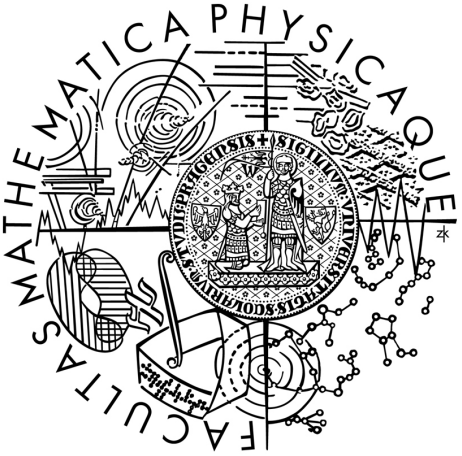
Charles University in Prague
Faculty of Mathematics and Physics

MASTER'S THESIS



Jan Heller

Stereo reconstruction from wide-angle images

Stereo-rekonstrukce z obrazů s vysokým FOV

Department of Software and Computer Science Education

Supervisor: Ing. Tomáš Pajdla, Ph.D.
Study plan: Informatics, Software systems

2007

First of all, I would like to thank my supervisor, Dr. Tomáš Pajdla, for introducing me to the world of omnidirectional vision. Futhermore, I would like to thank the Center for Machine Perception at the Czech Technical University for providing excellent research facilities.

Last but not least, I would like to thank my family and Tereza Marčíková for their support during my work on the Thesis.

I declare that I have written this Master's Thesis on my own and listed all used sources. I agree with lending of the Thesis.

Prague Jan Heller

Contents

Název práce:	Stereo-rekonstrukce z obrazů s vysokým FOV
Autor:	Jan Heller
Katedra (ústav):	Kabinet software a výuky informatiky
Vedoucí diplomové práce:	Ing. Tomáš Pajdla, Ph.D.
e-mail vedoucího:	pajdla@cmp.felk.cvut.cz
Abstrakt:	Tato práce prezentuje rektifikační metody párů obrázků pořízených všesměrovou kamerou. Je představeno několik metod včetně nové rektifikační metody nazvané stereografická rektifikace. V druhé části je prezentována metoda pro 3D rekonstrukci za pomoci páru rektifikovaného některou z uvedených metod. Dále je odvozena metoda sloužící k posuzování kvality esenciálních matic založená na stereografické rektifikaci. V závěru jsou ukázány příklady párů rektifikovaných za pomoci toolboxu pro prostředí MATLAB *OmniRect*, který byl vyvinut jako součást této práce.
Klíčová slova:	všesměrová kamera, rektifikace, 3D rekonstrukce

Title:	Stereo reconstruction from wide-angle images
Author:	Jan Heller
Department:	Department of Software and Computer Science Education
Supervisor:	Ing. Tomáš Pajdla, Ph.D.
Supervisor's e-mail address:	pajdla@cmp.felk.cvut.cz
Abstract:	The thesis presents rectification methods of an omnidirectional stereo pair. Several methods are developed and a novel rectification method called stereographic rectification is introduced. In the second part of the thesis, a method of 3D reconstruction from a stereo pair rectified using an arbitrary method of the adduced methods is presented. Further, a method of judging the accuracy of an essential matrix based on the stereographic rectification is devised. At the end of the thesis, several examples of image pairs rectified by *OmniRect* MATLAB toolbox developed as a part of this work, are shown.
Keywords:	omnidirectional camera, rectification, 3D reconstruction

1 Introduction

Since the introduction of omnidirectional cameras into computer vision community in late 90-ties, the omnidirectional cameras, i.e., cameras with a large field of view, remain subject of extensive study. Omnidirectional vision has proved useful for motion estimation and thus for stereo-reconstruction. The geometry of omnidirectional cameras as well as epipolar geometry of omnidirectional stereo pair is now well understood. Even state-of-the-art epipolar geometry calibration methods for omnidirectional cameras however still fail to produce satisfactory results in certain situation.

A pair of images is thought to be rectified when a parametrized pair of images is produced in the way that epipolar lines coincide. Rectification is typically a pre-step for methods of dense stereo matching and is mostly parametrized so that epipolar lines coincide with image scanlines. This type of rectification simplifies following dense stereo matching and various methods for scanline rectification have been developed. A pair of perspective images is commonly rectified by projecting epipoles into infinity using homographies [4]. Although this method performs well in case when epipoles are not present in original images, it produces infinitely large images in case when epipoles are present. Pollefeys et al. [7] proposed rectification method based on polar parametrization, producing finite area images even for cases when epipoles are located in the images.

However, in case of omnidirectional images epipolar lines become *epipolar curves*, thus homographies cannot be used. Another difference between perspective and omnidirectional epipolar geometries is the existence of a *second epipole* in a single image. As the closest equivalent to the method described in [7], spherical parametrization can be considered. In [1], Arifan and Frossard used spherical parametrization in connection with an energy minimization based approach to estimate dense disparities for omnidirectional images. In [2] Geyer and Daniilidis proved the existence of conformal rectification of omnidirectional stereo pairs, superposing bipolar coordinate system onto an image's two epipoles. Nonetheless, both spherical and bipolar parametrizations inherit a significant setback congenital to all types of scanline rectifications – a severely disproportional expansion of the area near epipole. Since at least one epipole is always present in an omnidirectional image, every scanline rectified omnidirectional image suffers from this blowout. This might not pose a problem in cases when rectified stereo pair is used as a pre-step for epipolar lines marching techniques, since the epipolar lines are parametrized anyway. However, it can be heavily counterproductive when techniques not primarily concerned with epipolar lines are employed.

The goal of our work is to present general methods of rectification of an omnidirectional stereo pair and implementation of these methods in computer. We also show how to use images rectified using presented methods for 3D reconstruction of the original scene. Further, we propose a rectification method based on stereographic projection. Using stereographic projection, scanline rectification cannot be achieved,

epipolar curves are again mapped into curves – circles. In exchange for such a mapping we get a parametrization that in certain sense minimizes distortion of original omnidirectional images as well as spacial distance between corresponding image points. Further, we devise a method based on this rectification to automatically judge quality of a essential matrix.

In Chapter 2 we review the theory of geometric image transformation. In Chapter 3 we summarize the theory of omnidirectional projection and and epipolar geometry of central omnidirectional cameras to a level necessary to our work. In Chapter 4 the stereographic projection is discussed, as it serves as a underlying method for various rectification methods presented in Chapter 5. In Chapter 7 we propose a method for measuring and comparison of quality of essential matrices. Appendix A covers the documentation of the MATLAB rectification toolbox and Appendix B presents examples of image pairs rectified using the toolbox.

1.1 Notation

$a, b, \ldots$	scalars
$A, B, \ldots$	sets
$\mathbf{a}, \mathbf{b}, \ldots, \mathbf{A}, \mathbf{B}, \ldots$	vectors
$\mathtt{A}, \mathtt{B}, \ldots$	matrices
$\mathbb{R}^n, \mathbb{P}^n, \ldots$	n-dimensional spaces
$\mathcal{A}, \mathcal{B}, \ldots$	transformations

1.2 Basic definitions

1.2.1 Vectors and matrices

All vectors in this work are *column* vectors and are considered 3×1 matrices when multiplied by a matrix. $\begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \ldots & \mathbf{a}_m \end{pmatrix}$ denotes a matrix $\mathtt{A} \in \mathbb{R}^{n \times m}$, which columns are vectors $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \in \mathbb{R}^n$.

Definition 1 *Let* $\mathbf{a} = (a_1, a_2, a_3)^\top \in \mathbb{R}^3 \backslash \{(0,0,0)^\top\}$*, then* $[\mathbf{a}]_\times$ *is a skew-symmetric matrix*

$$[\mathbf{a}]_\times \stackrel{\text{def}}{=} \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & -a_1 & 0 \end{pmatrix}.$$

Definition 2 *Let* $\mathbf{a} = (a_1, a_2, \ldots, a_n)^\top \in \mathbb{R}^n$*. Then*

$$\lfloor \mathbf{a} \rfloor \stackrel{\text{def}}{=} (\lfloor a_1 \rfloor, \lfloor a_2 \rfloor, \ldots, \lfloor a_n \rfloor),$$

where $\lfloor \cdot \rfloor$ *is the floor function.*

Definition 3 *Let* $\mathbf{a} \in \mathbb{R}^n$*,* $R \subset \mathbb{R}^n$*. Then*

$$\Delta(R, \mathbf{a}) \stackrel{\text{def}}{=} \inf_{\mathbf{b} \in R} \|\mathbf{a} - \mathbf{b}\|$$

is the distance between the vector $\mathbf{a}$ and the set R.

Definition 4 *Let* $\mathtt{A}, \mathtt{I} \in \mathbb{R}^{n \times n}$, $\mathtt{I}$ *the unit matrix. Then if*

$$\mathtt{A}^\top \mathtt{A} = \mathtt{A}\mathtt{A}^\top = \mathtt{I},$$

$\mathtt{A}$ *is called the* orthogonal matrix.

1.2.2 Unit sphere

Definition 5

$$\mathbb{S}^3 \stackrel{\text{def}}{=} \{\mathbf{v} \in \mathbb{R}^3 : ||\mathbf{v}|| = 1\}$$

is the unit sphere.

Definition 6 *Let* $\mathbf{a}, \mathbf{b} \in \mathbb{S}^3$. *Then*

$$\delta_{\mathbb{S}^3}(\mathbf{a}, \mathbf{b}) \stackrel{\text{def}}{=} \arccos(\mathbf{a}^\top \mathbf{b})$$

is the spherical distance between the vectors $\mathbf{a}$ and $\mathbf{b}$.

Definition 7 *Let* $\mathbf{a} \in \mathbb{S}^3$, $S \subset \mathbb{S}^3$. *Then*

$$\Delta_{\mathbb{S}^3}(S, \mathbf{a}) \stackrel{\text{def}}{=} \inf_{\mathbf{b} \in S} \delta_{\mathbb{S}^3}(\mathbf{a}, \mathbf{b})$$

is the spherical distance between the vector $\mathbf{a}$ and the set S.

2 Geometric image transformations

Since a stereo rectification is basically a geometric transformation of a pair of digital images, a rigorous definition of a digital image and geometric image transformations is needed. In this chapter we present theory of geometric transformations and image transformations, based on [3] and [14].

A digital image is a representation of a two-dimensional real world image – projection on an eye's retina, a photograph, a picture, or to put it more mathematically, continuous two-dimensional signal – as a finite set of values called *pixels*. In a computer, this set is typically represented by a continuous block of memory, with memory cells interpreted as various value types, such as `char` or `double`, depending on number of values a pixel can assume. One-dimensional pixel values are sufficient to represent a grayscale image – in order to represent a color image, higher dimensional values are needed. A 3D *RGB color model* is the most commonly used, however, in our definition, we allow for higher dimensions as well.

2.1 Discrete image

Definition 8 *Let* $\Phi = [0, a_1] \times [0, a_2] \subset \mathbb{R}^2$, $n \in \mathbb{N}$. *Then a mapping*

$$\mathcal{P}^c : \Phi \to \mathbb{R}^n$$

is called a continuous image, *where* $a_1 \in \mathbb{R}$ *is the width of the image and* $a_2 \in \mathbb{R}$ *is the height of the image.*

Such a continuous image, however mathematically appealing, is far from a data set acquired by a digital camera. Both range and domain of a digital image are finite. However, Definition 9 allows for an infinite domain. This is only a mathematical simplification. In the "real world", the domain of a digital image is always clamped to some conveniently representable interval, typically $[0, 255] \subset \mathbb{N}$.

Definition 9 *Let* $\Psi = [0, a_1] \times [0, a_2] \subset \mathbb{N}^2$, $n \in \mathbb{N}$. *Then a mapping*

$$\mathcal{P} : \Psi \to \mathbb{Z}^n$$

is called a discrete image, *where* $a_1 \in \mathbb{N}$ *is the width of the image and* $a_2 \in \mathbb{N}$ *is the height of the image.*

In the following, only discrete images are considered and terms *discrete image* and *digital image* denote the same type of mapping.

2.2 Geometric image transformations

Firstly, we have to distinguish two terms: *geometric transformation* and *image transformation based on a geometric transformation*. Geometric transformation, such as rotation and translation, is an intuitive term. Since we deal with two-dimensional images, we will only consider two-dimensional geometric transformations, i.e.,

$$\mathcal{G} : \mathbb{R}^2 \to \mathbb{R}^2.$$

Because we are equally interested in describing image rectification as a geometric transformation as in producing adequate images based on such geometric transformation, we introduce the term *image transformation based on a geometric transformation*. By the term we understand the process of creating new discrete image, i.e., defining a new map, by applying given geometric transformation to a given discrete image. In literature, these term often coincide or are distinguished implicitly. For convenience, we abbreviate the term *image transformation based on a geometric transformation* as *geometric image transformation*.

A geometric transformation of a digital image is not, in contrary to the underlying geometric transformation, an uniquely defined mathematical transformation, but rather a set of methods, which differ in computational complexity as well as in their results. Let's formulate a geometric image transformation based on image interpolation and inverse geometric transformation.

Definition 10 *Let* $\mathcal{P} : \Psi = [0, a_1] \times [0, a_2] \subset \mathbb{N}^2 \to \mathbb{Z}^n$ *be a discrete image,* $\Phi = [0, a_1] \times [0, a_2] \subset \mathbb{R}^2$*,* φ *an interpolating synthesis function. Then a mapping* $\mathcal{I}_{\mathcal{I}}^{\varphi} : \Phi \to \mathbb{Z}^n$ *such that*

$$\forall \mathbf{x} \in \Phi : \mathcal{I}_{\mathcal{P}}^{\varphi}(\mathbf{x}) \quad = \quad \left\lfloor \sum_{\mathbf{k} \in \Psi} \mathcal{P}(\mathbf{k}) \varphi(\mathbf{x} - \mathbf{k}) \right\rfloor,$$

is called interpolation of the image $\mathcal{P}$ using the synthesis function φ.

The floor function $\lfloor \cdot \rfloor$ guaranties that the domain of the interpolated image is again $\mathbb{Z}^n$. In a computer program, the floor function can be replaced by the round function or by simple retyping. The synthesis function $\varphi : \mathbb{R}^n \to \mathbb{R}^n$ is the only thing deciding the quality and the properties of the interpolation. The desired properties of a synthesis function are

- interpolating property

$$\forall \mathbf{k} \in \mathbb{Z}^n : \varphi(\mathbf{k}) = \begin{cases} 1 & \textit{for } \mathbf{k} = (0, \ldots, 0)^\top, \\ 0 & \textit{otherwise}, \end{cases}$$

- separability

$$\forall \mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n : \varphi(\mathbf{x}) = \prod_{i=1}^{n} \varphi_i(x_i),$$

- symmetry
$$\forall \mathbf{x} \in \mathbb{R}^n : \varphi(-\mathbf{x}) = -\varphi(\mathbf{x}),$$
- and partition of unity
$$\forall \mathbf{x} \in \mathbb{R}^n : 1 = \sum_{\mathbf{k} \in \mathbb{Z}^n} \varphi(\mathbf{x} - \mathbf{k}).$$

The most common interpolation techniques use in 2D computer graphics are the *nearest neighbor interpolation* and the *bilinear interpolation*. Let's define synthesis function that, together with Definition 10, lead to such interpolations.

Definition 11 *Let* $\varphi : \mathbb{R} \to \mathbb{R}$ *such that*
$$\forall x \in \mathbb{R} : \varphi(x) = \begin{cases} 1 & \text{for } |x| < \frac{1}{2}, \\ \frac{1}{2} & \text{for } |x| = \frac{1}{2}, \\ 0 & \text{else}, \end{cases}$$
then $\varphi_{NN} : \mathbb{R}^2 \to \mathbb{R}^2$ *such that*
$$\forall \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : \varphi_{NN}(\mathbf{x}) = \varphi(x_1)\,\varphi(x_2), \tag{2.1}$$
is called the nearest neighbor interpolating synthesis function.

Definition 12 *Let* $\varphi : \mathbb{R} \to \mathbb{R}$ *such that*
$$\forall x \in \mathbb{R} : \varphi(x) = \begin{cases} 1 - |x| & \text{for } |x| < 1, \\ 0 & \text{else}, \end{cases}$$
then $\varphi_{BL} : \mathbb{R}^2 \to \mathbb{R}^2$ *such that*
$$\forall \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : \varphi_{BL}(\mathbf{x}) = \varphi(x_1)\,\varphi(x_2), \tag{2.2}$$
is called the bilinear interpolating synthesis function.

The function φ_{BL} can be conveniently rewritten as
$$\varphi_{BL}(\mathbf{x}) = \max(1 - |x_1|, 0)\max(1 - |x_2|, 0).$$

Consequence 1 *Let* $\mathcal{P} : \Psi = [0, a_1] \times [0, a_2] \subset \mathbb{N}^2 \to \mathbb{Z}^n$ *be a discrete image,* $\mathcal{G} : \mathbb{R}^2 \to \mathbb{R}^2$ *an invertible geometric transformation,* $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ *an interpolating synthesis function. Then a discrete image* $\mathcal{P}_\mathcal{G} : \Psi' = [0, a_1'] \times [0, a_2'] \subset \mathbb{N}^2 \to \mathbb{Z}^n$*, such that*
$$\forall \mathbf{x} \in \Psi' : \mathcal{P}_\mathcal{G}(\mathbf{x}) = \mathcal{I}_\mathcal{P}^\varphi\left(\mathcal{G}^{-1}(\mathbf{x})\right),$$
is an image transformation of the image $\mathcal{P}$ *based on the geometric transformation* $\mathcal{G}$*.*

3 Geometry of Central Omnidirectional Cameras

Central omnidirectional camera is any panoramic camera having a single effective viewpoint. In this chapter the spherical model of central omnidirectional cameras is review, as well as the epipolar geometry of two omnidirectional cameras. Finally, a pair of transformations called epipolar alignment is derived.

3.1 Omnidirectional projection

Standard central perspective camera model is based on projective geometry and states that

$$\exists \alpha \neq 0 : \ \alpha \mathbf{x} = \mathtt{P}\mathbf{X}, \tag{3.1}$$

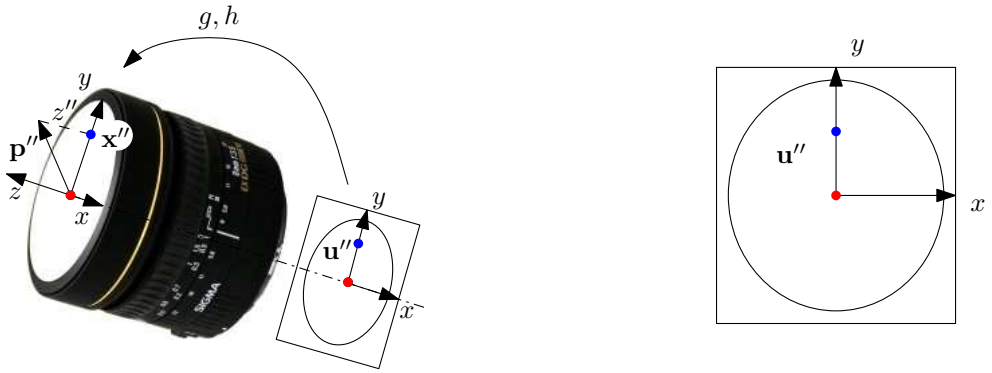
where $\mathbf{X} \in \mathbb{R}^4 \backslash \{(0,0,0,0)^\top\}$ is a scene point, $\mathtt{P} \in \mathbb{R}^{3\times 4}$ is a camera projection matrix and $\mathbf{x} \in \mathbb{R}^3 \backslash \{(0,0,0)^\top\}$ represents an image point [4]. In this model all scene points lying on the same line passing through the optical center of the camera – in front as well as behind the camera – are represented by one image point, see 3.1(a). This representation may be sufficient for directional cameras with field of view smaller than 180°, however, is unsuitable for modeling omnidirectional cameras, where points behind the camera and points in front of the camera are projected onto different image points. This issue is addressed by the *spherical model*, where lines are split into half-lines, see Figure 3.1(b). In this model, represented by unit vectors in $\mathbb{R}^3$, one vector represents one half-line, so that one image point represents all scene points lying on a half of a line passing through the center of the camera and another image point represents all scene points lying on the opposite half of the same line. This fact formulates as

$$\exists \alpha > 0 : \ \alpha \mathbf{x} = \mathtt{P}\mathbf{X}, \tag{3.2}$$

where $\mathbf{X}$ and $\mathtt{P}$ are the same as in equation 3.1 and $\mathbf{x} \in \mathbb{R}^3 \backslash \{0,0,0\}$ is a vector representing an image point.

(a) (b)

Figure 3.1: (a) Directional camera. Scene points are represented by straight lines. (b) Omnidirectional camera. Scene points are represented by half-lines. (Adopted from [6])



(a) (b)

Figure 3.2: Omnidirectional image formation: (a) Projection of a scene point to a sensor plane. (b) The sensor plane with field of view circle. (Adopted from [6])

3.2 Image formation and camera calibration

By image formation we will understand the formation of a digital image from a surrounding scene through an optical system and the process of digitization. Let us briefly summarize the mathematical formalism of image formation of central omnidirectional cameras, as described in [6], to a level significant to our work. In the next, it is assumed that the lenses and mirrors are

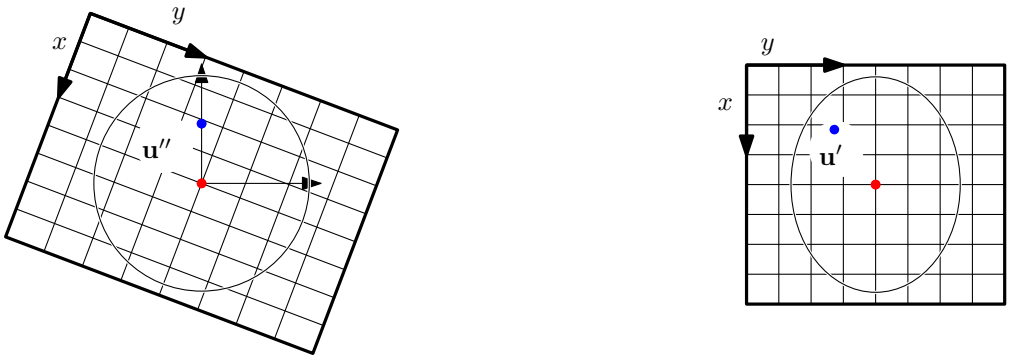
i, symmetric w.r.t. an axis and

ii, the axis of the lens, or the mirror, is perpendicular to a sensor plane.

Figure 3.3 shows the process of image formation. Using spherical model described in Section 3.1, the projection of a scene point $\mathbf{X}$ is represented by unit vector $\mathbf{q}'' \in \mathbb{S}^3$. From assumptions *i*, and *ii*, one infers that there always exists a vector $\mathbf{p}''=\left(\mathbf{x}''^{\top}, z''\right)^{\top} \in \mathbb{R}^3$ and a vector $\mathbf{u}'' \in \mathbb{R}^2$ in the sensor plane, for which following holds:

$$\begin{aligned} \exists \alpha \in \mathbb{R}^+ : \mathbf{p}'' &= \alpha \mathbf{q}'', \\ \exists \beta \in \mathbb{R}^+ : \mathbf{x}'' &= \beta \mathbf{u}'', \qquad (3.3) \\ \mathbf{p}'' &= \begin{pmatrix} h\left(||\mathbf{u}''||, \mathbf{a}''\right) \mathbf{u}'' \\ g\left(||\mathbf{u}''||, \mathbf{a}''\right) \end{pmatrix}. \qquad (3.4) \end{aligned}$$

Functions $h, g : \mathbb{R}^N \times \mathbb{R} \rightarrow \mathbb{R}$ are rotationally symmetric and depend on $||\mathbf{u}''||$, that is on the distance between the optical axis and $\mathbf{u}''$, and on a vector of parameters $\mathbf{a}'' \in \mathbb{R}^N$, where N is the number of parameters. The functions capture the type of a omnidirectional camera. Function g typically depend on the shape of the mirror for catadioptric omnidirectional cameras, function h captures the projection of the camera. Note that β from Equation (3.3), explicitly stating the collinearity between $\mathbf{u}''$ and $\mathbf{x}''$, equals $h\left(||\mathbf{u}''||, \mathbf{a}''\right)$ from Equation (3.4). Figures 3.2(a, b) show general relation between an image point $\mathbf{u}''$ and corresponding vector $\mathbf{p}''$, Figures 3.4(a, b) show the relation in case of a fish-eye lens.



