Panoramic and 3D Computer Vision

Akihiko Torii[1] and Reinhard Klette[2]

[1] Czech Technical University, Prague, Czech Republic
[2] The University of Auckland, Auckland, New Zealand

Abstract. Image sensor technology provides in recent years the tools for wide-angle and high-resolution 3D recording, analysis and modeling of static or dynamic scenes, ranging from small objects such as artifacts in a museum to large-scale 3D models of a castle or 3D city maps, also allowing real time 3D data acquisition from a moving platform, such as in vision-based driver assistance. The paper documents a particular subfield of these developments by providing mappings of omnidirectional images (catadioptric or dioptric images) into panoramic images on a cylinder. The mappings are specified by using the geometries of omnidirectional cameras. Image examples illustrate potentials for projects in arts, science or technology.

Keywords: Panoramic imaging, omnidirectional viewing.

Cusco mural by Juan Bravo (1992, length: about 17 m)

1 Introduction

Panoramic photography is needed for recording a painting as shown above. We briefly review the geometry of catadioptric and dioptric cameras, which are relatively new tools for panoramic photography. Omnidirectional camera systems [2] have been developed to observe a 360-degree field of view; see Fig. 1 for an example. Popular omnidirectional imaging systems are catadioptric or dioptric camera systems.[1] A catadioptric camera system is constructed as a combination of a quadric mirror and a conventional camera [8,9,10]. The dioptric camera system is constructed with a specially designed refractor, which controls the angles of rays passing through the lens [9], as the optical lens of the camera. Cylindric panoramas are, for example, generated by rotating line cameras [7]; we consider

[1] *Catadioptric*: pertaining to, or involving both the reflection and the refraction of light; *dioptric*: relating to the refraction of light

F. Huang and R.-C. Wang (Eds.): ArtsIT 2009, LNICST 30, pp. 9–16, 2010.

Fig. 1. Omnidirectional cameras. Left: a digital camera with a hyperboloidal shaped mirror. Right: a fish-eye camera. [1,2]

single-projection center cylindric panoramas in the following (i.e., where the off-axis distance equals zero). Omnidirectional images are either of the catadioptric or dioptric kind.

Omnidirectional imaging systems are used for controlling mobile robots [1] or in vision-based driver assistance [5]. Therefore, image analysis is also required for omnidirectional images. Basically, analysis may happen directly in those catadioptric or dioptric images, or after a transform of those into traditional formats (preferably cylindric panoramas).

Besides robotics, another significant use of catadioptric or dioptric cameras is the generation of panoramic views from captured images. Panoramic images are essential for virtual reality applications, such as Google Earth [6]. Furthermore, panoramic images play an important role in the creation of artistic photographs [12]. In this paper, we identify mappings of catadioptric and dioptric images into images on a cylinder (i.e., cylindric images). See Fig. 2 for three examples.

Fig. 2. Upper row: original fish-eye images (180-degree fields of view, showing Prague castle and a group of people). Lower row: resulting panoramic images.

Omnidirectional images are required to be pre-processed before applying computer vision or image recognition tools (e.g., detectors and trackers of pedestrians, cars, and other objects [3,4]), because available tools are typically established and trained on perspective images. The aim is often to generate perspective cutouts from omnidirectional images, and to adapt common tools without special modifications. However, such cutouts completely lose the benefit of the large field of view (e.g., surveying a large area, without a need to switch between multiple images). Consequently, a mapping of the captured 360-degree field of view into a cylindric panorama is often the best solution, because single-center cylindric images possess perspective-like appearances while suppressing circular distortion due to catadioptric or dioptric projection.

The structure of this short paper is as follows. In Section 2, we first introduce central hyperbolic and parabolic camera models for characterizing common catadioptric cameras, and then we derive mappings from hyperbolic and parabolic images into cylindric images. In Section 3, we describe a fish-eye camera model as an example of dioptric cameras, and derive a mapping of fish-eye images into cylindric images. Finally, we show examples of cylindric images obtained from recorded hyperbolic or fish-eye images.

