CENTER FOR MACHINE PERCEPTION

CZECH TECHNICAL UNIVERSITY

RESEARCH REPORT

Estimation of Mobile Robot Position Using Panoramic Cameras—Thesis Proposal

Tomáš Svoboda

svoboda@cmp.felk.cvut.cz

N° K335/98/152

January 16, 1998

Acknowledgement

This research was supported by the Czech Ministry of Education grant VS96049, the internal grant of the Czech Technical University 3097472/335, the Grant Agency of the Czech Republic, grants 102/95/1378, 102/97/0480, 102/97/0855 and European Union grant Copernicus CP941068.

Czech Technical University, Faculty of Electrical Engineering
Department of Control Engineering, Center for Machine Perception
121 35 Prague 2, Karlovo náměstí 13, Czech Republic
fax +420 2 24357385, phone +420 2 24357465, http://cmp.felk.cvut.cz

ftp://cmp.felk.cvut.cz/pub/cmp/articles/svoboda/TR-K335-98-152.ps.gz.

Contents

1 Panoramic cameras, are they useful? 1

2 Task formulation 3

3 State of the art 4
3.1 Perspective camera and its model 4
3.2 Epipolar geometry for perspective cameras 7
3.3 Estimation of robot motion 8
3.4 Ego-motion via optical flow 10
3.5 Panoramic cameras 10

4 Our original contribution 12
4.1 Central panoramic cameras 12
4.1.1 Shape of the mirror 12
4.1.2 Model of a panoramic camera with a hyperbolic mirror 13
4.1.3 Design of a useful hyperbolic mirror 15
4.2 Epipolar geometry for panoramic cameras 16
4.3 Experiments 18
4.4 Summary 18

5 Thesis plan 19

Abstract

This report is dedicated to serve as a thesis proposal. It describes the state of the art of the observer motion estimation using a camera. The mathematical model of the camera is given, the epipolar geometry as a fundamental tool dealing with two images is presented, the ego-motion algorithm using point correspondences is explained. The more continuous approach employing an optical flow is also sketched. Previous works dealing with panoramic (omnidirectional) vision is overviewed.

As an original contribution, we present foundations of panoramic stereo vision by presenting the analysis of epipolar geometry for panoramic cameras. We show that the panoramic cameras with convex hyperbolic or parabolic mirrors, so called central panoramic cameras, allow the same epipolar geometry as perspective cameras. We work out the model of image formation by a central panoramic camera. It is shown that the epipolar curves in central panoramic images are conics and their equation is derived. The approach for designing a perspective camera–hyperbolic mirror system is also sketched. The theory is demonstrated by a simulated experiment which corroborates the conclusions drawn from the theory.

Finally, the thesis plan is outlined.

keywords: computer vision, robot vision, omnidirectional vision, panoramic vision, motion analysis, stereo vision.

Chapter 1

Panoramic cameras, are they useful?

Autonomous mobile robots are one of the important areas of applications of computer vision. A vehicle that can navigate without human supervision has many advantages in wide area of applications, providing access to hazardous industrial environments for instance. The tasks required for successful autonomous navigation by a mobile robot can be broadly classified as (1) sensing the environment; (2) building its own representation of the environment; (3) locating itself with respect to the environment; and (4) planning and executing efficient routes in this environment. Our research refers to the first and third task.

Our research is primarily motivated to develop an efficient tool for estimation of a mobile robot's position. Many approaches emerged recently dealing with. Some of the different sensors considered by previous researches are visual sensors (both monocular and binocular stereo), infrared sensors, ultrasonic sensors, and laser range finders. We deal with *visual sensors* in our research.

The standard visual sensor is TV camera. The *perspective*, sometimes called *pinhole*, mathematical model is usually used to describe such a camera. However, when a perspective camera is used for estimation of robot position, several principal problems emerge. The camera has a limited field of view. It is well known that motion estimation algorithms cannot, in some cases, well distinguish a small pure translation of the camera from a small rotation. An example is a translation parallel to the image plane and a rotation around an axis perpendicular to the direction of the translation [5]. The confusion becomes dominant when depth variation in the scene are small or if the field of view is narrow.