(a) (b)

Figure 3.3: Omnidirectional image formation: (a, b) Digitization process – affine transformation of the field of view circle. (Adopted from [6])

The next step in the image formation process is digitization. Process of transforming sensor plane point $\mathbf{u}''$ into digital image point $\mathbf{u}'$ can be modeled by an affine transformation

$$\mathbf{u}'' = \mathtt{A}'\mathbf{u}' + \mathbf{t}', \tag{3.5}$$

where $\mathbf{A}' \in \mathbb{R}^{2\times 2}$ is a regular matrix, $\mathbf{t}' \in \mathbb{R}^2$ is a translation vector and $\mathbf{u}'$ is a point in a digital image. The digitization process is depicted in Figure 3.3(a, b). By plugging the image formation process into Equation (3.2) the complete projection equation for omnidirectional cameras can be written as

$$\exists \alpha > 0 : \ \alpha \mathbf{p}'' = \alpha \begin{pmatrix} h\left(||\mathbf{u}''||, \mathbf{a}''\right)\mathbf{u}'' \\ g\left(||\mathbf{u}''||, \mathbf{a}''\right) \end{pmatrix} = \alpha \begin{pmatrix} h\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right)(\mathtt{A}'\mathbf{u}' + \mathbf{t}') \\ g\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right) \end{pmatrix} = \mathtt{P}\mathbf{X}. \tag{3.6}$$

The objective of the camera calibration is to find mapping from a digital image point $\mathbf{u}'$ to a corresponding 3D ray represented by a vector $\mathbf{q}''$ for a given camera. This means, that affine transformation from a digital image to the camera's sensor plane, as well as the mapping from camera's sensor plane to scene rays have to be recovered. The process of calibration of various central omnidirectional cameras is beyond the scope of our work and can be found in great detail in [6].

Definition 13 *Let $g, h, \mathtt{A}, \mathbf{u}'', \mathbf{t}', \mathbf{a}$ have the same meaning as in Equation (3.6) for an omnidirectional camera C. Then a map $\mathcal{C}_C : \mathbb{R}^2 \to \mathbb{S}^3$ so that*

$$\forall \mathbf{u}' \in \mathbb{R}^2 : \ \mathcal{C}_C(\mathbf{u}') = \frac{\begin{pmatrix} h\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right)(\mathtt{A}'\mathbf{u}' + \mathbf{t}') \\ g\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right) \end{pmatrix}}{\left\| \begin{pmatrix} h\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right)(\mathtt{A}'\mathbf{u}' + \mathbf{t}') \\ g\left(||\mathtt{A}'\mathbf{u}' + \mathbf{t}'||, \mathbf{a}''\right) \end{pmatrix} \right\|}$$

realizing the mapping from a digital image point $\mathbf{u}'$ to a corresponding 3D ray for the omnidirectional camera C is called the calibration transformation of a camera C.

(a) (b)

Figure 3.4: Mapping of a scene point $\mathbf{X}$ into a sensor plane point $\mathbf{u}''$ for a fish-eye lens. Since orthographic camera projection is assumed, $h = 1$. (Adopted from [6])

Since a calibration function of a camera is always a compromise between correctness and computability, simple definitions of g, h are preferred, with invertibility in mind. In our work the existence of an inverse mapping $\mathcal{C}_C^{-1}$ for a camera C is always assumed. Further, orientation such that $\mathcal{C}_C((0,0))^\top = (0,0,-1)^\top$ is assumed.

3.3 Epipolar geometry

The epipolar geometry is motivated by stereo matching, i.e., by searching for the projections of a scene point $\mathbf{X}$ in two different views of the same rigid scene. Let's suppose that $\mathbf{X} \in \mathbb{R}^3$ projects onto $\mathbf{u}_1 \in \mathbb{P}^2$ in the first view and onto $\mathbf{u}_2 \in \mathbb{P}^2$ in the second, see Figure 3.5. From the fundamental properties of central projection follows that the centers of the cameras $\mathbf{C}_1$, $\mathbf{C}_2$ and the points $\mathbf{u}_1$,$\mathbf{u}_2$ and $\mathbf{X}$ are coplanar. Scene points together with baseline $\overline{\mathbf{C}_1\mathbf{C}_2}$ create a pencil of planes called *epipolar planes*. Every of such planes intersects the projective planes of the two views in straight line – *epipolar line*. Epipolar lines again form a pencils of lines in their respective projective planes, that intersect in two respective points, called *epipoles*. Epipoles can be equivalently described as projections of camera centers into the image planes of the opposite view. An example of epipolar geometry of two perspective cameras is given in Figure 3.5(a).

The fact that if a scene point $\mathbf{X}$ is projected onto an point $\mathbf{u}_1$in the first view, then the image of the point $\mathbf{X}$ in the second image $\mathbf{u}_2$ must lie on a epipolar line $\mathbf{l}' \in \mathbb{P}^2$, which is the projection of the epipolar plane corresponding to the point $\mathbf{u}_1$, is called the *epipolar constraint*. The epipolar constraint can be algebraically written as

$$\begin{pmatrix} \mathbf{u}_2^\top & 1 \end{pmatrix} \mathtt{F} \begin{pmatrix} \mathbf{u}_1 \\ 1 \end{pmatrix} = 0,$$

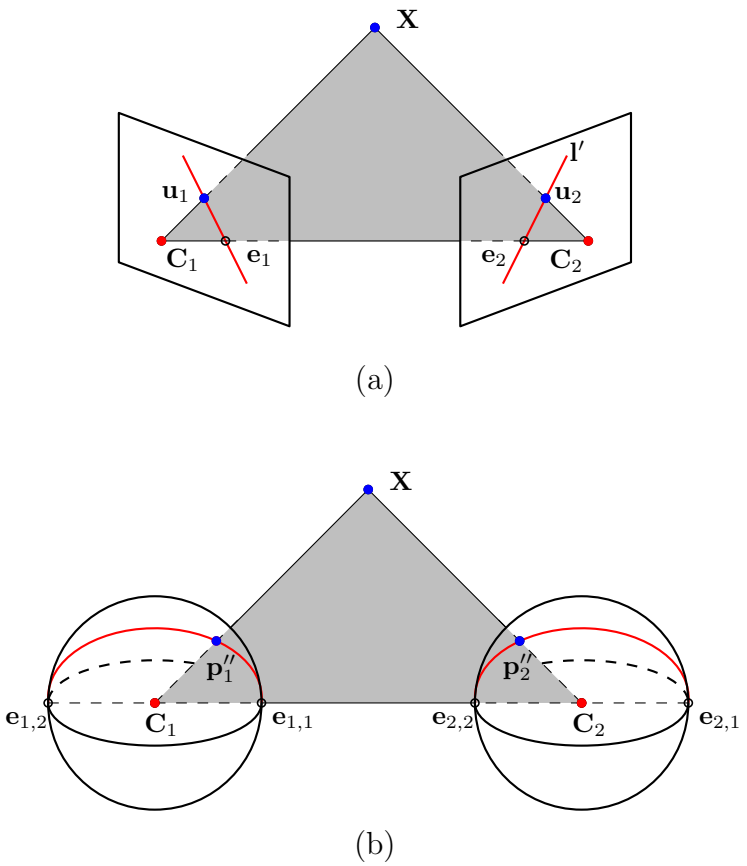


Figure 3.5: (a) Epipolar geometry of standard perspective cameras. (b) Epipolar geometry of omnidirectional central cameras. (Adopted from [6])

where $\mathtt{F} \in \mathbb{R}^{3\times3}$ is a *fundamental matrix*. The fundamental matrix realizes the mapping $\mathbf{u}_1 \mapsto \mathbf{l}'$, i.e.,

$$\mathbf{l}' = \mathtt{F}\mathbf{u}_1,$$

from which follows that $\mathrm{rank}(\mathtt{F}) = 2$.

An analogy to the epipolar geometry of central perspective cameras for central omnidirectional cameras can be formulated likewise. The difference between directional and omnidirectional cameras is the shape of the retinas as well as the distinguishability of the rays orientations. The pencil of planes intersect the spherical retinas of the spherical model in great circles, which are projected into sensor plane as *epipolar curves*, intersecting the $\overline{\mathbf{C}_1\mathbf{C}_2}$ baseline in *two epipoles*, $\mathbf{e}_{1,1}, \mathbf{e}_{1,2}$ in the first view, $\mathbf{e}_{2,1}, \mathbf{e}_{2,2}$ in the second view, see Figure 3.5(b). The epipolar curves are conics for quadric catadioptric cameras [10], more general curves for fish-eye lenses [6]. Since vectors $\mathbf{p}''_1$ and $\mathbf{p}''_2$, as depicted in Figure 3.5(b), create an epipolar plane, an epipolar geometry can be formulated for them. The epipolar constraint for a pair of omnidirectional images reads as

$$\mathbf{p}''^{\top}_2 \mathtt{F}'' \mathbf{p}''_1 = 0, \tag{3.7}$$

where $\mathtt{F}'' \in \mathbb{R}^{3\times3}$ is an analogy to the fundamental matrix called *essential matrix*, mapping one-dimensional subspace, $\mathbf{p}''_1$, to a two-dimensional subspace in $\mathbb{R}^3$, the epipolar plane containing $\mathbf{p}''_2$, from which again follows that $\mathrm{rank}(\mathtt{F}'') = 2$. Figures 3.6(a, b,

c, d) show example of two image pairs with denoted epipolar geometries. The main difference between epipolar geometries of perspective directional cameras and omnidirectional cameras lies in the fact that whereas the epipolar constraint of directional cameras applies directly to image points, the epipolar constrain of omnidirectional cameras applies to 3D vectors acquired by camera calibration using functions g, h.

Since for any vector $\mathbf{p}''$ other than $\mathbf{e}_{1,1}, \mathbf{e}_{1,2}$ the epipolar plane specified by the normal vector $\mathbf{n} = \mathtt{F}''\mathbf{p}''$ contains epipoles $\mathbf{e}_{2,1}, \mathbf{e}_{2,2}$, equations

$$\forall \mathbf{p}'' \in \mathbb{R}^3 : \ \mathbf{e}_{2,1}^\top(\mathtt{F}''\mathbf{p}'') = 0 \ \& \ \mathbf{e}_{2,2}^\top(\mathtt{F}''\mathbf{p}'') = 0$$

hold true. It follows, that $\mathbf{e}_{2,1}^\top\mathtt{F}'' = \mathbf{e}_{2,2}^\top\mathtt{F}'' = 0$, i.e., $\mathbf{e}_{2,1} = -\mathbf{e}_{2,2}$ is the left null-space of $\mathtt{F}''$. Analogously $\mathtt{F}''\mathbf{e}_{1,1} = \mathtt{F}''\mathbf{e}_{1,2} = 0$, , i.e., $\mathbf{e}_{1,1} = -\mathbf{e}_{1,2}$ is the right null-space of $\mathtt{F}''$. Given an essential matrix $\mathtt{F}''$, epipoles are standardly computed using singular value decomposition (SVD) of $\mathtt{F}''$, see Result 1.

Result 1 *Let* $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_2)\,\mathrm{diag}(1,1,0)\,(\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)^\top$ *be SVD of an essential matrix* $\mathtt{F}'' \in \mathbb{R}^{3\times3}, \mathrm{rank}(\mathtt{F}'') = 2$. *Then* $\mathtt{F}''\mathbf{e}_1 = 0, \mathbf{e}_2^\top\mathtt{F}'' = 0$, *and* $\mathbf{e}_1, -\mathbf{e}_1, \mathbf{e}_2, -\mathbf{e}_2$ *are the respective epipoles of the epipolar geometry specified by* $\mathtt{F}''$.

Following result explicitly states the ambiguity of epipoles computed from $\mathtt{F}''$.

Result 2 *Let* $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_2)\,\mathrm{diag}(1,1,0)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)^\top$ *be SVD of an essential matrix* $\mathtt{F}'' \in \mathbb{R}^{3\times3}, \mathrm{rank}(\mathtt{F}'') = 2$. *Then*

$$\begin{array}{rll} (\mathbf{u}_1, \mathbf{u}_2, -\mathbf{e}_2) & \mathrm{diag}(1,1,0) & (\mathbf{v}_1, \mathbf{v}_1, -\mathbf{e}_1)^\top, \\ (\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_2) & \mathrm{diag}(1,1,0) & (\mathbf{v}_1, \mathbf{v}_2, -\mathbf{e}_1)^\top, \\ (\mathbf{u}_1, \mathbf{u}_2, -\mathbf{e}_2) & \mathrm{diag}(1,1,0) & (\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)^\top, \end{array}$$

are also SVD of $\mathtt{F}''$.

In the following, epipole orientation such that $\mathbf{e}_{1,1}, \mathbf{e}_{2,1}$ are always directions to the same scene point $\mathbf{X}$ and such that if only one epipole is visible in the first view, $\mathbf{e}_{1,1}$ is that epipole is assumed.

Theorem 1 *Let* $\mathtt{F}$ *be an essential matrix and* $\mathbf{e}_{1,i}$, $\mathbf{e}_{2,i}$, $i = 1, 2$ *the respective epipoles,* $\mathbf{u} \in \mathbb{R}^3\backslash\{(0,0,0)^\top$ *such that* $\mathbf{u}$ *and* $\mathbf{e}_{1,2}$ *are lineary independent. Then vectors* $\mathbf{u}$ *and* $[\mathbf{e}_{2,2}]_\times\,\mathtt{F}\mathbf{u}$ *lie in the same epipolar plane.*

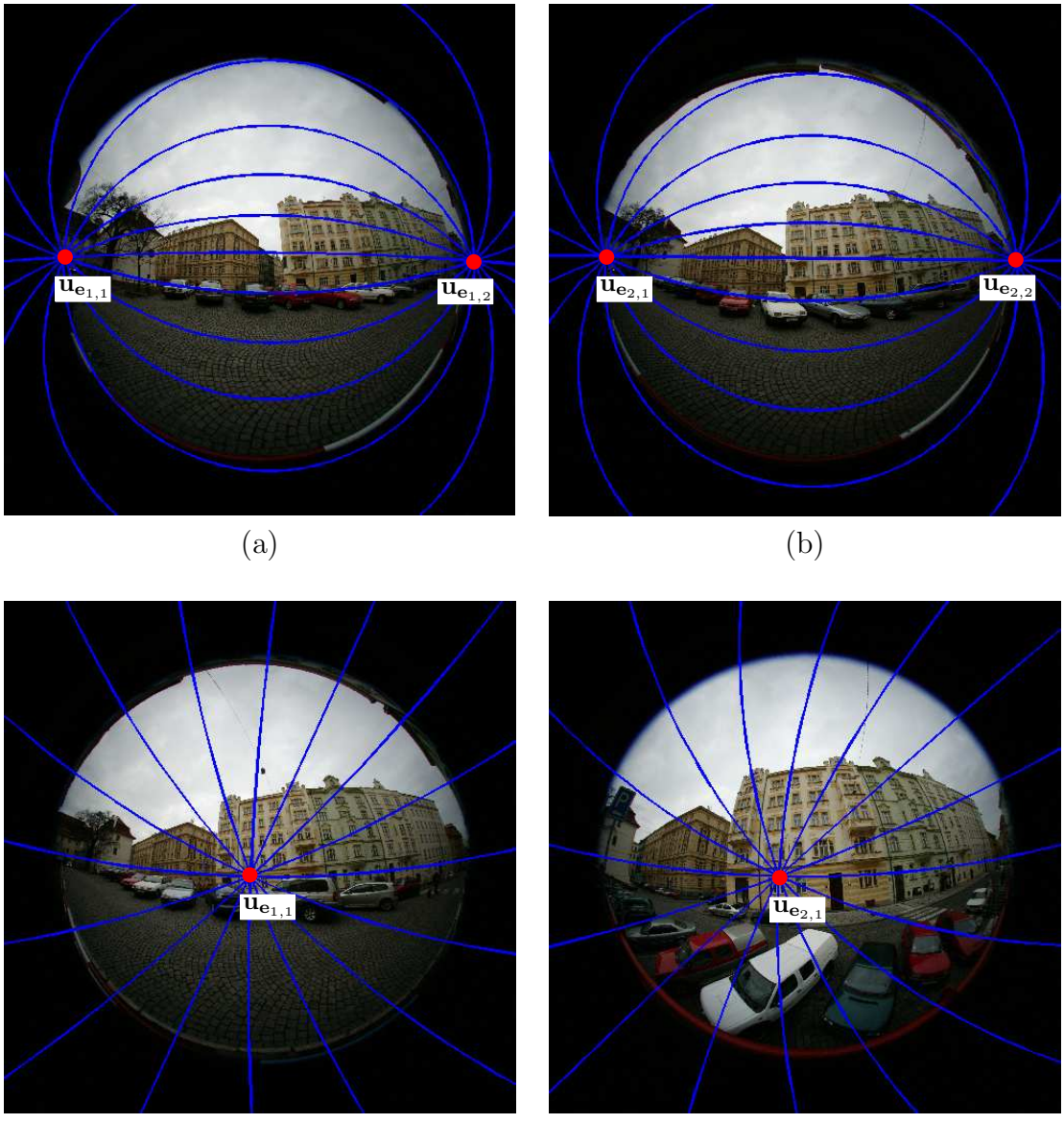
Proof. Seeing that $\mathbf{n}' = \mathtt{F}\mathbf{u}$ is the normal to an epipolar plane $\mathbf{P}' = \{\mathbf{v} \in \mathbb{R}^3 : \mathbf{v}^\top\mathbf{n}' = 0\}$ in which $\mathbf{u}$ is lying, it holds that

$$\forall \mathbf{u}' \in \mathbf{P}' : \ \mathbf{u}' \times \mathbf{n}' = [\mathbf{u}]_\times\,\mathbf{n}' \in \mathbf{P}'.$$

Since $\mathbf{e}_{2,2}$ lies in every epipolar plane, $\mathbf{e}_{2,2}$ lies in $\mathbf{P}'$ as well. Thus

$$\mathbf{e}_{2,2} \times \mathbf{n}' = [\mathbf{e}_{2,2}]_\times\,\mathbf{n}' = [\mathbf{e}_{2,2}]_\times\,\mathtt{F}\mathbf{u} \in \mathbf{P}'.$$

□



(a) (b)

(c) (d)

Figure 3.6: Two image pairs acquired by a fish-eye lens with field of view of 180° with respective epipolar geometries. Images were transformed so it would appear as like they had been acquired by para-catadioptric camera in order to transform epipolar curves into circles. (a) An image pair resulting from by lateral move of the camera. Both epipoles are visible. (b) An image pair resulting from forward move of the camera. Only one epipole is visible.

3.4 Epipolar alignment

Given two calibrated images of the same rigid scene and an essential matrix describing the epipolar geometry of the image pair, the goal of this section is to derive transformation $\mathcal{A}_1$ from the coordinate system of the first camera C_1 and transformation $\mathcal{A}_2$ from the coordinate system of the second camera C_2 to the world coordinate system, so the respective epipole pairs $\mathbf{e}_{1,i}$, $\mathbf{e}_{2,i}$, $i = 1, 2$ coincide with the z axis and the corresponding epipolar circles are superimposed, see Figure 3.7. The pair $[\mathcal{A}_1, \mathcal{A}_2]$ will be called the *epipolar alignment* of an image pair. Is is a simple observation, that transformations $\mathcal{A}_1, \mathcal{A}_2 : \mathbb{R}^3 \to \mathbb{R}^3$ are linear automorphisms and as such algebraically expressed as matrix multiplications

$$\forall \mathbf{q} \in \mathbb{R}^3 : \mathcal{A}_1(\mathbf{q}) = \mathtt{A}_1\mathbf{q},\ \mathcal{A}_2(\mathbf{q}) = \mathtt{A}_2\mathbf{q},$$

where $\mathtt{A}_1, \mathtt{A}_2 \in \mathbb{R}^{3\times3}$. Since $\mathcal{A}_1, \mathcal{A}_2$ are automorphisms, there always exist inverse linear transformations $\mathcal{A}_1^{-1}, \mathcal{A}_2^{-1}$, such that

$$\forall \mathbf{q} \in \mathbb{R}^3 : \mathcal{A}_1^{-1}(\mathbf{q}) = \mathtt{A}_1^{-1}\mathbf{q},\ \mathcal{A}_2^{-1}(\mathbf{q}) = \mathtt{A}_2^{-1}\mathbf{q},$$

mapping z axis of the world coordinate system onto the respective epipoles.

Definition 14 *Let* $\mathbf{e}$ *be an epipole in an image from an omnidirectional stereo pair. Then a coordinate system* $\Sigma_{\mathbf{e}}^{\mathbf{u}} = [\mathbf{x}, \mathbf{y}, \mathbf{e}]$*, so that*

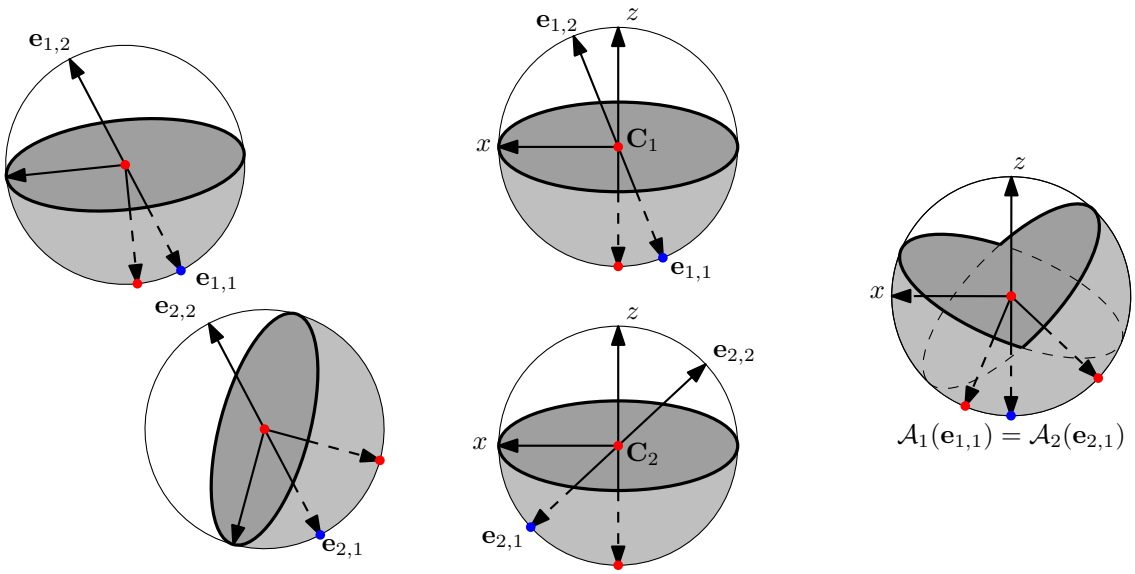
$$\begin{aligned} \mathbf{u} \in \mathbb{R}^3 \backslash \{(0,0,0)^\top\} \quad &: \quad \neg\left(\exists \alpha \in \mathbb{R} : \alpha\mathbf{u} = \mathbf{e}\right), \\ \mathbf{x} \quad &= \quad \frac{[\mathbf{e}]_\times \mathbf{u}}{\left\|[\mathbf{e}]_\times \mathbf{u}\right\|}, \\ \mathbf{y} \quad &= \quad \frac{[\mathbf{x}]_\times \mathbf{e}}{\left\|[\mathbf{x}]_\times \mathbf{e}\right\|}, \end{aligned}$$

is called the epipolar coordinate system incident to the epipole $\mathbf{e}$ with up-vector $\mathbf{u}$.

Let $\mathtt{F}$ be an essential matrix and $\mathbf{e}_{1,i}$, $\mathbf{e}_{2,i}$, $i = 1, 2$ the respective epipoles, as described in section 3.3, $\Omega = \left[(1,0,0)^\top, (0,1,0)^\top, (0,0,1)^\top\right]$ the world coordinate system. Transformation from the ordered basis $\Sigma_{\mathbf{e}_{1,2}}^{\mathbf{u}_1}$ to the ordered basis Ω and transformation from the ordered basis $\Sigma_{\mathbf{e}_{2,2}}^{\mathbf{u}_2}$ to the ordered basis Ω, for $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{R}^3$, where $\mathbf{u_1}$ is not collinear with $\mathbf{e}_{1,2}$ and $\mathbf{u_2}$ is not collinear with $\mathbf{e}_{2,2}$, would solve the goal of superimposing epipoles with z axis, however, in order to superimpose epipolar circles as well, another constraint to these mappings must be introduced. In order ensure superposition of epipolar circles, up-vectors $\mathbf{u}_1, \mathbf{u}_2$ have to "select" the same epipolar circle, i.e., lie in the same epipolar plane. From Theorem 1 follows that $\mathbf{u_2} = [\mathbf{e}_{2,2}]_\times \mathtt{F}\mathbf{u}_1$ is a sufficient condition for $\mathbf{u_1}, \mathbf{u}_2$ to lie in the same epipolar plane.