2 Catadioptric into Cylindric Images

Hyperbolic or parabolic mirrors are two commonly used in catadioptric cameras, and further mirrors are under consideration (e.g., to ensure a single projection center and uniform image resolution at the same time).

2.1 Hyperbolic into Cylindric Images

A central hyperbolic camera consists of a hyperboloidal shaped mirror and a conventional digital camera [8,10]. Such a hyperbolic camera has a unique center of projection at the focus of the hyperboloidal mirror by locating the projection center of the digital camera at the other focus of the hyperboloid, as illustrated on the left in Fig. 3. In this configuration, light rays which pass through the focus of the hyperboloid are reflected on the mirror, and the reflected rays pass through the other focus of the hyperboloid.

For deriving the transformation equation, we consider at first the case as illustrated on the left in Fig. 3. Let $\mathbf{C} = (0, 0, -2e)$ be the center of the pinhole camera. Letting the focus $\mathbf{F}$ of the hyperboloid C^2 be the origin of the world coordinate system, the hyperboloid is expressed in the quadric form:

$$(x, y, z, 1) \begin{pmatrix} \frac{1}{a^2} & 0 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{b^2} & -\frac{e}{b^2} \\ 0 & 0 & -\frac{e}{b^2} & -\frac{e^2}{b^2} + 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0$$

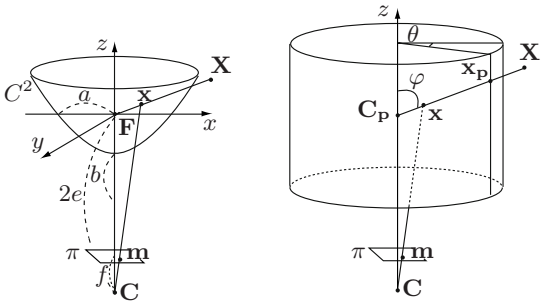

Fig. 3. Mapping from hyperbolic to cylindric images. Left: hyperbolic camera system. Right: hyperbolic and cylindric camera systems combined.

where $e = \sqrt{a^2 + b^2}$. The projection of a point $\mathbf{X} = (X, Y, Z)^\top$ in 3D space into a point $\mathbf{x} = (x, y, z)^\top$ on the hyperboloid C^2 is then expressed as $\mathbf{x} = \chi\mathbf{X}$ where

$$\chi = \frac{a^2}{b|\mathbf{X}| - eZ}$$

The projection of point $\mathbf{x}$ into a corresponding point $\mathbf{m} = (u, v)^\top$ on the image plane π is expressed as

$$\begin{pmatrix} \mathbf{m} \\ 1 \end{pmatrix} = \frac{1}{z + 2e} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$$

Accordingly, the total mapping, from $\mathbf{X}$ to $\mathbf{m}$, is formulated as

$$u = \frac{fa^2 X}{(a^2 - 2e^2)Z + 2be|\mathbf{X}|} \quad \text{and} \quad v = \frac{fa^2 Y}{(a^2 - 2e^2)Z + 2be|\mathbf{X}|} \tag{1}$$

The hyperbolic to cylindric image transform is illustrated on the right in Fig. 3. Let $\mathbf{C}_p$ be the center of the cylindric projection. By setting $\mathbf{C_p} = \mathbf{F}$, a point $\mathbf{x_p}$ on the cylindric image and a point $\mathbf{x}$ on the hyperboloid lie on a line, connecting point $\mathbf{X}$ in 3D space with the focus $\mathbf{F}$ of the hyperboloid. The cylindric coordinate system expresses a point $\mathbf{x_p} = (\mathbf{x_p}, \mathbf{y_p}, \mathbf{z_p})$ on the cylindric surface as

