. The estimated shape of the target is shown in Figure 13. In the proposed method, the optimal order of the polynomial surface fitted to the target is automatically determined by geometric AIC [14]. In this experiment, the order is 5 and 4 for the left and right pages, respectively. The unwrapped mosaic images before and after shadow removal are shown in Figure 14(a) and (b), respectively. The resolution of the mosaic image is 3200×2192. As can be seen, the distortion on the target has been removed in the resultant image. The performance of the system on the desktop PC for this sequence is as follows: 27 seconds for initial 3-D reconstruction, 71 seconds for camera parameter refinement, and 113 seconds for generating the final mosaic image.

Fig. 10. Conference proceedings used as a curved target.



1st frame 66th frame 123th frame 200th frame

Fig. 11. Sampled frames of input images and tracked features (curved target).



(a) frontal view (b) side view

Fig. 12. Estimated extrinsic camera parameters and 3-D positions of features (curved target).

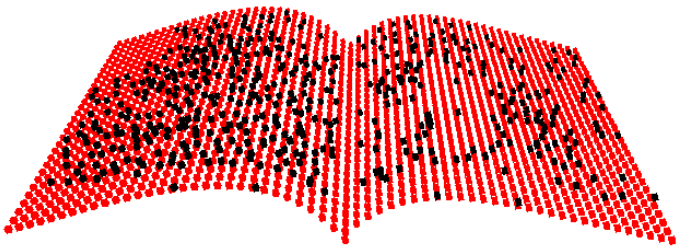
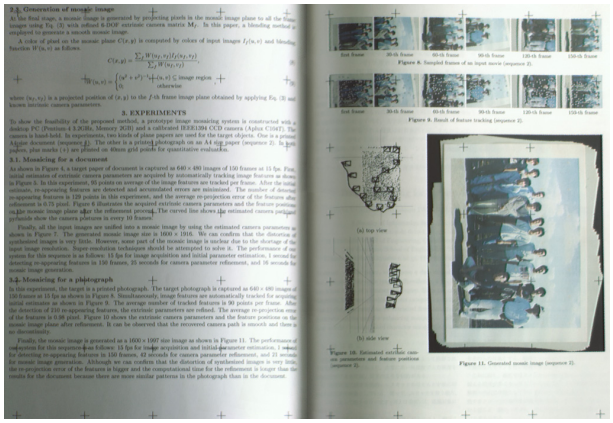
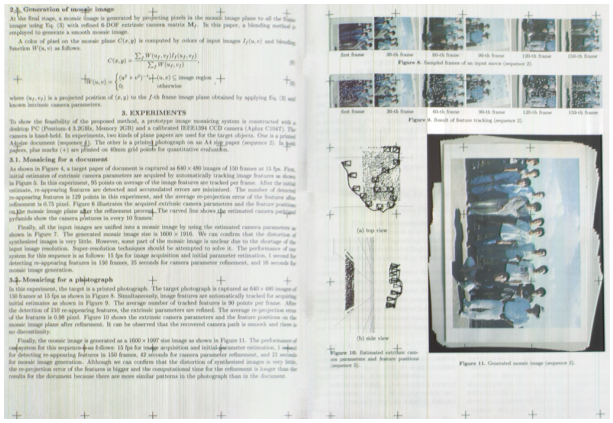


Fig. 13. Target shape estimated from 3-D point cloud.



(a) Before shadow removal.



(b) After shadow removal.

Fig. 14. Unwarpped mosaic image (curved target).

4 Conclusion

A novel video mosaicing method for generating a high-resolution and distortion-free mosaic image for flat and curved documents has been proposed. With this method based on 3-D reconstruction, the 6-DOF camera motion and the shape of target document are estimated without any other device besides a web-cam. In experiments, a prototype system of mobile video mosaicing has been developed and has been successfully demonstrated. For the flat target, the mosaic image without the slanted effect is successfully generated. Although the curved mode has not been implemented on the current version of mobile system yet, we have shown the off-line result for curved documents generated by desktop PC. In the curved mode, assuming the curve of target lies along one direction, the shape

model is fitted to the feature point cloud and unwrapped image without shadow is automatically generated. Our future work is to reduce the computational cost in the curved mode and to improve the accuracy of 3-D reconstruction around the spine of the book.

References

1. R. Szeliski: "Image Mosaicing for Tele-Reality Applications," Proc. IEEE Workshop on Applications of Computer Vision, pp. 230–236, 1994.
2. N. Chiba, H. Kano, M. Higashihara, M. Yasuda and M. Osumi: "Feature-based Image Mosaicing," Proc. IAPR Workshop on Machine Vision Applications, pp. 5–10, 1998.
3. S. Takeuchi, D. Shibuichi, N. Terashima and H. Tominage: "Adaptive Resolution Image Acquisition Using Image Mosaicing Technique from Video Sequence," Proc. IEEE Int. Conf. on Image Processing, vol. I, pp. 220–223, 2000.
4. C.T. Hsu, T.H. Cheng, R.A. Beuker and J.K. Hong: "Feature-based Video Mosaicing," Proc. IEEE Int. Conf. on Image Processing, Vol. II, pp. 887–890, 2000.
5. M. Lhuillier, L. Quan, H. Shum and H.T. Tsui: "Relief Mosaicing by Joint View Triangulation," Proc. IEEE Int. Conf. on Computer Vision and Pattern Recognition, vol. I, pp. 785–790, 2001.
6. P.F. McLauchlan and A. Jaenicke: "Image Mosaicing Using Bundle Adjustment,", Image and Vision Computing, 20, pp. 751–759, 2002.
7. D.W. Kim and K.S. Hong: "Fast Global Registration for Image Mosaicing," Proc. IEEE Int. Conf. on Image Processing, II, pp. 295–298, 2003.
8. U. Bhosle, S. Chaudhuri and S.D. Roy: "A Fast Method for Image Mosaicing Using Geometric Hashing," IETE J. of Research, Special Issue on Visual Media Processing, Vol. 48, No. 3-4, pp. 317–324, 2002.
9. P. Grattono and M. Spertino: "A Mosaicing Approach for the Acquisition and Representation of 3D Painted Surfaces for Conservation and Restoration purpose," Machine Vision and Applications, Vol. 15, No. 1, pp. 1–10, 2003.
10. T. Sato, M. Kanbara, N. Yokoya and H. Takemura: "Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo Using a Hand-held Video Camera," Int. J. of Computer Vision, Vol. 47, No. 1-3, pp. 119–129, 2002.
11. C. Harris and M. Stephens: "A Combined Corner and Edge Detector," Proc. Alvey Vision Conf., pp. 147–151, 1988.
12. M.A. Fischler and R.C. Bolles: "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography,", Communications of the ACM, Vol. 24, No. 6, pp. 381–395, 1981.
13. B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: "Bundle Adjustment a Modern Synthesis," Proc. Int. Workshop on Vision Algorithms, pp. 298–372, 1999.
14. K. Kanatani: "Geometric Information Criterion for Model Selection,", Int. J. of Computer Vision, Vol. 26, No. 3, pp. 171–189, 1998.
15. R.Y. Tsai: "An Efficient and Accurate Camera Calibration Technique for 3D Machine Vision," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 364–374, 1986.