Let us derive $\mathtt{A}_1^{-1}$ realizing transformation from ordered basis Ω to the ordered basis $\Sigma_{\mathbf{e}_{1,2}}^{\mathbf{u}_1} = [\mathbf{x}, \mathbf{y}, \mathbf{e}_{1,2}]$. After specifying that

$$\mathtt{A}_1^{-1} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix},$$



(a) (b) (c)

Figure 3.7: Epipolar alignment. Red dots denote camera centers and vectors incident to the respective centers of the fields of view. Grey areas represent vectors in the fields of view of the respective cameras. (a) An example of epipolar geometry of an image pair. (b) Positions of the epipoles as computed by camera and epipolar calibration, i.e., from SVD of an essential matrix. (c) Positions of the epipoles after the epipolar alignment.

it holds that

$$\begin{aligned} \mathbf{x} &= \mathtt{A}_1^{-1}(1,0,0)^\top = (a_{1,1}, a_{2,1}, a_{3,1})^\top, \\ \mathbf{y} &= \mathtt{A}_1^{-1}(0,1,0)^\top = (a_{1,2}, a_{2,2}, a_{3,2})^\top, \\ \mathbf{e}_{1,2} &= \mathtt{A}_1^{-1}(0,0,1)^\top = (a_{1,3}, a_{2,3}, a_{3,3})^\top, \end{aligned}$$

that is

$$\mathtt{A}_1^{-1} = \begin{pmatrix} \mathbf{x} & \mathbf{y} & \mathbf{e}_{1,2} \end{pmatrix}.$$

By analogy, matrix $\mathtt{A}_2^{-1}$ realizing the transformation from the ordered basis Ω to the ordered basis $\Sigma^{\mathbf{u}_2}_{\mathbf{e}_{2,2}} = [\mathbf{x}', \mathbf{y}', \mathbf{e}_{2,2}]$, where $\mathbf{u_2} = [\mathbf{e}_{2,2}]_\times \mathtt{F}\mathbf{u}_1$, reads as

$$\mathtt{A}_2^{-1} = \begin{pmatrix} \mathbf{x}' & \mathbf{y}' & \mathbf{e}_{2,2} \end{pmatrix}.$$

Finally, we can derive the pair of transformations $\mathcal{A}_1, \mathcal{A}_2$ forming *the epipolar alignment* of an image pair connected by essential matrix $\mathtt{F}$ as

$$\forall \mathbf{q} \in \mathbb{R}^3 : \mathcal{A}_1(\mathbf{q}) = \mathtt{A}_1\mathbf{q} = \begin{pmatrix} \mathbf{x} & \mathbf{y} & \mathbf{e}_{1,2} \end{pmatrix}^{-1}\mathbf{q}, \; \mathcal{A}_2(\mathbf{q}) = \mathtt{A}_2\mathbf{q} = \begin{pmatrix} \mathbf{x}' & \mathbf{y}' & \mathbf{e}_{2,2} \end{pmatrix}^{-1}\mathbf{q}, \tag{3.8}$$

where $[\mathbf{x}, \mathbf{y}, \mathbf{e}_{1,2}] = \Sigma^{\mathbf{u}_1}_{\mathbf{e}_{1,2}}$, $[\mathbf{x}', \mathbf{y}', \mathbf{e}_{2,2}] = \Sigma^{\mathbf{u}_2}_{\mathbf{e}_{2,2}}$, $\mathbf{u_2} = [\mathbf{e}_{2,2}]_\times \mathtt{F}\mathbf{u}_1$ and $\mathbf{e}_{1,i}$, $\mathbf{e}_{2,i}$, $i = 1, 2$ are the respective epipole pairs.

4 Stereographic projection

Stereographic projection is a well known transformation of the surface of a sphere onto the surface of a plane. Since it is a crucial transformation to several rectification methods, this chapter discusses it in detail.

4.1 Stereographic projection

In the canonical definition of the transformation, the unit sphere centered in the origin and plane $z = 0$ are considered. The sphere is mapped onto the plane by means of *central projection*, where the center of the projection is the North Pole $\mathbf{N} = (0, 0, 1)^\top$, see Figure 4.1. Since the North Pole itself is not projected onto the plane, it is customary to add a new point, called ∞, to the plane, and to complete the map by mapping the North Pole onto ∞. This step turns stereographic projection into a *bijection*, and leads to the following definition [5]:

Definition 15 *Let* $\mathbf{q} = (q_x, q_y, q_z)^\top \in \mathbb{S}^3$ *be a point on the surface of the unit sphere. Then the* stereographic projection $\mathcal{S} : \mathbb{S}^3 \to \mathbb{R}^2 \cup \{\infty\}$ *maps* $\mathbf{q}$ *onto a point* $\mathbf{u} \in \mathbb{R}^2 \cup \{\infty\}$ *on the plane* $z = 0$ *extended by* ∞ *so that*

$$\mathbf{u} = \begin{cases} \infty & \text{for } \mathbf{q} = (0, 0, 1)^\top, \\ \begin{pmatrix} \frac{q_x}{1-q_z} \\ \frac{q_y}{1-q_z} \end{pmatrix} & \text{else.} \end{cases} \tag{4.1}$$

Consequence 2 *Let* $\mathbf{u} = (u_x, u_y) \in \mathbb{R}^2 \cup \{\infty\}$ *be a point on the plane* $z = 0$ *extended by* ∞*. Then the* inverse stereographic projection $\mathcal{S}^{-1} : \mathbb{R}^2 \cup \{\infty\} \to \mathbb{S}^3$ *maps* $\mathbf{u}$ *onto the unit sphere onto a point* $\mathbf{q} = (q_x, q_y, q_z)^\top \in \mathbb{S}^3$ *on the unit sphere* $\mathbb{S}^3$ *so that*

$$\mathbf{q} = \begin{cases} (0, 0, 1)^\top & \text{for } \mathbf{u} = \infty, \\ \begin{pmatrix} \frac{2u_x}{1+u_x^2+u_y^2} \\ \frac{2u_y}{1+u_x^2+u_y^2} \\ \frac{-1+u_x^2+u_y^2}{1+u_x^2+u_y^2} \end{pmatrix} & \text{else.} \end{cases} \tag{4.2}$$

4.2 Changing the center of projection

In the canonical definition of stereographic projection, the center of projection is the North Pole. That is, $\mathcal{S}(\mathbf{N}) = \infty$, $\mathcal{S}(-\mathbf{N}) = (0, 0)^\top$. How about a projection from an arbitrary point on the unit sphere?

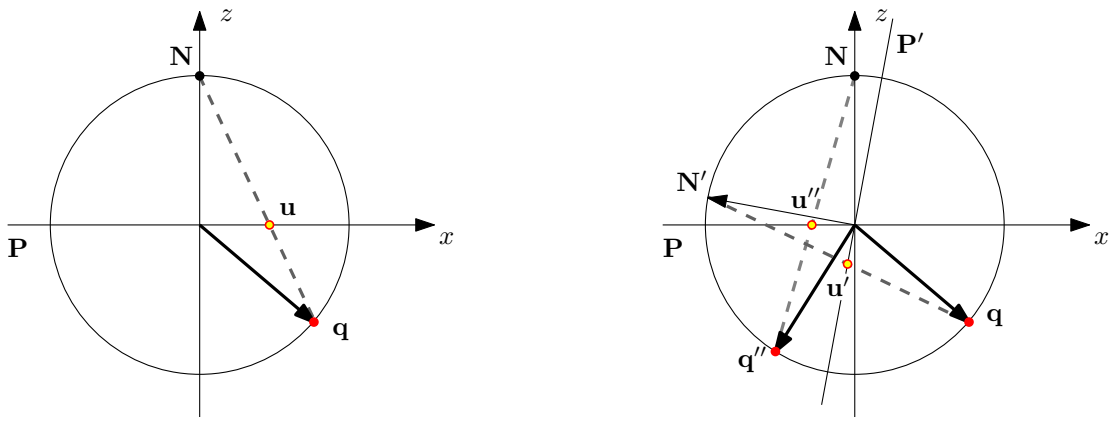


Figure 4.1: Stereographic projection: (a) Projection of a vector $\mathbf{q} \in \mathbb{S}^3$ onto a vector $\mathbf{u} \in \mathbb{R}^2$ by means of central projection from the North Pole $\mathbf{N}$ to the plane $\mathbf{P}$. (b) Projection of a vector $\mathbf{q} \in \mathbb{S}^3$ from $\mathbf{N}'$ to the plane $\mathbf{P}'$ is equivalent to the projection of a vector $\mathbf{q}''$ to a vector $\mathbf{u}''$ under the canonical stereographic projection, i.e., $\mathbf{u}'_{\mathbf{P}'} = \mathbf{u}''_{\mathbf{P}}$.

Definition 16 *The central projection* $\mathcal{S}_{\mathbf{N}'} : \mathbb{S}^3 \to \mathbb{R}^2 \cup \{\infty\}$ *of the unit sphere* $\mathbb{S}^3$ *with the center of projection* $\mathbf{N}' \in \mathbb{S}^3$ *to the plane* $\mathbf{N}' \cdot (x, y, z)^\top = 0$ *is called* stereographic projection from the point $\mathbf{N}'$.

Figure 4.1(b) depicts such a mapping.

Note, that the Definition 16 is not an uniquely defined transformation, but rather a set of plausible transformations, i.e., for an arbitrary vector $\mathbf{N} \in \mathbb{S}^3$ there exist more that one non-identical transformations that fit the definition of the stereographic projection from the point $\mathbf{N}$.

Theorem 2 *Let* $\mathtt{R} \in \mathbb{R}^{3\times3}$ *be an orthogonal matrix,* $\mathbf{N}' = \mathtt{R}\mathbf{N} = \mathtt{R}(0, 0, 1)^\top$. *Then*

$$\forall \mathbf{q} \in \mathbb{S}^3 : \ \mathcal{S}_{\mathbf{N}'}(\mathbf{q}) = \mathcal{S}(\mathtt{R}^{-1}\mathbf{q}).$$

Proof. Since $\mathtt{R}$ is an orthogonal matrix, $\mathtt{R}^{-1}$ always exits. The stereographic projection from the point $\mathbf{N}'$ of a point $\mathbf{q} \in \mathbb{S}^3$ is equivalent to finding the intersection of the ray emanating from the center of projection through the projected point,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = (\mathbf{N}' - \mathbf{q})\, u,$$

where $u \in \mathbb{R}$, and plane with normal vector $\mathbf{N}'$ and containing the origin $(0, 0, 0)^\top$,

$$\mathbf{N}' \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

Plugging the equation of the ray into the plane equation we get

$$\mathbf{N}'^{\top}\left(\mathbf{N}'-\mathbf{q}\right)u \;=\; 0. \tag{4.3}$$

From the assumption $\mathbf{N}'=\mathtt{R}\mathbf{N}$ and the orthogonality of the matrix $\mathtt{R}$

$$\begin{aligned}\left(\mathtt{R}\mathbf{N}\right)^{\top}\left(\mathtt{R}\mathbf{N}-\mathbf{q}\right)u &= 0\\ \mathbf{N}^{\top}\mathtt{R}^{-1}\left(\mathtt{R}\mathbf{N}-\mathbf{q}\right)u &= 0\\ \mathbf{N}^{\top}\left(\mathbf{N}-\mathtt{R}^{-1}\mathbf{q}\right)u &= 0\end{aligned} \tag{4.4}$$

which is the equation of the intersection of the plane $z=0$

$$\mathbf{N}\cdot\begin{pmatrix}x\\y\\z\end{pmatrix}=0$$

and the ray emanating from the North Pole through the point $\mathtt{R}^{-1}\mathbf{q}$,

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\left(\mathbf{N}-\mathtt{R}^{-1}\mathbf{q}\right),$$

Let's suppose $u'\in\mathbb{R}$ solves the identical Equations (4.3) and (4.4), then

$$\begin{aligned}\forall\mathbf{q}\in\mathbb{S}^3\backslash\{\mathbf{N}'\}:\ \mathcal{S}_{\mathbf{N}'}(\mathbf{q}) &= \mathtt{R}^{\top}\left(\mathbf{N}'-\mathbf{q}\right)u'\\ &= \mathtt{R}^{-1}\left(\mathbf{N}'-\mathbf{q}\right)u'\\ &= \mathtt{R}^{-1}\left(\mathtt{R}\mathbf{N}-\mathbf{q}\right)u'\\ &= \left(\mathtt{R}^{-1}\mathtt{R}\mathbf{N}-\mathtt{R}^{-1}\mathbf{q}\right)u'\\ &= \left(\mathbf{N}-\mathtt{R}^{-1}\mathbf{q}\right)u'\\ &= \mathcal{S}(\mathtt{R}^{-1}\mathbf{q})\end{aligned}$$

To complete the transformation,

$$\mathcal{S}_{\mathbf{N}'}(\mathbf{N}')=\mathcal{S}(\mathbf{N})=\infty.$$

□

Consequence 3 *Let $\mathtt{R}\in\mathbb{R}^{3\times3}$ be an orthogonal matrix, $\mathbf{N}'=\mathtt{R}\mathbf{N}=\mathtt{R}(0,0,1)^{\top}$. Then the inverse stereographic projection from the point $\mathbf{N}'$ reads as*

$$\forall\mathbf{u}\in\mathbb{R}^2:\ \mathcal{S}^{-1}_{\mathbf{N}'}(\mathbf{u})=\mathtt{R}\mathcal{S}(\mathbf{u}).$$

5 Rectification of an omnidirectional image pair

The theory of previous chapters allows us to present a general technique for rectification of an omnidirectional image pair.

Let $[\mathcal{P}_1, \mathcal{P}_2]$ be an omnidirectional image pair, such that $\mathcal{P}_1 : [0, a_1] \times [0, a_2] \rightarrow \mathbb{Z}^m, \mathcal{P}_2\,[0, a'_1] \times [0, a'_2] \rightarrow \mathbb{Z}^m$, i.e., $[a_1, a_2]$ and $[a'_1, a'_2]$ being the respective dimensions of the images. Let $\mathcal{C}_1$ be the calibration transformation of the camera C_1 that acquired the image $\mathcal{P}_1$, $\mathcal{C}_2$ the calibration transformation of the camera C_2 that acquired the image $\mathcal{P}_2$. Let $\mathtt{F}$ be an essential matrix describing the epipolar geometry of the cameras C_1 and C_2 and $[\mathcal{A}_1, \mathcal{A}_2]$ the epipolar alignment based on the essential matrix $\mathtt{F}$. Then a rectification of the omnidirectional image pair $[\mathcal{P}_1, \mathcal{P}_2]$ can be viewed as a pair of dependent geometric image transformations with the underlying geometric transformations connected by the epipolar alignment $[\mathcal{A}_1, \mathcal{A}_2]$. A pair of such underlying geometric transformations $[\mathcal{G}_1, \mathcal{G}_2]$ is called a *rectification method*. The inner structure of the geometric transformations is the following:

$$\begin{aligned} \mathcal{G}_1 &= \mathcal{C}_1 \circ \mathcal{A}_1 \circ \mathcal{T} \circ \mathcal{F}, \\ \mathcal{G}_2 &= \mathcal{C}_2 \circ \mathcal{A}_2 \circ \mathcal{T} \circ \mathcal{F}, \end{aligned}$$

where $\mathcal{T} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ is the *characteristic transformation* of a rectification method and $\mathcal{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ a final affine transformation. The first part $\mathcal{C}_{1,2} \circ \mathcal{A}_{1,2}$ is clearly mutual to all methods, thus to fully define a rectification method only the second part $\mathcal{T} \circ \mathcal{F}$ needs to be specified.

In order to derive a rectification of an image pair based on the image interpolation, inverse transformations

$$\begin{aligned} \mathcal{G}_1^{-1} &= \mathcal{F}^{-1} \circ \mathcal{T}^{-1} \circ \mathcal{A}_1^{-1} \circ \mathcal{C}_1^{-1}, \\ \mathcal{G}_2^{-1} &= \mathcal{F}^{-1} \circ \mathcal{T}^{-1} \circ \mathcal{A}_2^{-1} \circ \mathcal{C}_2^{-1}, \end{aligned}$$

need to be derived. Since we assume the existence of $\mathcal{C}_{1,2}^{-1}$ and previously derived $\mathcal{A}_{1,2}^{-1}$, again, only $\mathcal{F}^{-1} \circ \mathcal{T}^{-1}$ needs to be specified. The rectification of an image pair $[\mathcal{P}_1, \mathcal{P}_2]$ using method $[\mathcal{G}_1, \mathcal{G}_2]$ based on image interpolation using interpolating synthesis function φ then reads as

$$\left[\mathcal{I}_{\mathcal{P}_1}^{\varphi}\left(\mathcal{G}_1^{-1}\right), \mathcal{I}_{\mathcal{P}_2}^{\varphi}\left(\mathcal{G}_2^{-1}\right)\right].$$

In this chapter we present three scanline rectification methods, i.e., methods that transform epipolar curves to the scanlines of the resulting rectified image pair. The first two are based on spherical parametrization, the third is the conformal rectification method described in [2]. Further, we present a rectification method based on the stereographic projection. Finally, properties of the respective rectification methods are discussed using various examples of image pairs from Appendix B, rectified by the *OmniRect* MATLAB toolbox described in Appendix A.

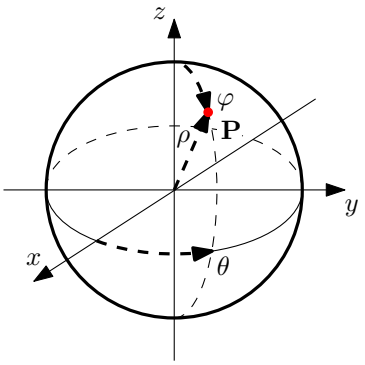


Figure 5.1: Spherical parametrization of a point **P** on the unit sphere.

5.1 Spherical rectification

The characteristic transformation of the spherical rectification (SR) is based on the spherical parametrization. The Spherical coordinates (ρ, φ, θ), see Figure 5.1, are obtained from Cartesian coordinates as [13]

$$\begin{aligned} \rho &= \sqrt{x^2 + y^2 + z^2}, \\ \varphi &= \arctan\left(\frac{\sqrt{x^2 + y^2}}{z}\right), \\ \theta &= \arctan\left(\frac{y}{x}\right), \end{aligned}$$

where $\rho \in [0, \infty)$, $\varphi \in [0, \pi]$, $\theta \in [0, 2\pi]$. The Cartesian coordinates can be recovered as

$$\begin{aligned} x &= \rho \sin\varphi \cos\theta, \\ y &= \rho \sin\varphi \sin\theta, \\ z &= \rho \cos\varphi. \end{aligned} \tag{5.1}$$

The characteristic transformation of the spherical rectification $\mathcal{T} : \mathbb{R}^3 \to \mathbb{R}^2$ consists of spherical parametrization of the unit sphere, thus ρ is always 1. Further, for epipolar circles – after epipolar alignment coinciding with meridians of the unit sphere – to be mapped to scanlines, φ must coincide with the first coordinate.

Definition 17 $\mathcal{T}_{SR} : \mathbb{R}^3 \to \mathbb{R}^2$ *such that*

$$\forall \mathbf{p} = (p_1, p_2, p_3) \in \mathbb{S}^3 : \mathcal{T}_{SR}(\mathbf{p}) = \begin{pmatrix} \arctan\left(\frac{\sqrt{p_1^2 + p_2^2}}{p_3}\right) \\ \arctan\left(\frac{p_2}{p_1}\right) \end{pmatrix}, \tag{5.2}$$

is the characteristic transformation of the spherical rectification.

The function arctan used in Definition 17 must be defined so that is takes into account the correct quadrant of $\frac{a}{b}$. The function atan2(a, b) available in various programming languages can be used. However, note that the range of the function atan2(a, b) is

$(-\pi, \pi]$. The inverse transformation to the characteristic transformation of the spherical rectification $\mathcal{T}_{SR}^{-1} : \mathbb{R}^2 \to \mathbb{R}^3$ then reads as

$$\forall \mathbf{u} = (\varphi, \theta) \in \mathbb{R}^2 : \mathcal{T}_{SR}^{-1}(\mathbf{u}) = \begin{pmatrix} \sin\varphi\cos\theta \\ \sin\varphi\sin\theta \\ \cos\varphi \end{pmatrix}. \tag{5.3}$$

Seeing that the effective range of $\mathcal{T}_{SR}$ is $[0, \pi] \times [0, 2\pi]$ the respective affine transformation $\mathcal{F}_{SR} : \mathbb{R}^2 \to \mathbb{R}^2$ for an image with width w and height h derives as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{F}_{SR}(\mathbf{u}) = \begin{pmatrix} \frac{w}{\pi} & 0 \\ 0 & \frac{h}{2\pi} \end{pmatrix} \mathbf{u}, \tag{5.4}$$

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{F}_{SR}^{-1}(\mathbf{u}) = \begin{pmatrix} \frac{\pi}{w} & 0 \\ 0 & \frac{2\pi}{h} \end{pmatrix} \mathbf{u}. \tag{5.5}$$

5.2 Swapped spherical rectification

Swapped spherical rectification (SSR) is almost identical to the spherical rectification, the only difference being that the domain intervals of the spherical coordinate system $[0, \pi] \times [0, 2\pi]$ are, indeed, swapped. This swap leads to somewhat different results from the spherical rectification.

Definition 18 $\mathcal{T}_{SSR} : \mathbb{R}^3 \to \mathbb{R}^2$ *such that*

$$\forall \mathbf{p} = (p_1, p_2, p_3) \in \mathbb{S}^3 : \mathcal{T}_{SSR}(\mathbf{p}) = \begin{pmatrix} \arctan\left(\frac{\sqrt{p_1^2 + p_2^2}}{p_3}\right) \\ \arctan\left(\frac{p_2}{p_1}\right) \end{pmatrix}, \tag{5.6}$$

is the characteristic transformation of the swapped spherical rectification.

Note that $\mathcal{T}_{SSR} = \mathcal{T}_{SR}$. The inverse transformation to the characteristic transformation of the spherical rectification $\mathcal{T}_{SSR}^{-1} : \mathbb{R}^2 \to \mathbb{R}^3$ again reads as

$$\forall \mathbf{u} = (\varphi, \theta) \in \mathbb{R}^2 : \mathcal{T}_{SSR}^{-1}(\mathbf{u}) = \begin{pmatrix} \sin\varphi\cos\theta \\ \sin\varphi\sin\theta \\ \cos\varphi \end{pmatrix}.$$

The difference between spherical and swapped spherical rectifications lie in the final affine transformation. Seeing that the range of $\mathcal{T}_{SSR}$ is $[0, 2\pi] \times [0, \pi]$ the respective affine transformation $\mathcal{F}_{SSR} : \mathbb{R}^2 \to \mathbb{R}^2$ for an image with width w and height h derives as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{F}_{SSR}(\mathbf{u}) = \begin{pmatrix} \frac{w}{2\pi} & 0 \\ 0 & \frac{h}{\pi} \end{pmatrix} \mathbf{u}, \tag{5.7}$$

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{F}_{SSR}^{-1}(\mathbf{u}) = \begin{pmatrix} \frac{2\pi}{w} & 0 \\ 0 & \frac{\pi}{h} \end{pmatrix} \mathbf{u}. \tag{5.8}$$

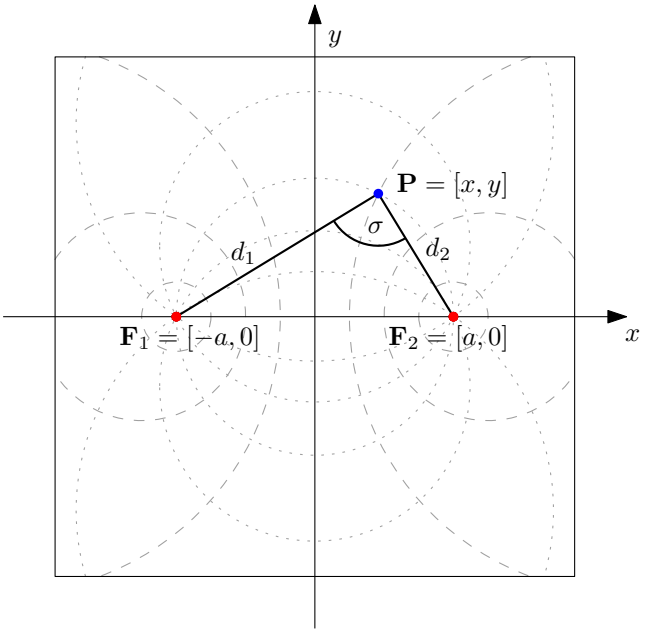


Figure 5.2: Bipolar coordinate system. A few σ isosurfaces are shown as dotted circles, dashed circles denote τ isosurfaces .

