MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Towards Bendable Augmented Maps

Sandy Martedi* and Hideo Saito †

Graduate School of Science and Technology
Keio University, Japan

Abstract

Three different kinds assumptions are often used for representing the shape of a paper:rigid, foldable and nonrigid. Nonrigid surface detection is intensively explored as challenging topics which addresses two problems: recovering the paper shape and estimating the camera pose. The state-of-the-art researches try to solve both problems robustly for real time purpose. We propose an augmented reality application that use a nonrigid detection method to recover the shape of the bendable paper using dots as keypoints and estimate the camera pose simultaneously. Our approach recovers the multi-planarity of the paper as the initial shape and iteratively approximates the surface shape. The multi-planarity is estimated by using the tracking by descriptor update method that uses the correspondence between captured and reference keypoints. We then optimize the shape using the progressive finite newton optimization method.

1 Introduction

Recently, combining paper maps and digital geographic data so called augmented maps have been intensively explored. Early augmented maps were based on overlaying 3D geographic models using ARToolKit [7, 3]. In these systems, the user can interact with the models by watching them from arbitrary views through a head-mounted display (HMD). In our previous work, we developed 2D maps with intersection dots printed in order to perform semantic registration between the maps and 3D geographic models [21].

A projector-camera-based table-top system focusing on 2D visualization had been developed by Reitmayr *et al.* [19]. In this system, geographic animations are projected onto the table and operated using a personal digital assistant (PDA). In outdoor use, positioning by global positioning system (GPS) is used as a trigger to find the user's position on a map [14, 15]. We have proposed a folded surface model and apply it in augmented maps to enrich its visualization and interaction [13]. In our proposal, we can fold a paper map and the system is able to overlay the 3D city model on the folded map.

In this paper, we present our preliminary work on developing bendable augmented maps. This work extends our previous work entitled folded augmented maps using standard paper maps and semantical dots [13]. First, keypoint correspondences between an input image and a reference map are first established. From these correspondences, multiple planes are detected by iterative homography computation because

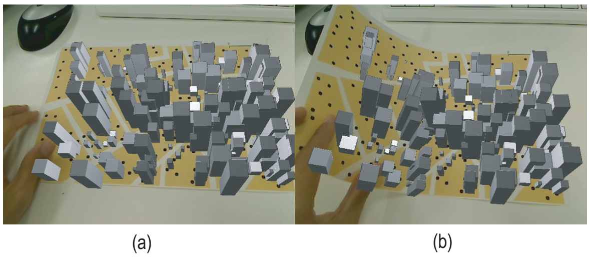

Figure 1. Application overview. (a). An augmented map (b). Bendable augmented map. The 3D building model is overlaid on top of bended paper map.

we use multiple planes as the initial shape of the optimization. We then optimize our shape by applying the progressive finite newton optimization [22]. We then overlay the 3D building model onto the regularized map using the piecewise homography computation as shown in Fig. 1.

We design the setup of our application using a monocular camera and a display. It is difficult to recover the surface shape using only one camera and in the same time to estimate the camera pose. Moreover, the 3D deformation costs computation time and memory since the parameterizing the surface in in three dimensional domain. The challenge in our application will be the accuracy of the surface shape registration, estimating the camera pose and augmenting the 3D models on the 2D deformation in real time. Despite the existence of many researches on bended surface detection, our proposal is based on the tracking method that depends on the content of the paper which can be considered as a novel application on augmented reality.

The rest of the paper is organized as follows: the details of the previous augmented maps and augmentation on bended surfaces are explained in Section 2. Section 3 describes how we adapt our previous folded surface detection method to handle a bended surface. The implementation using a set of map data is presented in Section 4, and Section 5 concludes this paper.