The confusion can be overcome if a camera with large field of view is used [5]. Ideally one would like to use a panoramic camera which has complete 360° field of view and sees to all directions. It can be imagined as a pinhole camera with a spherical imaging surface (instead of planar one as it is usual) centered at the focal point of the pinhole camera. *Panoramic camera* can, in principle, obtain correspondences from everywhere independently of the motion direction. The uncertainty of the motion estimation will therefore also become independent from the direction of motion. The above intuitive reasoning has been corroborated by the result of Brodský et al. [4] who shows that the motion estimation is almost never ambiguous if the spherical imaging surface is assumed.

Chapter 2

Task formulation

Out intention is to propose an efficient visual sensor for mobile robots with appropriate algorithms for estimation of a robot motion.

The algorithm for motion estimation from a pair of perspective images has to (1) find the corresponding points, (2) estimate the Fundamental matrix describing the epipolar geometry, (3) calibrate the camera, and (4) decompose the Fundamental matrix into a translation and rotation. All these steps were extensively studied in the past and new algorithms improving older ones still appear. We describe the principles of this approach in the chapter 3.

Motion estimation from panoramic images solves similar problems. It requires to

1. design a practical panoramic camera with a simple mathematical model,
2. propose a method for its calibration,
3. develop the epipolar geometry for panoramic images, and
4. work out an algorithm for motion estimation.

Panoramic stereo vision also needs efficient search for the correspondences in panoramic images which calls:

5. for the analysis of the shape of the epipolar curves in order to constraint the location of corresponding points and for
6. the study of epipolar alignment of the panoramic images in order to speed up the search.

Overview of the works most related to panoramic vision is given in the section 3.5. Our new contribution to panoramic, sometimes called omnidirectional, vision is described in the chapter 4.

Chapter 3

State of the art

This chapter covers the fundamentals of ego-motion estimation from consecutive images. It consists of five main parts.

1. We describe the mathematic model of the perspective camera in section 3.1. Perspective (pinhole) camera is mostly used for capturing images in computer vision.

2. Epipolar geometry is the fundamental tool dealing with two images. Section 3.2.

3. We briefly describe an ego-motion estimation approach using two views in section 3.3.

4. Having very dense sequence of consecutive images, an approach using optical flow can be used. Fundamentals are given in the section 3.4.

5. Panoramic cameras that have emerged recently are described in the section 3.5. Our most recent results in this field are mentioned, too.

3.1 Perspective camera and its model

Since the goal of our work is to develop methods for performing metric measurement from images, we have to define accurate quantitative model of the used cameras. We can build a geometric model of the *pinhole camera* as indicated in Figure 3.1. It consists of the *retina*, or *projection plane*, plane in which the image is formed through an operation called *perspective projection*: an optical center C, located at distance f, the focal length of the optical system is used to form the image $\mathbf{u}$ in the retina of the 3-D point X as the intersection of the line CX with the retina. The *optical axis* is the line going through the optical center C and perpendicular to retina, which it pierces at a point c.

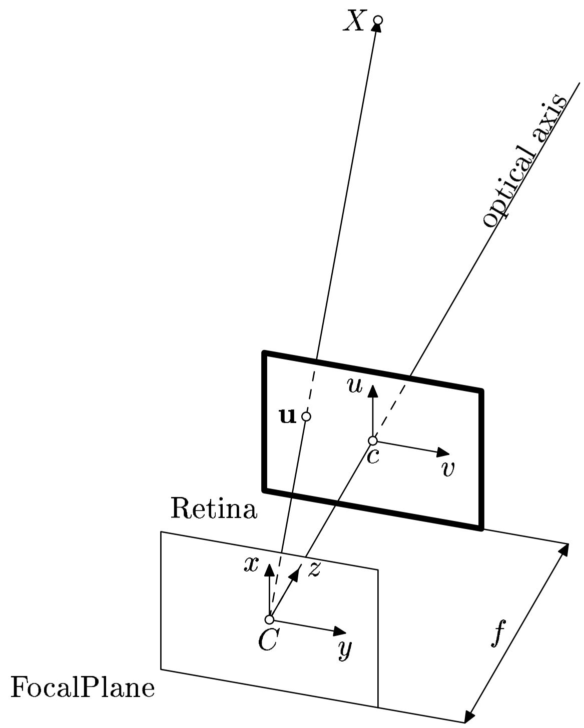


Figure 3.1: The mathematical model of the perspective (pinhole) camera.