5.3 Bipolar rectification

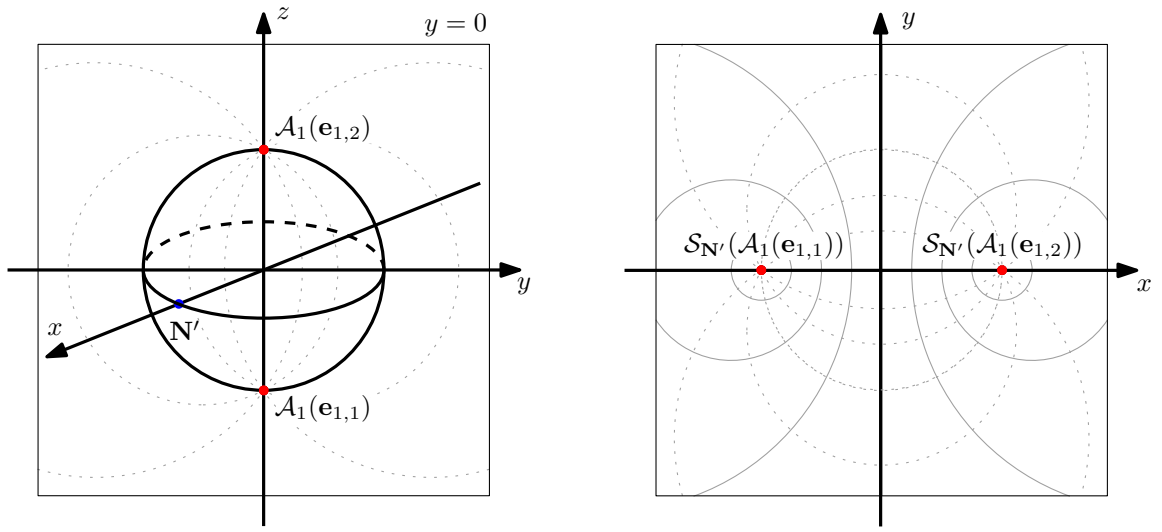
The bipolar rectification (BR) is based on a conformal rectification described in [2]. The work describes rectification of a stereo pair acquired by a para-catadioptric camera based on direct bipolar parametrization, where the foci of the parametrization are identified with epipoles. Since para-catadioptric projection is equivalent to stereographic projection, a method based on the bipolar parametrization of the stereographic projection of the spherical model can be used for every omnidirectional central camera. However, it remains a conformal rectification only for images acquired by para-catadioptric camera.

The Bipolar coordinates (σ, τ) of a point $\mathbf{P} = (x, y)$, see Figure 5.2, are defined as [12]

$$\begin{aligned} x &= a\,\frac{\sinh\tau}{\cosh\tau - \cos\sigma}, \\ y &= a\,\frac{\sin\sigma}{\cosh\tau - \cos\sigma}, \end{aligned}$$

where $\sigma \in [-\pi, \pi]$ is the angle $\mathbf{F}_1\mathbf{P}\mathbf{F}_2$ and $\tau \in (-\infty, \infty)$ the natural logarithm of distances $d_1 = |\mathbf{F}_1\mathbf{P}|$ and $d_2 = |\mathbf{F}_1\mathbf{P}|$,

$$\begin{aligned} \sigma &= \arccos\left(\frac{x^2 + y^2 - a^2}{\sqrt{(a-x)^2 + y^2}\sqrt{(a+x)^2 + y^2}}\right), \\ \tau &= \frac{1}{2}\ln\left(\frac{(a-x)^2 + y^2}{(a+x)^2 + y^2}\right). \end{aligned}$$



(a) (b)

Figure 5.3: Bipolar rectification. (a) Stereographic projection from $\mathbf{N}' = (1,0,0)^\top$ to the plane $y = 0$. A few epipolar circles are shown to be projected onto dotted circles in the plane $y = 0$. (b) The unit sphere after stereographic projection. The epipolar circles are mapped onto circles going through $(-1,0)^\top$ and $(1,0)$, or to put it alternatively, onto σ isosurfaces in bipolar coordinate system with $\mathbf{F}_1 = (-1,0)^\top$ and $\mathbf{F}_2 = (1,0)^\top$.

The bipolar rectification is based on the fact that the spherical rectification projects circles onto circles. After the epipolar alignment the epipoles are situated in $(0,0,1)^\top$ and $(0,0,-1)^\top$ respectively, thus the canonical stereographic projection is not suitable. To obtain the projection so that

$$\begin{aligned} \mathcal{A}_1(\mathbf{e}_{1,1}) = \mathcal{A}_2(\mathbf{e}_{2,1}) = (0,0,-1)^\top &\mapsto (0,-1)^\top, \\ \mathcal{A}_1(\mathbf{e}_{1,2}) = \mathcal{A}_2(\mathbf{e}_{2,2}) = (0,0,1)^\top &\mapsto (0,1)^\top, \end{aligned}$$

stereographic projection from $\mathbf{N}' = (1,0,0)^\top$ has to be used, see Figure 5.2. The projection plane of such a projection is $y = 0$. An orthogonal matrix

$$\mathtt{R} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

satisfies the condition of Theorem 2, that is $\mathtt{R}(0,0,1)^\top = (1,0,0)^\top = \mathbf{N}'$, and we get

$$\forall \mathbf{q} \in \mathbb{R}^3 : \ \mathcal{S}_{\mathbf{N}'}(\mathbf{q}) = \mathcal{S}(\mathtt{R}^{-1}\mathbf{q}),$$

as well as

$$\begin{aligned} \mathcal{S}_{\mathbf{N}'}(\mathcal{A}_1(\mathbf{e}_{1,1})) = \mathcal{S}_{\mathbf{N}'}(\mathcal{A}_2(\mathbf{e}_{2,1})) = \mathcal{S}_{\mathbf{N}'}((0,0,-1)^\top) = \mathcal{S}(\mathtt{R}^{-1}(0,0,-1)^\top) &= (-1,0)^\top, \\ \mathcal{S}_{\mathbf{N}'}(\mathcal{A}_1(\mathbf{e}_{1,2})) = \mathcal{S}_{\mathbf{N}'}(\mathcal{A}_2(\mathbf{e}_{2,2})) = \mathcal{S}_{\mathbf{N}'}((0,0,1)^\top) = \mathcal{S}(\mathtt{R}^{-1}(0,0,1)^\top) &= (1,0)^\top. \end{aligned}$$

To finalize the bipolar rectification, the resulting stereographic projection has to be converted into bipolar coordinates, so that the foci of the bipolar coordinate system

are $\mathbf{F}_1 = (-1, 0)^\top$ and $\mathbf{F}_2 = (1, 0)^\top$, see Figure 5.3(b). Seeing that the epipolar lines – that we want to map onto scanlines – are σ isosurfaces, positions of σ and τ in the canonical definition of bipolar transformation need to be swapped.

Definition 19 $\mathcal{B} : \mathbb{R}^2 \to \mathbb{R}^2$, *such that*

$$\forall \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2 : \mathcal{B}(\mathbf{u}) = \begin{pmatrix} \arccos\left(\frac{u_1^2+u_2^2-1}{\sqrt{(1-u_2)^2+u_1^2}\sqrt{(1+u_2)^2+u_1^2}}\right) \\ \frac{1}{2}\ln\left(\frac{(1-u_2)^2+u_1^2}{(1+u_2)^2+u_1^2}\right) \end{pmatrix}$$

is called bipolar transformation.

The inverse bipolar transformation reads as

$$\forall \mathbf{u} = (u_1, u_2) \in \mathbb{R}^2 : \mathcal{B}^{-1}(\mathbf{u}) = \begin{pmatrix} \frac{\sinh u_1}{\cosh u_1 - \cos u_2} \\ \frac{\sin u_2}{\cosh u_1 - \cos u_2} \end{pmatrix}.$$

Now, we can define the characteristic transformation of the bipolar rectification.

Definition 20 *Let*

$$\mathtt{R} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then $\mathcal{T}_{BR} : \mathbb{R}^3 \to \mathbb{R}^2$ *such that*

$$\forall \mathbf{q} \in \mathbb{R}^3 : \mathcal{T}_{BR}(\mathbf{q}) = \mathcal{B}(\mathcal{S}(\mathtt{R}^{-1}\mathbf{q})) \tag{5.9}$$

is the characteristic transformation of the bipolar rectification.

Using Consequence 3, inverse characteristic transformation of the bipolar rectification reads as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{T}_{BR}^{-1}(\mathbf{u}) = \mathtt{R}\mathcal{S}^{-1}(\mathcal{B}^{-1}(\mathbf{u})). \tag{5.10}$$

The range of $\mathcal{B}$ is $(-\infty, \infty) \times [-\pi, \pi]$ and not even the final affine transformation can map the unlimited interval to the limited image. Using an additional parameter δ, the visible range of $\mathcal{B}$ can to be limited to $[-\delta, \delta] \times [-\pi, \pi]$. Now the final affine transformation transforming the origin to the center of the image and limiting the visible range using parameter δ formulates as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{F}_{SR}^{\delta}(\mathbf{u}) = \begin{pmatrix} \frac{w}{2\delta} & 0 \\ 0 & \frac{h}{2\pi} \end{pmatrix} \mathbf{u} + \begin{pmatrix} \frac{w}{2} \\ \frac{h}{2} \end{pmatrix}, \tag{5.11}$$

where w is the width and h the height of the resulting image. The inverse transformation derives as

$$\forall \mathbf{u} \in \mathbb{R}^2 : {\mathcal{F}_{SR}^{\delta}}^{-1}(\mathbf{u}) = \begin{pmatrix} \frac{2\delta}{w} & 0 \\ 0 & \frac{2\pi}{h} \end{pmatrix} \mathbf{u} - \begin{pmatrix} \delta \\ \pi \end{pmatrix}. \tag{5.12}$$

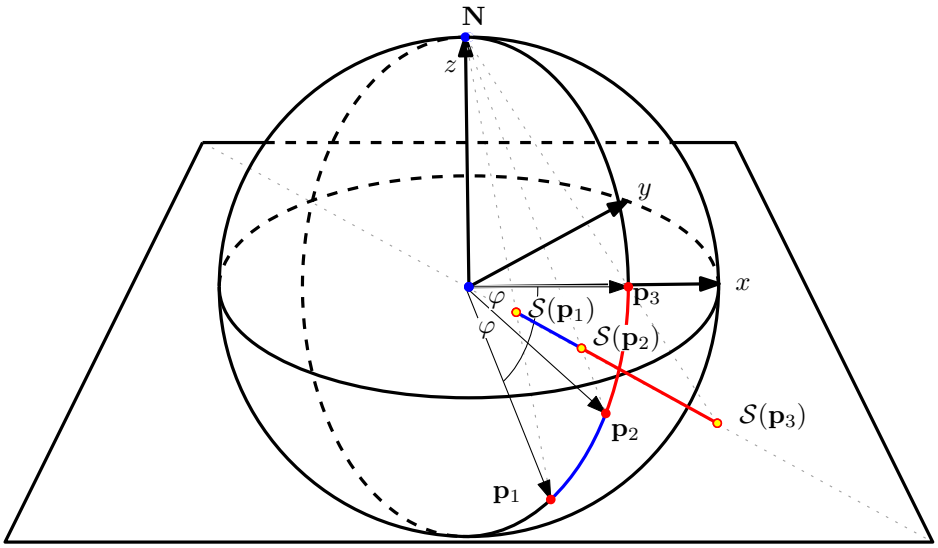


Figure 5.4: Spherically equidistantly positioned vectors $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$ and their respective stereographic transformations $\mathcal{S}(\mathbf{p}_1), \mathcal{S}(\mathbf{p}_2), \mathcal{S}(\mathbf{p}_3)$.

5.4 Stereographic rectification

The stereographic rectification (SGR) is a stereographic projection from a certain point of the stereo pair after elipolar alignment. The basic difference between stereographic rectification and previously described rectification methods is the fact that since it only employs stereographic projection in its characteristic transformation, it does not map epipolar curves onto scanlines but onto general circles.

The inherent setback of the stereographic projection is the fact, that areas near the center of projection on the sphere are projected onto "disproportionally" large areas on the projection plane, see Figure 5.4, to be more precise, the ratio of the area on the sphere and its stereographic projection is non-linear. The Theorem 3 formally substantiates this assertion.

Theorem 3

$$\neg\left(\exists\alpha\in\mathbb{R}\,\forall\mathbf{p}\in\mathbb{S}^3\backslash\{(0,0,\pm1)^\top\}:\ \frac{\delta_{\mathbb{S}^3}((0,0,-1)^\top,\mathbf{p})}{\|\mathcal{S}(\mathbf{p})\|}=\alpha\right)$$

Proof. The assertion becomes apparent after transformation using modified spherical coordinates

$$\mathbf{p}=\begin{pmatrix}\sin(\pi-\varphi)\cos\theta\\ \sin(\pi-\varphi)\sin\theta\\ \cos(\pi-\varphi)\end{pmatrix}, \tag{5.13}$$

where $\varphi\in[0,\pi], \theta\in[0,2\pi]$. The only difference to the canonical spherical coordinate transformation from Equation (5.1) is the fact that the meridians are parametrized starting in the South Pole and the fact that since we parametrize unit length vectors only, there is no need for the parameter ρ. The domain $\mathbb{S}^3\setminus\{(0,0,1)^\top$ becomes $[0,\pi)\times[0,2\pi]$.

Starting with $\delta_{\mathbb{S}^3}((0,0,-1)^\top, \mathbf{p})$, from the definition we have

$$\delta_{\mathbb{S}^3}((0,0,-1)^\top, \mathbf{p}) \;=\; \arccos((0,0,-1)\mathbf{p}).$$

After reparametrization using Equation (5.13), we have

$$\begin{aligned}\delta_{\mathbb{S}^3}(\mathbf{p}) &= \arccos((0,0,-1)(\sin(\pi-\varphi)\cos\theta, \sin(\pi-\varphi)\sin\theta, \cos(\pi-\varphi))^\top),\\ &= \arccos(-\cos(\pi-\varphi)),\\ &= \arccos(\cos(\varphi)),\end{aligned}$$

and finally since $\varphi \in (0,\pi)$

$$\delta_{\mathbb{S}^3}(\mathbf{p}) \;=\; \varphi.$$

After denoting $\mathbf{p} = (p_1, p_2, p_3)^\top$ the denominator $\|\mathcal{S}(\mathbf{p})\|$ reads as

$$\|\mathcal{S}(\mathbf{p})\| \;=\; \sqrt{\left(\frac{p_1}{1-p_3}\right)^2 + \left(\frac{p_2}{1-p_3}\right)^2}.$$

Using Equation (5.13) again, we have, after a few trigonometric recombinations,

$$\begin{aligned}\|\mathcal{S}(\mathbf{p})\| &= \sqrt{\left(\frac{\sin(\pi-\varphi)\cos(\theta)}{1-\cos(\pi-\varphi)}\right)^2 + \left(\frac{\sin(\pi-\varphi)\sin(\theta)}{1-\cos(\pi-\varphi)}\right)^2},\\ &= \sqrt{\frac{1-\cos(2\varphi)}{3+4\cos\varphi+\cos(2\varphi)}},\\ &= \sqrt{\frac{1-\cos(\varphi)}{1+\cos(\varphi)}},\end{aligned}$$

and finally, again since $\varphi \in (0,\pi)$,

$$\|\mathcal{S}(\mathbf{p})\| \;=\; \tan\left(\frac{\varphi}{2}\right).$$

The assertion can now be rewritten as

$$\neg\left(\exists\alpha\in\mathbb{R}\;\forall\varphi\in(0,\pi):\; \frac{\varphi}{\tan\left(\frac{\varphi}{2}\right)} = \alpha\right).$$

Indeed, the function $\frac{\varphi}{\tan\left(\frac{\varphi}{2}\right)} = \varphi\cot\frac{\varphi}{2}$ is clearly non-linear. □

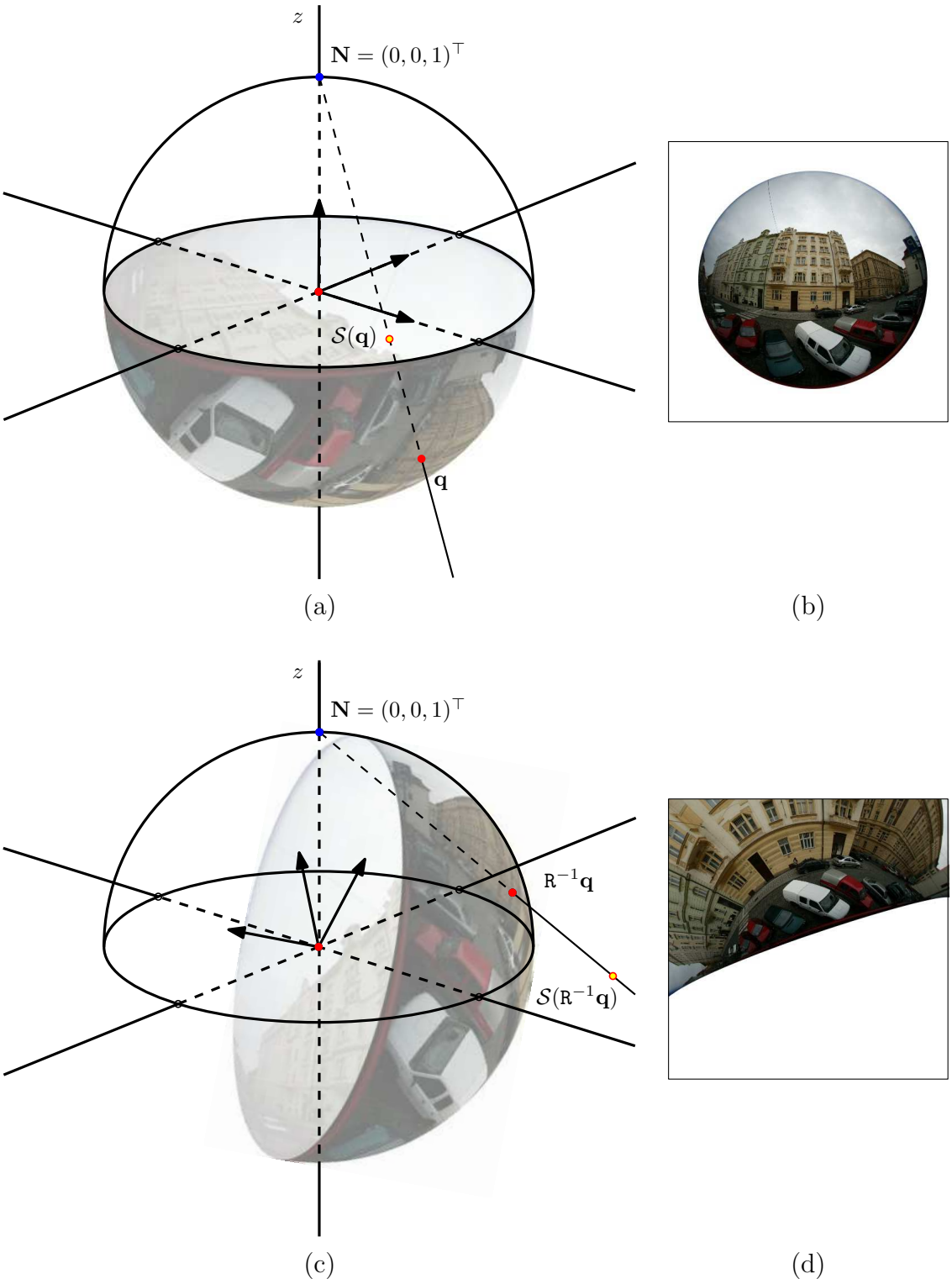


Figure 5.5: The effect of the change of the center of stereographic projection on a calibrated omnidirectional image. (a) The unit sphere colored using calibration transformation. (b) Stereographic projection of the calibrated image. (c) The unit sphere transformed by an arbitrary orthogonal matrix $\mathtt{R}^{-1}$. (d) Stereographic projection of the transformed sphere.

A consequence of the Theorem 3 is the fact that in order to obtain as equally dimensioned stereographic projection of an area on the unit sphere as possible, the center of projection has to be on the unit sphere as far as possible. To put it formally, for an area $\Omega \subset \mathbb{S}^3$, the center of projection so that $\mathcal{S}_{\mathbf{N}'}(\Omega)$ is as "large" – or distortless – as Ω as possible, is

$$\mathbf{N}' = \arg\max_{\mathbf{p}\in\mathbb{S}^3} \Delta_{\mathbb{S}^3}(\Omega, \mathbf{p}),$$

where $\Delta_{\mathbb{S}^3}$ is the spherical distance between $\mathbf{p}$ and Ω, see Definition 7. If more that one area is considered, a compromise in distortions introduced to the respective areas can be achieved by projecting from a point maximizing the distance from the union of the areas,

$$\mathbf{N}' = \arg\max_{\mathbf{p}\in\mathbb{S}^3} \Delta_{\mathbb{S}^3}(\bigcup_{i\in I} \Omega_i, \mathbf{p}).$$

In the case of the stereographic rectification the areas of interest are the respective field on views. If we denote $V_i, i = 1, 2$ as sets of pixels lying in the fields of view of the respective images, i.e., the field of view ellipses, the "visible" directions in the coordinate systems of the respective cameras reads as

$$\Omega_i' = \{\mathbf{p} : \ \mathbf{p} \in \mathbb{S}^3 \,\&\, \mathcal{C}_i^{-1}(\mathbf{p}) \in V_i\},$$

where $i = 1, 2$ and $\mathcal{C}_i$ are the respective calibration transformations. After the epipolar alignment $\Omega_i', i = 1, 2$ transform to

$$\Omega_i = \{\mathbf{p} : \ \mathbf{p} \in \mathbb{S}^3 \,\&\, \mathcal{C}_i^{-1}(\mathcal{A}_i^{-1}(\mathbf{p})) \in V_i\}.$$

Now the characteristic transformation of the stereographic rectification can be formulated.

Definition 21 *Let* $\Omega_i = \{\mathbf{p} : \ \mathbf{p} \in \mathbb{S}^3 \,\&\, \mathcal{C}_i^{-1}(\mathcal{A}_i^{-1}(\mathbf{p})) \in V_i\}$, $\mathbf{N}' = \arg\max_{\mathbf{p}\in\mathbb{S}^3} \Delta_{\mathbb{S}^3}(\bigcup_{i\in I} \Omega_i, \mathbf{p})$, $i = 1, 2$. *Then* $\mathcal{T}_{SGR} : \mathbb{R}^3 \to \mathbb{R}^2$, *such that*

$$\forall \mathbf{q} \in \mathbb{R}^3 : \ \mathcal{T}_{SGR}(\mathbf{q}) = \mathcal{S}_{\mathbf{N}'}(\mathbf{q})$$

is called the characteristic transformation of the stereographic rectification.

Although the Definition 21 is a mathematically correct definition of a transformation, it says little about its actual enumeration. The obstacle lies in the computation of the center of the stereographic projection $\mathbf{N}'$, which in its general formulation poses a rather difficult optimization problem. However, with some supplementary assumptions, this can be heavily simplified, at least in the case of two sets. Let's assume that the visible angle of both cameras is 180°. Then $\Omega_i', i = 1, 2$ become exactly the southern hemisphere $\Omega_i' = \{\mathbf{p} : \ \mathbf{p} \in \mathbb{S}^3 \,\&\, (0, 0, 1)\mathbf{p} \leq 0\}$, where $\mathbf{N}_i' = (0, 0, 1)^\top$ are the hemisphere "normals". After the epipolar alignment $\Omega_i', i = 1, 2$ transform into arbitrary positioned hemispheres $\Omega_i = \{\mathbf{p} : \ \mathbf{p} \in \mathbb{S}^3 \,\&\, \mathbf{N}_i^\top \mathbf{p} \leq 0\}$, where $\mathbf{N}_i = \mathcal{A}_i((0, 0, 1)^\top)$ are the new normals, see Figure 5.6. Another plausible simplifying assumption is the restriction $1 > \mathbf{N}_1^\top \mathbf{N}_2 > 0$. This assumption constrains the angle between the normals to be less than 180°. The opposite case would result in hemispheres facing in almost

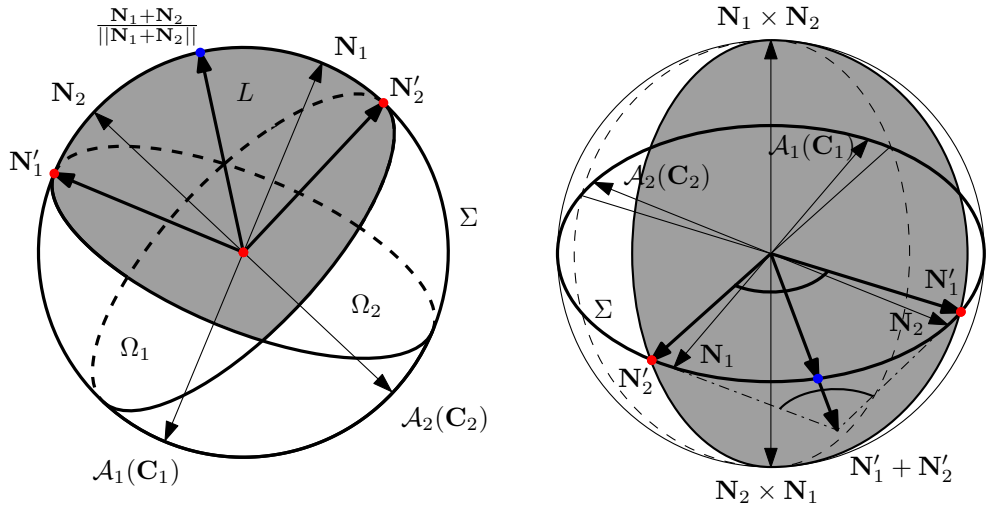


Figure 5.6: The epipolar alignment of $\Omega_i, i = 1, 2$ with the assumption that the visible angle of both cameras is 180°. Both figure (a) and (b) depict the identical situation, only from a different viewpoints

opposite direction, "seeing" almost different scenes, which is not a reasonable setup for stereo matching. In the case that $\mathbf{N}_1^\top \mathbf{N}_2 = 1$, the optimal projection point is, indeed, $\mathbf{N}' = \mathbf{N}_1 = \mathbf{N}_2$. Using these constraints, Theorem 4 explicitly states the optimal projection center.