$$x_p = r\cos\theta, \;\; y_p = r\sin\theta, \;\; z_p = r\tan\varphi \tag{2}$$

where $0 \leq \theta < 2\pi$ and $-\pi/2 \leq \varphi < \pi/2$. Hereafter, without loss of generality, we set $r = 1$. The mapping between the hyperbolic image $I(u, v)$ and the cylindric image $I_P(\theta, \varphi)$ can then be formulated as

$$u = \frac{fa^2\cos\theta}{(a^2 \mp 2e^2)\tan\varphi \pm 2be\sqrt{1 + \tan\varphi^2}}$$
$$v = \frac{fa^2\sin\theta}{(a^2 \mp 2e^2)\tan\varphi \pm 2be\sqrt{1 + \tan\varphi^2}}$$

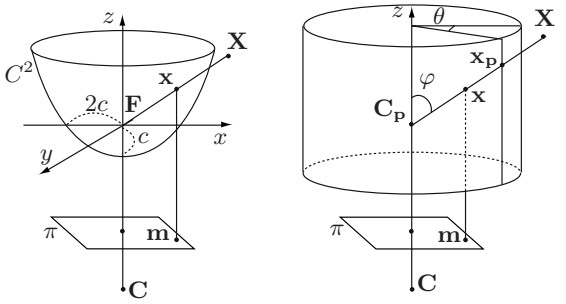

Fig. 4. Mapping from parabolic to cylindric images. Left: parabolic camera system. Right: parabolic and cylindric camera systems combined.

We thus derived a one-to-one correspondence between points on a hyperbolic image and points on a cylindric image. This allows to transform these images using common interpolation techniques, such as bilinear, cubic convolution, or B-spline interpolation. Note that substituting Equ. (2) into Equ. (1) derives a mapping with central cylindric projection. Cylindric images can also be derived for different projections (e.g., setting $z_p = r\varphi$ achieves equi-rectangular projection).

2.2 Parabolic into Cylindric Images

We consider at first the case illustrated on the left in Fig. 4. Let $\mathbf{C} = (0, 0, -\infty)$ be the center of the orthographic camera. Letting the focus $\mathbf{F}$ of the paraboloid C^2 be the origin of the world coordinate system, the paraboloid C^2 is expressed in the quadric form:

$$(x, y, z, 1) \begin{pmatrix} \frac{1}{4c} & 0 & 0 & 0 \\ 0 & \frac{1}{4c} & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} = 0$$

The projection of a point $\mathbf{X} = (X, Y, Z)^\top$ in 3D space into a point $\mathbf{x} = (x, y, z)^\top$ on the paraboloid is then expressed as $\mathbf{x} = \chi\mathbf{X}$ where

$$\chi = \frac{2c}{|\mathbf{X}| - \mathbf{Z}}$$

The projection of point $\mathbf{x}$ into a point $\mathbf{m} = (u, v)^\top$ in the image plane π is expressed as

$$\begin{pmatrix} \mathbf{m} \\ 1 \end{pmatrix} = \begin{pmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,0\,1 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$$

Accordingly, the mapping from $\mathbf{X}$ into $\mathbf{m}$ is formulated as

$$u = \frac{2cX}{|\mathbf{X}| - \mathbf{Z}}, \quad v = \frac{2cY}{|\mathbf{X}| - \mathbf{Z}}$$

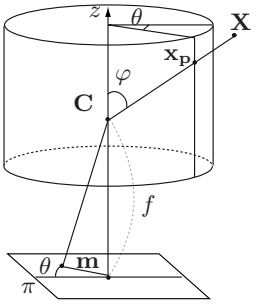

Fig. 5. Transform of a fish-eye into a cylindric image

Furthermore, as show on the right in Fig. 4, by setting $\mathbf{C_p} = \mathbf{F}$, a point $\mathbf{x_p}$ on the panoramic image and a point $\mathbf{x}$ on the paraboloid lie on a line, connecting point $\mathbf{X}$ in 3D space with the focal point $\mathbf{F}$ of the paraboloid. Let $\mathbf{x}_p = (x_p, y_p, z_p)$ be a point on the cylindric surface as in Equ. (2). A mapping between the parabolic image $I(u, v)$ and the cylindric image $I_P(\theta, \varphi)$ is given as follows:

$$u_h = \frac{2c\cos\theta}{\sqrt{1+\tan\varphi^2}-\tan\varphi} \quad \text{and} \quad v_h = \frac{2c\sin\theta}{\sqrt{1+\tan\varphi^2}-\tan\varphi}$$

