Shape from Contour for the Digitization of Curved Documents

Frédéric Courteille, Jean-Denis Durou and Pierre Gurdjos

IRIT, Toulouse, France
E-mail: {courteil,durou,gurdjos}@irit.fr

Abstract. We are aiming at extending the basic digital camera functionalities to the ability to simulate the flattening of a document, by virtually acting like a flatbed scanner. Typically, the document is the warped page of an opened book. The problem is stated as a computer vision problem, whose resolution involves, in particular, a 3D reconstruction technique, namely shape from contour. Assuming that a photograph is taken by a camera in arbitrary position or orientation, and that the model of the document surface is a generalized cylinder, we show how the corrections of its geometric distortions, including perspective distortion, can be achieved from a single view of the document. The performances of the proposed technique are assessed and illustrated through experiments on real images.

1 Introduction

The digitization of documents currently knows an increasing popularity, because of the expansion of Internet browsing. The traditional process, which uses a flatbed scanner, is satisfactory for flat documents, but is unsuitable for curved documents like for example a thick book, since some defects will appear in the digitized image. Several specific systems have been designed, but such systems are sometimes intrusive with regard to the documents and, before all, they cannot be referred to as consumer equipments. An alternative consists in *simulating* the flattening of curved documents *i.e.*, in correcting the defects of images provided by a flatbed scanner or a digital camera. In this paper, we describe a new method of simulation of document flattening which uses one image taken from an arbitrary angle of view, and not only in frontal view (as this is often the case in the literature). The obtained results are very encouraging.

In Section 2, different techniques of simulation of document flattening are reviewed. In Section 3, a new 3D-reconstruction method based on the so-called shape-from-contour technique is discussed. In Section 4, this method is applied to the flattening simulation of curved documents. Finally, Section 5 concludes our study and states several perspectives.

2 Techniques of Simulation of Document Flattening

One can address a purely 2D-deformation of the image, in order to correct its defects according to an *a priori* modelling of the flattened document [1, 2]. In [3–

5], the characters orientation is estimated, so as to straighten them out. In all these papers, the results are of poor quality because, if the lines of text are rather well uncurved, the narrowing of the characters near the binding is not well corrected. In [6], a judicious 2D-deformation is introduced, which considers that the contour of each page becomes rectangular after flattening: the results are nice, but a “paper checkerboard pattern” must be placed behind the document to force both them having the same 3D-shape, and this makes the process rather complicated. In order to successfully simulate the document flattening, it is necessary to compute its surface shape.

Stereoscopy aims at reconstructing the shape of a document from several photographs taken from different angles of view. In [7], the CPU time is very high when dealing with two images of size 2048×1360. On the other hand, this technique works well only if the stereo ring has been intrinsically and extrinsically calibrated. **Stereophotometry** requires several photographs taken from the same angle of view, but under different lightings. A modelling linking the image greylevel to the orientation of the surface is then used. It has been implemented by Cho *et al.* [8]. The results are of mean quality since, for photographs taken at close range, perspective should be taken into account. **Structured-lighting systems** make also use of two photographs taken under two different lightings, knowing that, for one of the photographs, a pattern is projected onto the document [9, 10]. The deformation of the pattern in this image gives some information on the surface shape. A second photograph is required, in order to avoid possible artefacts of the pattern in the flattening simulation. The best results using this technique are presented in [11], but they make use of a dedicated imaging system. **Shape-from-texture** has also been used. An *a priori* knowledge on the document assumes that the text is printed along parallel lines, which is the case for most documents. Hence, the shape of the document surface may be deduced from the deformation of the lines of text in the image. This technique, that has been implemented by Cao *et al.* [12] on photographs of cylindrical books taken in frontal view, works well and quickly (some seconds on an image of size 1200×1600). Its crucial step consists in extracting the lines of text. In [13], it is generalized to any angles of view. Nevertheless, the latter work assumes that the lines of text are also equally spaced. The oldest contribution to the simulation of document flattening uses the **shape-from-shading** technique. Wada *et al.* [14] take advantage of the greylevel gradation in the non-inked areas of a scanned image, in order to estimate the slope of the document surface. This idea has been resumed and improved by Tan *et al.* [15], whose results are of good quality, and also by Courteille *et al.* [16]. The latter paper provides two noticeable improvements: a digital camera replaces the flatbed scanner, so as to accelerate the digitization process; a new modelling of shape-from-shading is stated, that takes perspective into account. Finally, the **shape-from-contour** technique may be used *i.e.*, the deformation of the contours in the image provides information on the surface shape. This technique has been implemented in [17, 6, 18] on photographs of cylindrical books taken in frontal view. In [19], it is generalized to any applicable surfaces: the results are of mean quality, but this last contribution is worth of mention, since it reformulates the problem elegantly, as a set of differential equations.