Theorem 4 *Let $i = 1, 2$, $\mathcal{A}_i$ be the epipolar alignment of a stereo pair,$\mathbf{N}_i = \mathcal{A}_i((0, 0, 1)^\top)$ such that $1 > \mathbf{N}_1^\top \mathbf{N}_2 > 0$, $\Omega_i = \{\mathbf{p} : \mathbf{p} \in \mathbb{S}^3 \,\&\, (0, 0, 1)\mathcal{A}_i^{-1}(\mathbf{p}) \leq 0\}$. Then*

$$\mathbf{N}' = \arg\max_{\mathbf{v} \in \mathbb{S}^3} \Delta_{\mathbb{S}^3}(\bigcup_{i \in \{1,2\}} \Omega_i, \mathbf{v}) = \frac{\mathbf{N}_1 + \mathbf{N}_2}{\|\mathbf{N}_1 + \mathbf{N}_2\|}.$$

Proof. Since $\Omega_i, i = 1, 2$ are hemispheres, it follows that the boundaries $\partial\Omega_i, i = 1, 2$ are great circles on the unit sphere and

$$L = \mathbb{S}^3 \setminus \bigcup_{i \in \{1,2\}} \Omega_i$$

is a *lune* on the unit sphere, see Figure 5.6. Now the unknown $\mathbf{N}'$ can be rewritten as

$$\mathbf{N}' = \arg\max_{\mathbf{v} \in L} \Delta_{\mathbb{S}^3}(\partial L, \mathbf{v}).$$

Let $P_\Sigma = \{\mathbf{p} : \mathbf{p} \in \mathbb{R}^3 \,\&\, \mathbf{p}^\top(\mathbf{N}_1 \times \mathbf{N}_2) = 0\}$ be a plane containing both $\mathbf{N}_1$ and $\mathbf{N}_2$, $\Sigma = \{\mathbf{p} : \mathbf{p} \in \mathbb{S}^3 \,\&\, \mathbf{p} \in P_\Sigma\}$ a great circle defined by the plane P_Σ, see Figure 5.6(b). Let $P_{\boldsymbol{N}_1} = \{\mathbf{p} : \mathbf{p} \in \mathbb{R}^3 \,\&\, \mathbf{p}^\top \mathbf{N}_1 = 0\}$, a plane defined by the normal vector $\mathbf{N}_1$, $P_{\mathbf{N}_2} = \{\mathbf{p} : \mathbf{p} \in \mathbb{R}^3 \,\&\, \mathbf{p}^\top \mathbf{N}_2 = 0\}$, a plane defined by the normal vector $\mathbf{N}_2$. Let

$$\begin{aligned} \mathbf{N}_1' &= P_{\mathbf{N}_1} \cap \Sigma \cap \partial L, \\ \mathbf{N}_2' &= P_{\mathbf{N}_2} \cap \Sigma \cap \partial L, \end{aligned}$$

see Figure 5.6(b). It is an easily conceivable geometrical fact, that the vector $\mathbf{N}' \in L$ maximizing the distance $\Delta_{\mathbb{S}^3}(\partial L, \mathbf{N}')$ lies on the "equator" of L, precisely between $\mathbf{N}'_1$ and $\mathbf{N}'_2$. A vector dissecting the equator between $\mathbf{N}'_1$ and $\mathbf{N}'_2$ is constructed as $\mathbf{N}'_1 + \mathbf{N}'_2$ and since $\mathbf{N}' \in \mathbb{S}^3$ is required,

$$\mathbf{N}' = \frac{\mathbf{N}'_1 + \mathbf{N}'_2}{\|\mathbf{N}'_1 + \mathbf{N}'_2\|}.$$

$\mathbf{N}'_1$ explicitly quantifies as

$$\mathbf{N}'_1 = \frac{(\mathbf{N}_1 \times \mathbf{N}_2) \times \mathbf{N}_1}{\|(\mathbf{N}_1 \times \mathbf{N}_2) \times \mathbf{N}_1\|},$$

since $((\mathbf{N}_1 \times \mathbf{N}_2)^\top \mathbf{N}'_1 = 0) \Rightarrow \mathbf{N}'_1 \in \Sigma$ and $(\mathbf{N}_1 \times \mathbf{N}'_1 = 0) \Rightarrow \mathbf{N}'_1 \in P_{N_1}$. The requirement $\mathbf{N}'_1 \in \partial L$ is guarantied by the order of the members of the cross products. By analogy,

$$\mathbf{N}'_2 = \frac{(\mathbf{N}_2 \times \mathbf{N}_1) \times \mathbf{N}_2}{\|(\mathbf{N}_2 \times \mathbf{N}_1) \times \mathbf{N}_2\|}.$$

Using Lagrange's formula, the fact that $\mathbf{N}_1, \mathbf{N}_2$ are unit vectors and the commutativity of the scalar product, we have

$$\begin{aligned}(\mathbf{N}_1 \times \mathbf{N}_2) \times \mathbf{N}_1 &= -(\mathbf{N}_1(\mathbf{N}_1 \cdot \mathbf{N}_2) - \mathbf{N}_2(\mathbf{N}_1 \cdot \mathbf{N}_1)) = \mathbf{N}_2 - \mathbf{N}_1(\mathbf{N}_1^\top \mathbf{N}_2),\\ (\mathbf{N}_2 \times \mathbf{N}_1) \times \mathbf{N}_2 &= -(\mathbf{N}_2(\mathbf{N}_2 \cdot \mathbf{N}_1) - \mathbf{N}_1(\mathbf{N}_2 \cdot \mathbf{N}_2)) = \mathbf{N}_1 - \mathbf{N}_2(\mathbf{N}_1^\top \mathbf{N}_2).\end{aligned}$$

The addition $\mathbf{N}'_1 + \mathbf{N}'_2$ now reads as

$$\begin{aligned}\mathbf{N}_2 - \mathbf{N}_1(\mathbf{N}_1^\top \mathbf{N}_2) + \mathbf{N}_1 - \mathbf{N}_2(\mathbf{N}_1^\top \mathbf{N}_2) &= (\mathbf{N}_1 + \mathbf{N}_2) - (\mathbf{N}_1 + \mathbf{N}_2)(\mathbf{N}_1^\top \mathbf{N}_2)\\ &= (\mathbf{N}_1 + \mathbf{N}_2)(1 - (\mathbf{N}_1^\top \mathbf{N}_2)).\end{aligned}$$

To finally express the vector $\mathbf{N}'$, exploiting the assumption that $0 < \mathbf{N}_1^\top \mathbf{N}_2 < 1$,

$$\mathbf{N}' = \frac{\mathbf{N}'_1 + \mathbf{N}'_2}{\|\mathbf{N}'_1 + \mathbf{N}'_2\|} = \frac{(1 - (\mathbf{N}_1^\top \mathbf{N}_2))}{\left|(1 - (\mathbf{N}_1^\top \mathbf{N}_2))\right|} \frac{(\mathbf{N}_1 + \mathbf{N}_2)}{\|(\mathbf{N}_1 + \mathbf{N}_2)\|} = \frac{\mathbf{N}_1 + \mathbf{N}_2}{\|\mathbf{N}_1 + \mathbf{N}_2\|}.$$

□

Since $0 < \mathbf{N}_1^\top \mathbf{N}_2 < 1$ is assumed, $\exists \alpha \forall \mathbf{p}, \mathbf{q} \in \bigcup_{\imath \in I} \Omega_i : \ \|\mathcal{S}(\mathbf{p}) - \mathcal{S}(\mathbf{q})\| < \alpha$, i.e., the image of the union of the areas of interest – $\mathcal{S}\left(\bigcup_{i \in I} \Omega_i\right)$ – will be always a limited set. However, a boundary of such a set is not easily computable for general areas. A final affine transformation based on a parameter δ, cropping the range of $\mathcal{T}_{SGR}$ to $[-\delta, \delta] \times [-\delta, \delta]$, can be derived to circumvent the problem.

The final affine transformation with the parameter δ, $\mathcal{F}_{SR} : \mathbb{R}^2 \to \mathbb{R}^2$ for an image with width w and height h reads as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \ \mathcal{F}_{GSR}(\mathbf{u}) = \begin{pmatrix} \frac{w}{2\delta} & 0 \\ 0 & \frac{h}{2\delta} \end{pmatrix} \mathbf{u} + \begin{pmatrix} \frac{w}{2} \\ \frac{h}{2} \end{pmatrix}. \tag{5.14}$$

The inverse mapping then derives as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \ \mathcal{F}^{-1}_{GSR}(\mathbf{u}) = \begin{pmatrix} \frac{2\delta}{w} & 0 \\ 0 & \frac{2\delta}{h} \end{pmatrix} \mathbf{u} - \begin{pmatrix} \delta \\ \delta \end{pmatrix}. \tag{5.15}$$

5.5 Discussion

This section discusses the presented methods using examples of rectified image pairs presented in Appendix B. The focus of the discussion is the distorsion produced by the respective rectification methods.

The first sequence presented in the Appendix B is the 'Street' sequence from Figure B.1. It documents behaviour of the presented rectification methods in cases when an epipole is located near the boundary of the field of view – lateral move – and cases when an epipole is located near the center of the field of view – forward move. All scanline rectification methods perform reasonably well in the case of a lateral move, see Figures B.2(a, b), B.3(a, b), B.4(a, b), B.5(a, b), yet there is noticable area enlargement around epipoles in the case of bipolar rectification. The measure of this enlargement is directly dependent of the size of the parameter δ of the final affine transformation. However, all scanline rectification methods produce heavily distorted results in the case of a forward move, see Figures B.2(c, d), B.3(c, d), B.4(c, d), B.5(c, d). Again, in the case of bipolar rectification the areas near the epipoles are heavily enlarged at the expense of the rest of the image. On the other hand, stereographic rectification performs well for both lateral and forward move, see the rectification overlay in Figure B.6 – the images produced using stereographic rectification feature the least distorsion from the original images.

The second sequence presented in the Appendix B is the 'Office' sequence from Figure B.8. It documents behaviour of the presented rectification methods in cases when an epipole is present between the center of the field of view and its boundary. In case of images with the field of view up to 180°, there is always at most one epipole visible in the field of view. The closer the epipole is to the center of the field of view, the greater distorsion is introduced by all scanline rectification methods, see B.2(a, b, c, d), B.3(a, b, c, d), B.4(a, b, c, d). The stereographic rectification produces heavy distorsion as well, see Figures B.5(a, b, c, d). However, as expected, the areas near the center of the rectified images are mainly unaffected. On the contrary, the matching features lie closer to each other than in the original image pair, see Figures B.6(a, b) vs. B.6(c, d). The area enlargement produced by the stereographic rectification, see Figure B.6(e, f), is limited to the parts of the respective fields of view that do not overlap and is thus no problem for stereo matching.

6 3D Reconstruction

The rectification of a stereo pair is usually used as a pre-step to a dense stereo correspondence algorithm. Such an algorithm tries to determine for every pixel $\mathbf{u}_1$ in the first image a pixel $\mathbf{u}_2$ in the second image, that correspond to the same scene point as pixel $\mathbf{u}_1$. Let's denote the relation as $\mathbf{u}_1 \leftrightarrow \mathbf{u}_2$. Various methods have been proposed to solve the dense stereo correspondence problem [8].

Definition 22 *Let $P = [\mathcal{P}_1, \mathcal{P}_2]$ be a rectified stereo pair, where $D = [0, w] \times [0, h]$ is the domain of $\mathcal{P}_1, \mathcal{P}_2$. Then a map $\mathcal{D}_P : D \to \mathbb{R}^2$ is called the* disparity map *for the stereo pair P.*

Definition 23 *Let $P = [\mathcal{P}_1, \mathcal{P}_2]$ be a rectified stereo pair, where $D = [0, w] \times [0, h]$ is the domain of $\mathcal{P}_1, \mathcal{P}_2$. Then a map $\mathcal{D}_P^T :\to \mathbb{R}^2$ such that*

$$\forall \mathbf{u} \in D : \left(\mathcal{D}_P^T(\mathbf{u}) = \mathbf{v} \Longleftrightarrow \mathbf{u} \leftrightarrow (\mathbf{u} + \mathbf{v})\right),$$

is called the ground truth disparity map *for the stereo pair P.*

The result of running a dense stereo correspondence algorithm on a rectified stereo pair $P = [\mathcal{P}_1, \mathcal{P}_2]$ is usually a *disparity map*, see Definition 22, such that $\mathcal{D}_P(\mathbf{u}) = \mathbf{v}'$, if the algorithm flags, possibly not correctly, the pair $\mathbf{u}, \mathbf{u} + \mathbf{v}$ as corresponding to the same scene point. Let's denote this relation $\mathbf{u} \leftrightarrow_{\mathcal{D}_P} \mathbf{u} + \mathbf{v}$. The goal of the algorithm is to produce a disparity map as close to the *ground truth*, see Definition 23, as possible. The accuracy of the resulting disparity maps varies from method to method. In case that the pre-step rectification was a scanline rectification, the disparity map will have the following property:

$$\forall \mathbf{u} \in D : \left(\mathcal{D}_P(\mathbf{u}) = \mathbf{v} \Longrightarrow \mathbf{v} = (v_x, 0)^\top\right).$$

Let's suppose that a stereo pair $P = [\mathcal{P}_1, \mathcal{P}_2]$ was rectified using rectification method $\mathcal{G} = [\mathcal{G}_1, \mathcal{G}_2]$,

$$\begin{aligned} \mathcal{G}_1 &= \mathcal{C}_1 \circ \mathcal{A}_1 \circ \mathcal{T} \circ \mathcal{F}, \\ \mathcal{G}_2 &= \mathcal{C}_2 \circ \mathcal{A}_2 \circ \mathcal{T} \circ \mathcal{F}. \end{aligned}$$

Further, let's suppose that a dense stereo correspondence algorithm produced a disparity map $\mathcal{D}_P$ for the rectification of the stereo pair $\left[\mathcal{I}^{\varphi}_{\mathcal{P}_1}\left(\mathcal{G}_1^{-1}\right), \mathcal{I}^{\varphi}_{\mathcal{P}_2}\left(\mathcal{G}_2^{-1}\right)\right]$. The disparity map applies to the pixel position in the rectified images, however, the relation "$\leftrightarrow_{\mathcal{D}_P}$" can be directly transformed into a relation of 3D vectors as

$$\forall \mathbf{p}, \mathbf{q} \in \mathbb{S}^3 : \left(\mathbf{p} \underset{\mathcal{D}_P}{\leftrightarrow} \mathbf{q} \Longleftrightarrow \mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_1\left(\mathbf{q}\right)\right)\right) \underset{\mathcal{D}_P}{\leftrightarrow} \mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_2\left(\mathbf{q}\right)\right)\right)\right).$$

To elaborate, according to the disparity map $\mathcal{D}_P$, a vector $\mathbf{p} \in \mathbb{R}^3$ in the the Cartesian coordinate system of the first camera and a vector $\mathbf{q} \in \mathbb{R}^3$ in the Cartesian coordinate system of the second camera point the the direction of the same scene point $\mathbf{X}$, iff $\mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_1\left(\mathbf{q}\right)\right)\right) \leftrightarrow_{\mathcal{D}_P} \mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_2\left(\mathbf{q}\right)\right)\right)$.

Having vectors $\mathbf{p} = (p_1, p_2, p_3), \mathbf{q} = (q_1, q_2, q_3)$ such that $\mathbf{p} \leftrightarrow_{\mathcal{D}_P} \mathbf{q}$, the observed scene point $\mathbf{X} \in \mathbb{R}^3$ can be reconstructed up to scale using a simple linear triangulation method presented in [6]. Remind from Equation (3.2) that

$$\exists \alpha > 0 : \ \alpha \mathbf{x} = \mathtt{P}\mathbf{X}, \tag{6.1}$$

where $\mathtt{P} \in \mathbb{R}^{3\times4}$ is a projection matrix. Let $\mathtt{P}_1, \mathtt{P}_2 \in \mathbb{R}^{3\times4}$ be the respective projection matrices for cameras that acquired the image pair P, then it holds that

$$\alpha_1 \mathbf{p} = \mathtt{P}_1 \mathbf{X}, \tag{6.2}$$
$$\alpha_2 \mathbf{q} = \mathtt{P}_2 \mathbf{X}. \tag{6.3}$$

After specifying matrices $\mathtt{P}_1, \mathtt{P}_2$ as

$$\mathtt{P}_1 = \begin{pmatrix} \mathbf{r}_1^1 & \mathbf{r}_1^2 & \mathbf{r}_1^3 \end{pmatrix}^\top, \mathtt{P}_2 = \begin{pmatrix} \mathbf{r}_2^1 & \mathbf{r}_2^2 & \mathbf{r}_2^3 \end{pmatrix}^\top,$$

i.e., vectors $\mathbf{r}_i^j, i = 1, 2, j = 1, 2, 3$ are the rows of the respective matrices, Equations (6.2) and (6.3) rewrite as

$$\begin{aligned}
\alpha_1 p_1 &= \mathbf{r}_1^{1\top} \mathbf{X}, \\
\alpha_1 p_2 &= \mathbf{r}_1^{2\top} \mathbf{X}, \\
\alpha_1 p_3 &= \mathbf{r}_1^{3\top} \mathbf{X}, \\
\alpha_2 p_1 &= \mathbf{r}_2^{1\top} \mathbf{X}, \\
\alpha_2 p_2 &= \mathbf{r}_2^{2\top} \mathbf{X}, \\
\alpha_2 p_3 &= \mathbf{r}_2^{3\top} \mathbf{X}.
\end{aligned}$$

The respective scales α_1, α_2 can be eliminated by mutual division of the equation. After recombining the six equations as

$$\mathtt{A}\mathbf{X} = 0, \tag{6.4}$$

where

$$\mathtt{A} = \begin{pmatrix}
p_1 \mathbf{r}_1^{3\top} - p_3 \mathbf{r}_1^{1\top} \\
p_1 \mathbf{r}_1^{2\top} - p_2 \mathbf{r}_1^{1\top} \\
p_2 \mathbf{r}_1^{3\top} - p_3 \mathbf{r}_1^{2\top} \\
q_1 \mathbf{r}_2^{3\top} - q_3 \mathbf{r}_2^{1\top} \\
q_1 \mathbf{r}_2^{2\top} - q_2 \mathbf{r}_2^{1\top} \\
q_2 \mathbf{r}_2^{3\top} - q_3 \mathbf{r}_2^{2\top}
\end{pmatrix},$$

the Equation (6.4) can be standardly solved for $\mathbf{X}$ using SVD. In the case that $\mathbf{p} \leftrightarrow \mathbf{q}$, i.e., noiseless correspondence, $\mathrm{rank}(\mathtt{A}) = 3$.

7 Measuring the accuracy of an essential matrix

The standard method of derivation of an essential matrix for a stereo image pair is based on a computation of putative correspondences followed by RANSAC-based estimation of the essential matrix [4, 6]. The goal of this chapter is to devise a method to judge the accuracy of an already computed essential matrix based on a rectification method and a partially defined disparity map.

7.1 Accuracy of an essential matrix

Let's suppose that a stereo pair $P = [\mathcal{P}_1, \mathcal{P}_2]$ was rectified as a pair of images $R = \left[\mathcal{I}^{\varphi}_{\mathcal{P}_1}\left(\mathcal{G}_1^{-1}\right), \mathcal{I}^{\varphi}_{\mathcal{P}_2}\left(\mathcal{G}_2^{-1}\right)\right]$ using an essential matrix $\mathtt{F}$ and using rectification method $\mathcal{G} = [\mathcal{G}_1, \mathcal{G}_2]$,

$$\begin{aligned} \mathcal{G}_1 &= \mathcal{C}_1 \circ \mathcal{A}_1 \circ \mathcal{T} \circ \mathcal{F}, \\ \mathcal{G}_2 &= \mathcal{C}_2 \circ \mathcal{A}_2 \circ \mathcal{T} \circ \mathcal{F}. \end{aligned}$$

After the epipolar alignment $[\mathcal{A}_1, \mathcal{A}_2]$ the 2D epipolar curves of the original stereo pair are aligned with the meridians of the unit sphere and after the characteristic rectification transformation $\mathcal{T}$ followed by the final affine transformation $\mathcal{F}$ are the epipolar curves again transformed onto 2D curves, in the case of scanline rectification method onto lines parallel to $y = 0$, in case of stereographic projection onto circles. Let $\varepsilon_{\mathbf{v}}$ be a 2D epipolar curve after rectification containing a point $\mathbf{v} \in \mathbb{R}^2$,

$$\varepsilon_{\mathbf{v}} = \left\{\mathbf{v}' : \exists \mathbf{n} \in \mathbb{S}^3 \left(((0,0,1)\mathbf{n} = 0) \ \& \ \left(\mathbf{n}^\top \mathbf{v}' = 0\right) \ \& \ \left(\mathbf{n}^\top \mathcal{T}^{-1}(\mathcal{F}^{-1}(\mathbf{v})) = 0\right)\right)\right\}.$$

then a set E containing all epipolar curves reads as

$$E = \left\{\varepsilon_{\mathbf{v}} : \mathbf{v} \in \mathbb{R}^2\right\}.$$

As a consequence of the Equation (3.7) and the construction of the rectification method $\mathcal{G}$, following holds for vectors $\mathbf{p}, \mathbf{q} \in \mathbb{R}^3$

$$\mathbf{q}^\top \mathtt{F} \mathbf{p} = 0 \implies \exists \varepsilon \in E : \mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_1\left(\mathbf{p}\right)\right)\right) \in \varepsilon \ \& \ \mathcal{F}\left(\mathcal{T}\left(\mathcal{A}_2\left(\mathbf{p}\right)\right)\right) \in \varepsilon.$$

Suppose there is a method capable of producing, possibly not inaccurate, disparity map $\mathcal{D}_P : D \to \mathbb{R}^2$, see Definition 22, derived from the rectified image pair R. Let's call such a method *disparity method*. Then the accuracy of the essential matrix $\mathtt{F}$ with respect to the disparity map $\mathcal{D}_p$, can be quantified by the amount of diversion of the vectors defined by the disparity matrix from the epipolar lines E defined by $\mathtt{F}$. Such a as measure leads to the Definition 24.

Definition 24 *Let P be an stereo image pair, $\mathcal{G} = [\mathcal{C}_1 \circ \mathcal{A}_1 \circ \mathcal{T} \circ \mathcal{F}, \mathcal{C}_2 \circ \mathcal{A}_2 \circ \mathcal{T} \circ \mathcal{F}]$ a rectification method, $R = [\mathcal{R}_1, \mathcal{R}_2]$ the rectification of P using $\mathcal{G}$ and and an essential matrix $\mathtt{F}$, such that D' is the range of $\mathcal{R}_1, \mathcal{R}_2$. Let $\mathcal{D}_P : D \subset D' \to \mathbb{R}^2$ be a disparity map based on R, $|D|$ the cardinality of D. Then*

$$\mathcal{E}_{\mathcal{D}_P}(\mathtt{F}) = \frac{\sum_{\mathbf{v} \in D} \Delta(\varepsilon_{\mathbf{v}}, \mathbf{v} + \mathcal{D}_P(\mathbf{v}))}{|D|}, \tag{7.1}$$

is the accuracy of the essential matrix $\mathtt{F}$ with respect to the disparity map $\mathcal{D}_P$.

Since the function Δ is always non-negative, the measure $\mathcal{E}_{\mathcal{D}_P}$ is also always non-negative. For a perfectly accurate disparity map $\mathcal{E}_{\mathcal{D}_P}\left(\mathtt{F}^T\right) = 0$ for perfectly accurate essential matrix $\mathtt{F}^T$.

Suppose an essential matrix $\mathtt{F}' \neq \mathtt{F}$ aiming to describe the same epipolar geometry of the stereo pair P. Let R' be the rectification of P based on the rectification method $\mathcal{G}$ and the essential matrix $\mathtt{F}'$. Let $\mathcal{D}'_P : D' \to \mathbb{R}^2$ be a disparity map derived by the same method as $\mathcal{D}_P$ and based on the rectification R'. Then if $\mathcal{E}_{\mathcal{D}'_P}(\mathtt{F}') < \mathcal{E}_{\mathcal{D}_P}(\mathtt{F})$, $\mathtt{F}'$ can be judged to be better describing the epipolar geometry of the image pair P as $\mathtt{F}$, with respect to the disparity maps $\mathcal{D}_P, \mathcal{D}'_P$.

