社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2006－CVIM－154（30）
2006／5／18

画像理解のための中心カメラ系の解析

鳥居 秋彦 † 　**井宮 淳** ††

† **千葉大学 自然科学研究科**
†† **千葉大学 総合メディア基盤センター**
〒 263-8522 千葉市稲毛区弥生町 1-33

あらまし　本論文では，画像理解を実現するための中心カメラ系システムを幾何学的に解析する．中心カメラとは，空間の光線を一点に集めるカメラである．現在広く普及しているアナログ，デジタルカメラ，全方位カメラ，魚眼レンズカメラは，幾何学的にこの中心カメラというクラスに属する．現在まで，画像理解の数理解析を行う上で，一般的なカメラを表す数理モデルとして，ピンホールカメラモデルが用いられてきた．しかし，ピンホールカメラモデルでは，幾何学的に同じクラスに属する，全方位カメラや魚眼レンズカメラを表すことは不可能である．そこで，統一的に画像理解のための数理解析を行うためには，より広い枠組みでカメラを表す数理モデルが必要である．本論文では，はじめに，様々なカメラを幾何，代数的に解析し，中心カメラ全般を表現する標準的数理モデルとして球面カメラモデルを導く．そして，球面カメラモデルを基に，画像理解において基本となる三次元復元手法を提案する．球面カメラモデルを用いて三次元復元を行う上で，不可欠な問題は，球面画像解析，多視点幾何学である．球面画像解析の手法として，ハフ変換を用いた球面上大円検出を提案した．さらに，これらの手法を基に三次元復元を行うために，複数の球面カメラが空間に配置された時に生じる幾何学的拘束を導出し，多視点幾何の定式化を行った．これらの球面カメラモデルに基づいた手法は，異なるカメラを組み合わせて構成されたカメラシステムに対して，統一的に適用可能である．

Analysis of Central Camera Systems for Computer Vision

Akihiko TORII†　and　Atsushi IMIYA††

†Graduate School of Science and Technology, Chiba University
††Institute of Media and Information Technology, Chiba University
Yayoi-cho 1-33, Inage-ku, 263-8522, Chiba, Japan

Abstract　This paper aims to geometrically analyze central camera systems for computer vision. In the sence of geometry, a central camera is a collection of all rays incident to one point. Generally speaking, most cameras that are widely-used in the field of computer vision and robot vision are designed to fall into the class of central cameras. This is because geometric and algebraic expressions of central cameras are simple compared to ones of non-central cameras. In the history of computer vision, the conventional cameras are modeled as pinhole cameras. However, the pinhole-camera model can not express catadioptric, dioptric, and panoramic cameras that also belong to the class of central cameras. In this paper, we propose a spherical-camera model as a standard camera model expressing all central cameras. Using the spherical-camera model as a standard camera model of central cameras, we develop algorithms for computer vision. One of the fundamental studies in the field of computer vision is the three-dimensional reconstruction from multiple images. As a fundamental tool for spherical image analysis and as a fundamental pre-processing for three-dimensional reconstruction, we develop an algorithm for detecting great circles, which correspond to lines in a space, on a sphere. On the basis of the spherical image analysis, the formulation of multiple view geometry is then established. The analysis enables us to develop unified algorithms which can be applied to all central cameras and to realize camera systems involving many different cameras.

1　Introduction

Computer Vision is the study for constructing systems that are capable of acquiring and recognizing 3D information [2, 5, 6, 7, 11, 15, 16, 17]. Historically, in the field of computer vision, images obtained by conventional cameras have been used as input information for acquiring 3D information. Since the images are projected from 3D space, reconstruction of 3D information from images are the most fundamental study in computer vision. The essential 3D information, which should be reconstructed, are positions and orientations of cameras to acquire 3D information in the scene. For the reconstruction of the 3D information, it is necessary to mathematically describe relationships between cameras, images and objects in a space. In the study of computer vision, a conventional camera is generally modeled as a *pinhole camera* for expressing such mathematical relationships.

Definition 1.1 *A pinhole camera consists of a camera center and a plane in a space. The pinhole camera collects rays in a space at the center of the camera and generates an image on the plane in a space.*

A *camera model* geometrically and algebraically expresses projection from 3D scene onto an image. Most problems of computer vision has been solved based on the pinhole-camera model.

Mathematically, a pinhole camera can have maximally 2π-steradian field of view since it has an image on a plane. Practically, a conventional pinhole camera generally generates a $\pi/2$-steradian field-of-view image. For the 3D reconstruction from 2D images, it is necessary to share a common region in a space among the images. Therefore, the development of cameras which practically have a large field of view compared to conventional pinhole cameras has been one of the important studies in the field of computer vision, robot vision and image processing [3]. Some researchers developed cameras constructed by the combination of a quadric-shaped mirror and a pinhole camera[1] [1, 19, 23]. These cameras are called *catadioptric cameras*. The word *"catadioptric"* means that pertaining to or involving both the reflection and the refraction of light. In Figure 1, we show an example of a catadioptric camera, its image and a spherical image transformed from the catadioptric camera image. On the other hand, it is also possible to construct the cameras, which can observe omni-directional scene, using the refraction of light. These cameras are called *dioptric cameras*. The word *"dioptric"* means that relating to the refraction of light. The dioptric camera is constructed with a specially designed refractor as the optical lens of the camera. For example, fish-eye-lens cameras [21] are the typical ones of dioptric cameras. In Figure 2, we show an example of a dioptric camera, its image and a spherical image transformed from the dioptric camera image. In this paper, we call both catadioptric and dioptric cameras *omni-directional cameras*. The significant advantage of omni-directional cameras is to acquire a large field of view with real time using a single camera. Omni-directional cameras are widely-used in the robot navigation and video surveillance [8, 14, 24, 27].

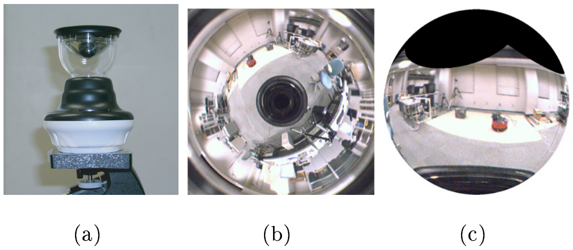


Figure 1: An example of a catadioptric camera, its image and a spherical image. (a) Hyperbolic camera system, SOIOS-55 CAM. (b) Image acquired with this hyperbolic camera system. (c) Spherical image transformed from (b).

According to the development of many types of cameras, recently, the new geometric concept of cameras is proposed [20, 22].

Definition 1.2 *A central camera is a collection of all rays incident to one point. The point is called the camera center.*

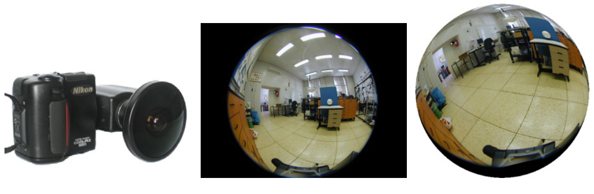
Definition 1.3 *A non-central camera is a collection of rays which are not incident to one point.*

On the basis of these definitions, we geometrically classify cameras and images. Figure 3 shows the classification of cameras based on these definitions. The conventional pinhole camera is classified into the class of central cameras. Depending on the shapes of the mirror and lenses, omni-directional cameras are classified into the classes of the central and non-central cameras. The typical non-central cameras are the orthographic camera and push-broom camera [10]. Generally speaking, most cameras that are widely-used in the field of computer vision and robot vision are designed to fall into the class of central cameras. This is because geometric and algebraic expressions of central cameras are simple compared to ones of non-central cameras. In this paper, we therefore focus on the mathematical analysis of central camera systems for computer vision.

In the computer-vision communities, traditional algorithms and their applications are developed based on the pinhole-camera model. However, it is impossible to express omni-directional cameras using the pinhole-camera model theoretically and practically since the pinhole camera has an image on a plane. For the generation of an image which practically expresses an omni-directional scene in a space, a camera must project the scene onto a quadric surface. For the development of theories and algorithms that are applicable to all central cameras, it is necessary to use a unified and standard camera model, otherwise, the proposed

---

[1] One of the catadioptric cameras uses an orthographic camera that collects light rays perpendicular to an image plane.



(a) (b) (c)

Figure 2: An example of a dioptric camera, its image and a spherical image. (a) Fish-eye-lens camera, Nikon Coolpix 950 digital camera with Nikon Fisheye converter FC-E8. (b) Image acquired by this fish-eye-lens camera system. (c) Spherical image transformed from (b).

algorithms become dependent on selection of the cameras. For the standardization of central cameras, in this paper, we mathematically classify central cameras and define the camera models. These mathematical clarifications of central cameras naturally leads to the spherical-camera model as a standard camera model expressing all central cameras. The spherical camera is defined as follows.

Definition 1.4 *A spherical camera consists of a camera center and a surface of a unit sphere whose center is the camera center. The spherical camera collects rays in a space at the camera center and generates an image on the surface of the unit sphere.*

The central catadioptric and dioptric cameras geometrically collect light rays at a single point in a space. Therefore, it is possible to transform the images acquired by these cameras to images on a sphere, *spherical images* in abbreviated form, when the appropriate factors of the cameras, such as the parameters of the quadric surface in the catadioptric system and the method of refraction in the dioptric system, are known [2][18, 22]. Obviously, it is also possible to transform the pinhole-camera images to the spherical images. These geometric properties of central cameras lead that analysis on central-camera images are converged to spherical image analysis. In other words, theories and algorithms proposed for spherical image analysis are applicable to all central-camera image analysis.

In Section 2, we mathematically classify cameras in the class of central cameras and define these camera models. We then define the spherical-camera model in Section 3. For the standard use of the spherical-camera model instead of each central-camera model, it is essential to transform central-camera images to spherical images. The transformations are uniquely defined when a central camera center and a spherical camera center are geometrically configured at the same point in a space. On the basis of this geometric configuration of the cameras, in Section 3, we formulate the transformations from central-camera images to spherical images. Using the spherical-camera model as a standard camera model of central cameras, we develop algorithms for computer vision. One of the fundamental studies in the field of computer vision is the 3D reconstruction from multiple images. As a preprocessing of 3D reconstruction using the spherical-camera model, it is essential to analyze the images on a sphere. In Section 4, as a fundamental tool for spherical image analysis and as a fundamental preprocessing for 3D reconstruction, we develop an algorithm for detecting great circles, which correspond to lines in a space, on a sphere. On the basis of the spherical image analysis, the formulation of multiple view geometry is then established in Section 5. In Section 5, we formulate multiple view geometry for 3D reconstruction based on the spherical-camera model using the combinations of point and great-circle correspondences. The mathematical analysis of central cameras enables us to observe underlying mathematical relationship among cameras, images and 3D information beyond the pinhole cameras. Practically, the analysis enables us to develop unified algorithms which can be applied to all central cameras and to realize camera network systems involving many different cameras.