The method of simulation of document flattening that we discuss in this paper uses the shape-from-contour technique to compute the shape of the document from one photograph taken from an arbitrary angle of view.

3 3D-reconstruction using Shape-from-contour

In the most general situation, shape-from-contour (SFC) is an ill-posed problem: the same contours may characterize a lot of different scenes. To make the problem well-posed, it is necessary to make some assumptions. In [20], the scene is supposed to be cylindrically-symmetrical. In the present work, we suppose that its surface is a generalized cylinder having straight meridians.

3.1 Notations and Choice of the Coordinate Systems

The photographic bench is represented in Fig. 1: f refers to the focal length and C to the optical center; the axis Cz coincides with the optical axis, so that the equation of the image plane Π_i is $z = f$. The digital camera is supposed to lie in an arbitrary position with regard to the book, apart from the fact that the optical axis must be non-parallel to the binding. Hence, the vanishing point F of the binding direction can be located at infinity, but it is separate from the principal point O. Thus, we can define the axis Ox by the straight line FO, in such a way that $\overline{FO} > 0$. Furthermore, we complete Ox by an axis Oy such that Oxy is an orthonormal coordinate system of the image plane Π_i (cf. Fig. 1). It is convenient to define two 3D orthonormal coordinate systems: $\mathcal{R}_o = Cxyz$ and $\mathcal{R}_p = Cuvw$, where Cu coincides with the straight line FC, which is parallel to the binding, and where Cv coincides with Cy. Since Cu intersects Π_i at $F \neq O$, it follows that Cw intersects Π_i at a point Ω which also lies on the axis Ox. We introduce the orthonormal coordinate system $\mathcal{R}_i = \Omega xy$ of Π_i. The angle between Cz and Cw is denoted α (cf. Fig. 1). The case $\alpha = 0$ corresponds to the frontal orientation of the camera. Denoting $c = \cos\alpha$, $s = \sin\alpha$ and $t = \tan\alpha$, it can easily be stated that $\overline{O\Omega} = t\,f$, $\overline{FO} = f/t$ and $\overline{F\Omega} = f/(c\,s)$. Finally, we denote Π_r the plane orthogonal to Cw and containing the binding, whose equation is $w = \delta$.

3.2 Relations between Object and Image

Let P be an object point, whose coordinates are (X, Y, Z) w.r.t. $\mathcal{R}_o$ and (U, V, W) w.r.t. $\mathcal{R}_p$. The transformation rules between these two sets of coordinates are:

$$X = c\,U + s\,W, \tag{1a}$$

$$Y = V, \tag{1b}$$

$$Z = -s\,U + c\,W. \tag{1c}$$

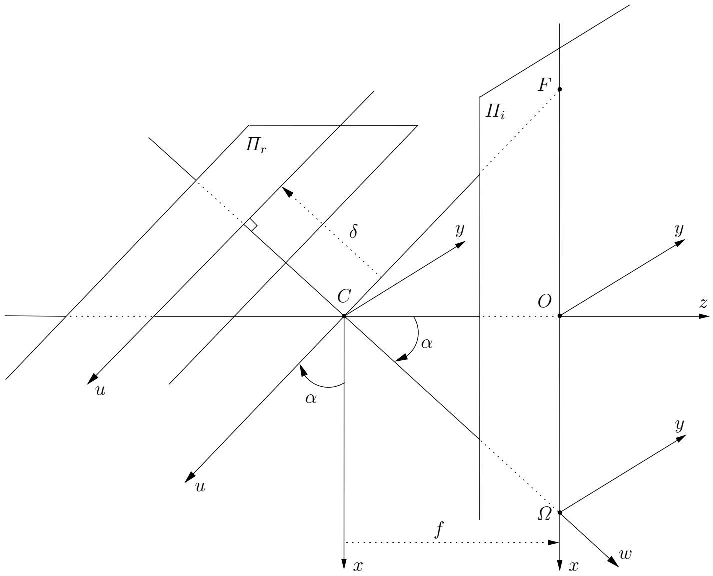


Fig. 1. Representation of the photographic bench.