It is obvious that such a method for judging and comparison of quality of essential matrices stands and falls with – so far undisclosed – method of producing sparse but fairly accurate disparity maps of rectified image pairs. Obviously, dense stereo correspondence algorithms discussed in Chapter 6 are useless as a disparity method, since such algorithms are designed to search for disparities only exactly on the epipolar lines, i.e., $\mathcal{E}_{\mathcal{D}_P}(\mathtt{F}) = 0$ for every $\mathcal{D}_P$ produced by such a algorithm. An usefull disparity method has to be based on another assumptions. Method [9] is a candidate for a reasonable disparity method. The disparity method as well as the methodology of measuring the accuracy of an essential matrix remain the subject of futher study.

7.2 Δ based on stereographic rectification

The stereographic rectification was designed specifically as a rectification method suitable for measuring the accuracy of an essential matrix. It introduces the least distortion to the original stereo pair from the rectification methods discussed in this work and does not suffer from the distortion of the area near the epipoles and thus is suitable as a pre-procesing step of a disparity method. Further, the epipolar curves in the rectified stereo pair are easily parametrized as circles or lines. In this section, the explicit formula for the function Δ from the Equation (24) is derived for the case when the stereographic rectification is the underlying rectification method.

Let $\mathcal{G} = [\mathcal{C}_1 \circ \mathcal{A}_1 \circ \mathcal{T} \circ \mathcal{F}, \mathcal{C}_2 \circ \mathcal{A}_2 \circ \mathcal{T} \circ \mathcal{F}]$ be the stereographic rectification method based on an essential matrix $\mathtt{F}$. In the case that none of the epipoles is projected onto infinity, i.e., $\mathbf{e} = (e_1, e_2) = \mathcal{F}\left(\mathcal{T}\left((0,0,1)^\top\right)\right) \in \mathbb{R}^2, \dot{\mathbf{e}} = (\dot{e}_1, \dot{e}_2) = \mathcal{F}\left(\mathcal{T}\left((0,0,-1)^\top\right)\right) \in \mathbb{R}^2$, and a point $\mathbf{v} \in \mathbb{R}^2$, not collinear with $\mathbf{e}$ and $\dot{\mathbf{e}}$, the epipolar curve containing $\mathbf{v}$ is a circle and can be parametrized using the formula for a circle defined by three points [11] as

$$\varepsilon_{\mathbf{v}} = \left\{\mathbf{u} :\ \mathbf{u} = (x, y)^\top \in \mathbb{R}^2 \ \&\ |\mathtt{A}| = 0\right\},$$

where

$$\mathtt{A} = \begin{pmatrix} x^2+y^2 & x & y & 1 \\ v_1^2+v_2^2 & v_1 & v_2 & 1 \\ e_1^2+e_2^2 & e_1 & e_2 & 1 \\ \dot{e}_1^2+\dot{e}_2^2 & \dot{e}_1 & \dot{e}_2 & 1 \end{pmatrix}.$$

The center and the radius of the epipolar circle $\varepsilon_{\mathbf{v}}$ then reads as

$$\mathbf{c}_{\varepsilon_{\mathbf{v}}} = \left(-\frac{\beta}{2\alpha}, -\frac{\gamma}{2\alpha}\right)^{\top}, r_{\varepsilon_{\mathbf{v}}} = \sqrt{\frac{\beta^2+\gamma^2}{4\alpha^2} + \frac{\delta}{\alpha}},$$

where

$$\begin{aligned} \alpha &= \begin{vmatrix} v_1 & v_2 & 1 \\ e_1 & e_2 & 1 \\ \dot{e}_1 & \dot{e}_2 & 1 \end{vmatrix}, \\ \beta &= -\begin{vmatrix} v_1^2+v_2^2 & v_2 & 1 \\ e_1^2+e_2^2 & e_2 & 1 \\ \dot{e}_1^2+\dot{e}_2^2 & \dot{e}_2 & 1 \end{vmatrix}, \\ \gamma &= \begin{vmatrix} v_1^2+v_2^2 & v_1 & 1 \\ e_1^2+e_2^2 & e_1 & 1 \\ \dot{e}_1^2+\dot{e}_2^2 & \dot{e}_1 & 1 \end{vmatrix}, \\ \delta &= -\begin{vmatrix} v_1^2+v_2^2 & v_1 & v_2 \\ e_1^2+e_2^2 & e_1 & e_2 \\ \dot{e}_1^2+\dot{e}_2^2 & \dot{e}_1 & \dot{e}_2 \end{vmatrix}. \end{aligned}$$

The function Δ for quantifying the distance of the vector $\mathbf{v}+\mathcal{D}_P(\mathbf{v})$ and the epipolar circle $\varepsilon_{\mathbf{v}}$ reads as

$$\Delta(\varepsilon_{\mathbf{v}}, \mathbf{v}+\mathcal{D}_P(\mathbf{v})) = \left| \|\mathbf{c}_{\varepsilon_{\mathbf{v}}} - (\mathbf{v}+\mathcal{D}_P(\mathbf{v}))\| - r_{\varepsilon_{\mathbf{v}}} \right|.$$

In the case that $\mathbf{v}$ is collinear with $\mathbf{e}$ and $\dot{\mathbf{e}}$, or, without loss of generality, $\dot{\mathbf{e}} = \infty$, the epipolar line containing the point $\mathbf{v}$ is a straight line, parametrized as

$$\varepsilon_{\mathbf{v}} = \left\{ \mathbf{u} : \ \mathbf{u} \in \mathbb{R}^2 \ \& \ (\exists \alpha \in \mathbb{R} : \ \mathbf{e} + \alpha(\mathbf{e}-\mathbf{v}) = \mathbf{u}) \right\}.$$

The function Δ for quantifying the distance of the point $\mathbf{v}+\mathcal{D}_P(\mathbf{v})$ and the epipolar line $\varepsilon_{\mathbf{v}}$ reads as

$$\begin{aligned} \Delta(\varepsilon_{\mathbf{v}}, \mathbf{v}+\mathcal{D}_P(\mathbf{v})) &= \left| (\mathbf{v} - (\mathbf{v}+\mathcal{D}_P(\mathbf{v}))) - \left((\mathbf{v} - (\mathbf{v}+\mathcal{D}_P(\mathbf{v}))) \cdot \frac{(\mathbf{e}-\mathbf{v})}{\|\mathbf{e}-\mathbf{v}\|} \right) \frac{(\mathbf{e}-\mathbf{v})}{\|\mathbf{e}-\mathbf{v}\|} \right|, \\ &= \left| \frac{(\mathbf{e}-\mathbf{v})^{\top} \mathcal{D}_P(\mathbf{v})}{\|\mathbf{e}-\mathbf{v}\|^2} (\mathbf{e}-\mathbf{v}) - \mathcal{D}_P(\mathbf{v}) \right|. \end{aligned}$$

8 Conclusion

The thesis presented rectification methods of an omnidirectional stereo pair. Several methods were developed and a novel rectification method called stereographic rectification was introduced. A MATLAB toolbox for rectification of a stereo pair using the methods developed in this work was implemented as a part of this thesis. The appendices of this work include the documentation of the toolbox as well as examples of the results obtained using the toolbox.

In the second part of the thesis we presented a method of 3D reconstruction of a stereo pair rectified using an arbitrary of the adduced methods based on linear triangulation method and an externally supplied disparity map. Further, a method of judging the accuracy of an essential matrix based on the stereographic rectification was devised.

Bibliography

[1] Zafer Arican and Pascal Frossard. Dense disparity estimation from omnidirectional images. In *2007 IEEE International Conference on Advanced Video and Signal based Surveillance*, 2007.

[2] C. Geyer and K. Daniilidis. Conformal rectification of an omnidirectional stereo pair. June 2003.

[3] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, Upper Saddle River, NJ, 2nd edition, 2002.

[4] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University, Cambridge, 2nd edition, 2003.

[5] Michael Henle. *Modern Geometries : Non-Euclidean, Projective, and Discrete*. Prentice Hall, Upper Saddle River, New Jersey, USA, 2nd edition, 2001.

[6] Branislav Micusik and Tomas Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7):1135–1149, 2006.

[7] Marc Pollefeys and Luc Van Gool. A simple and efficient rectification method for general motion. 2001.

[8] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. 2002.

[9] A. Shekhovtsov, I. Kovtun, and V. Hlavac. Efficient MRF deformation model for non-rigid image matching. In *IEEE CVPR*, 2007.

[10] Hlavac V. Soboda T., Pajdla T. Epipolar geometry for panoramic cameras. 1998.

[11] Eric W Weisstein. Circle. from mathworld – a wolfram web resource, 2007.

[12] Wikipedia. Bipolar coordinates — wikipedia, the free encyclopedia, 2007.

[13] Wikipedia. Spherical coordinate system — wikipedia, the free encyclopedia, 2007.

[14] Jiří Žára, Bedřich Beneš, Jiří Sochor, and Petr Felkel. *Moderní počítačová grafika*. Computer Press, Brno, 2004.

A OmniRect toolbox

OmniRect toolbox is a direct MATLAB implementation of the rectification methods presented in this work. The implementation is for speed optimization purposes divided between two parts written MATLAB and C programming languages respectively. This appendix describes version 1.0.0 of the toolbox. The toolbox is known to work with MATLAB 7.5 and `gcc` 4.1.2.

A.1 OmniRect distribution

The toolbox distribution can be found as archive `omnirect-1.0.0.tar.bz2` on the CD-ROM. Following listing shows the content of the distribution:

```
omnirect.m
orect.c
demo.m
data/
```

Following list explains function of the individual files:

- `omnirect.m` – Implementation of the MATLAB function `omnirect`. The function `omnirect` only prepares data for the function `orec`.
- `orect.c` – Implementation of the C function `orect`. The function `orect` does the actual computing.
- `demo.m` – The `demo` script demonstrates the usage of the toolbox on several examples.
- `data` – The directory `data` contains files necessary for the `demo` demonstration script.

A.2 Usage

The only entry point to the toolbox is the function `omnirect`. The input of the function is the image pair and various parameters, output is the rectified image pair. The following shows and explains how to call the MATLAB function `omnirect`:

```
[rim1, rim2] = omnirect(im1, im2, calib1, calib2, F, ...
                                                rectfunc, affine, intfunc);
```

- `im1` – The first image from the stereo pair. The format of parameter `im1` is a $width \times height \times 3$ matrix of type `uint8`, i.e., RGB values in the range $[0, 255]$. The type is compatible with `imread` function from the MATLAB's Image Processing Toolbox.
- `im2` – The second image from the stereo pair. The fomat of parameter `im2` is identical to parameter `im1`.
- `calib1` – An instance of the `calib` structure for the first image, see Section A.2.1.
- `calib2` – An instance of the `calib` structure for the second image, see Section A.2.1.
- `F` – Essential matrix describing the epipolar geometry of `im1` and `im2`, see Section 3.3.
- `rectfunc` – Selects the desired rectification method, see Section A.2.2.
- `affine` – Parameters of final affine transformation of the rectification method selected by the parameter `rectfunc`, see Section A.2.2.
- `intfunc` – Selects an interpolation method to be used, see Section A.2.3.
- `rim1`, `rim2` – The resulting rectified image pair.

A.2.1 Calibration

Two calibration transformations are supported: one-parametric linear model of a fisheye lens and two-parametric non-linear model of a fisheye lens. The calibration transformation of the one-parametric model is defined as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{C}_a = \begin{pmatrix} \mathbf{u} \\ \frac{\|\mathbf{u}\|}{\tan(a\,\|\mathbf{u}\|)} \end{pmatrix}. \tag{A.1}$$

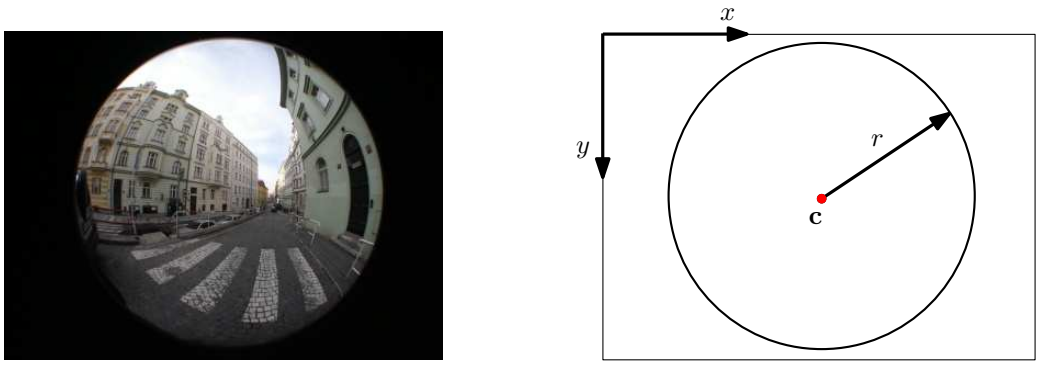
The transformation of the two-parametric non-linear model is defined as

$$\forall \mathbf{u} \in \mathbb{R}^2 : \mathcal{C}_{a,b} = \begin{pmatrix} \mathbf{u} \\ \frac{\|\mathbf{u}\|}{\tan\left(\frac{1}{b}\arcsin\left(\frac{b\|\mathbf{u}\|}{a}\right)\right)} \end{pmatrix}. \tag{A.2}$$

The properties of both models are discussed in detail in [6].

The `calib` structure provides necessary data for computation of the calibration transformation. However, not a general affine transformation as proposed in Equation (3.5) is implemented, only scaling and translation is supported at the moment, see Figure A.2.1. The following lists members of the `calib` structure.

- `cam` – Selects the calibration transformation, set to '`fisheye1p`' for one-parametric model, set to '`fisheye2p`' for the two-parametric non-linear model.
- `a` – Parameter a of the one-parametric model, see Equation (A.1), and the two-parametric model, see Equation (A.2), respectively.



(a) (b)

Figure A.1: The meaning of the `calib.c` and `calib.r` parameters.

- `b` – Parameter b of the two-parametric model, see Equation (A.2). Undefined for `calib.cam` $== fisheye1p$.
- `c` – 2×1 vector denoting the center of the field of view in pixels, see Figure A.2.1.
- `r` – Radius of the circle bounding the field of view in pixels, see Figure A.2.1.
- `e` – Parameter `e` denotes the first epipole of the epipole pair. Let

  $$(\mathbf{u}_1, \mathbf{u}_2, \mathbf{e}_2)\,\mathrm{diag}(1, 1, 0)(\mathbf{v}_1, \mathbf{v}_2, \mathbf{e}_1)^\top$$

  be the SVD decomposition of an essential matrix `F` passed to the function `omnirect`, as computed by standard MATLAB function `svd`, $i = 1, 2$. Then `calib`i`.e` $== 0$ selects $\mathbf{e}_i$ as the first epipole, i.e., $\mathbf{e}_{i,1}$, `calib`i`.e` $== 1$ selects $-\mathbf{e}_i$ as the first epipole for the respective image.

Note that since two instances of `calib` structures are passed to the `omnirect` function, it is possible to use different calibration transformations for one image pair.

A.2.2 Rectification

The rectification method is selected by the parameter `rectfunc`, which can assume one the following values:

- '`sr`' – Spherical rectification, see Section 5.1.
- '`ssr`' – Swapped-spherical rectification, see Section 5.2.
- '`br`' – Bipolar rectification, see Section 5.3.
- '`sgr`' – Stereographic rectification, see Section 5.4.

Although all of the final affine transformations from Chapter 5 correctly map the reasonable range of the respective characteristic transformations to the dimensions of final images, *OmniRect* toolbox introduces a modification to the final transformations

to produce visually more appealing results without changing the desirable properties. If

$$\mathcal{F}(\mathbf{u}) = \mathtt{A}\mathbf{u} + \mathbf{t}$$

is the final affine transformation for transformation into domain $[0, w] \times [0, h] \in \mathbb{R}^2$, when

$$\mathcal{F}'(\mathbf{u}) = \mathbf{v} + \mathtt{R}\left(\mathtt{A}\mathbf{u} + \mathbf{t} + \mathbf{p}\right), \tag{A.3}$$

where $\mathbf{o} = (o_1, o_2)^\top \in \{0, 1\} \times \{0, 1\}$, $\mathbf{v} = (o_1 w, o_2 h)^\top$, $\mathbf{p} \in \mathbb{R}^2$ and

$$\mathtt{R} = \mathtt{R}^{-1} = \begin{pmatrix} 1 - 2o_1 & 0 \\ 0 & 1 - 2o_2 \end{pmatrix}.$$

is the modified final affine transformation used in *OmniRect*. The vectors $\mathbf{o}, \mathbf{p}$ are user supplied. The vector $\mathbf{o} = (o_1, o_2)^\top$ "flips" the images along the respective axis if set to 1. The vector $\mathbf{p}$ is an arbitrary translation vector. The vector $\mathbf{p}$ is particular useful for spherical and swapped spherical rectifications since the characteristic functions of these rectification methods are periodical functions. The inverse transformation $\mathcal{F}'^{-1}$ reads as

$$\mathcal{F}'^{-1}(\mathbf{u}) = \mathtt{A}^{-1}\left(\mathtt{R}\left(\mathbf{u} - \mathbf{v}\right) - \mathbf{t} - \mathbf{p}\right).$$

The parameters specific to the selected rectification method are passed through the `affine` structure. The following lists members of the `affine` structure.

- `width` – Width of the resulting images.
- `height` – Height of the resulting images.
- `p` – see Equation (A.3).
- `o` – see Equation (A.3).
- `param` – Parameter referred to in the final affine transformations of the respective rectification methods as δ. Undefined in case the δ parameter is not featured in the final affine transformation.

A.2.3 Interpolation

Two interpolation methods are supported: nearest neighbor interpolation and bilinear interpolation. The following lists allowed values of the parameter `intfunc`:

- '`nn`' – Nearest neighbor interpolation, see Equation (2.1).
- '`bl`' – Bilinear interpolation, see Equation (2.2).

B Comparison of rectification methods

This appendix presents results obtained using the rectification methods described in this work through the *OmniRect* MATLAB toolbox. The images are direct output of the `demo` script included in the *OmniRect* toolbox.

B.1 'Street' sequence

'Street' sequence was acquired using Canon EOS 1Ds with Sigma 8mm-f4-EX fish-eye lens. Rectification of two image pairs

$$P_L = \left[\mathcal{P}_L^1, \mathcal{P}_L^2\right], P_F = \left[\mathcal{P}_F^1, \mathcal{P}_F^2\right]$$

is performed using all of the presented methods, see Figure B.1. The image pair P_L resulted from a lateral move of the camera between the shots. Essential matrix used for the image pair P_L

$$\mathtt{F}_L = \begin{pmatrix} 0.0002 & 0.0350 & -0.0090 \\ -0.0502 & 0.0622 & 0.9968 \\ 0.0052 & -0.9974 & 0.0623 \end{pmatrix}$$

Calibration structure `calib` for $\mathcal{P}_L^1$

$$\begin{aligned} \mathtt{a} &= 1.5708 \\ \mathtt{b} &= 0 \\ \mathtt{c} &= (1050, 680)^\top \\ \mathtt{r} &= 641 \\ \mathtt{e} &= 0 \\ \mathtt{cam} &= fisheye1p \end{aligned}$$

Calibration structure `calib` for $\mathcal{P}_L^2$

$$\begin{aligned} \mathtt{a} &= 1.5708 \\ \mathtt{b} &= 0 \\ \mathtt{c} &= (1050, 680)^\top \\ \mathtt{r} &= 641 \\ \mathtt{e} &= 1 \\ \mathtt{cam} &= fisheye1p \end{aligned}$$

The image pair P_F resulted from a forward move of the camera between the shots. Essential matrix used for the image pair P_F

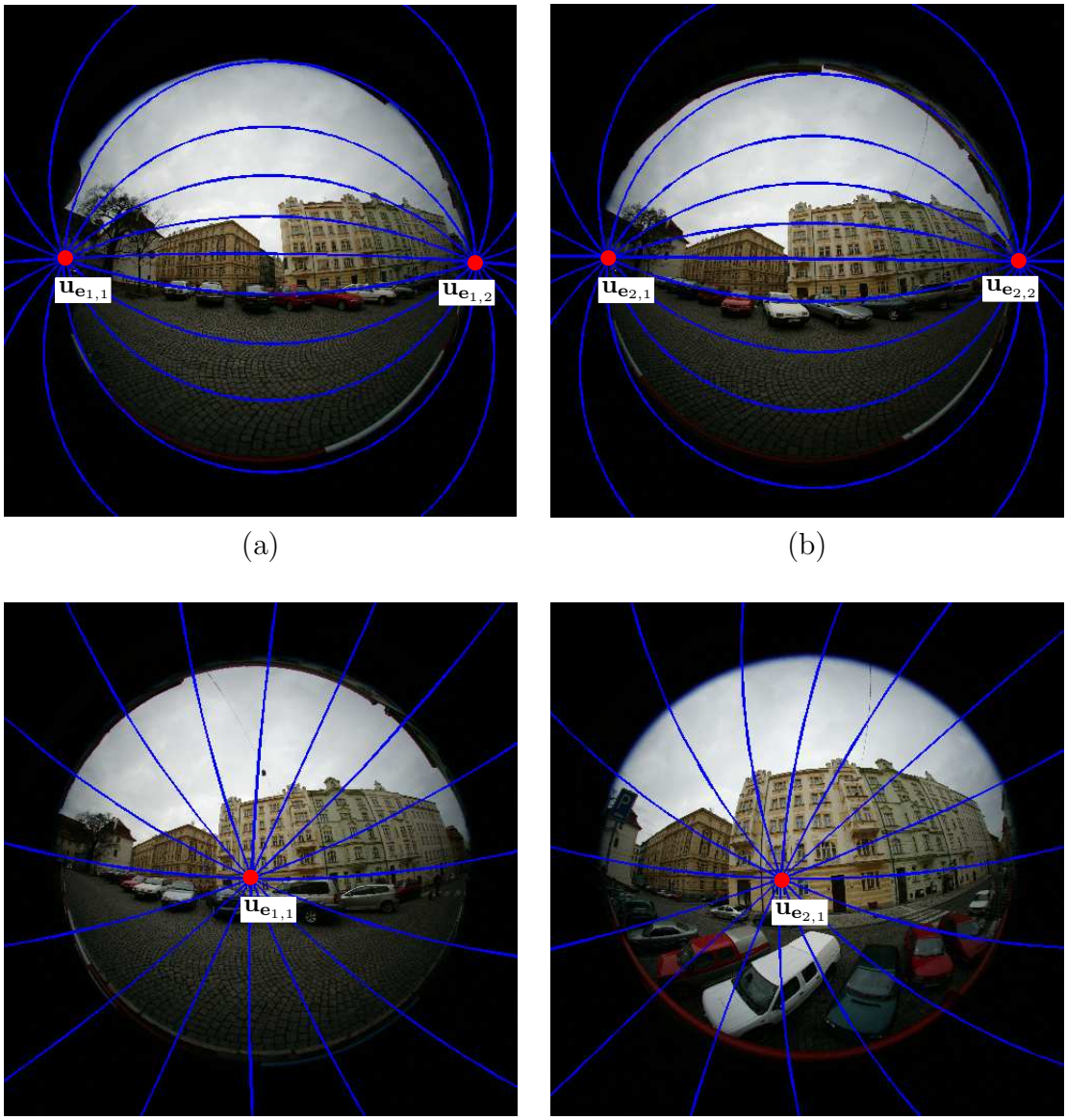
$$\mathtt{F}_F = \begin{pmatrix} -0.0212 & 0.9516 & 0.1913 \\ -0.9782 & 0.0093 & 0.0832 \\ -0.1900 & -0.2375 & -0.0315 \end{pmatrix}$$

Calibration structure `calib` for $\mathcal{P}_F^1$

$$\begin{aligned} \mathtt{a} &= 1.5708 \\ \mathtt{b} &= 0 \\ \mathtt{c} &= (1050, 680)^\top \\ \mathtt{r} &= 641 \\ \mathtt{e} &= 0 \\ \mathtt{cam} &= fisheye1p \end{aligned}$$

Calibration structure `calib` for $\mathcal{P}_F^2$

$$\begin{aligned} \mathtt{a} &= 1.5708 \\ \mathtt{b} &= 0 \\ \mathtt{c} &= (1050, 680)^\top \\ \mathtt{r} &= 641 \\ \mathtt{e} &= 1 \\ \mathtt{cam} &= fisheye1p \end{aligned}$$

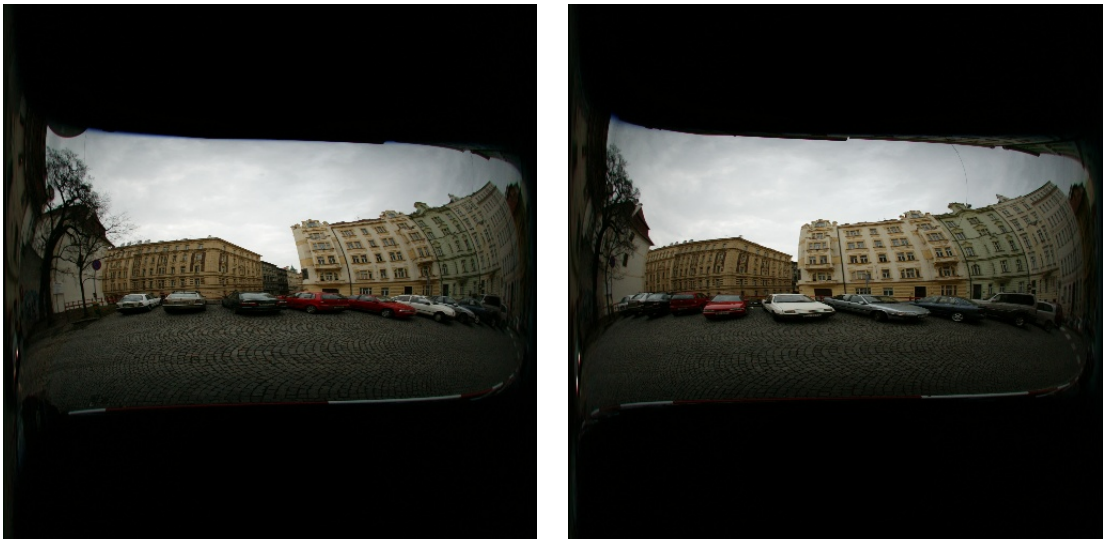


(a) (b)

(c) (d)

Figure B.1: (a, b) The image pair $P_L = [\mathcal{P}_L^1, \mathcal{P}_L^2]$ acquired by a lateral motion of the camera. (b, c) The image pair $P_F = [\mathcal{P}_F^1, \mathcal{P}_F^2]$ acquired by a forward motion of the camera.