The coordinate system (C, x, y, z) is called the *standard coordinate system* of the camera. In the retina, we introduce the *image coordinate system* (c, u, v). From Figure 3.1 it is clear that the relationship between image coordinates u, v and 3-D space coordinates $X = [x, y, z]^T$ can be written as

$$\frac{f}{z} = \frac{u}{x} = \frac{v}{y}, \tag{3.1}$$

which can be rewritten linearly as

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \tag{3.2}$$

where $u = U/S$, $v = V/S$ if $S \neq 0$. If $S = 0$, the corresponding point $[u, v]^T$ lies in infinity. This relationship can be rewritten in matrix form as

$$\mathbf{u} = \frac{1}{z} MX. \tag{3.3}$$

The perspective camera camera can be considered as a system that performs a linear projective transformations from the projective space $\mathcal{P}^3$ into the projective space $\mathcal{P}^2$.

Now, we describe how the matrix M changes when the retina system is transformed. We will introduce *intrinsic camera parameters*, called *calibration parameters*. The transformation is illustrated in Figure 3.2. The new

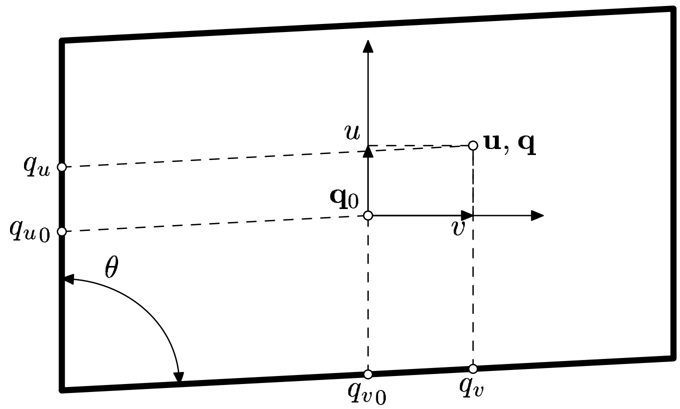


Figure 3.2: Projection plane. The transformation from image coordinates into pixel coordinates.

system into which the image coordinates u, v are to be transformed need not to be orthogonal. Let the scale between u and q_u be denoted k_u resp. k_v be the scale between v and q_v. First, the coordinate system is skewed and scaled. Second, it translates, since the coordinate system (c, u, v) has the origin in the *principal point*. From Figure 3.2 follows

$$q_v = \frac{k_v v}{\sin(\theta)} + q_{v0}, \quad (3.4)$$

$$q_u = k_u u - k_u v \frac{\cos(\theta)}{\sin(\theta)} + q_{u0}. \quad (3.5)$$

Equations (3.4)(3.5) can be rewritten in a matrix form as

$$\mathbf{q} = \begin{bmatrix} q_u \\ q_v \\ 1 \end{bmatrix} = H\mathbf{u}. \quad (3.6)$$

The 3×3 matrix H is given by

$$H = \begin{bmatrix} k_u & k_u \cot(\theta) & q_{u0} \\ 0 & \frac{k_v}{\sin(\theta)} & q_{v0} \\ 0 & 0 & 1 \end{bmatrix}. \quad (3.7)$$

According to the equation (3.3) we have

$$\mathbf{q} = \frac{1}{z} HMX. \quad (3.8)$$

Introducing matrix K, we can finally write

$$\mathbf{q} = \frac{1}{z} K X. \tag{3.9}$$

With regard to equations (3.7)(3.8), the matrix K can be written as

$$K = \begin{bmatrix} f k_u & k_u \cot(\theta) & q_{u0} \\ 0 & \frac{f k_v}{\sin(\theta)} & q_{v0} \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.10}$$

The parameters f, k_u, k_v, θ, q_{u0} and q_{v0} do not depend on the position and orientation of the camera in the space, and they are thus called *intrinsic* ones.