Let Q be the image point conjugated to P, whose coordinates are (u, v) w.r.t. $\mathcal{R}_i$. Using the perspective projection rules, we obtain:

$$u = \frac{f}{Z}\,X - t\,f, \tag{2a}$$

$$v = \frac{f}{Z}\,Y. \tag{2b}$$

Denoting $f' = f/c$, the equations (1a), (1b), (1c), (2a) and (2b) give:

$$u = f'\,\frac{U}{-s\,U + c\,W}, \tag{3a}$$

$$v = f\,\frac{V}{-s\,U + c\,W}. \tag{3b}$$

Let us define the "pseudo-image" $\overline{Q}$ of P as the image of the orthogonal projection $\overline{P}$ of P on Π_r. As the coordinates of $\overline{P}$ are (U, V, δ) w.r.t. $\mathcal{R}_p$, the coordinates $(\overline{u}, \overline{v})$ of $\overline{Q}$ w.r.t. $\mathcal{R}_i$ are:

$$\overline{u} = f'\,\frac{U}{-s\,U + c\,\delta}, \tag{4a}$$

$$\overline{v} = f\,\frac{V}{-s\,U + c\,\delta}. \tag{4b}$$

Dividing (4a) by (3a) and (4b) by (3b), we find:

$$\frac{\overline{u}}{u} = \frac{\overline{v}}{v}. \tag{5}$$

This equality means that the image points Q, $\overline{Q}$ and Ω are aligned *i.e.*, Ω is the vanishing point of the direction orthogonal to Π_r. In a general way, the knowledge of an image point Q does not allow us to compute its conjugated object point P. But, if we also know the pseudo-image $\overline{Q}$ associated to P, then we can compute the coordinates of P. Actually, it can be deduced from (3a), (3b) and (4a):

$$U = \delta \, \frac{c \, \overline{u}}{f' + s \, \overline{u}}, \tag{6a}$$

$$V = \delta \, \frac{\overline{u}}{u} \, \frac{v}{f' + s \, \overline{u}}, \tag{6b}$$

$$W = \delta \, \frac{\overline{u}}{u} \, \frac{f' + s \, u}{f' + s \, \overline{u}}. \tag{6c}$$

Note that (6c) gives $W = \delta$ when $\overline{u} = u$ *i.e.*, for image points such that $\overline{Q} = Q$. For a given image point Q, if the location of the associated pseudo-image $\overline{Q}$ is known, the coordinates (U, V, W) of the conjugated object point P can be computed using (6a), (6b) and (6c). Nevertheless, in a general way, the location of the pseudo-image $\overline{Q}$ on the straight line ΩQ is unknown.

3.3 Additional Assumptions

Within the framework of our application, the scene is a book. We make two additional assumptions:

- A_1 - The flattened pages are rectangular.
- A_2 - The pages of the book are curved in a such way that they form a generalized cylinder.

As the surface of the book is a generalized cylinder, the lower and upper contours are located in two planes which are orthogonal to the binding. Thanks to this property, the SFC problem becomes well-posed.

As the binding belongs to the plane Π_r, its image and its pseudo-image coincide. Under the assumptions A_1 and A_2, it is easy to predict that the pseudo-image of the upper and lower contours of the book (whose images are called C_u and C_l) are the straight lines L_u and L_l, parallel to Ωy and passing through the ends B_u and B_l of the image B of the binding (cf. Fig. 2). Let Q be an image point of coordinates (u, v) w.r.t. $\mathcal{R}_i = \Omega xy$. We call θ the polar angle in the coordinate system Fxy, and Q_u and Q_l the two image points located on C_u and C_l, which have the same polar angle θ as Q (cf. Fig. 2). Considering the assumptions A_1 and A_2, and knowing that F is the vanishing point of the binding direction, the object point P conjugated to Q has the same coordinate W as the two object points conjugated to Q_u and Q_l. We denote $u_u(\theta)$ and $u_l(\theta)$ the abscissas of Q_u and Q_l w.r.t. $\mathcal{R}_i$. Finally, we denote θ_B the polar angle of B. According to these notations, $u_u(\theta_B)$ and $u_l(\theta_B)$ are the abscissas of the