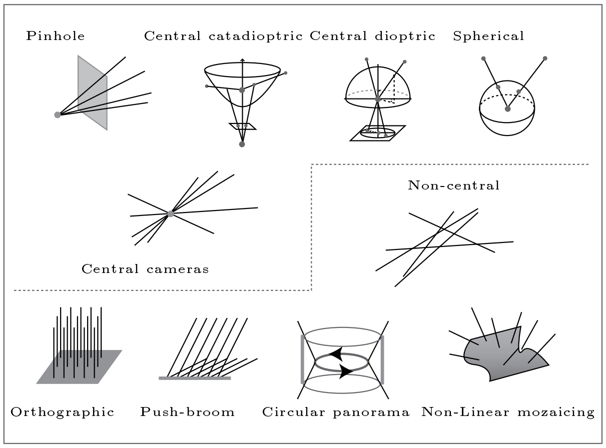


Figure 3: Central and non-central cameras. Figures show examples of cameras classified based on **Definition 1.2** and **Definition 1.3**.

---

[2] For the practically accurate transformation from omnidirectional images to spherical images, it is required to calibrate the omni-directional images. In the calibration, it is necessary to estimate the parameters of image distortion as same as the calibration of conventional pinhole cameras and to re-estimate the appropriate factors of the cameras. The well-established methods are proposed by [9, 18, 22].

2 Central Camera Models

2.1 Mathematical Classification of Central Cameras

A pinhole camera collects light rays and generates an image on a plane. As shown in Figure 4 (a), a point $\boldsymbol{X} \in \mathbb{R}^3$ is mapped to a point $\boldsymbol{p} \in \mathbb{R}^2$ on the pinhole-camera image by the perspective projection P:

$$P : \boldsymbol{X} \rightarrow \boldsymbol{p}. \tag{1}$$

As shown in Figure 4 (b), a catadioptric camera generates an image following the two step. A point $\boldsymbol{X} \in \mathbb{R}^3$ is transformed to a point $\boldsymbol{x} \in C^2$ on a quadric surface by the nonlinear function f:

$$f : \boldsymbol{X} \rightarrow \boldsymbol{x}. \tag{2}$$

The point $\boldsymbol{x} \in C^2$ is projected by a pinhole or orthographic camera to a point $\boldsymbol{m} \in \mathbb{R}^2$ on a planar image:

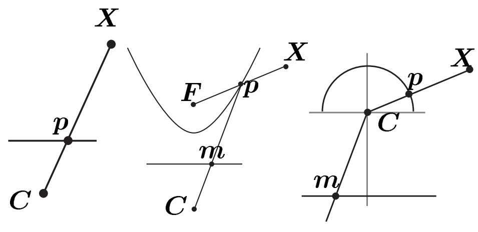
$$P : \boldsymbol{x} \rightarrow \boldsymbol{m}. \tag{3}$$

As shown in Figure 4 (c), a dioptric camera collects light rays to a point in a space. This point is the center of the dioptric camera. The dioptric camera generates an image according to the following two steps. A point $\boldsymbol{X} \in \mathbb{R}^3$ is transformed to a point $\boldsymbol{x} \in S^2$ on the unit sphere by the nonlinear function f:

$$f : \boldsymbol{X} \rightarrow \boldsymbol{x}. \tag{4}$$

The point $\boldsymbol{x} \in S^2$ is transformed to a point $\boldsymbol{m} \in \mathbb{R}^2$ on a planar image by the nonlinear function g:

$$g : \boldsymbol{x} \rightarrow \boldsymbol{m}. \tag{5}$$



(a) (b) (c)

Figure 4: Central cameras. (a) Pinhole camera. (b) Catadioptric camera. (c) Dioptric camera.

2.2 Pinhole Camera

2.2.1 Pinhole-Camera Model

On the basis of Definition 1.1, we describe the pinhole-camera model algebraically. Hereafter, we call the center of the pinhole camera the *pinhole-camera center* and the image plane of the pinhole camera the *pinhole-camera image.* We formulate the relation between a point on a pinhole-camera image and a point in a space. As shown in Figure 5, setting a pinhole-camera center $\boldsymbol{C} \in \mathbb{R}^3$ to be the origin of the world coordinate system, the pinhole camera collects a light ray from a point $\boldsymbol{X} \in \mathbb{R}^3$ to the pinhole-camera center $\boldsymbol{C}$. The z-axis is perpendicular to the pinhole-camera-image plane. The intersection of the light ray and pinhole-camera-image plane yields a point $\boldsymbol{p} = (u, v)^\top \in \mathbb{R}^2$. Using the homogeneous expression $\tilde{\boldsymbol{p}}$ and $\tilde{\boldsymbol{X}}$ of $\boldsymbol{p}$ and $\boldsymbol{X}$, we have the equation between $\boldsymbol{p}$ and $\boldsymbol{X}$, for a given focal length f,

$$\tilde{\boldsymbol{p}} = \frac{1}{Z} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tilde{\boldsymbol{X}}. \tag{6}$$

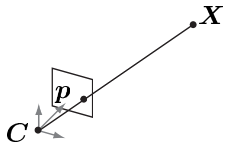


Figure 5: Pinhole-camera model.

2.3 Catadioptric Camera

We express a quadric surface in the homogeneous form as

$$\tilde{\boldsymbol{x}}^\top \mathbf{A} \tilde{\boldsymbol{x}} = 0, \tag{7}$$

where $\tilde{\boldsymbol{x}} = (x, y, z, 1)^\top$, and $\mathbf{A} = \{a_{ij}\}$ for $i, j = 1, 2, 3, 4$. The matrix $\mathbf{A}$ satisfies the relation $\mathbf{A}^\top = \mathbf{A}$. A quadric surface is also expressed as

$$\boldsymbol{x}^\top \mathbf{A}_0 \boldsymbol{x} + 2\boldsymbol{b}^\top \boldsymbol{x} + a_{44} = 0, \tag{8}$$

where $\boldsymbol{x} = (x, y, z)^\top$, $\mathbf{A}_0 = \{a_{ij}\}$, $i, j = 1, 2, 3$, and $\boldsymbol{b} = (a_{41}, a_{42}, a_{43})^\top$. We set λ_m and σ_n for $m = 1, 2, 3, 4$ and $n = 1, 2, 3$ are the eigenvalues of the matrix $\mathbf{A}$ and $\mathbf{A}_0$, respectively. If λ_m and σ_n satisfy the two conditions, the quadric surface is the revolution surface of quadratic curve, that is, an ellipsoid of revolution, a hyperboloid of two sheets, a paraboloid of revolution. One is that the signs of λ_i are three positives and one negative, and vice versa. The other is $\sigma_1 = \sigma_2$ and $\sigma_3 \in \mathbb{R}$. Such a quadric surface has two focal points. For the construction of central-catadioptric cameras, we use a quadric-shaped mirror which has two focal point. When a light ray passes through the focal point of the quadric surface, the light ray reflected on the quadric surface passes through the other focal point. If we locate the pinhole-camera center at this focal point, it is possible to construct central-catadioptric cameras. [3].

---

[3] In case of $\sigma_3 = 0$, a camera center must be at the point at infinity for the catadioptric camera to be fallen into the class

Furthermore, assuming that the focal point of the quadric surface is the origin of the world coordinate system, we set that a light ray has the direction $\boldsymbol{p} \in \mathbb{R}^3$ and passes through a point $\boldsymbol{q} \in \mathbb{R}^3$ in a space. The intersection of the light ray and the quadric surface is expressed as

$$\boldsymbol{x} = \mu \boldsymbol{p} + \boldsymbol{q}. \tag{9}$$

μ is computed from the substitution of Eq. (9) to Eq. (8), such that,

$$\mu = \frac{-\beta \pm \sqrt{\beta^2 - \alpha\gamma}}{\alpha}, \tag{10}$$

where $\alpha = \sum_{j=1}^4 \sum_{i=1}^4 p_j a_{ij} p_i$, $\beta = \sum_{j=1}^4 \sum_{i=1}^4 p_j a_{ij} q_i$, $\gamma = \sum_{j=1}^4 \sum_{i=1}^4 q_j a_{ij} q_i$, for $p_4 = 0$ and $q_4 = 1$. The sign of μ depends on the geometric configuration of the quadric surface and the light ray. These algebraic representations enable us to express the correspondence between a point on the quadric surface and any point in a space in a unified way.

2.3.1 Example of Catadioptric Camera Model

A central hyperbolic camera is constructed from the combination of a hyperboloidal shaped mirror and the conventional pinhole camera [1, 19, 23]. The hyperbolic camera collects light rays in a space at the focus of the hyperboloid as shown in Figure 6. The light rays, which pass through the focus of the hyperboloid, are reflected on the mirror. Then, the reflected rays pass through the other focus of the hyperboloid. For the construction of the central hyperbolic camera, we locate the pinhole-camera center at the focus, which collects the reflected rays. The pinhole camera collects the light rays which are reflected from the mirror surface at the pinhole-camera center and generates an image on a plane.

As shown in Figure 6, the focus $\boldsymbol{F}$ of the hyperboloid C^2 is located at the origin of the world coordinate system. The hyperbolic-camera axis $\boldsymbol{l}$ is the line which connects $\boldsymbol{C}$ and $\boldsymbol{F}$. The center of the pinhole camera is located at the point $\boldsymbol{C} = (0, 0, -2e)$. The pinhole-camera image plane is perpendicular to the hyperbolic-camera axis $\boldsymbol{l}$. Using this geometric configuration of the camera and mirror, it is possible to construct the central hyperbolic camera. We set the hyperboloid C^2 :

$$\tilde{\boldsymbol{x}}^\top \mathbf{A} \tilde{\boldsymbol{x}} = 0, \tag{11}$$

where

$$\mathbf{A} = \begin{pmatrix} \frac{1}{a^2} & 0 & 0 & 0 \\ 0 & \frac{1}{a^2} & 0 & 0 \\ 0 & 0 & -\frac{1}{b^2} & -\frac{e}{b^2} \\ 0 & 0 & -\frac{e}{b^2} & -\frac{e^2}{b^2}+1 \end{pmatrix}, \tag{12}$$

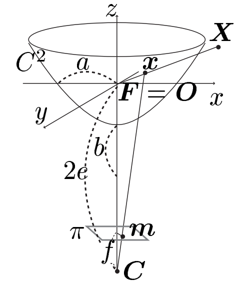


Figure 6: Hyperbolic camera model. A point $\boldsymbol{X}$ in a space is transformed to the point $\boldsymbol{x}$ on the hyperboloid. and $\boldsymbol{x}$ is mapped to the point $\boldsymbol{m}$ on image plane.

and $e = \sqrt{a^2 + b^2}$. We set a point $\boldsymbol{X} = (X, Y, Z)^\top$ in a space, a point $\boldsymbol{x} = (x, y, z)^\top$ on the hyperboloid C^2, and $\boldsymbol{m} = (u, v)^\top$ on the image plane π. The nonlinear transform in Eq. (2) is expressed as

$$\boldsymbol{x} = \chi \boldsymbol{X}, \tag{13}$$

where

$$\chi = \frac{a^2}{b|\boldsymbol{X}| - eZ}. \tag{14}$$

The projection in Eq. (3) is expressed as

$$\begin{pmatrix} \boldsymbol{m} \\ 1 \end{pmatrix} = \frac{1}{z+2e} \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{x} \\ 1 \end{pmatrix} \tag{15}$$

Accordingly, a point $\boldsymbol{X}$ in a space is transformed to a point $\boldsymbol{m}$ as

$$u = \frac{fa^2 X}{(a^2 - 2e^2)Z + 2be|\boldsymbol{X}|},\ v = \frac{fa^2 Y}{(a^2 - 2e^2)Z + 2be|\boldsymbol{X}|}. \tag{16}$$

2.4 Dioptric Camera

2.4.1 Example of Dioptric Camera Model

The well-established dioptric camera is the fish-eye-lens camera. Depending on the methods of refraction, fish-eye-lens cameras generate images according to the stereographic, equi-solid angle, orthogonal, and equi-distance projection. As shown in Figure 7, first, setting the center of the fish-eye-lens camera to be at the origin of the world coordinate system, the fish-eye-lens camera collects a point $\boldsymbol{X} = (X, Y, Z)^\top$ to a point $\boldsymbol{x} = (x, y, z)^\top$ on S^2. This collection of the light ray is the nonlinear transform in Eq. (4). The nonlinear transform is expressed as

$$\boldsymbol{x} = \boldsymbol{X}/|\boldsymbol{X}|. \tag{17}$$

Next, we set $\boldsymbol{m} = (u, v)^\top$ and $\boldsymbol{\xi} = (\theta, \varphi)^\top$ to be a point on an image acquired by fish-eye-lens camera and the

---

of central cameras. Therefore, it is necessary to combine a paraboloidal mirror and an orthographic camera.