3 Fish-Eye into Cylindric Images

A fish-eye camera generates an image on the basis of a stereographic, equi-solid angle, orthogonal, and equi-distant projection. This projection is illustrated in Fig. 5. Let $\mathbf{m} = (u, v)^\top$ and $\mathbf{X} = (X, Y, Z)^\top$ be a point on an image acquired by fish-eye camera and a point in a 3D space, respectively. Depending on the projection model, the mapping of $\mathbf{X}$ into $\mathbf{m}$ can be stated as follows:

$$\begin{aligned}
\text{stereographic}: &\ u = 2fa\tan(\cos^{-1}(c)/2), && v = 2fb\tan(\cos^{-1}(c)/2)\\
\text{equi-solid angle}: &\ u = 2fa\sin(\cos^{-1}(c)/2), && v = 2fb\sin(\cos^{-1}(c)/2)\\
\text{orthogonal}: &\ u = fa\sin(\cos^{-1}(c)), && v = fb\sin(\cos^{-1}(c))\\
\text{equi-distant}: &\ u = fa\cos^{-1}(c), && v = fb\cos^{-1}(c)
\end{aligned}$$

where

$$a = \frac{X}{\sqrt{X^2+Y^2}}, \quad b = \frac{Y}{\sqrt{X^2+Y^2}}, \quad c = Z/|X|$$

Next, let $\mathbf{x}_p = (x_p, y_p, z_p)$ be a point on the cylindric surface as in Equ. (2). A mapping between the fish-eye camera image $I(u, v)$ and the cylindric image $I_P(\theta, \varphi)$ is then as follows (where $\gamma = \pi/2 - \varphi$):

$$\begin{aligned}
\text{stereographic}: &\ u = 2f\tan(\gamma/2)\cos\theta, & v = 2f\tan(\gamma/2)\sin\theta \\
\text{equi-solid angle}: &\ u = 2f\sin(\gamma/2)\cos\theta, & v = 2f\sin(\gamma/2)\sin\theta \\
\text{orthogonal}: &\ u = f\sin(\gamma)\cos\theta, & v = f\sin(\gamma)\sin\theta \\
\text{equi-distant}: &\ u = f(\gamma)\cos\theta, & v = f(\gamma)\sin\theta
\end{aligned}$$

4 Experiments

We show some examples of cylindric images transformed either from hyperbolic or fish-eye images into this format. The top left and the middle left in Fig. 6 show two hyperbolic images with 360-degree fields of view, acquired with a hyperbolic camera as shown on the left of Fig. 1. Using the mapping derived in Section 2.1, the original images are instantly transformed into cylindric images as shown on the top right and middle right of Fig. 6. The bottom row in Fig. 6 shows the Inca shield of Cusco (part of Bravo's mural as shown above) and its cylindric image, assuming that the shield was created by imagining a 'hyperbolic camera'.

Furthermore, the top row in Fig. 2 shows three original images recorded with a hemispherical field of view, acquired with a fish-eye camera as show on the right of Fig. 1. The resulting panoramic images are shown in the bottom row of Fig. 2.