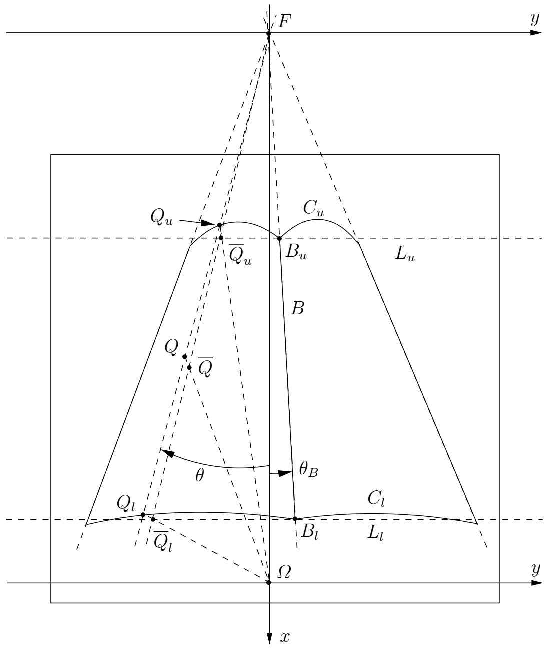


Fig. 2. Geometric construction of the pseudo-image $\overline{Q}$ associated to an image point Q.

pseudo-images $\overline{Q}_u$ and $\overline{Q}_l$ associated to Q_u and Q_l, w.r.t. $\mathcal{R}_i$. Hence, if we denote $f'' = f'/s$, the equation (6c) gives, when applied to Q_u and to Q_l:

$$W = \delta \, \frac{u_u(\theta_B)}{u_u(\theta)} \, \frac{f'' + u_u(\theta)}{f'' + u_u(\theta_B)}, \tag{7a}$$

$$W = \delta \, \frac{u_l(\theta_B)}{u_l(\theta)} \, \frac{f'' + u_l(\theta)}{f'' + u_l(\theta_B)}. \tag{7b}$$

From one of both these expressions of W, we can deduce the other coordinates U and V of P, solving the system of two equations (3a) and (3b).

Considering the expressions (7a) and (7b) of W and the equations (3a) and (3b), it appears that the computation of the shape of the document requires the knowledge of some parameters: the focal length f, the viewing angle α and the location of the principal point O. On the other hand, the parameter δ can be chosen arbitrarily, because the shape of the document can be computed only up to a scale factor.

4 Application to the Simulation of Document Flattening

Due to lack of space, we do not show any result on synthetic images, but only on real images. The left column of Fig. 3 shows three photographs of the same

Fig. 3. Three photographs of the same book taken from three different angles of view, and the flattening simulations obtained from each of them: (a-b) $\alpha = 1.5°$; (c-d) $\alpha = 20.38°$; (e-f) $\alpha = 40.54°$.