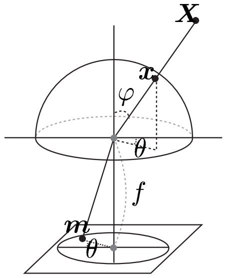


Figure 7: Fish-eye-lens camera model. A fish-eye-lens camera acquires an image by stereographic, equi-solid angle, orthogonal and equi-distance projection.

point mapped onto the unit sphere in the spherical coordinates, respectively. The point $\boldsymbol{\xi}$ is transformed to the point $\boldsymbol{m}$ on the fish-eye-lens camera image depending on the methods of projections in Eq. (5) as follows,

$$u = 2f\tan(\varphi/2)\cos\theta,\quad v = 2f\tan(\varphi/2)\sin\theta, \tag{18}$$

$$u = 2f\sin(\varphi/2)\cos\theta,\quad v = 2f\sin(\varphi/2)\sin\theta, \tag{19}$$

$$u = f\sin\varphi\cos\theta,\quad v = f\sin\varphi\sin\theta, \tag{20}$$

$$u = f\varphi\cos\theta,\quad v = f\varphi\sin\theta. \tag{21}$$

3 Standardization of Central Cameras

3.1 Standard Camera Model for Central Cameras

3.1.1 Spherical-Camera Model

On the basis of Definition 1.4, we define the spherical-camera model. Hereafter, we call the center of spherical camera the *spherical camera center* and the surface of a unit sphere the *spherical image*. As shown in Figure 8, a spherical camera center $\boldsymbol{C}$ is located at the origin of the world coordinate system. The z-axis corresponds to the north pole of the unit sphere. The spherical camera collects a light ray from a point $\boldsymbol{X} \in \mathbb{R}^3$ to the spherical camera center $\boldsymbol{C}$. The intersection of the light ray and spherical image yields a point to the point $\boldsymbol{x} = (x, y, z)^\top$ on S according to the formulation,

$$\boldsymbol{x} = \boldsymbol{X}/|\boldsymbol{X}|. \tag{22}$$

The spherical coordinate system also expresses a point $\boldsymbol{x} = (x, y, z)^\top$ on the unit sphere as

$$\boldsymbol{x} = (\cos\theta\sin\varphi, \sin\theta\sin\varphi, \cos\varphi)^\top \tag{23}$$

where $0 \leq \theta < 2\pi$ and $0 \leq \varphi < \pi$. Therefore, the spherical image is also expressed as $I(\theta, \varphi)$.

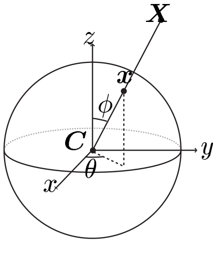


Figure 8: spherical-camera model.

3.2 Spherical Image Transform for Standardization

As shown in Figure 9 (a), we locate the spherical camera center at the pinhole-camera center. A nonlinear function h expresses the one-to-one mapping from a point $\boldsymbol{p} \in \mathbb{R}^2$ to a point $\boldsymbol{\xi} \in S^2$ on the unit sphere:

$$h : \boldsymbol{p} \rightarrow \boldsymbol{\xi}. \tag{24}$$

Since $\boldsymbol{p}$ and $\boldsymbol{\xi}$ satisfies one-to-one correspondence, there exists an inverse function h^{-1}:

$$h^{-1} : \boldsymbol{\xi} \rightarrow \boldsymbol{p}. \tag{25}$$

This nonlinear function is the pinhole-to-spherical image transform.

As shown in Figure 9 (b), locating the spherical camera center at the focus of the quadric surface, a nonlinear function h expresses the one-to-one mapping from a point $\boldsymbol{x} \in C^2$ on the quadric surface to a point $\boldsymbol{\xi} \in S^2$ on the unit sphere:

$$h : \boldsymbol{x} \rightarrow \boldsymbol{\xi}. \tag{26}$$

Since $\boldsymbol{x}$ and $\boldsymbol{\xi}$ satisfies one-to-one correspondence, there exists an inverse function h^{-1}:

$$h^{-1} : \boldsymbol{\xi} \rightarrow \boldsymbol{x}. \tag{27}$$

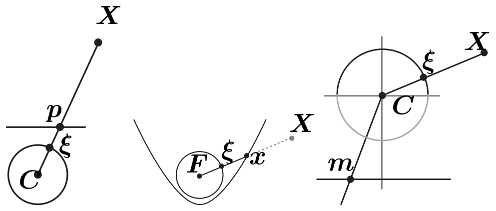
This nonlinear function is the catadioptric-to-spherical image transform.

In Eq. (4), the dioptric camera collects rays at a single point in a space and transforms a point in a space onto a point on a sphere. Therefore, as shown in Figure 9 (c), the function in Eq. (5) directly expresses the one-to-one mapping from a point $\boldsymbol{\xi} \in S^2$ to $\boldsymbol{m} \in \mathbb{R}^2$. This nonlinear function is the dioptric-to-spherical image transform. Since $\boldsymbol{\xi}$ and $\boldsymbol{m}$ satisfies one-to-one correspondence, there also exists an inverse function g^{-1}:

$$g^{-1} : \boldsymbol{m} \rightarrow \boldsymbol{\xi}. \tag{28}$$

3.2.1 Pinhole-to-Spherical Image Transform

In this section, we formulate the transform from an image acquired by a conventional pinhole camera to



(a) (b) (c)

Figure 9: Geometric configuration of the central cameras and the spherical camera. In this configuration, there exist a one-to-one correspondence. (a) Correspondence between a point on a pinhole-camera image and a point on a spherical image. (b) Correspondence between a point on a quadric mirror of a catadioptric camera and a point on a spherical image. (c) Correspondence between a point on a quadric mirror of catadioptric camera and a point on a spherical image (c).

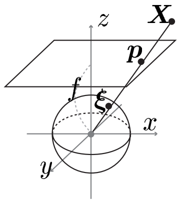


Figure 10: Image transform of pinhole-to-spherical cameras. In this geometrical configuration of the cameras, a point $\boldsymbol{\xi}$ on the spherical image and a point $\boldsymbol{p}$ on the pinhole-camera image lie on a line connecting a point $\boldsymbol{X}$ in a space and the camera center.

an image on a sphere. As shown in Figure 10, setting the pinhole-camera center $\boldsymbol{C}$ to be the origin of the world coordinate system. A point $\boldsymbol{X} = (X, Y, Z)^\top$ in a space is projected to the point $\boldsymbol{p} = (u, v)^\top$ on the pinhole-camera-image plane. For the pinhole-camera-to-spherical image transform, as shown in Figure 10, we set the center of a unit sphere $\boldsymbol{C}_s = \boldsymbol{C}$. The point $\boldsymbol{p}$ on the pinhole-camera image and the point $\boldsymbol{\xi}$ on the spherical image in the spherical coordinates satisfy the relation

$$u = f \tan\varphi \cos\theta, \quad v = f \tan\varphi \sin\theta. \tag{29}$$

Therefore, the pinhole-camera image $I(u, v)$ is transformed to the images $I_S(\theta, \phi)$ on a sphere, that is,

$$I(u, v) = I(f \tan\varphi \cos\theta, f \tan\varphi \sin\theta) = I_S(\theta, \phi). \tag{30}$$

3.2.2 Catadioptric-to-Spherical Image Transform

In this section, we show the uniform algebraic expression of the transform from an image acquired by central-catadioptric cameras to spherical images. Locating the center of the unit sphere at the focus of quadric mirror, all the rays which pass through the focus of quadric mirror and the center of the sphere are identical. In this geometric configuration of camera centers, it is possible to set $\boldsymbol{p} = \boldsymbol{\xi}$ and $\boldsymbol{q} = \boldsymbol{0}$ in Eq. (9). Therefore, the nonlinear transform h^{-1} in Eq. (27) is expressed as:

$$\boldsymbol{x} = \mu \boldsymbol{\xi}, \tag{31}$$

where μ is computed from the substitution of Eq. (31) to Eq. (8).

Example of Catadioptric-to-Spherical Image Transform

We show the hyperbolic-to-spherical image transform. As shown in Figure 11, Setting $\boldsymbol{\xi} = (\xi_x, \xi_y, \xi_z)^\top$ to be a point on the unit sphere, the spherical-camera center $\boldsymbol{C}_s$ and the the focal point $\boldsymbol{F}$ of the hyperboloid C^2 are $\boldsymbol{C}_s = \boldsymbol{F} = \boldsymbol{0}$. Therefore, $\boldsymbol{q} = \boldsymbol{0}$ in Eq. (9). Furthermore, $\boldsymbol{l}_s$ denotes the axis connecting $\boldsymbol{C}_s$ and north pole of the spherical surface. For the axis $\boldsymbol{l}_s$ and the hyperbolic-camera axis $\boldsymbol{l}$ we set $\boldsymbol{l}_s = \boldsymbol{l} = k(0, 0, 1)^\top$ for $k \in \mathbb{R}$, that is, the directions of $\boldsymbol{l}_s$ and $\boldsymbol{l}$ are the direction of the z axis. For the configuration of the spherical camera and the hyperbolic camera which share axes $\boldsymbol{l}_s$ and $\boldsymbol{l}$ as shown in Figure 11, the nonlinear function in Eq. (27) is expressed as:

$$\boldsymbol{x} = \mu \boldsymbol{\xi}, \tag{32}$$

where

$$\mu = \frac{\pm a^2}{b \mp e\xi_z}. \tag{33}$$

Applying the spherical coordinate systems, a point $\boldsymbol{m}$ on the hyperbolic image and a point $\boldsymbol{\xi}$ on the sphere derives the equations:

$$u = \frac{f a^2 \cos\theta \sin\varphi}{(a^2 - 2e^2)\cos\varphi + 2be}, v = \frac{f a^2 \sin\theta \sin\varphi}{(a^2 - 2e^2)\cos\varphi + 2be}. \tag{34}$$

Setting $I(u, v)$ and $I_S(\theta, \varphi)$ to be the hyperbolic-camera image and the spherical image, respectively, the hyperbolic-to-spherical image transform is expressed as follows:

$$I_S(\theta, \varphi) = \\ I\left(\frac{f a^2 \cos\theta \sin\varphi}{(a^2 - 2e^2)\cos\varphi + 2be}, \frac{f a^2 \sin\theta \sin\varphi}{(a^2 - 2e^2)\cos\varphi + 2be}\right), \tag{35}$$

for the hyperbolic-camera image $I(u, v)$.