2 Related Works

The conventional methods on augmented maps are based on the planar surface detection and fiducial markers [7, 3, 15, 9, 20]. Previously, we initiated a research on augmented maps that apply geographic data (scattered dots) as the marker [21]. However, we only consider a planar surface map. We continued to use the geographic data as the marker and proposed a foldable surface model [13]. We applied the folded surface model for developing foldable augmented maps. Our model is categorized between rigid and nonrigid sur-

*e-mail: sandy@hvrl.ics.keio.ac.jp
†e-mail: saito@hvrl.ics.keio.ac.jp

face model. However, our model can not be applied to a bended paper.

Depending on the material, a paper can deform and then can be considered as nonrigid surface. The early work on shape modeling paper-like surface by taking into account the boundary points of the surface has been done by Kergosien *et al.* [10]. Bo and Wang used the geodesic of the surface instead of the boundary points to estimate the shape more accurately and intuitively [2]. Bellile *et al.* developed a method for estimating the nonrigid shape in occluded conditions [6]. Lee *et al.* took a different approach by placing LED markers on the surface and used a tracker camera to project virtual contents on the surface [11].

Pilet *et al.* demonstrated a fast registration of nonrigid surface and illumination handling for augmented reality application for 2D deformation [18]. Their work requires many feature points for determining the deformation of a surface from the planar form. To collect an enough number of feature points, they apply a wide baseline matching method so called randomized tree [12]. This method requires more iteration on the optimization process than the method that was proposed by Zhu *et al.* [22] which use finite optimization.

Hoi and Lyu introduced a method to approximate a surface shape without predefining a triangulated mesh model and use texture information instead [8]. Their method can handle a sharply folding and severely bended surfaces. Nonrigid or paper-like surface had been applied for many virtual reality and computer vision applications [5, 4, 1, 16, 17]. However, none of them handle the 3D deformable surface and overlaying 3D contents over the surface.

Our proposed method is an the extension of foldable augmented maps [13] because we assume our folded model can be applied for approximating the bending. The final goal of this research is to be able to approximate the bended surface and augment 3D model on top of a bended surface especially for paper maps in real time.

3 Method

We propose a method that combine our folded surface model in Foldable Augmented Maps [13] with the regularization using the progressive finite newton optimization [22] . We use the folded surface model as the initial shape for the optimization process. We choose the progressive finite newton optimization method because the number of iteration is fixed and applicable for developing a real time application. The overview of our method is illustrated on Fig. 2.

3.1 Initial shape

We modify the folded surface detection into a bended surface detection method by adding a shape optimization process. We use the same flow of our previous method from the keypoint extraction to the multiple plane detection. However, we exclude the folded and nonfolded state detection. Instead, in every frame when two planar surfaces are detected (Fig. 3), the folded surface is then estimated. Because the two detected planes are in the different coordinate system, we unify them into one coordinate system. Here, we choose the coordinate system from the first detected plane.

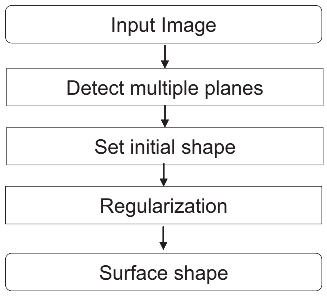

Figure 2. Proposed method in the application. An initial shape is formed by detecting two planes. The initial shape is then regularized by the progressive finite newton optimization method.

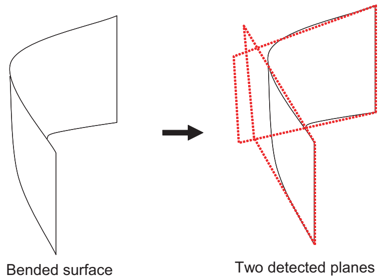

Figure 3. Detecting two planes from bended surface. From a bended surface, we detect two nonparallel planar surfaces.

We create a mesh over the folded surface as illustrated in Fig. 4. This mesh is the initial shape for the regularization step of the fast nonrigid surface detection method.