This is a *reconstruction problem.* Recovering of the translation between views is the *ego-motion problem*, which we deal with.

3.2 Epipolar geometry for perspective cameras

Epipolar geometry is the fundamental mathematical tool for dealing with pair of images. Epipolar geometry of two perspective cameras [6, 14], see Figure 3.3, assigns to each point $\mathbf{q}_1$ in one image an epipolar line $\mathbf{l}_2$ in the second image. All epipolar lines in each image intersect in the epipoles $\mathbf{e}_1$ and $\mathbf{e}_2$. Using geometry, the epipolar line can be defined as

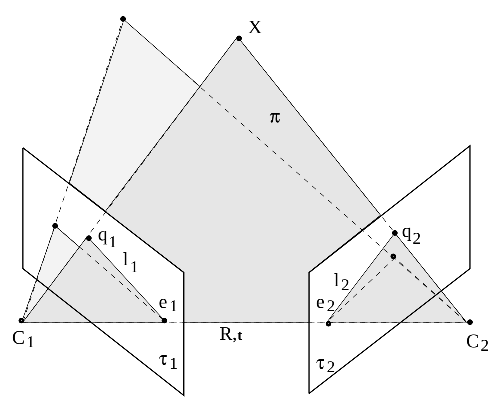


Figure 3.3: The epipolar geometry of two perspective cameras.

$$\mathbf{l}_1 = \mathbf{q}_1 \wedge \mathbf{e}_1, \tag{3.11}$$

where the symbol $\wedge$ denotes cross product. This equation can be rewritten in a matrix form as

$$\mathbf{l}_1 = B_{(\mathbf{e}_1)} \mathbf{q}_1. \tag{3.12}$$

The vector $\mathbf{l}_1$ is perpendicular to the line joining points $\mathbf{e}_1$ and $\mathbf{q}_1$ and the points on this line satisfy $\mathbf{l}_1\mathbf{x} = 0$. Since the epipolar lines $\mathbf{l}_1$ and $\mathbf{l}_2$ are coplanar, the transformation between them is collineation, defined by 3×3 matrix A.

$$\mathbf{l}_2 = A\mathbf{l}_1. \tag{3.13}$$

Association of equations (3.13) and (3.12) together, leaves us:

$$\mathbf{l}_2 = AB_{(\mathbf{e}_1)}\mathbf{q}_1 = Q\mathbf{q}_1. \tag{3.14}$$

The mathematical expression of the epipolar geometry

$$\mathbf{q}_2^T Q\mathbf{q_1} = 0, \tag{3.15}$$

says that the vector $\mathbf{t}$ connecting the centers of the cameras C_1 and C_2 is coplanar with the corresponding vectors $\mathbf{q}_1$ and $\mathbf{q}_2$. Q represents the *fundamental matrix* [6].

3.3 Estimation of robot motion

Suppose we have two or more images of one scene captured by a camera from different viewpoints. The two views case is sketched in Figure 3.3. We want to estimate the displacement between cameras. If the camera fixed on the mobile robot this problem is the same as the estimation of robot motion. Given two images, the estimation process encompasses two main problems.

1. For an image point in the first image, we have to decide which point in the second image corresponds to it. *Correspond* means that they are the images of the same physical point. This is the *correspondence problem.*

2. Given the number of corresponding points, we want to estimate the displacement between positions of the cameras. This is the *ego-motion problem.*

Many works dealing with these problems have emerged recently. The correspondence problem is one of the most essential ones in computer vision. An approach based on the recovering the epipolar geometry is presented in [24], related software called `image-matching` is available free on web[1]. A linear algorithm for scene reconstruction and ego-motion estimation was firstly described in [13]. An extensive work dealing with ego-motion and estimation of uncertainty is [21]. We developed a method how the uncertainty in the motion parameters can be predicted knowing uncertainty in camera calibration [20].