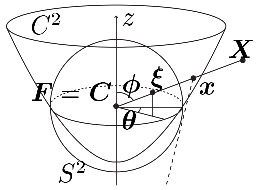


Figure 11: Image transform of hyperbolic-to-spherical cameras. In this geometrical configuration of the cameras, a point $\boldsymbol{\xi}$ on the spherical image and a point $\boldsymbol{x}$ on the hyperboloid lie on a line connecting a point $\boldsymbol{X}$ in a space and the focal point $\boldsymbol{F}$ of the hyperboloid.

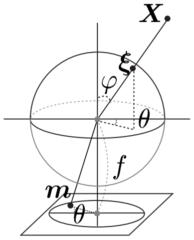


Figure 12: Image transform of a dioptric-to-spherical cameras. Since the fish-eye-lens cameras acquire images by stereographic, equi-solid angle, orthogonal and equi-distance projection, it is possible to transform fish-eye-lens camera images to spherical images directly.

3.2.3 Dioptric-to-Spherical Image Transform

Fish-Eye-Lens-to-Spherical Image Transform

In Eq. (4) the fish-eye-lens camera collects rays at a single point in a space and generates an image on a sphere. When the projection method is known, as shown in Figure 12, using the relationship expressed by the function g in Eq. (5) between the point $\boldsymbol{m} \in \mathbb{R}^2$ and the point $\boldsymbol{\xi} \in S^2$ it is possible to derive the image transform as follows:

$$\begin{aligned}
&I_S(\theta, \phi) = \\
&\quad I(2f\tan(\pi/4 - \varphi/2)\cos\theta, 2f\tan(\pi/4 - \varphi/2)\sin\theta), \\
&I_S(\theta, \phi) = \\
&\quad I(2f\sin(\pi/4 - \varphi/2)\cos\theta, 2f\sin(\pi/4 - \varphi/2)\sin\theta), \\
&I_S(\theta, \phi) = \\
&\quad I(f\sin(\pi/2 - \varphi)\cos\theta, f\sin(\pi/2 - \varphi)\sin\theta), \\
&I_S(\theta, \phi) = \\
&\quad I(f(\pi/2 - \varphi)\cos\theta, f(\pi/2 - \varphi)\sin\theta),
\end{aligned}$$

for the fish-eye-lens-camera image $I(u, v)$.

4 Feature Detection on Spherical Images

In this section, as a fundamental pre-processing for 3D reconstruction, we introduce a method based on the randomized Hough transform [26, 25], which is the well-established method of using voting and accumulation, for the detection of great circles from a collection of points on a sphere. In the field of pattern recognition and computer vision, lines in a space are considered as one of the fundamental features for the recognition of three-dimensional space. Elementary mathematics in spherical geometry states that a great circle on a sphere corresponds to a line on a plane in a space. Therefore, great-circle detection is an essential problem in spherical image analysis.

4.1 Duality of Point and Great Circle on Sphere

A great circle is a common curve with S^2 and a plane which passes through the origin. This geometrical property of the great circle implies that great-circle detection on S^2 is achieved by detecting the plane that passes through the origin, since a plane which passes through the origin is expressed as

$$\boldsymbol{a}^\top \boldsymbol{x} = 0, \quad \boldsymbol{a} \in S^2_+, \quad \boldsymbol{x} \in \mathbb{R}^3. \tag{36}$$

For spherical images, $\boldsymbol{x}$ also lies on S^2, that is, $|\boldsymbol{x}| = 1$. Therefore, for $\boldsymbol{x} = (x, y, z)^\top \in S^2$ and $\boldsymbol{a} = (a, b, c)^\top \in S^2_+$ the equation

$$\boldsymbol{a}^\top \boldsymbol{x} = 0, \quad \boldsymbol{a} \in S^2_+, \quad \boldsymbol{x} \in S^2 \tag{37}$$

defines a great circle on S^2 when we fix $\boldsymbol{a} \in S^2_+$. Conversely, Eq. (37) defines a great circle on S^2_+ when we fix $\boldsymbol{x} \in S^2$. This geometric duality provides us the standard Hough transform, SHT in abbreviated form, and the randomized Hough transform [26, 25], RHT in abbreviated form, for great-circle detection.

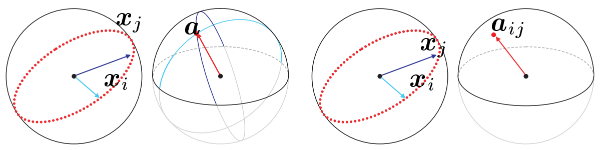
4.2 SHT and RHT on Sphere

On the basis of the duality of a point on a sphere and a great circle on a hemisphere, it is possible to achieve SHT for great-circle detection. Setting $\boldsymbol{P} = \{\boldsymbol{x}_i = (x_i, y_i, z_i)^\top \in S^2\}_{i=1}^n$ to be samples on a spherical image, Eq. (37) leads to the transform from a collection of points $\boldsymbol{P}$ on S^2 to great circles on S^2_+. Let the function $u(\tau)$ be

$$u(\tau) = \begin{cases} 1, & \text{if } \tau = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{38}$$

The voting is then expressed by

$$h(\boldsymbol{a} \in S^2_+) = \sum_{\boldsymbol{x}_i \in \boldsymbol{P}} u(\boldsymbol{a}^\top \boldsymbol{x}_i). \tag{39}$$



(a) (b) (c) (d)

Figure 13: Duality and Hough transform. For point-to-great-circle voting for SHT, (a) points $\{\boldsymbol{x}_i \in S^2\}_{i=1}^n$ yield great circles on a dual parameter space S^2_+, and (b) the intersection of the great circles expresses the parameter of the great circle on the original spherical image. For points-to-point voting for RHT, (c) randomly selected two points on a spherical image are sampled, and (d) the cross product of the two points $\boldsymbol{x}_i, \boldsymbol{x}_j \in S^2$ yields a point $\boldsymbol{a}_{ij}$ on a dual parameter space S^2_+. This point $\boldsymbol{a}_{ij}$ expresses the parameter of the great circle on the original spherical image.

This is the mathematical expression of point-to-great-circle voting for SHT. Figures 13 (a) and (b) show this point-to-great-circle voting.

On the other hand, since the duality of a point and a great circle also defines the transform from a pair of points on S^2 to a point on S^2_+, it is possible to achieve RHT for great-circle detection. Setting $\boldsymbol{P} = \{\boldsymbol{x}_i = (x_i, y_i, z_i)^\top \in S^2\}_{i=1}^n$ to be samples on a spherical image, the solution of the system of equations,

$$\boldsymbol{a}^\top \boldsymbol{x}_i = 0, \quad \boldsymbol{a}^\top \boldsymbol{x}_j = 0, \tag{40}$$

where $\boldsymbol{a} = (a, b, c)^\top \in S^2_+$ and $i \neq j$, defines a point as the common point of a pair of great circles on S^2_+. This system of equations is equivalent to the system of equations,

$$u(\boldsymbol{a}^\top \boldsymbol{x}_i) = 1, \quad u(\boldsymbol{a}^\top \boldsymbol{x}_j) = 1, \tag{41}$$

where $u(\tau)$ is the function defined in Eq. (38). The solution of the systems of Eqs. (40) and (41) is equivalently computed by the cross-product of $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, such that

$$\boldsymbol{a}_{ij} = \lambda \frac{\boldsymbol{x}_i \times \boldsymbol{x}_j}{|\boldsymbol{x}_i \times \boldsymbol{x}_j|}, \quad \lambda \in \{-1, 1\}, \tag{42}$$

where λ is selected so that vector $\boldsymbol{a}_{ij}$ lies on S^2_+. These mathematical properties derive the voting for RHT expressed by

$$v(\boldsymbol{a} \in S^2_+) = \sum_{\boldsymbol{x}_i, \boldsymbol{x}_j \in \boldsymbol{P}} u(\boldsymbol{a} - \boldsymbol{a}_{ij}), \tag{43}$$

for $i \neq j$. This is the mathematical expression of points-to-point voting for RHT. Figures 13 (c) and (d) show this points-to-point voting. Since the randomized Hough transform involves a random sampling process [26, 25], this is the fast version of SHT. Furthermore, in the phase of practical implementation, the points-to-point voting requires a small memory space as the Hough space compared with the point-to-great-circle voting. This is the main reason why we use RHT on a sphere for great-circle detection.

Table 1: **Algorithm 1: RHT for great-circle detection on a sphere**

```
i. Repeat for a predetermined number of iterations $N$,
        Operate the voting using Eqs. (42) and (45) and accumulate the voting to the cells.
ii. Detect peaks of the accumulation.
iii. Output the peaks as great circles.
```

In RHT on a sphere, voting is accumulated to a cell generated by tessellating the dual hemispherical space S^2_+. Parameters $\boldsymbol{a}$ of great circles are determined by detecting peaks of the accumulation. For practical operation of RHT on a sphere, the dual hemispherical space S^2_+ is tessellated to a collection of cells $\boldsymbol{D}$. Hereafter, we call this discrete hemispherical space $\boldsymbol{D}$ *the discrete Hough space*. Let the function $\hat{u}(\tau)$ be

$$\hat{u}(\tau) = \begin{cases} 1, & \text{if } |\tau| = T, \\ 0, & \text{otherwise,} \end{cases} \tag{44}$$

where T is a given real constant determined by the size of the cell. Therefore, the voting expressed in Eq. (43) is modified as

$$v(\boldsymbol{a} \in \boldsymbol{D}) = \sum_{\boldsymbol{x}_i, \boldsymbol{x}_j \in \boldsymbol{P}} \hat{u}(\boldsymbol{a} - \boldsymbol{a}_{ij}). \tag{45}$$

An algorithm of RHT with tessellated cell parameter space is implemented as shown in Table 1.

4.3 RHT-based Method for Great-Circle Detection

In this section, we propose a RHT-based method with a continuous Hough space. In this method, for a predetermined number of iterations, we repeat the random sampling and points-to-point mapping using Eq. (42) instead of voting and accumulation on RHT. Note that on the basis of geometric duality, points on a unique great circle define a unique point on a hemisphere. Therefore, if there is no noise for the samples, which corresponds to a unique great circle on the original image, points computed from a randomly selected pair of samples using Eq. (42) always denote a unique point in the parameter space S^2_+. However, because of the noise in the original image, points computed using Eq.