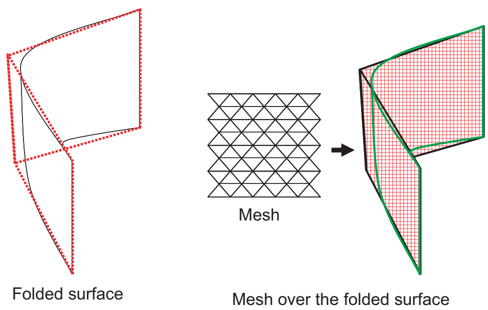

Figure 4. Folded mesh. We create a mesh over the folded surface. The mesh is then used as the initial shape for the optimization method (regularization).

3.2 Regularization using the progressive finite newton optimization [22]

We apply the regularization step of the nonrigid surface detection using features to deform the mesh (initial shape) so that it fits to the paper shape. Suppose we have vector S that represents a mesh of a bended surface, we minimize the two energies: regularization term and correspondence term as the object function of S defined as

$$\varepsilon(S) = \lambda_D \varepsilon_D(S) + \varepsilon_C(S), \tag{1}$$

where $\varepsilon_D(S)$ represents the regularization term that constrains the length of a line along the surface should be constant in any deformation, $\varepsilon_C(S)$ is the correspondence term that takes into account of the difference between the keypoints in the input image and the keypoints in bended surface and λ_D is a constant.

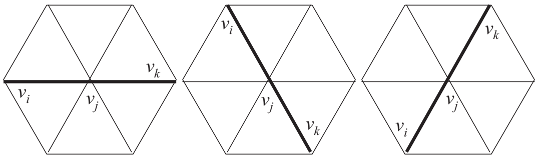

Figure 5. Index triplets. Set of index triplets are collected from the mesh. These triplets keep the shape of the surface.

To regularize the surface, we model the surface as collection points (mesh) and we collect the set of triplets E which consists of many index triplets (v_i, v_j, v_k) from the mesh (Fig. 5) and apply the following condition

$$\forall(i, j, k) \in E : v_i - v_j = v_j - v_k, \tag{2}$$

$\varepsilon_D(S)$ can be written in matrix form as

$$\varepsilon_D(S) = \frac{1}{2}(X^T K'^T K' X + Y^T K'^T K' Y), \tag{3}$$

where K' is a matrix $m \times n$. m is the number of triplets and n is the number of points in the mesh. K' is filled with 0 except the columns that are associated with the index triplets are filled with 1, -2 and 1 as shown in the following example

$$K' = \begin{bmatrix} 1 & 0 & -2 & 1 & 0 & 0 \ldots \\ 0 & 1 & 0 & -2 & 0 & 1 \ldots \\ \ldots & & & & & \end{bmatrix}. \tag{4}$$

For computing the correspondence term, we use the correspondence points that are retrieved by the keypoints extraction and matching. We then build a matrix for storing the barycentric coordinate (ξ_i, ξ_j, ξ_k) of each correspondence points. The number of element in the matrix is the same as the number of points in the mesh. For each correspondence points, the barycentric coordinate of reference keypoint is calculated and inserted into the matrix t using the following values $t_i = \xi_i$, $t_j = \xi_j$, $t_k = \xi_k$. The other values is set into zero.

For the sake of the simplicity, we put the final simplification formula of the optimization $\varepsilon(S)$ in Eq. 5 and 6. In each iteration, the mesh is updated using the result of finite newton step written as the following equations

$$s_x = (\lambda_r K + A)^{-1} b_x, \tag{5}$$

and

$$s_y = (\lambda_r K + A)^{-1} b_y \tag{6}$$

which can be solved using LU decomposition. The $A \in R^{N \times N}$ and $b \in R^{2N}$ are matrices that computed in each step using the following equations:

$$A = \sum_{m \in M_1} \frac{1}{\sigma^n} t t^\top \tag{7}$$

and

$$b = \begin{bmatrix} b_x \\ b_y \end{bmatrix} = \sum_{m \in M_1} \frac{1}{\sigma^n} \begin{bmatrix} ut \\ vt \end{bmatrix} \tag{8}$$

which σ value is divided by 2 in every step. The correspondence points are filtered for the next step by calculating the error mapping after the deformation. The points of which the error mapping is bigger than σ^2 are removed.