B.1.1 Spherical rectification



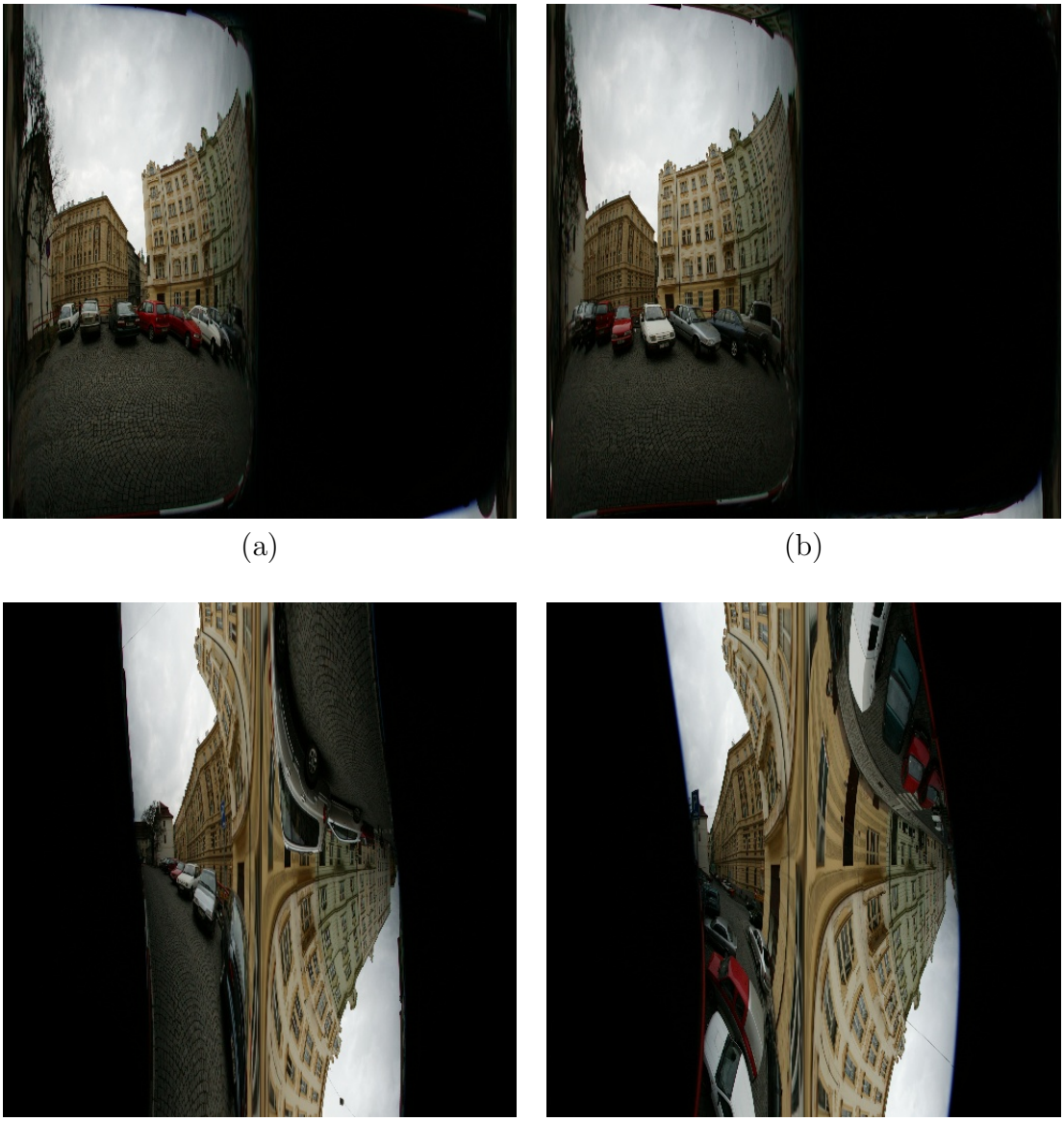
(a) (b)

(c) (d)

Figure B.2: (a, b) Spherical rectification of the image pair P_L. (b, c) Spherical rectification of the image pair P_F.

The `affine` structure for P_L	The `affine` structure for P_F
`width` = 500 `height` = 500 `o` = $(0,1)^\top$ `p` = $(0,0)^\top$ `param` = 0	`width` = 500 `height` = 500 `o` = $(1,0)^\top$ `p` = $(0,-\texttt{height}/2)^\top$ `param` = 0

B.1.2 Swapped spherical rectification

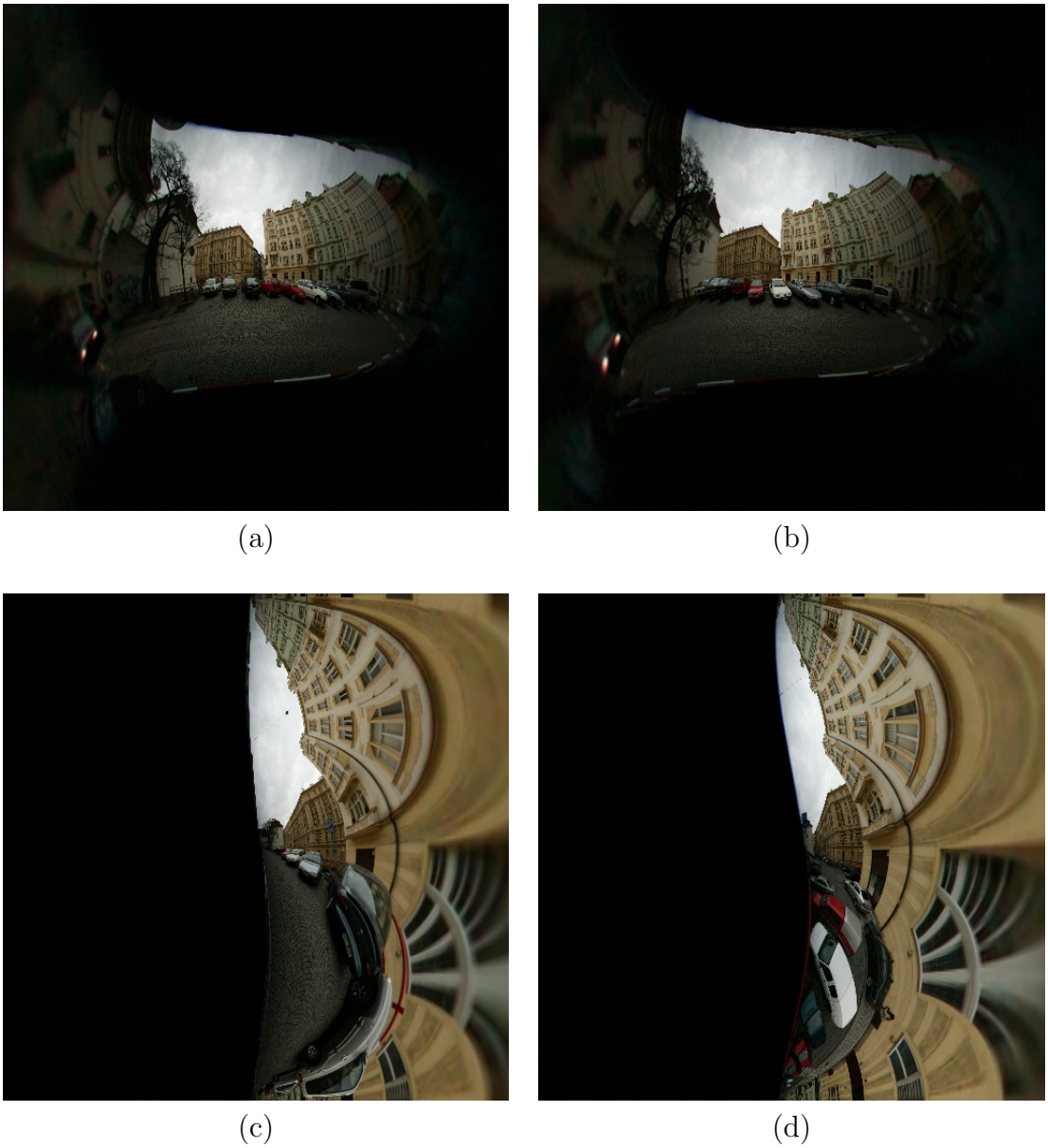


(a) (b) (c) (d)

Figure B.3: (a, b) Swapped spherical rectification of the image pair P_L. (b, c) Swapped spherical rectification of the image pair P_F.

The `affine` structure for P_L	The `affine` structure for P_F
$\texttt{width} = 500$ $\texttt{height} = 500$ $\texttt{o} = (1,1)^\top$ $\texttt{p} = (0, -\texttt{height}/2)^\top$ $\texttt{param} = 0$	$\texttt{width} = 500$ $\texttt{height} = 500$ $\texttt{o} = (1,0)^\top$ $\texttt{p} = (-\texttt{width}/2, -\texttt{height}/2)^\top$ $\texttt{param} = 0$

B.1.3 Bipolar rectification



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

Figure B.4: (a, b) Bipolar rectification of the image pair P_L. (b, c) Bipolar rectification of the image pair P_F.

The `affine` structure for P_L	The `affine` structure for P_F
`width` = 500 `height` = 500 `o` = $(1,1)^\top$ `p` = $(0,-\texttt{height}/2)^\top$ `param` = 5	`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 0

B.1.4 Stereographic rectification

(a)

(b)

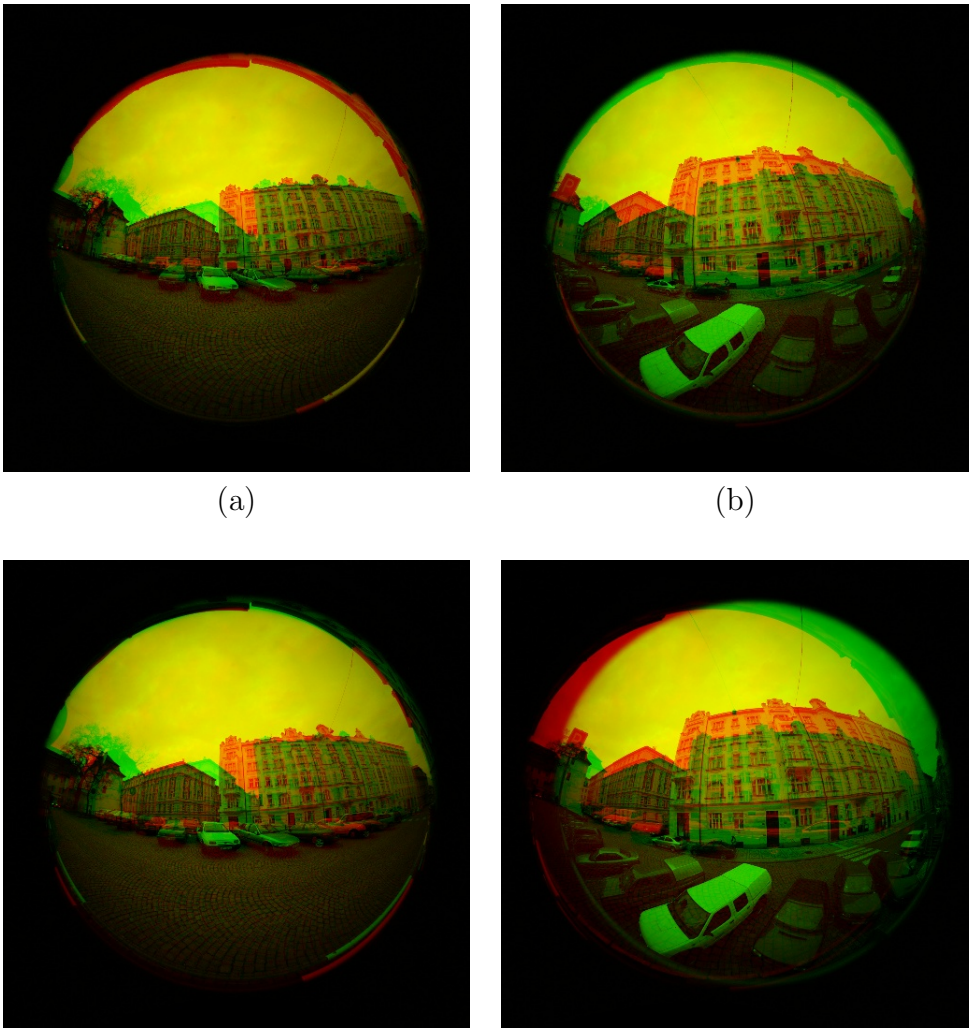
(c)

(d)

Figure B.5: (a, b) Stereographic rectification of the image pair P_L. (b, c) Stereographic rectification of the image pair P_F.

The `affine` structure for P_L	The `affine` structure for P_F
`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 1.2	`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 1.2

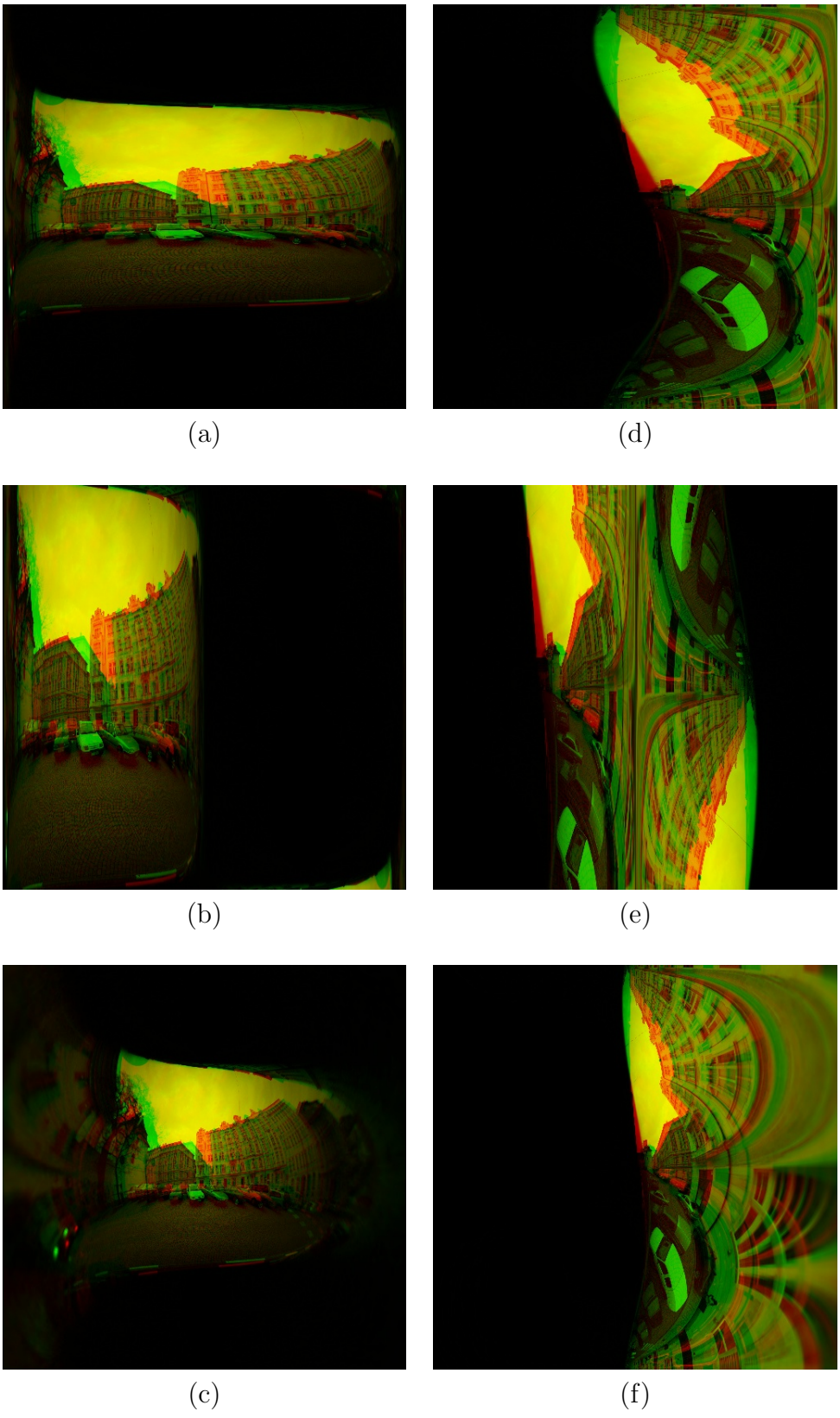
B.1.5 Rectification overlay



(a) (b)

(c) (d)

Figure B.6: (a) Overlay of P_L,$\mathcal{P}_L^1$ in the red channel, $\mathcal{P}_L^2$ in the green channel. (b) Overlay of P_F,$\mathcal{P}_F^1$ in the red channel, $\mathcal{P}_F^2$ in the green channel. (c) Overlay of stereographic rectification of P_L. (d) Overlay of stereographic rectification of P_L.



(a) (d)

(b) (e)

(c) (f)

Figure B.7: (a, b, c) Overlay of spherical, swapped spherical and bipolar rectification of P_L respectively. (d, e, f) Overlay of spherical, swapped spherical and bipolar rectification of P_F respectively.

B.2 'Office' sequence

'Office' sequence was acquired using KYOCERA Finecam M410R with custom mounted Nikon FC-E8 fish-eye lens. Rectification of two image pairs

$$P_{O_1} = \left[\mathcal{P}^1_{O_1}, \mathcal{P}^2_{O_1}\right], P_{O_2} = \left[\mathcal{P}^1_{O_2}, \mathcal{P}^2_{O_2}\right],$$

is performed using all of the presented methods, see Figure B.1. The image pair P_{O_1} resulted from moving the camera along a 1/4 of a circle with diameter of about 2m, the view direction facing the center of the circle. Essential matrix used for the image pair P_{O_2}

$$\mathtt{F}_{O_1} = \begin{pmatrix} 0.0258 & 0.6456 & 0.0007 \\ 0.6904 & -0.0152 & 0.7232 \\ -0.0186 & -0.7630 & 0.0110 \end{pmatrix}$$

Calibration structure `calib` for both $P^1_{O_1}$ and $P^2_{O_1}$

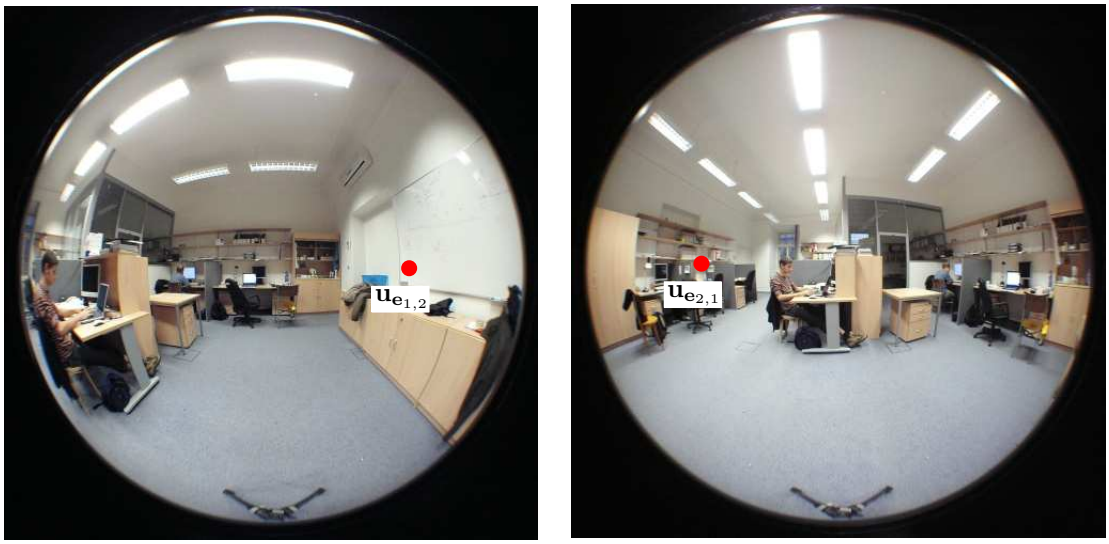
$$\begin{aligned} \mathtt{a} &= 1.5693 \\ \mathtt{b} &= -0.0446 \\ \mathtt{c} &= (1166, 846)^\top \\ \mathtt{r} &= 794 \\ \mathtt{e} &= 1 \\ \mathtt{cam} &= fisheye2p \end{aligned}$$

The image pair P_{O_1} resulted from moving the camera along a 1/8 of a circle with diameter of about 2m, the view direction facing the center of the circle. Essential matrix used for used for the image pair P_{O_2}

$$\mathtt{F}_{O_2} = \begin{pmatrix} -0.0001 & 0.3735 & -0.0034 \\ 0.3575 & -0.0063 & 0.9339 \\ -0.0055 & -0.9276 & -0.0065 \end{pmatrix}$$

Calibration structure `calib` for both $P^1_{O_2}$ and $P^2_{O_2}$

$$\begin{aligned} \mathtt{a} &= 1.5693 \\ \mathtt{b} &= -0.0446 \\ \mathtt{c} &= (1166, 846)^\top \\ \mathtt{r} &= 794 \\ \mathtt{e} &= 1 \\ \mathtt{cam} &= fisheye2p \end{aligned}$$

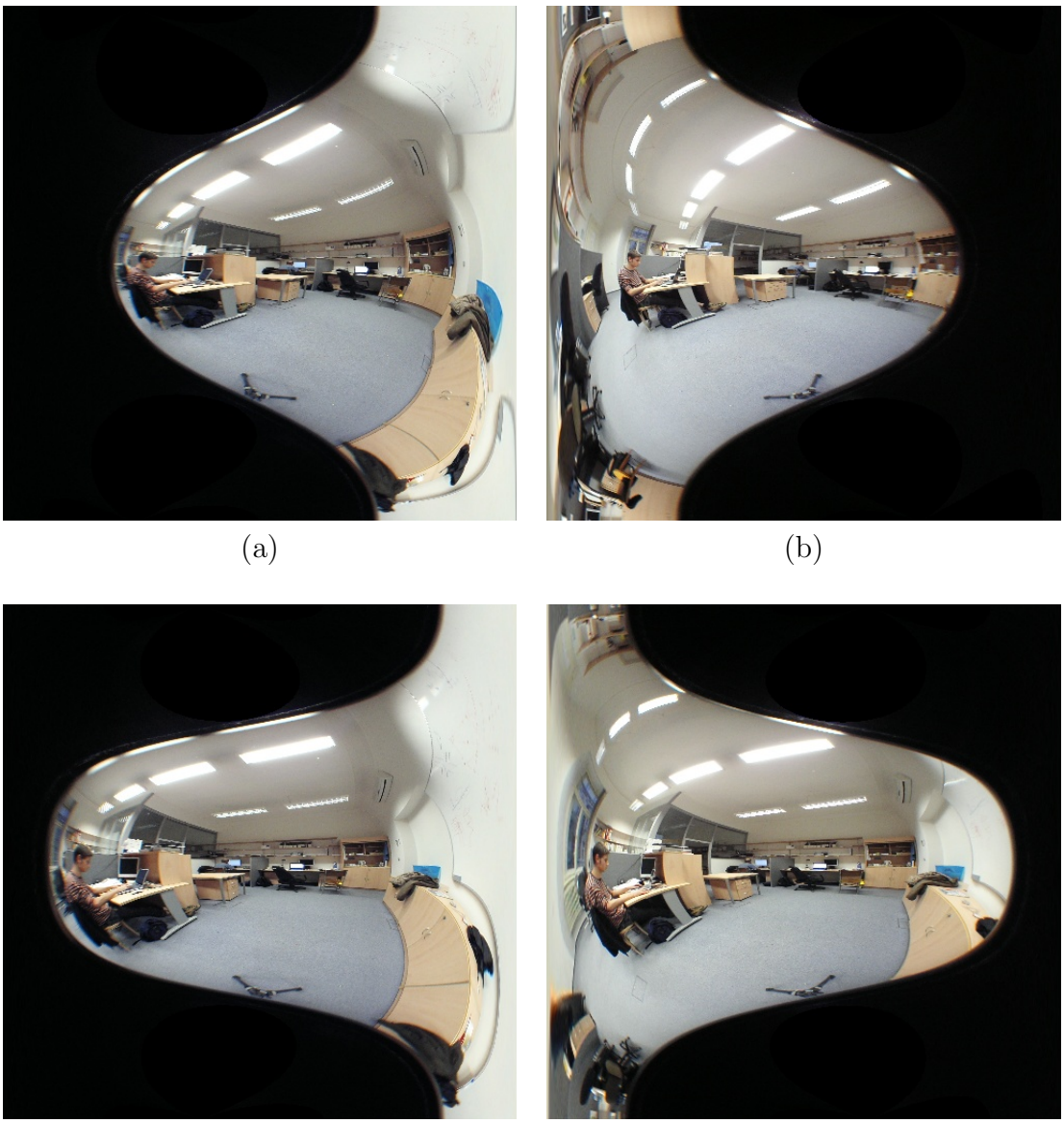


(a) (b)

(c) (d)

Figure B.8: (a, b) The image pair $P_{O_1} = \left[\mathcal{P}^1_{O_1}, \mathcal{P}^2_{O_1}\right]$. (b, c) The image pair $P_{O_2} = \left[\mathcal{P}^1_{O_2}, \mathcal{P}^2_{O_2}\right]$. Note that $\mathcal{P}^1_{O_1} = \mathcal{P}^1_{O_2}$.