Table 2: **Algorithm 2: RHT-based method for great-circle detection on a sphere**

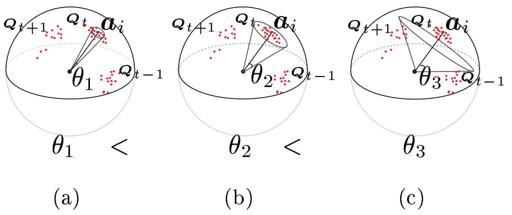
i. *Repeat for a predetermined number of iterations N,*

   For a randomly selected pair of points, operate the points-to-point mapping using Eq. (42).

ii. *Operate clustering of a point cloud on* S^2_+.

iii. *Repeat M times for the number of clusters M,*

   Compute a centroid of a cluster and output the centroid as a parameter of a great circle on the original spherical image.



(a) (b) (c)

Figure 14: Clustering of point clouds. (a) When the opening angle, which defines the size of the region $\Omega_{\boldsymbol{a_i}}(\theta_k)$, is small, the region $\Omega_{\boldsymbol{a_i}}(\theta_k)$ does not include a sufficient collection of points. (b) We select the the opening angle θ_2 when the cardinality of points $\overline{\boldsymbol{P}}$ in the region $\Omega_{\boldsymbol{a_i}}(\theta_k)$ becomes locally constant. Using this region, it is possible to extract a collection of points $\boldsymbol{Q}_t$ from the point clouds $\boldsymbol{Q}$. (c) If the opening angle of a cone is large, the cone includes other collections of points.

(42), which should correspond to a great circle in the original image, are distributed in the neighborhood of the true point on S^2_+. Therefore, the points yielded by random sampling and points-to-point mapping are distributed as a point cloud on S^2_+. Furthermore, when there exist many lines in the original image, this random sampling and points-to-point mapping procedure yields a number of point clouds on S^2_+. We need a process for clustering point clouds so that each cluster represents a great circle. Such an algorithm is summarized in Table 2.

4.3.1 Clustering of Point Cloud on Sphere

To cluster the point cloud on S^2_+, we set a collection of points $\boldsymbol{Q} = \{\boldsymbol{a}\}_{i=1}^m$ that expresses the point cloud. As shown in Figure 14, first, we generate a region $\Omega_{\boldsymbol{a_i}}(\theta)$ on S^2_+ defined by the intersection of a cone and a hemisphere, that is, the inside of a small circle on a sphere. The cone has the opening angle θ and the central axis of the cone passes through a point $\boldsymbol{a}_i \in S^2_+$ arbitrarily selected from $\boldsymbol{Q}$. Second, for θ_k, we compute the cardinality of points $\boldsymbol{Q}$ in the region $\Omega_{\boldsymbol{a_i}}(\theta_k)$ such that

$$\mathcal{C}(\theta_k;\boldsymbol{a}_i) = |\boldsymbol{Q} \cap \Omega_{\boldsymbol{a_i}}(\theta)|. \tag{46}$$

Third, by increasing the opening angle θ_k as $\theta_{k+1} = \theta_k + \kappa$, where κ is a given constant and $\theta_0 = \kappa$, we expand a region $\Omega_{\boldsymbol{a_i}}(\theta_k)$. Then, we detect the opening angle θ_k when the cardinality $\mathcal{C}(\theta_k;\boldsymbol{a}_i)$ of points in the region $\Omega_{\boldsymbol{a_i}}(\theta_k)$ becomes locally constant. We define this opening angle as the minimum opening angle $\theta_{\min}$ of $\boldsymbol{a}_i$. Finally, using the minimum opening angle $\theta_{\min}$, we extract a collection of points $\boldsymbol{Q}_t$ from $\boldsymbol{Q}$. Then, we compute the centroid of the collection of points $\boldsymbol{Q}_t$ as the parameter of a great circle on the original spherical image.

For this clustering, there does not exist any mathematical criterion for the selection of $\boldsymbol{a}_i \in S^2_+$ from $\boldsymbol{Q}$. Therefore, for all points in $\boldsymbol{Q}$, we first compute the minimum opening angle $\theta_{\min}$ of $\boldsymbol{a}_i$ based on the cardinality $\mathcal{C}(\theta_k;\boldsymbol{a}_i)$. Second, we select the point that has the largest cardinality in the region $\Omega_{\boldsymbol{a_i}}(\theta_{min})$. Third, we compute the centroid of the collection of points $\boldsymbol{Q}_t$ and remove the collection of points that lie in the region $\Omega_{\boldsymbol{a_i}}(\theta_{min})$ from $\boldsymbol{Q}$. These operations merge the collection of points, which may express the same line, in the continuous Hough space. Finally, we repeat these operations for the remaining points $\boldsymbol{Q}\backslash\boldsymbol{Q}_t$ in the parameter space. We summarize the RHT-based algorithm for great-circle detection in Table 2.

4.4 Numerical Examples

4.4.1 Edge Detection

In this section, we show numerical examples of the detection of great circles from spherical images. As a pre-process to great-circle detection, we extract sample points on a sphere by edge detection. Two methods are applicable. The first is a method in which an input original image is transformed onto a spherical image and edges are extracted as sample points on the spherical image. In the second method, the edges on an original central-camera image are extracted using the standard image processing technique and then the points extracted as sample points are mapped to a sphere. We use the second method in this numerical experiment.

4.4.2 Prescreening of Meaningless Sampling

The points-to-point mapping based on the random sampling of a pair of points yields a meaningless point in the continuous Hough space if the selected two points do not lie on the same great circle. To eliminate

Table 3: **Algorithm 3: RHT-based method for great-circle detection with prescreening based on collinearity**

```
i   Repeat for a predetermined number of iterations $N$,
        For three randomly selected points, test the collinearity.
        If the three points are collinear, operate the points-to-point mapping
        using Eqs. (42) and (48)
ii  Operate clustering of a point cloud on $S^2_+$.
iii Repeat $M$ times for the number of clusters $M$,
        Compute a centroid of a cluster and output the centroid as a parameter of a great circle
        on the original spherical image.
```

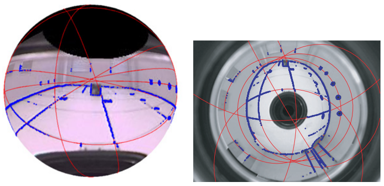
meaningless sampling, we adopt the idea of three-point Hough transform [4]. This prescreening process tests the collinearity of three sample points on the original image. The collinearity is evaluated by the scalar triple product for the three sample points $\boldsymbol{x}_i$, $\boldsymbol{x}_j$ and $\boldsymbol{x}_k \in S^2$ as

$$\mathcal{G}(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{x}_k) = |(\boldsymbol{x}_i \times \boldsymbol{x}_j)^\top \boldsymbol{x}_k|. \qquad (47)$$

If the selected three points are collinear, we map a point $\boldsymbol{a}_{ijk}$ to the continuous Hough space S^2_+ as

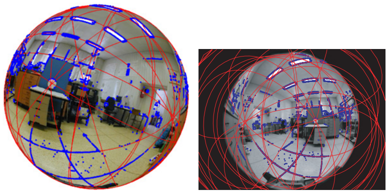
$$\boldsymbol{a}_{ijk} = \frac{1}{3}(\boldsymbol{a}_{ij} + \boldsymbol{a}_{jk} + \boldsymbol{a}_{ki}). \qquad (48)$$

For $\alpha, \beta \in \{i, j, k\}$, $\boldsymbol{a}_{\alpha\beta}$ is computed using Eq. (42). This prescreening process for eliminating meaningless sampling leads to the algorithm in Table 3.



(a) (b)

Figure 15: Results of numerical experiments using the spherical image transformed from a hyperbolic-camera image. (a) Using edge points extracted by the Laplacian filter, eight great circles are detected on the spherical image. (b) The great circles back-projected onto the original image.



(a) (b)

Figure 16: Results of numerical experiments using the spherical image transformed from a fish-eye-lens-camera image. (a) Using edge points extracted by the Laplacian filter, 20 great circles detected on the spherical image. (b) The great circles back-projected onto the original image.

4.4.3 Numerical Examples

We employed **Algorithm 3** for the detection of great circles on a sphere. Figure 15 shows the results of great-circle detection on a sphere. Sample points on a sphere are prepared using a hyperbolic-camera image. The original hyperbolic-camera image is acquired using a SOIOS-55 CAM whose mirror is hyperboloidal. The original catadioptric image has 640 × 480 pixels. We extract 2236 edge points from the original image using a Laplacian filter. The extracted edge points are transformed onto a unit sphere as sample points. Figure 15 (a) shows the sample points and the eight great circles detected by our RHT-based method for great-circle detection. To evaluate our result on the original image, the detected great circles are back-projected to the original image, as shown in Fig. 15 (b).

Figure 16 shows the results of great-circle detection on a sphere. Sample points on a sphere are prepared using a fish-eye-lens-camera image. The original fish-eye-lens-camera image is acquired using Nikon Coolpix 950 and a Nikon Fisheye converter FC-E8, which is a fish-eye-lens converter. We extract 5572 edge points from the original image using a Laplacian filter. The extracted edge points are transformed onto a unit sphere as sample points. Figure 16 (a) shows the sample points and the 20 great circles detected by our RHT-based method on the spherical image. In the same manner as in the experiment for the catadioptric image, to evaluate our result on the original image, the detected great circles are back-projected to the original image, as shown in Fig. 16 (b).

These experimental results lead to the conclusion that our RHT-based method is applicable to different types of central-camera images since the algorithm is achieved on the basis of images on a sphere. Furthermore, our method semi-automatically detects the number of clusters, that is, the numbers of great circles. Moreover, the results show that the great circles are robustly distinct from background noise.

5 Multiple View Geometry for Spherical Cameras

In this section, we formulate multiple-view geometry based on the spherical-camera model defined in Section 3. We deal with two-view, three-view and four-view geometry since there exist multilinear constraints up to four views [11]. The multiple view geometry based on the spherical-camera model naturally represents geometric constraints in the dual space. On the basis of the analysis of the geometric constraints in the dual space, we derive the epipolar circumcircle constraint.

5.1 Two-View Geometry for Spherical Cameras

5.1.1 Epipolar Geometry for Spherical Cameras

Setting $\boldsymbol{C}$ and $\boldsymbol{C}'$ to be centers of spherical cameras, points $\boldsymbol{X} \in \mathbb{R}^3$ and $\boldsymbol{X}' \in \mathbb{R}^3$ are mapped to points $\boldsymbol{x}$ and $\boldsymbol{x}' \in S^2$ on the spherical camera images as shown in Figure 17. We assume that $\boldsymbol{C} = \boldsymbol{O}$, and R and $\boldsymbol{t}$ express the transform of the camera coordinates between $\boldsymbol{C}$ and $\boldsymbol{C}'$. The mappings are expressed as follows,

$$\lambda \boldsymbol{x} = \boldsymbol{X}, \quad \lambda \in \mathbb{R} \tag{49}$$

$$\lambda' \boldsymbol{x}' = \boldsymbol{X}' = \mathrm{R}\boldsymbol{X} + \boldsymbol{t}, \quad \lambda' \in \mathbb{R}. \tag{50}$$

Therefore, we have the equation,

$$\lambda' \boldsymbol{x}' - \lambda \mathrm{R}\boldsymbol{x} - \boldsymbol{t} = 0. \tag{51}$$

In the same manner as the epipolar geometry based on a pinhole camera model [5, 11], since the vectors $\boldsymbol{x}'$, $\mathrm{R}\boldsymbol{x}$ and $\boldsymbol{t}$ are coplanar, we obtain the epipolar constraint on the spherical-camera model, that is,

$$\boldsymbol{x}'\mathrm{E}\boldsymbol{x} = 0, \quad \mathrm{E} = [\boldsymbol{t}]_\times \mathrm{R}, \tag{52}$$

where $[\boldsymbol{t}]_\times$ is a skew-symmetric matrix. The 3×3 matrix E is linearly computed using eight pairs of corresponding points on the spherical images.