The method works for 2D deformation. To augment the 3D model, we estimate the local orientation using the piecewise homography orientation. The surface is divided into smaller local region. The homography of each region is estimated using the correspondence of from points in the mesh. The camera pose then can be computed from the homography.

4 Implementation

4.1 Map data

Our map is generated from Geographical Information System (GIS) data. It is a database of vector for storing geographical data such as road, building, landmark, terrestrial, and satellite data for making maps. Our map contains the background data and geographic symbols printed on a map as illustrated in Fig. 6.

The offline procedure for multiple detection and tracking is similar to the foldable augmented maps [13]. The reader is suggested to refer the foldable augmented maps method for the implementation detail.

Figure 6. A part of the map from GIS. We extract the center of each black symbol as keypoints using color segmentation.

4.2 Results

After the regularization step, the surface shape is bended and forms curved along the bended area (Fig. 7). In Fig. 7, the red line shows that the mesh deforms along the surface shape and forms a curved surface. The 3D building model then can be overlaid on top of the bended surface as shown in Fig. 8.

In the implementation, the value of parameter λ_r is 1×10^{-4} and the initial σ is 80. The initial shape estimation, the regularization, and 3D model augmentation are performed in approximately 4fps. This computation cost can be reduced to achieve our final goal for building real-time application. More robust and fast improvement is necessary for achieving bendable augmented maps in real time. We also realize that we can improve our folding approach to handle more complicated surface in the future.

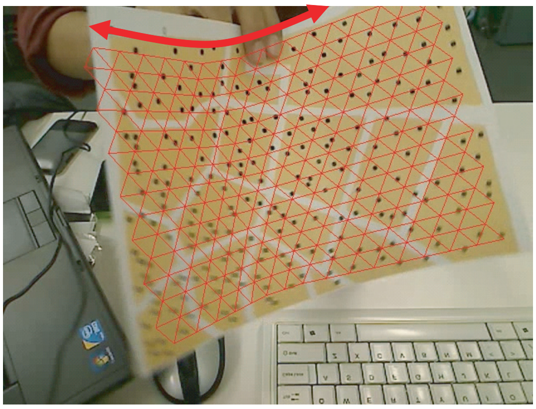

Figure 7. Bended shape. The initial shape is regularized.

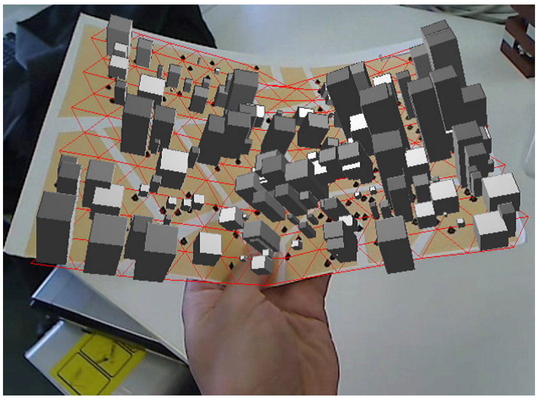

Figure 8. 3D building model is overlaid onto the bended map.

5 Conclusions and Future Works

We presented our work on developing augmented maps application on a bended surface. In our application, we apply a folded surface model as the initial shape for the regularization of nonrigid surface detection.

We presented our implementation using a map data and showed bended surface results. We are planning to optimize our method to make it more robust and reduce the computational cost. A set of experiments is also necessary to estimate the accuracy of our propose method.

Acknowledgment

This work is supported in part by a Grant-in-Aid for the Global Center of Excellence for high-Level Global Cooperation for Leading-Edge Platform on Access Spaces from the Ministry of Education, Culture, Sport, Science, and Technology in Japan and Grant-in-Aid for JSPS Fellows.