[1]`http://www.inria.fr/robotvis/personnel/zzhang/softwares.html`

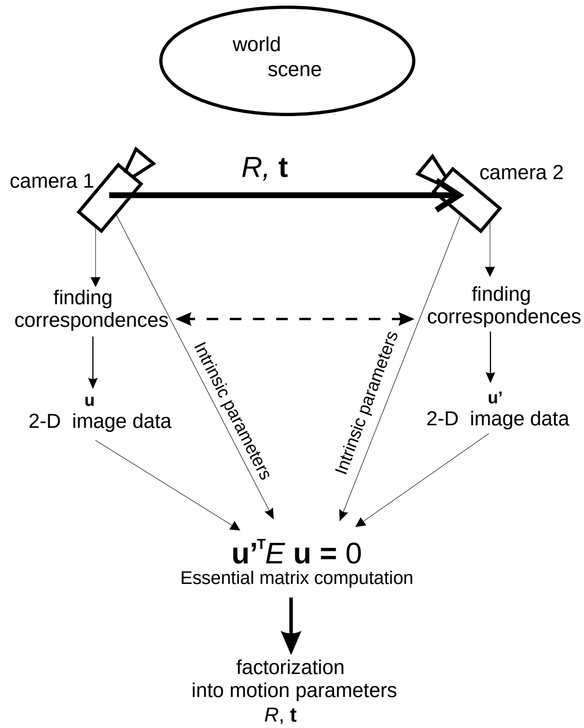


Figure 3.4: The relative displacement of the cameras can be computed from the pairs of corresponding points.

We briefly sketch the approach for the ego-motion estimation. Let the motion between two positions of the camera be given by the rotation matrix R and the translation vector $\mathbf{t}$. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be normalized image coordinates of corresponding point in the first resp. second image. Then the coplanarity constraint of vectors $\mathbf{t}$, $\mathbf{u}_1$ and $\mathbf{u}_2$ can be written in the coordinate system of the second camera as:

$$\mathbf{u}_2^T R(\mathbf{t} \wedge \mathbf{u}_1) = 0. \tag{3.16}$$

Introducing the antisymmetric matrix S

$$S = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix}, \tag{3.17}$$

we can rewrite the coplanarity constraint (3.16) in matrix form as

$$\mathbf{u}_2^T E \mathbf{u}_1 = 0. \tag{3.18}$$

Matrix $E = RS$ stands for the *essential matrix*. The essential matrix E can be used here instead of a fundamental matrix Q, since vectors $\mathbf{u}_i$ are metric entities. Several methods for essential matrix estimation have been published recently [6, 11, 13, 21]. We proposed an efficient approach for essential matrix computation in [17]. Knowing E the motion parameters R and $\mathbf{t}$ can be recovered using method [10], for instance.

3.4 Ego-motion via optical flow

In the previous section we discussed the problem of motion estimation from point correspondences, which technique treats the images as a samples of the scene taken at discrete times. Having very dense sequence of images, we can determine the *optical flow*. Optical flow is the velocity field in the image plane that arises due to projection of moving patterns in the scene onto the image plane [1, 9]. Since we do not work with the optical flow in our research, only basic equations are given here.

We denote by $[\dot{u}, \dot{v}]^T$ the velocity of the image point $[u, v]^T$, which is the projection of the space point $[x, y, z]^T$. By the geometry of perspective projection, we have

$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{3.19}$$

To change equation (3.19) into optical flow equation, we take time derivatives of both sides:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \frac{f}{z^2} \begin{bmatrix} z\dot{x} - x\dot{z} \\ z\dot{y} - y\dot{z} \end{bmatrix}. \tag{3.20}$$

Solving equation (3.19) for x, y and substituting into equation (3.20) yields the fundamental optical flow equation:

$$\begin{bmatrix} \dot{u} \\ \dot{v} \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 & -u \\ 0 & f & -v \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{bmatrix}. \tag{3.21}$$

Now suppose we observe an N-point optical flow field $\{[u_n, v_n, \dot{u}_n, \dot{v}_n]^T\}_{n=1}^N$ of a corresponding set of unknown three-dimensional points, all of which are assumed to be moving with the same but unknown velocity $[\dot{x}, \dot{y}, \dot{z}]^T$. Then up to a multiplicative constant, the position of all points and their common velocity can be determined from the optical flow solving set of equations (3.21) [1]. When the scene points are rigid and the camera moving, then the estimated velocity is equals to the velocity of the camera (ego-motion).