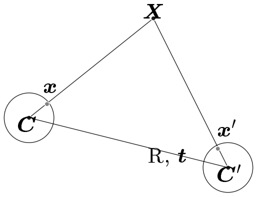
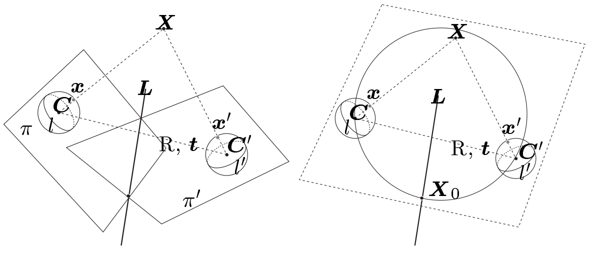


Figure 17: Epipolar geometry for spherical cameras. For a pair of points $\boldsymbol{x} \in S^2$ and $\boldsymbol{x}' \in S^2$ corresponding to a point $\boldsymbol{X} \in \mathbb{R}^3$, the points $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{X}$ and the spherical camera centers $\boldsymbol{C}$ and $\boldsymbol{C}'$ are coplanar.



(a) (b)

Figure 18: Epipolar circumcircle for spherical cameras. (a) The geometric duality of points and great circles on a sphere derives that points $\boldsymbol{x}$ and $\boldsymbol{x}' \in S^2$ draw great circles C and C', respectively. The intersection of the planes π and π', that include $\boldsymbol{C}$ and $\boldsymbol{C}'$, yields a line in $\mathbb{R}^3$. (b) The six points $\boldsymbol{X}_0$, $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{C}$, $\boldsymbol{C}'$ and $\boldsymbol{X}$ exist on the same epipolar plane. The four points $\boldsymbol{X}_0$, $\boldsymbol{C}$, $\boldsymbol{C}'$ and $\boldsymbol{X}$ are on a circle on the epipolar plane.

5.1.2 Geometric Aspect of Two-View Geometry Based on Duality

The geometric duality of points and great circles on a sphere derives that points $\boldsymbol{x}$ and $\boldsymbol{x}' \in S^2$ draw great circles l and l' on spherical images, respectively. Since l and l' are on planes π and π' in $\mathbb{R}^3$ that pass through the centers of cameras $\boldsymbol{C}$ and $\boldsymbol{C}'$, the intersection of these planes yields a line $\boldsymbol{L}$ in $\mathbb{R}^3$ as shown in Figure 18 (a) [4]. Furthermore, we set $\boldsymbol{X}_0$ to be a point of intersection of the line $\boldsymbol{L}$ and the epipolar plane on which $\boldsymbol{x}$, $\boldsymbol{x}'$ and $\boldsymbol{X}$ are included. Since we select the points $\boldsymbol{x}$ and $\boldsymbol{x}'$ corresponding to the point $\boldsymbol{X}$, the six points $\boldsymbol{X}_0$, $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{C}$, $\boldsymbol{C}'$ and $\boldsymbol{X}$ exist on the same epipolar plane as shown in Figure 18 (b). Moreover, since $\boldsymbol{X}_0$ is the intersection of lines, which pass through $\boldsymbol{C}$ and $\boldsymbol{C}'$ and are perpendicular to $\boldsymbol{x}$ and $\boldsymbol{x}'$, on the epipolar plane, the four points $\boldsymbol{X}_0$, $\boldsymbol{C}$, $\boldsymbol{C}'$ and $\boldsymbol{X}$ are on a circle on the epipolar plane as shown in Figure 18 (b). On the other hand, the convex polygon formed by these four points has a circumcircle. We call this circle as the *epipolar circumcircle.* This circle has a center $e_0 = \frac{1}{2}(\boldsymbol{X}_0 + \boldsymbol{X})$ and a radius $e_r = \frac{1}{2}|(\boldsymbol{X}_0 + \boldsymbol{X})|$. The epipolar circumcircle clarifies the geometric relation of a point at infinity and the epipole. If $\boldsymbol{X}$ is a point at infinity, $\boldsymbol{X}_0$ is on the epipole and π and π' become an identical plane. On the other hand, if $\boldsymbol{X}_0$ is a point at infinity, $\boldsymbol{X}$ is on the epipole and π and π' are parallel.

---

[4] Notice that the direction vector of this line $\boldsymbol{L}$ corresponds to the normal vector of the epipolar plane. Therefore, we have the relation,

$$\boldsymbol{p} = \frac{\boldsymbol{x}' \times \mathrm{R}\boldsymbol{x}}{|\boldsymbol{x}' \times \mathrm{R}\boldsymbol{x}|} = \frac{\boldsymbol{t} \times \mathrm{R}\boldsymbol{x}}{|\boldsymbol{t} \times \mathrm{R}\boldsymbol{x}|} = \frac{\boldsymbol{x}' \times \boldsymbol{t}}{|\boldsymbol{x}' \times \boldsymbol{t}|}. \tag{53}$$

5.2 Three-View Geometry for Spherical Cameras

5.2.1 Trifocal Tensor for Spherical Cameras

We express a line in a space by parametric form $\boldsymbol{L} : \boldsymbol{X}_I + \eta \boldsymbol{p} \in \mathbb{R}^3$ where $\boldsymbol{X}_I \in \mathbb{R}^3$, $\boldsymbol{p} \in \mathbb{R}^3$ and $\eta \in \mathbb{R}$. Setting $\boldsymbol{C}$, $\boldsymbol{C}'$ and $\boldsymbol{C}''$ to be centers of spherical cameras, lines $\boldsymbol{L}$, $\boldsymbol{L}'$ and $\boldsymbol{L}'' \in \mathbb{R}^3$ are mapped as great circles l, l' and $l'' \in S^2$ on the spherical camera images as shown in Figure 19. We assume that $\boldsymbol{C} = \boldsymbol{O}$ at the origin of world coordinates. R' and $\boldsymbol{t}'$, and R'' and $\boldsymbol{t}''$ express the transforms of the camera coordinates among $\boldsymbol{C}$ and $\boldsymbol{C}'$, and $\boldsymbol{C}$ and $\boldsymbol{C}''$, respectively.

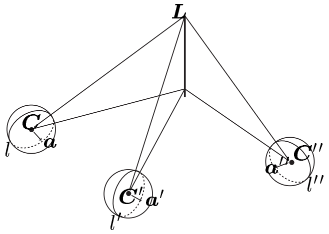


Figure 19: Three view geometry for spherical cameras. A line in a space is mapped as great circles on spherical images.

According to the geometric duality between a great circle and a point on a sphere, the great circles l, l' and l'' equivalently denote points $\boldsymbol{a}$, $\boldsymbol{a}'$ and $\boldsymbol{a}'' \in S^2$. It is possible to express the relations among the lines $\boldsymbol{L}$, $\boldsymbol{L}'$ and $\boldsymbol{L}''$ in a space and the points $\boldsymbol{a}$, $\boldsymbol{a}'$ and $\boldsymbol{a}''$ on a sphere, such that,

$$\boldsymbol{a}^\top \boldsymbol{L} = 0, \tag{54}$$

$$\boldsymbol{a}'^\top \boldsymbol{L}' = \boldsymbol{a}'^\top (\mathrm{R}' \boldsymbol{L} + \boldsymbol{t}') = 0, \tag{55}$$

$$\boldsymbol{a}''^\top \boldsymbol{L}'' = \boldsymbol{a}''^\top (\mathrm{R}'' \boldsymbol{L} + \boldsymbol{t}'') = 0. \tag{56}$$

Using 3×4 matrix M, we can express these equations as the following form,

$$\mathrm{M}\tilde{\boldsymbol{L}} = \mathbf{O}, \tag{57}$$

where

$$\mathrm{M} = \begin{bmatrix} \boldsymbol{a}^\top & 0 \\ \boldsymbol{a}'^\top \mathrm{R}' & \boldsymbol{a}'^\top \boldsymbol{t}' \\ \boldsymbol{a}''^\top \mathrm{R}'' & \boldsymbol{a}''^\top \boldsymbol{t}'' \end{bmatrix}, \tag{58}$$

and $\tilde{\boldsymbol{L}}^\top = (\boldsymbol{L}^\top, 1)^\top$. When this equation implies a line in a space, rank(M) = 2. Setting $\mathrm{M}^\top = (\boldsymbol{m}_1, \boldsymbol{m}_2, \boldsymbol{m}_3)$,

$$\boldsymbol{m}_1 = \alpha \boldsymbol{m}_2 + \beta \boldsymbol{m}_3. \tag{59}$$

Since $m_{14} = 0$, $\alpha = k\boldsymbol{t}''^\top \boldsymbol{a}''$ and $\beta = -k\boldsymbol{t}'^\top \boldsymbol{a}'$ where $k \in \mathbb{R}$. Again, in the same manner as the trifocal tensor based on the pinhole camera model [11], up to a homogeneous scale factor, we can express the relations among $\boldsymbol{a}$, $\boldsymbol{a}'$ and $\boldsymbol{a}''$ using tensor notation, such that,

$$\begin{aligned} \boldsymbol{a} &= \alpha \mathrm{R}'^\top \boldsymbol{a}' + \beta \mathrm{R}''^\top \boldsymbol{a}'' \\ &= (\boldsymbol{t}''^\top \boldsymbol{a}'')\mathrm{R}'^\top \boldsymbol{a}' - (\boldsymbol{t}'^\top \boldsymbol{a}')\mathrm{R}''^\top \boldsymbol{a}'' \\ &= (\boldsymbol{a}''^\top \boldsymbol{t}'')\mathrm{R}'^\top \boldsymbol{a}' - (\boldsymbol{a}'^\top \boldsymbol{t}')\mathrm{R}''^\top \boldsymbol{a}''. \end{aligned} \tag{60}$$

For ith element of this vector,

$$\begin{aligned} a_i &= \boldsymbol{a}''^\top (\boldsymbol{t}'' \mathrm{R}_{\mathrm{i}}'^\top) \boldsymbol{a}' - \boldsymbol{a}'^\top (\boldsymbol{t}' \mathrm{R}_{\mathrm{i}}''^\top) \boldsymbol{a}'' \\ &= \boldsymbol{a}'^\top (\mathrm{R}_{\mathrm{i}}'^\top \boldsymbol{t}'') \boldsymbol{a}'' - \boldsymbol{a}'^\top (\boldsymbol{t}' \mathrm{R}_{\mathrm{i}}''^\top) \boldsymbol{a}'', \end{aligned} \tag{61}$$

where R_{i}' and $\mathrm{R}_{\mathrm{i}}''$ are the ith column vectors of R' and R''. We have the trifocal tensor $[\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3]$ which are the set of matrices $\mathrm{T}_{\mathrm{i}} = \mathrm{R}_{\mathrm{i}}'^\top \boldsymbol{t}'' - \boldsymbol{t}' \mathrm{R}_{\mathrm{i}}''^\top$. Rewriting Eqs. (61) and (60), we have the equations,

$$a_i = \boldsymbol{a}'^\top \mathrm{T}_{\mathrm{i}} \boldsymbol{a}'', \tag{62}$$

$$\boldsymbol{a}^\top = \boldsymbol{a}'^\top [\mathrm{T}_1, \mathrm{T}_2, \mathrm{T}_3] \boldsymbol{a}''. \tag{63}$$