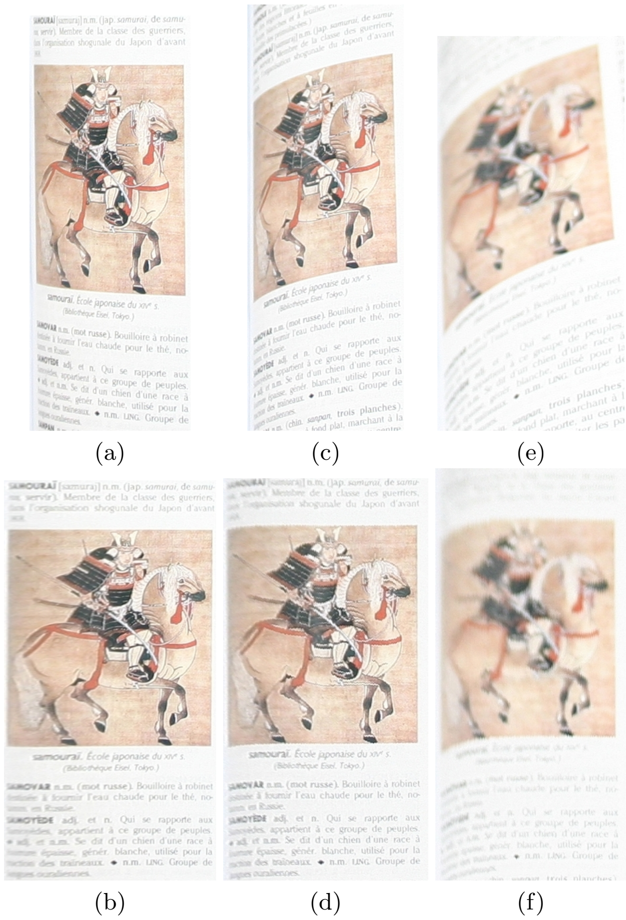
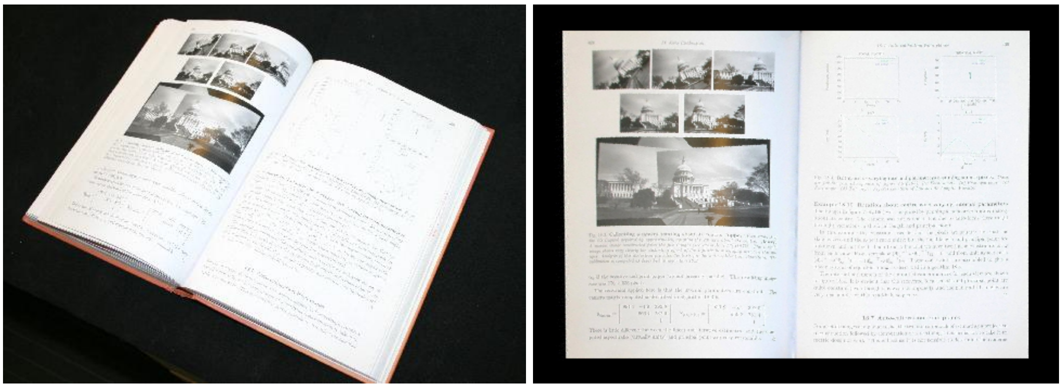


Fig. 4. Zooms on a common area of the six images of Fig. 3.

book taken from three different angles of view. For each of them, the shape of the document is computed using the method described in Section 3. The flattening simulations are shown on the right column of Fig. 3, knowing that a generalized cylinder is particularly easy to flatten. The same six images are zoomed on the area near the binding which contains a picture representing a samurai (cf. Fig. 4). It appears that when the angle of view increases, the quality of the flattening simulation decreases but, even for a strong angle of view (cf. Fig. 3-e), the result remains acceptable (cf. Fig. 3-f).

A second result proves that our method works well even with the most general case of camera pose *i.e.*, when the optical axis is also tilted in the direction perpendicular to the axis of the cylinder (cf. Fig. 5).



(a) (b)

Fig. 5. Most general case of camera pose: (a) original and (b) flattened images.

5 Conclusion and Perspectives

In this paper, we generalize the 3D-shape reconstruction of a document from its contours, as it had previously been stated in frontal view in [17, 6, 18], to the case of an arbitrary view. We validate this result by simulating the flattening of curved documents taken from different angles of view. Even when the angle of view noticeably increases, the quality of the result remains rather good, in comparison with other results in the literature that are obtained under similar conditions.

In the present state of our knowledge, the focal length and the location of the principal point have to be known. As a first perspective, we aim at generalizing the 3D-shape reconstruction to the case of an uncalibrated camera. In addition, when the angle of view is too large, then focusing blur occurs, which inevitably restricts the quality of the flattening simulation. Rather than enduring this defect, it could be interesting to correct it, knowing that the 3D-shape of the document could allow us to predict the focusing blur magnitude.

References

1. Tang, Y.Y., Suen, C.Y.: Image Transformation Approach to Nonlinear Shape Restoration. IEEE Trans. Syst. Man and Cybern.**23**(1) (1993) 155–172
2. Weng, Y., Zhu, Q.: Nonlinear Shape Restoration for Document Images. In: Proc. IEEE Conf. Comp. Vis. and Patt. Recog., San Francisco, California, USA(1996) 568–573
3. Zhang, Z., Tan, C.L.: Recovery of Distorted Document Images from Bound Volumes. In: Proc. 6^{th} Int. Conf. Doc. Anal. and Recog., Seattle, Washington, USA(2001) 429–433