Figure 7 just illustrates a stereo pair of fish-eye images, recorded on a roof-rack of HAKA1, a research vehicle in the driver assistance project *.enpeda..* at Auckland University. Stereo images require geometric rectification before applying a

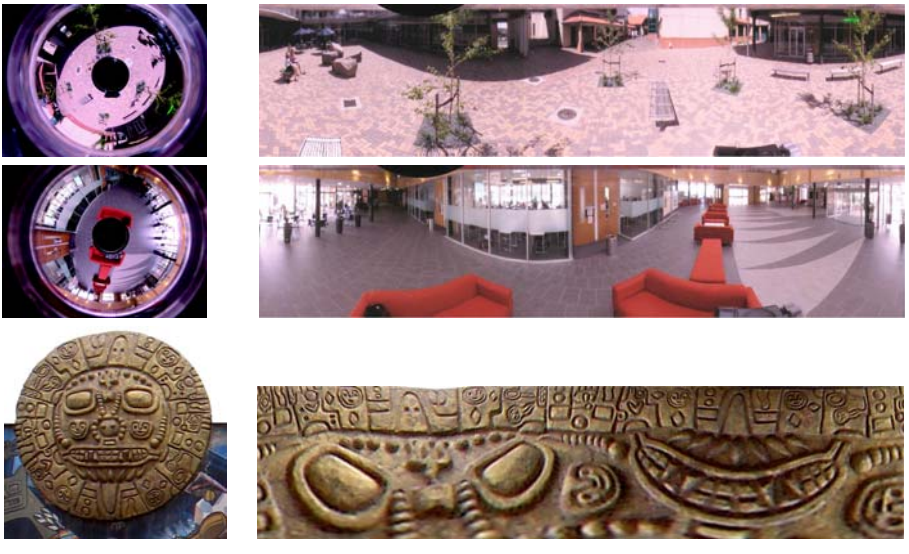

Fig. 6. Hyperbolic into panoramic images (top and middle row: Tamaki campus, The University of Auckland; bottom row: Inca shield of Cusco). Left: original images; right: resulting panoramic images.

Fig. 7. Stereo pair of fish-eye images, for ensuring a wide field of view in vision based driver assistance

stereo matching algorithm [5]. Prior to rectification, these images are mapped into cylindric format as described above.

5 Concluding Remarks

This brief paper accompanies a keynote talk at this conference by providing transformation formulas of catadioptric and dioptric images into cylindric images, thus detailing one particular subject in the large field of panoramic or 3D computer vision. The provided ideal geometric transforms still need to be "refined" by further analysis of potential impacts, such as optical distortions, to ensure resultant panoramas of 'perfect quality'.

References

1. Daniilidis, K., Papanikolopoulos, N. (eds.): Panoramic Robots. Special issue of IEEE Robotics Automation Magazine 11 (2004)
2. Daniilidis, K., Klette, R. (eds.): Imaging Beyond the Pinhole Camera. Springer, New York (2007)
3. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Proc. CVPR, vol. 1, pp. 886–893 (2005)
4. Ess, A., Leibe, B., Schindler, K., Van Gool, L.: A mobile vision system for robust multi-person tracking. In: Proc. *CVPR*, CD (2008)
5. Gehrig, S., Rabe, C., Krueger, L.: 6D vision goes fisheye for intersection assistance. In: Proc. Canadian Conf. Computer Robot Vision, pp. 34–41 (2008)
6. Google. Google earth (2004), `http://earth.google.com/`
7. Huang, F., Klette, R., Scheibe, K.: Panoramic Imaging. Wiley, Chichester (2008)
8. Nayar, S.K.: Catadioptric omnidirectional cameras. In: Proc. CVPR, pp. 482–488 (1997)
9. Ray, S.F.: Applied Photographic Optics : Lenses and Optical Systems for Photography, Film, Video, Electronic and Digital Imaging, 3rd edn. Focal Press, Oxford (2002)
10. Svoboda, T., Pajdla, T., Hlaváč, V.: Epipolar geometry for panoramic cameras. In: Proc. Czech Technical University Workshop 1998, Prague, pp. 177–178 (1998)
11. Wei, S.K., Huang, F., Klette, R.: Color anaglyphs for panorama visualizations. Technical Report, CITR-TR-19, The University of Auckland (1998)
12. Wikipedia. Panoramic photography (2009), `http://en.wikipedia.org/wiki/Panoramic_photography`