Since a point $\boldsymbol{x}$ on a great circle $\boldsymbol{a}$ satisfies

$$\boldsymbol{x}^\top \boldsymbol{a} = 0, \tag{64}$$

we have point-line-line correspondence, such that,

$$\boldsymbol{a}'^\top (\sum_i^3 x^i \mathrm{T}_{\mathrm{i}}) \boldsymbol{a}'' = 0. \tag{65}$$

Furthermore, setting $\boldsymbol{x}'$ and $\boldsymbol{y}'$, and $\boldsymbol{x}''$ and $\boldsymbol{y}''$ to be the points on great circles $\boldsymbol{a}'$ and $\boldsymbol{a}''$, respectively, we have the relations,

$$\boldsymbol{a}' = k' \boldsymbol{x}' \times \boldsymbol{y}' = k' [\boldsymbol{x}']_\times \boldsymbol{y}', \tag{66}$$

$$\boldsymbol{a}'' = k'' \boldsymbol{x}'' \times \boldsymbol{y}'' = k'' [\boldsymbol{x}'']_\times \boldsymbol{y}'', \tag{67}$$

where k' and $k'' \in \mathbb{R}$ are selected so that $\boldsymbol{a}'$ and $\boldsymbol{a}''$ are on S^2. Eq. (65) is now represented as

$$\boldsymbol{a}'^\top (\sum x^i \mathrm{T}_{\mathrm{i}}) [\boldsymbol{x}'']_\times \boldsymbol{y}'' = 0, \tag{68}$$

$$\boldsymbol{y}''^\top [\boldsymbol{x}''] (\sum x^i \mathrm{T}_{\mathrm{i}}) [\boldsymbol{x}'']_\times \boldsymbol{y}'' = 0. \tag{69}$$

Since $\boldsymbol{x}'$ and $\boldsymbol{x}''$ are independent to $\boldsymbol{y}'$ and $\boldsymbol{y}''$, respectively, we have point-point-line and point-point-point correspondence as follows,

$$\boldsymbol{a}'^\top (\sum x^i \mathrm{T}_{\mathrm{i}}) [\boldsymbol{x}'']_\times = \mathbf{0}_{1\times 3}, \tag{70}$$

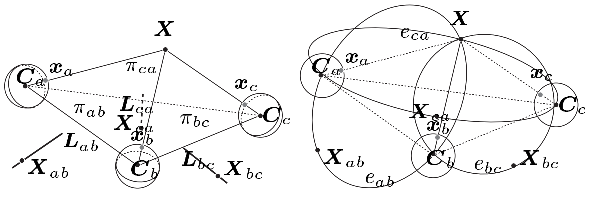
$$[\boldsymbol{x}''] (\sum x^i \mathrm{T}_{\mathrm{i}}) [\boldsymbol{x}'']_\times = \mathbf{0}_{3\times 3}. \tag{71}$$

Noting that, for the derivations of point-point-line and point-point-point correspondence, homography is not used because the homography among spherical images are not well-defined.

5.2.2 Geometric Aspect of Three-View Geometry Based on Duality

In this section, we describe the epipolar circumcircle constraint for the three-view geometry based on the spherical-camera model. For three cameras $\boldsymbol{C}_a$, $\boldsymbol{C}_b$ and $\boldsymbol{C}_c$, setting $\boldsymbol{x}_a$, $\boldsymbol{x}_b$ and $\boldsymbol{x}_c \in S^2$ to be a triplet of corresponding points to the point $\boldsymbol{X} \in \mathbb{R}^3$, respectively, there exist three epipolar planes π_{ab}, π_{bc} and π_{ca} as shown in Figure 20 (a). The duality of a point and a great circle on a sphere yields planes ξ_a, ξ_b and ξ_c for $\boldsymbol{x}_a$, $\boldsymbol{x}_b$ and $\boldsymbol{x}_c \in S^2$. The intersections of ξ_a and ξ_b, ξ_b and ξ_c, and ξ_c and ξ_a yield lines $\boldsymbol{L}_{ab}$, $\boldsymbol{L}_{bc}$ and $\boldsymbol{L}_{ca} \in \mathbb{R}^3$, respectively. Furthermore, we set $\boldsymbol{X}_{ab}$, $\boldsymbol{X}_{bc}$ and $\boldsymbol{X}_{ca}$ are the points of intersection of $\boldsymbol{L}_{ab}$ and π_{ab}, $\boldsymbol{L}_{bc}$ and π_{bc}, and $\boldsymbol{L}_{ca}$ and π_{ca}.

For the triplets $\boldsymbol{C}_a$, $\boldsymbol{C}_b$ and $\boldsymbol{X}_{ab}$, $\boldsymbol{C}_b$, $\boldsymbol{C}_c$ and $\boldsymbol{X}_{bc}$, and $\boldsymbol{C}_c$, $\boldsymbol{C}_a$ and $\boldsymbol{X}_{ca}$, there exist three epipolar circumcircles e_{ab}, e_{bc} and e_{ca}, respectively, as shown in Figure 20 (b). Obviously, since we assumed the triplet of points are corresponded, the three epipolar circumcircles intersect the point $\boldsymbol{X}$ in a space. This intersection of epipolar circumcircles is the geometric aspect of three-view geometry based on the duality.



(a) (b)

Figure 20: Epipolar circumcircle constraint for spherical cameras. (a) For three cameras $\boldsymbol{C}_a$, $\boldsymbol{C}_b$ and $\boldsymbol{C}_c$, the triplet of points $\boldsymbol{x}_a$, $\boldsymbol{x}_b$ and $\boldsymbol{x}_c \in S^2$ corresponding to a point $\boldsymbol{X} \in \mathbb{R}^3$ yields three epipolar planes π_{ab}, π_{bc} and π_{ca}. The intersections of ξ_a and ξ_b, ξ_b and ξ_c, and ξ_c and ξ_a yield lines $\boldsymbol{L}_{ab}$, $\boldsymbol{L}_{bc}$ and $\boldsymbol{L}_{ca} \in \mathbb{R}^3$, respectively. Furthermore, we set $\boldsymbol{X}_{ab}$, $\boldsymbol{X}_{bc}$ and $\boldsymbol{X}_{ca}$ are the points of intersection of $\boldsymbol{L}_{ab}$ and π_{ab}, $\boldsymbol{L}_{bc}$ and π_{bc}, and $\boldsymbol{L}_{ca}$ and π_{ca}. (b) For points $\boldsymbol{C}_a$, $\boldsymbol{C}_b$ and $\boldsymbol{X}_{ab}$, $\boldsymbol{C}_b$, $\boldsymbol{C}_c$ and $\boldsymbol{X}_{bc}$, and $\boldsymbol{C}_c$, $\boldsymbol{C}_a$ and $\boldsymbol{X}_{ca}$, there exist three epipolar circumcircles e_{ab}, e_{bc} and e_{ca}, respectively.

5.3 Four-View Geometry for Spherical Cameras

5.3.1 Quadrifocal Tensor for Spherical Cameras

In the same manner as the formulation of quadrilinear relationship and quadrifocal tensor on the pinhole-camera model [11, 13], we formulate the four-view geometry on the spherical-camera model.

Setting $\boldsymbol{C}$, $\boldsymbol{C}'$, $\boldsymbol{C}''$ and $\boldsymbol{C}'''$ to be centers of spherical cameras, $[\mathrm{R}, \boldsymbol{t}]$, $[\mathrm{R}', \boldsymbol{t}']$, $[\mathrm{R}'', \boldsymbol{t}'']$, and $[\mathrm{R}''', \boldsymbol{t}''']$ express the transform from the camera coordinates to the world coordinate systems. For a point $\boldsymbol{X} \in \mathbb{R}^3$, we have the four equations $\lambda \boldsymbol{x} = \mathrm{R}\boldsymbol{X} + \boldsymbol{t}$, $\lambda' \boldsymbol{x}' = \mathrm{R}'\boldsymbol{X} + \boldsymbol{t}'$, $\lambda'' \boldsymbol{x}'' = \mathrm{R}''\boldsymbol{X} + \boldsymbol{t}''$, $\lambda''' \boldsymbol{x}''' = \mathrm{R}'''\boldsymbol{X} + \boldsymbol{t}'''$, where $\lambda, \lambda', \lambda''$ and $\lambda''' \in \mathbb{R}$, and $\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{x}''$ and $\boldsymbol{x}''' \in S^2$ on the spherical images, respectively. We can re-write the equations,

$$
\left[\begin{array}{cccccc}
\mathbf{r}_1 & t_1 & u^1 & & & \\
\mathbf{r}_2 & t_2 & u^2 & & & \\
\mathbf{r}_3 & t_3 & u^3 & & & \\
\hline
\mathbf{r}'_1 & t'_1 & & u'^1 & & \\
\mathbf{r}'_2 & t'_2 & & u'^2 & & \\
\mathbf{r}'_3 & t'_3 & & u'^3 & & \\
\hline
\mathbf{r}''_1 & t''_1 & & & u''^1 & \\
\mathbf{r}''_2 & t''_2 & & & u''^2 & \\
\mathbf{r}''_3 & t''_3 & & & u''^3 & \\
\hline
\mathbf{r}'''_1 & t'''_1 & & & & u'''^1 \\
\mathbf{r}'''_2 & t'''_2 & & & & u'''^2 \\
\mathbf{r}'''_3 & t'''_3 & & & & u'''^3
\end{array}\right]
\begin{bmatrix} \boldsymbol{X} \\ 1 \\ -\lambda \\ -\lambda' \\ -\lambda'' \\ -\lambda''' \end{bmatrix} = 0, \tag{72}
$$

where $\mathbf{r}_i$, $\mathbf{r}'_i$, $\mathbf{r}''_i$ and $\mathbf{r}'''_i$ are the i-th row vectors of the matrices R, R$'$, R$''$ and R$'''$, t_i, t'_i, t''_i and t'''_i are the i-th element of the vectors $\boldsymbol{t}$, $\boldsymbol{t}'$, $\boldsymbol{t}''$ and $\boldsymbol{t}'''$, and u^i, u'^i, u''^i and u'''^i are the i-th element of the vectors $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{x}''$ and $\boldsymbol{x}'''$, respectively.