4. Lavialle, O., Molines, X., Angella, F., Baylou, P.: Active Contours Network to Straighten Distorted Text Lines. In: Proc. IEEE Int. Conf. Im. Proc. (vo. III), Thessaloniki, Greece(2001) 748–751
5. Wu, C.H., Agam, G.: Document Image De-Warping for Text/Graphics Recognition. In: Proc. Joint IAPR Int. Worksh. Syntact. Struct. Patt. Recog. and Stat. Patt. Recog.. Volume 2396 of LNCS., Windsor, Canada (2002) 348–357
6. Tsoi, Y.C., Brown, M.S.: Geometric and Shading Correction for Images of Printed Materials: A Unified Approach Using Boundary. In: Proc. IEEE Conf. Comp. Vis. and Patt. Recog. (vo. I), Washington, D.C., USA(2004) 240–246
7. Yamashita, A., Kawarago, A., Kaneko, T., Miura, K.T.: Shape Reconstr. and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System. In: Proc. 17^{th} Int. Conf. Patt. Recog. (vo. I), Cambridge, UK(2004) 482–485
8. Cho, S.I., Saito, H., Ozawa, S.: A Divide-and-conquer Strategy in Shape from Shading Problem. In: Proc. IEEE Conf. Comp. Vis. and Patt. Recog., San Juan, Puerto Rico(1997) 413–419
9. Doncescu, A., Bouju, A., Quillet, V.: Former books digital processing: image warping. In: Proc. IEEE Worksh. Doc. Im. Anal., San Juan, Puerto Rico(1997) 5–9
10. Brown, M.S., Seales, W.B.: Document Restoration Using 3D Shape: A General Deskewing Algorithm for Arbitrarily Warped Documents. In: Proc. 8^{th} IEEE Int. Conf. Comp. Vis. (vo. I), Vancouver, Canada (2001) 367–375
11. Sun, M., Yang, R., Yun, L., Landon, G., Seales, W.B., Brown, M.S.: Geometric and Photometric Restoration of Distorted Documents. In: Proc. 10^{th} IEEE Int. Conf. Comp. Vis. (vo. II), Beijing, China(2005) 1117–1123
12. Cao, H., Ding, X., Liu, C.: Rectifying the Bound Document Image Captured by the Camera: A model Based Approach. In: Proc. 7^{th} Int. Conf. Doc. Anal. and Recog., Edinburgh, UK(2003) 71–75
13. Liang, J., DeMenthon, D., Doermann, D.: Flattening Curved Documents in Images. In: Proc. IEEE Conf. Comp. Vis. and Patt. Recog. (vo. II), San Diego, California, USA(2005) 338–345
14. Wada, T., Ukida, H., Matsuyama, T.: Shape from Shading with Interreflections Under a Proximal Light Source: Distortion-Free Copying of an Unfolded Book. Int. J. Comp. Vis.**24**(2) (1997) 125–135
15. Tan, C.L., Zhang, L., Zhang, Z., Xia, T.: Restoring Warped Document Images through 3D Shape Modeling. IEEE Trans. PAMI**28**(2) (2006) 195–208
16. Courteille, F., Crouzil, A., Durou, J.D., Gurdjos, P.: Shape from Shading for the Digitization of Curved Documents. Mach. Vis. and Appl.(2007) (to appear).
17. Kashimura, M., Nakajima, T., Onda, N., Saito, H., Ozawa, S.: Practical Introduction of Image Processing Technology to Digital Archiving of Rare Books. In: Proc. Int. Conf. Sign. Proc. Appl. and Techn., Toronto, Canada (1998) 1025–1029
18. Courteille, F., Durou, J.D., Gurdjos, P.: Transform your Digital Camera into a Flatbed Scanner. In: Proc. 9^{th} Eur. Conf. Comp. Vis., 2^{nd} Works. Appl. Comp. Vis., Graz, Austria(2006) 40–48
19. Gumerov, N., Zandifar, A., Duraiswami, R., Davis, L.S.: Structure of Applicable Surfaces from Single Views. In: Proc. 8^{th} Eur. Conf. Comp. Vis. (vo. III). Volume 3022 of LNCS., Prague, Czech Republic(2004) 482–496
20. Colombo, C., Del Bimbo, A., Pernici, F.: Metric 3D Reconstruction and Texture Acquisition of Surfaces of Revolution from a Single Uncalibrated View. IEEE Trans. PAMI **27**(1) (2005) 99–114