References

[1] N. G. Ali, A. Z, R. Duraiswami, and L. S. Davis. Structure of applicable surfaces from single views. In *Proc. ECCV*, pages 482–496, 2004.

[2] P. Bo and W. Wang. Geodesic-controlled developable surfaces for modeling paper bending. In *Proc. Eurographics*, 2007.

[3] J. Bobrich and S. Otto. Augmented maps. In *ISPRS*, 2002.

[4] M. S. Brown and W. B. Seales. Image Restoration of Arbitrarily Warped Documents. *PAMI*, 26(10):1295–1306, 2004.

[5] M. Emori and H. Saito. Texture overlay onto deformable surface using HMD. In *Proc. VR*, pages 221–222. IEEE, 2004.

[6] V. Gay-Bellile, A. Bartoli, and P. Sayd. Direct estimation of nonrigid registrations with image-based self-occlusion reasoning. *PAMI*, 32:87–104, 2010.

[7] N. R. Hedley, M. Billinghurst, L. Postner, R. May, and H. Kato. Explorations in the use of augmented reality for geographic visualization. *Presence*, 11:119–133, 2002.

[8] S. Hoi and M. Lyu. Nonrigid shape recovery by Gaussian process regression. In *Proc. ICCV*, pages 1319–1326. Ieee, June 2009.

[9] H. Kato and M. Billinghurst. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In *Proc. IWAR*, pages 85–94, 1999.

[10] Y. L. Kergosien, H. Gotoda, and T. L. Kunii. Bending and creasing virtual paper. *CG&A*, 14(1):40–48, 1994.

[11] J. C. Lee, S. E. Hudson, and E. Tse. Foldable interactive displays. In *Proc. UIST*, pages 287–290, 2008.

[12] V. Lepetit, P. Lagger, and P. Fua. Randomized Trees for Real-Time Keypoint Recognition. In *Proc. CVPR*, volume 2, pages 775–781. Ieee, 2005.

[13] S. Martedi, H. Uchiyama, G. Enriquez, H. Saito, T. Miyashita, and T. Hara. Foldable Augmented Maps. In *Proc. ISMAR*, pages 65–72, 2010.

[14] A. Morrison, A. Oulasvirta, P. Peltonen, S. Lemmela, G. Jacucci, G. Reitmayr, J. Näsänen, and A. Juustila. Like bees around the hive: a comparative study of a mobile augmented reality map. In *Proc. CHI*, pages 1889–1898, 2009.

[15] V. Paelke and M. Sester. Augmented paper maps: Exploring the design space of a mixed reality system. *ISPRS*, 65:256–265, 2010.

[16] M. Perriollat and A. Bartoli. A Quasi-Minimal Model for Paper-Like Surfaces. In *Proc. CVPR*, volume 0, pages 1–7, 2007.

[17] M. Perriollat, R. Hartley, and A. Bartoli. Monocular Template-based Reconstruction of Inextensible Surfaces. In *Proc. BMVC*, 2008.

[18] J. Pilet, V. Lepetit, and P. Fua. Fast Non-Rigid Surface Detection, Registration and Realistic Augmentation. *IJCV*, 76(2):109–122, 2008.

[19] G. Reitmayr, E. Eade, and T. Drummond. Localisation and interaction for augmented maps. In *Proc. ISMAR*, pages 120–129, 2005.

[20] M. Rohs, J. Schoning, A. Kruger, and B. Hecht. Towards real-time markerless tracking of magic lenses on paper maps. In *adjunct Proc. Pervasive*, pages 69–72, 2007.

[21] H. Uchiyama, H. Saito, M. Servieres, and G. Moreau. AR GIS on a physical map based on map image tetrieval using LLAH tracking. In *Proc. MVA*, pages 382–385, 2009.

[22] J. Zhu, M. R. Lyu, and T. S. Huang. A fast 2D shape recovery approach by fusing features and appearance. *PAMI*, 31(7):1210–24, July 2009.