Since $\boldsymbol{x}$, $\boldsymbol{x}'$, $\boldsymbol{x}''$ and $\boldsymbol{x}'''$ correspond to the same point $\boldsymbol{X}$, this equation has a solution. The 12×8 matrix has rank at most 7. Therefore, all determinants of the 8×8 matrices which consist of eight rows selected from the 12×8 matrix are zero. We set that ϵ_{ijk} is the tensor as follows:

$$
\epsilon_{ijk} = \begin{cases} 0 & \text{unless } i,j \text{ and } k \text{ are distinct} \\ +1 & \text{if } ijk \text{ is an even permutation of 123} \\ -1 & \text{if } ijk \text{ is an odd permutation of 123.} \end{cases} \tag{73}
$$

Expansion of the determinant leads to a quadrilinear relationship

$$
u_i u'_j u''_k u'''_l \epsilon_{ipw} \epsilon_{jqx} \epsilon_{kry} \epsilon_{lsz} Q^{pqrs} = 0_{wxyz} \tag{74}
$$

where 0_{wxyz} is a zero tensor with four indices w, x, y and z, and the rank-4 tensor Q^{pqrs} is

$$
Q^{pqrs} = \det \begin{bmatrix} \mathbf{r}_p & t_p \\ \mathbf{r}'_q & t'_q \\ \mathbf{r}''_r & t''_r \\ \mathbf{r}'''_s & t'''_s \end{bmatrix}. \tag{75}
$$

Q^{pqrs} is the quadrifocal tensor for the correspondence of four points on spherical images. Finally, we set that a_i, a'_i, a''_i and a'''_i are the i-th element of great circles $\boldsymbol{a}$, $\boldsymbol{a}'$, $\boldsymbol{a}''$ and $\boldsymbol{a}'''$ on the four spherical images. The quadrifocal tensor for great circles is expressed as

$$
a_p a'_q a''_r a'''_s Q^{pqrs} = 0. \tag{76}
$$

5.3.2 Numerical Examples

Computation of Epipolar Geometry

We compute the epipolar geometry between two spherical images. Figure 21 (a) shows a spherical image transformed from a conventional pinhole-camera image in Figure 21 (b). Figure 21 (c) shows a spherical image transformed from a fish-eye-lens-camera image in Figure 21 (d). Since Eq. (52) is algebraically equivalent to the epipolar constraint on the pinhole-camera models, it is possible to compute the essential matrix E in the same manner with the classical methods [11] for the computation of epipolar geometry based on the pinhole-camera model. For the given eight pairs of point correspondences between the two spherical images, we compute the essential matrix linearly. The great circles on the spherical image in Figure 21 (c) are the epipolar circles, which are the epipolar lines in the pinhole-camera geometry.

Computation of Trifocal Tensor

We compute the trifocal tensor using great-circle correspondences among three spherical images. Figures 22 (a), (b) and (c) shows spherical images transformed from from fish-eye-lens-camera images in Figures 22 (d), (e) and (f), respectively. Since Eq. (63) is algebraically equivalent to the trifocal tensor on the pinhole-camera models, it is possible to compute the trifocal tensor T in Eq. (63) in the same manner with the classical methods [11, 12]. For the computation of the trifocal tensor using great-circle correspondence, we detect great circles on the spherical images using the RHT-based method proposed in Section 4. Then, using the given 13 pairs of great-circle correspondences, we compute the trifocal tensor from the three spherical images linearly. Finally, the two essential matrices are extracted from the computed trifocal tensor [11]. Using the extracted essential matrices, it is possible to draw the epipolar circles. We select some points on the spherical image in Figures 22 (a) and draw the epipolar circles on the spherical images in Figures 22 (b) and (c).

These experimental results lead to the conclusion that it is possible to compute the epipolar geometry and the trifocal tensor based on the spherical cameras. The standardization of different types of central cameras to the spherical cameras enables us to implement the computation of multiple view geometry in a unified method. Furthermore, the use of the large-field-of-view cameras enables us to obtain sufficient numbers of great-circle correspondences for the computation of the trifocal tensor practically.

6 Conclusions

In this paper, we mathematically analyzed central camera systems for computer vision. The classification of central cameras and mathematical definitions of camera models imply to define a spherical camera model as a standard camera model for all central camera systems. Using the spherical camera model as a standard camera model of central camera systems, we developed the basic algorithms for computer vision. The numerical examples of 3D reconstruction promise that the spherical cameras are the essential tool for 3D reconstruction. Furthermore, the proposed algorithm for detecting great circles on a sphere is a fundamental pre-processing for 3D reconstruction. Moreover, it is possible to implement the proposed algorithms for the camera systems which are constructed from different types of central cameras.

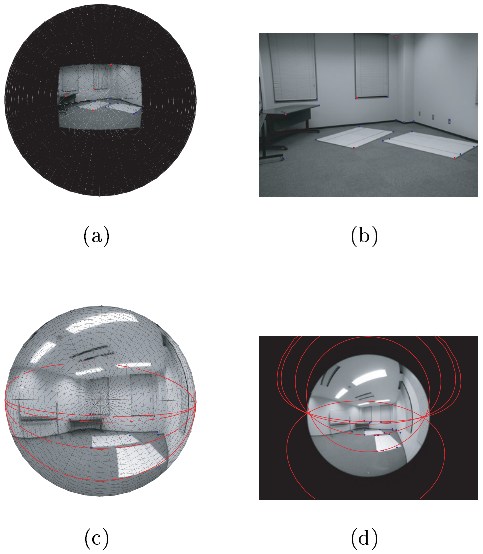

(a) (b)
(c) (d)

Figure 21: Computation of essential matrix. (a) shows a spherical image transformed from a conventional pinhole-camera image (b). (c) shows a spherical image transformed from a fish-eye-lens-camera image (d). The great circles on (c) are the epipolar circles which correspond to the points selected on (a).

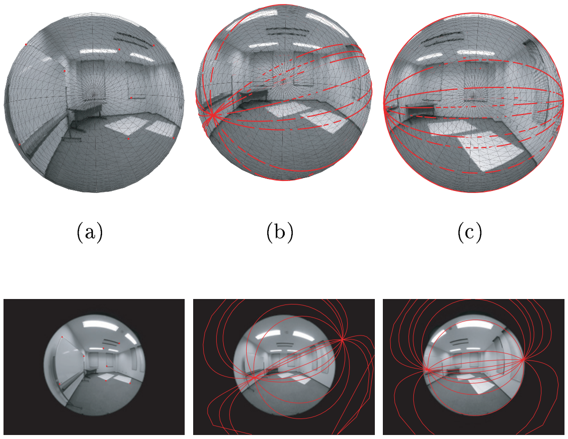
References

[1] S. Baker and S. K. Nayar. A theory of catadioptric image formation. In *ICCV*, pages 35–42, 1998.

[2] D. Ballard and C. M. Brown. Computer vision. *Prentice-Hall; New Jersey*, 1982.

[3] R. Benosman and S.B. Kang. *Panoramic Vision: Sensors, Theory and Applications.* Springer Verlag, New York, 2001.

[4] T. C. Chen and K. L. Chung. A new randomized algorithm for detecting lines. *Real-Time Imaging*, 7(6):473–481, 2001.



(a) (b) (c)

(d) (e) (f)

Figure 22: Computation of trifocal tensor. (a), (b) and (c) shows spherical images transformed from from fish-eye-lens-camera images (d), (e) and (f), respectively. We compute the trifocal tensor using the great-circle correspondence. The essential matrices are then extracted from the trifocal tensor. The great circles on (b) and (c) are the epipolar circles which correspond to the points selected on (a).

[5] O. Faugeras, Q. T. Luong, and T. Papadopoulou. *The Geometry of Multiple Images: The Laws That Govern The Formation of Images of A Scene and Some of Their Applications.* MIT Press, Cambridge, MA, USA, 2001.

[6] O. D. Faugeras. *Three-dimensional Computer Vision: A Geometric Viewpoint.* MIT Press, Cambridge, MA, 1993.

[7] D. A. Forsyth and J. Ponce. *Computer Vision: A Modern Approach.* Prentice Hall, 2002. FOR d 02:1 1.Ex.

[8] J. Gaspar, N. Winters, and J. Santos-Victor. Vision-based navigation and environmental representations with an omnidirectional camera. *In IEEE Transactions on Robotics and Automation*, 16(6):890–898, 2000.

[9] C. Geyer and K. Daniilidis. Para-cata-dioptric calibration. *PAMI*, 24:687–695, 2002.

[10] R. Gupta and R. I. Hartley. Linear pushbroom cameras. *PAMI*, 19(9):963–975, 1997.

[11] R. Hartley and A. Zisserman. *Multiple view geometry in computer vision.* Cambridge University, Cambridge, 2nd edition, 2003.

[12] R. I. Hartley. Lines and points in three views and the trifocal tensor. *IJCV*, 22(2):125–140, 1997.

[13] Richard I. Hartley. Computation of the quadrifocal tensor. In *ECCV*, volume 1, pages 20–35, 1998.

[14] S. Hrabar and G. S. Sukhatme. A comparison of two camera configurations for optic-flow based navigation of a uav through urban canyons. *In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2673–2680, 2004.

[15] K. Kanatani. *Geometric computation for machine vision.* Oxford University Press, Inc., New York, NY, USA, 1993.

[16] R. Klette, K. Schluens, and A. Koschan. *Computer Vision.* Springer-Verlag, 1998.

[17] S. Maybank. *Theory of Reconstruction from Image Motion.* Springer-Verlag, 1993.

[18] B. Mičušík and T. Pajdla. Para-catadioptric camera auto-calibration from epipolar geometry. In Ki-Sang Hong and Zhengyou Zhang, editors, *Proc. of the Asian Conference on Computer Vision (ACCV)*, volume 2, pages 748–753, Seoul, Korea South, January 2004. Asian Federation of Computer Vision Societies.

[19] S.K. Nayar. Catadioptric omnidirectional cameras. In *CVPR97*, pages 482–488, 1997.

[20] T. Pajdla. Epipolar geometry of some non-classical cameras. In B Likar, editor, *Proceedings of Computer Vision Winter Workshop*, pages 223–233, Ljubljana, Slovenia, February 2001. Slovenian Pattern Recognition Society.

[21] S. F. Ray. *Applied photographic optics : Lenses and optical systems for photography, film, video, electronic and digital imaging.* Focal Press, Oxford, 3rd edition, 2002.

[22] P. Sturm and S. Ramalingam. A generic concept for camera calibration. In LNCS 3021-3024, editor, *Proceedings of the European Conference on Computer Vision, Prague, Czech Republic*, volume 2, pages 1–13. Springer, May 2004.

[23] T. Svoboda, T.Pajdla. Epipolar Geometry for Central Catadioptric Cameras. *IJCV*, 49(1):23-37, Kluwer, August, 2002.

[24] N. Winters, J. Gaspar, G. Lacey, and J. Santos-Victor. Omni-directional vision for robot navigation. In *OMNIVIS*, 2000.

[25] L. Xu and E. Oja. Randomized Hough transform (rht): Basic mechanism, algorithm, and computational complexities. *CVGIP: Image Understanding*, 57:131–154, 1993.

[26] L. Xu, E. Oja, and P. Kultanen. A new curve detection method: randomized Hough transform (rht). *Pattern Recognition Letters*, 11(5):331–338, 1990.

[27] K. Yamazawa, Y. Yagi, and M. Yachida. 3d line segment reconstruction by using hyperomni vision and omnidirectional Hough transforming. In *ICPR*, volume 3, pages 487–490, 2000.