B.2.1 Spherical rectification



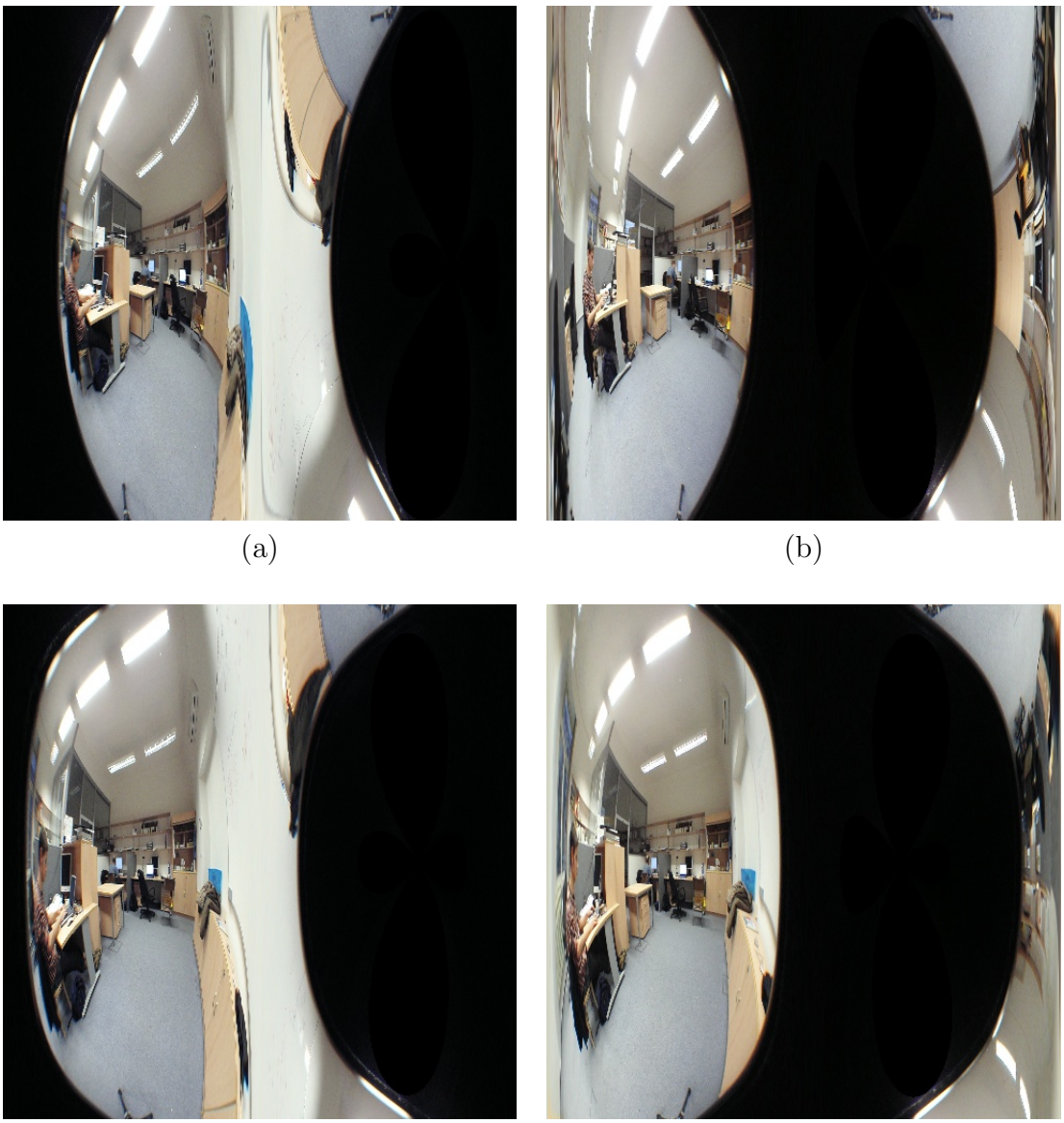
(a) (b)

(c) (d)

Figure B.9: (a, b) Spherical rectification of the image pair P_{O_1}. (b, c) Spherical rectification of the image pair P_{O_2}.

The `affine` structure for P_{O_1}	The `affine` structure for P_{O_2}
`width` = 500 `height` = 500 `o` = $(0,1)^\top$ `p` = $(0,0)^\top$ `param` = 0	`width` = 500 `height` = 500 `o` = $(0,1)^\top$ `p` = $(0,0)^\top$ `param` = 0

B.2.2 Swapped spherical rectification

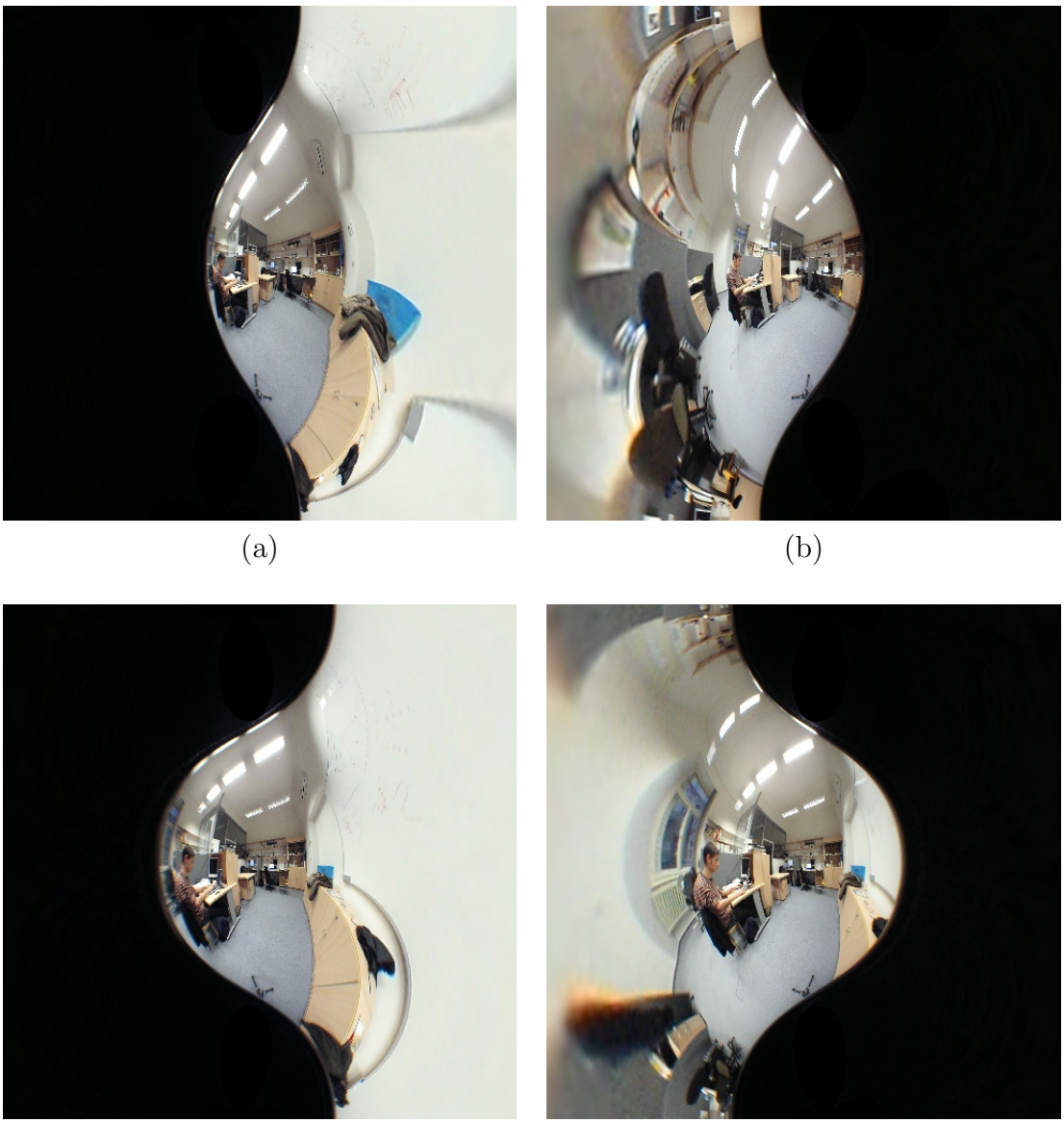


(a) (b) (c) (d)

Figure B.10: (a, b) Swapped spherical rectification of the image pair P_{O_1}. (b, c) Swapped spherical rectification of the image pair P_{O_2}.

The `affine` structure for P_{O_1}	The `affine` structure for P_{O_2}
`width` = 500 `height` = 500 `o` = $(0,1)^\top$ `p` = $(0, \texttt{height}/2)^\top$ `param` = 0	`width` = 500 `height` = 500 `o` = $(0,1)^\top$ `p` = $(0, \texttt{height}/2)^\top$ `param` = 0

B.2.3 Bipolar rectification

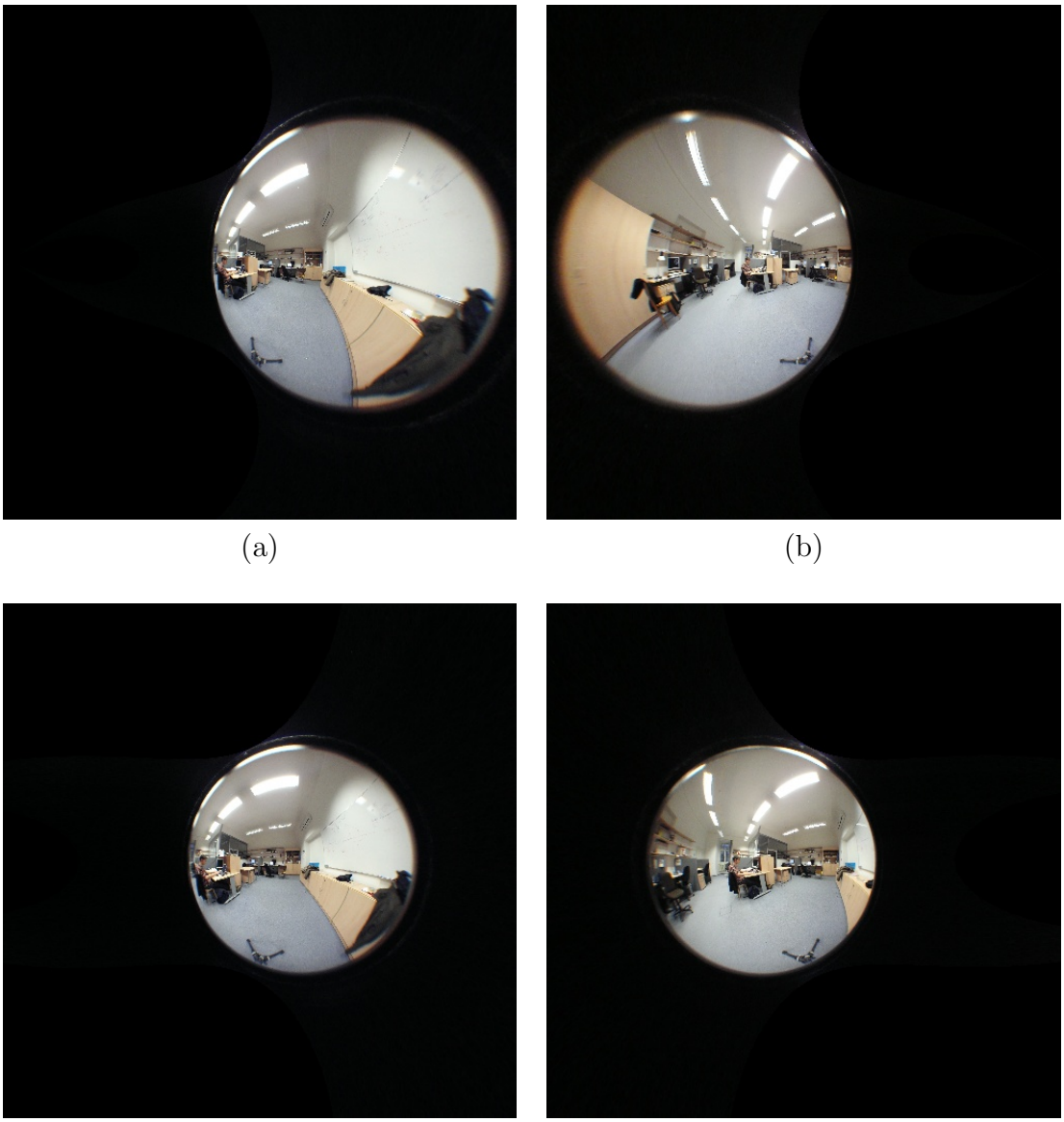


(a) (b) (c) (d)

Figure B.11: (a, b) Bipolar rectification of the image pair P_{O_1}. (b, c) Bipolar rectification of the image pair P_{O_2}.

The `affine` structure for P_{O_1}	The `affine` structure for P_{O_2}
$\texttt{width} = 500$ $\texttt{height} = 500$ $\texttt{o} = (1,1)^\top$ $\texttt{p} = (0,\texttt{height}/2)^\top$ $\texttt{param} = 5$	$\texttt{width} = 500$ $\texttt{height} = 500$ $\texttt{o} = (1,1)^\top$ $\texttt{p} = (0,\texttt{height}/2)^\top$ $\texttt{param} = 5$

B.2.4 Stereographic rectification



(a) (b)

(c) (d)

Figure B.12: (a, b) Stereographic rectification of the image pair P_{O_1}. (b, c) Stereographic rectification of the image pair P_{O_2}.

The `affine` structure for P_{O_1}	The `affine` structure for P_{O_2}
`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 2.7	`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 2.7

(a) (b)

(c) (d)

Figure B.13: (a, b) Stereographic rectification of the image pair P_{O_1}. (b, c) Stereographic rectification of the image pair P_{O_2}.

The `affine` structure for P_{O_1}	The `affine` structure for P_{O_2}
`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 1	`width` = 500 `height` = 500 `o` = $(0,0)^\top$ `p` = $(0,0)^\top$ `param` = 1

B.2.5 Rectification overlay

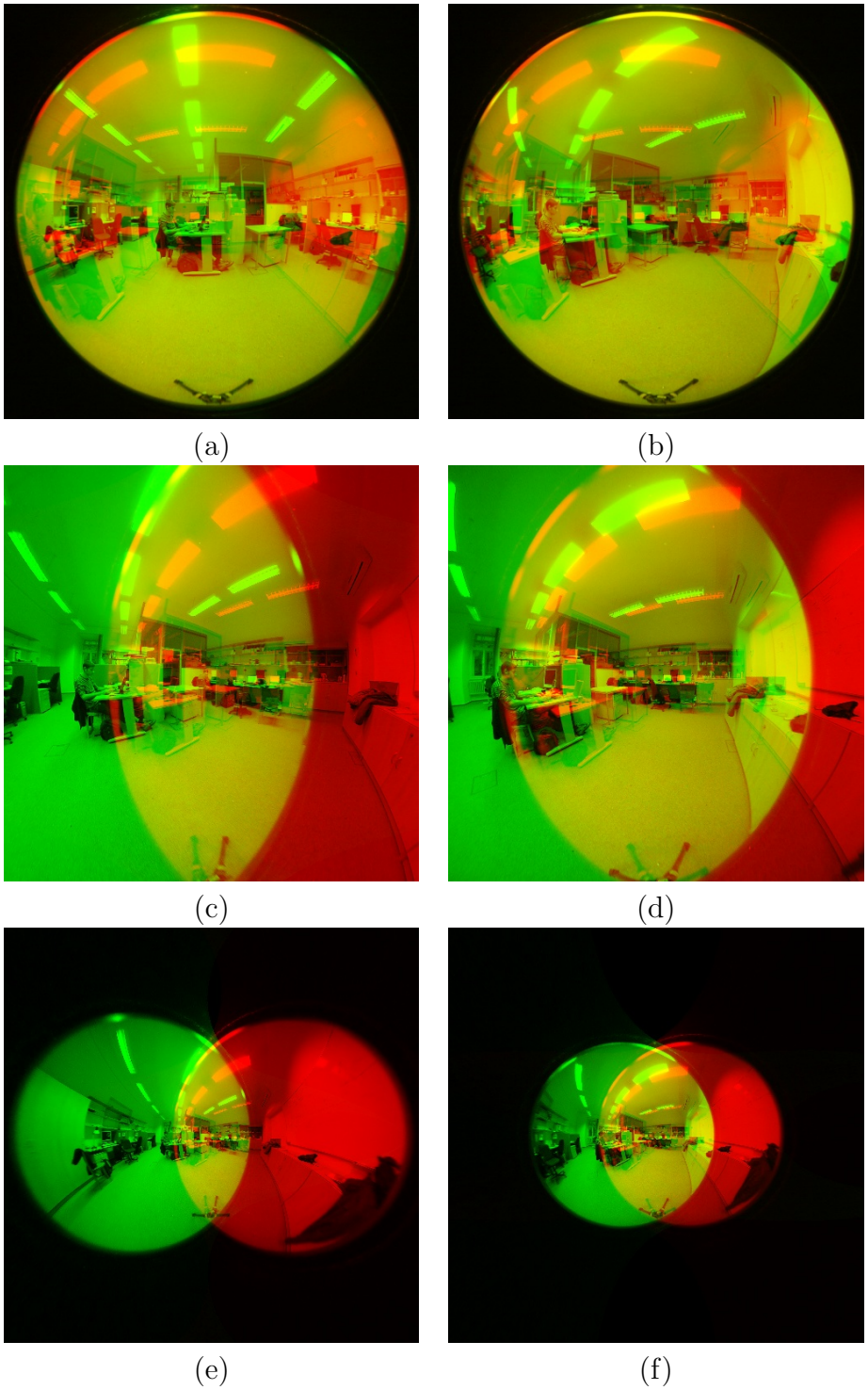
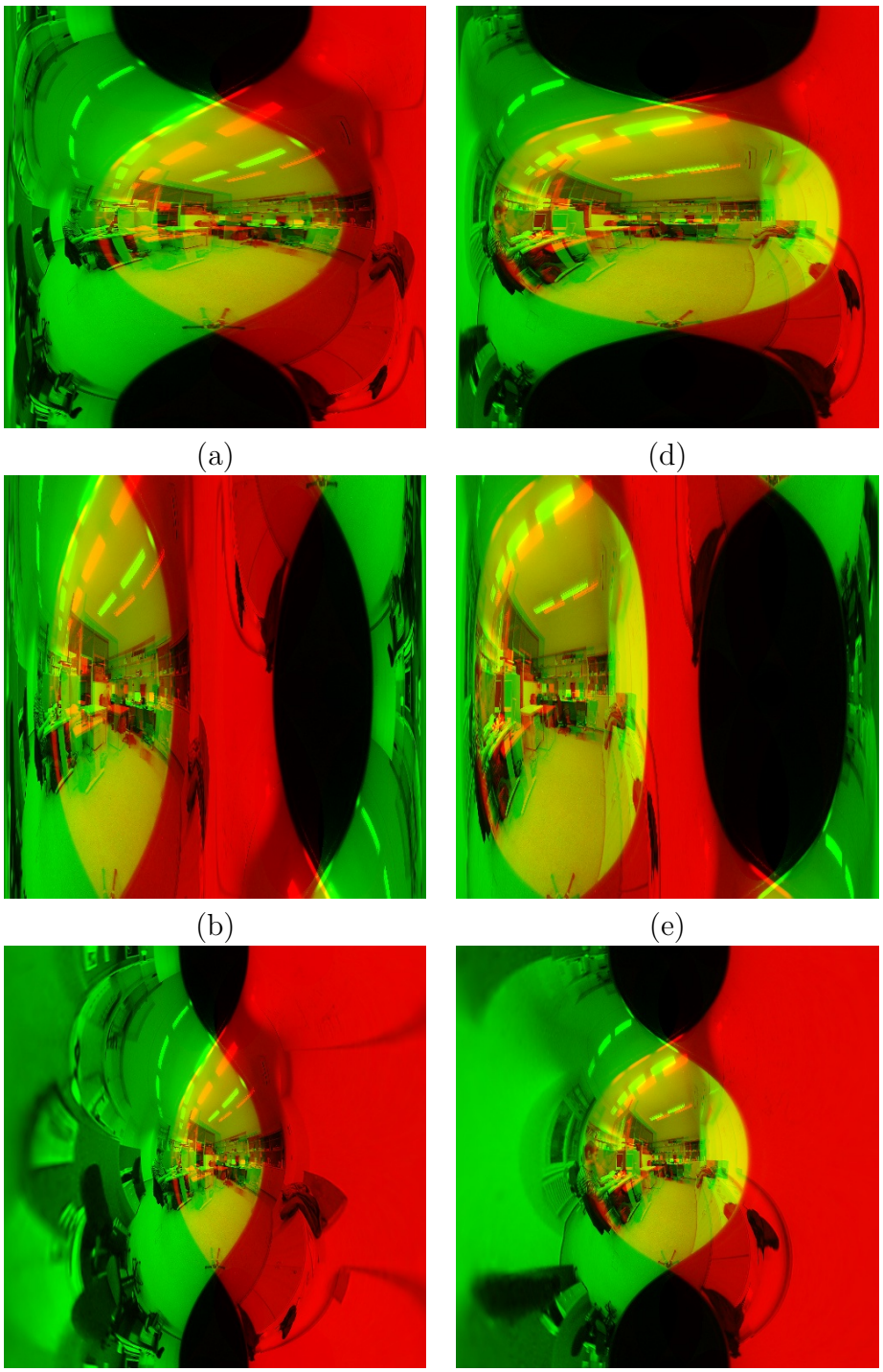


Figure B.14: (a) Overlay of P_{O_1},$\mathcal{P}^1_{O_1}$ in the red channel, $\mathcal{P}^2_{O_1}$ in the green channel. (b) Overlay of P_{O_2},$\mathcal{P}^1_{O_2}$ in the red channel, $\mathcal{P}^2_{O_2}$ in the green channel. (c,d) Overlay of stereographic rectification of P_{O_1} and P_{O_2} respectively, `param` $= 1$. (e,f) Overlay of stereographic rectification of $P_{O_1}$and P_{O_2} respectively, `param` $= 2.7$.



(a) (d) (b) (e) (c) (f)

Figure B.15: (a, b, c) Overlay of spherical, swapped spherical and bipolar rectification of P_{O_1} respectively. (d, e, f) Overlay of spherical, swapped spherical and bipolar rectification of P_{O_2} respectively.