3.5 Panoramic cameras

Since the using of panoramic cameras is the main topic of our recent research we mention almost all works yet emerged.

Though other researchers have already realized that the use of panoramic cameras improves motion estimation, no attention has been paid to developing the epipolar geometry of panoramic cameras. The most related works by others are by Benosman [3], Yagi [22, 23], and by Southwell [16]. Benosman uses two 1024×1 line cameras rotating around a vertical axis. He gets two panoramic images but does not calculate epipolar geometry since the corresponding features lie trivially in the same column. The disadvantage is in complicated construction of the sensor. Yagi [22] developed a panoramic camera with a conic mirror. He uses it for detection of an azimuth of vertical lines. He integrates an acoustics sensor with the optical one and finds the trajectory of a mobile robot. In [23] he uses a hyperbolic mirror, however he detects swiveling motion analyzing an optical flow and develops no epipolar geometry. Southwell [16] proposes an idea of the stereo with one camera and concentric double lobed mirror but he does not present solid mathematical background.

There is also a number of works [7, 8, 12, 15] and most recently [2], which use panoramic cameras for fast visualization of a complete surroundings of the observer or as a source of images in order to construct a scene representation for virtual reality.

Ishiguro [12] describes how to create panoramic representation from images captured by a single swiveling camera. Fleck [7] studies several imaging models using wide-angle imaging geometry which can be used to create a panoramic image covering a hemisphere. Nayar [15] uses *orthographic camera* looking on the convex hyperbolic mirror. Orthographic camera is the limited case of the pinhole one when the focal length f goes to the infinity; all optical rays are parallel to the optical axis.

The most related work is the most recent by Baker and Nayar [2]. They describe various mirrors on which the perspective camera looks for preserving the unique center of projection. They also do the study on the defocus blur problem.

We proposed a new epipolar geometry for *central panoramic cameras* [19]. Extended version of this submitted paper with more theory about the design of useful mirror is [18].

Chapter 4

Our original contribution

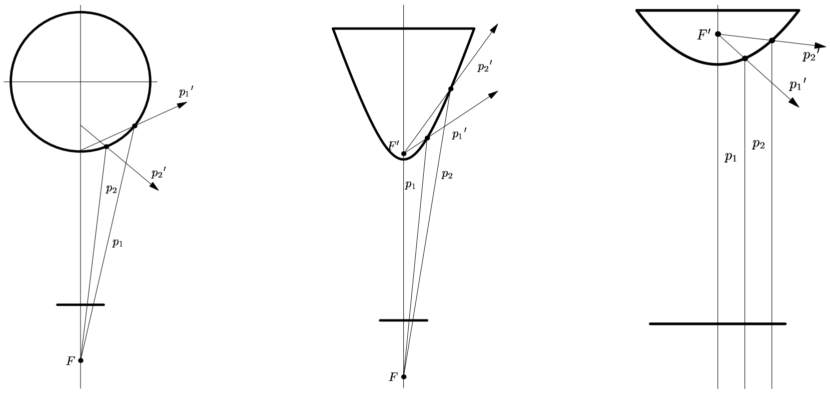
This chapter describes our most recent research results. It gives the foundations of panoramic stereo vision by presenting the analysis of epipolar geometry [6, 14] for panoramic cameras and explains how to design a real panoramic vision sensor. It has been primarily motivated by looking for an improvement of the motion estimation from a pair of images [6, 11, 13, 21] but its results are also applicable for structure from panoramic stereo images.

4.1 Central panoramic cameras

The cameras based on convex mirrors seem to be the most practical ones. They offer large field of view (approx. $360^\circ \times 150^\circ$), instant image acquisition (video rate), compact design, cheap production, and the freedom to choose the shape of the mirror in order to obtain a nice mathematical model of the camera.

4.1.1 Shape of the mirror

A panoramic camera with a mirror, Figure 4.1, consists of a perspective camera looking into a convex mirror. The ray p_1 going from (or coming into) the camera is reflected by the mirror into a ray p'_1. Each ray p has to pass through the focal point F of the perspective camera. Reflected rays p' can but need not to intersect themselves at the same point. Figure 4.1(a) shows a panoramic camera with a spherical mirror. In this case, the reflected rays do not intersect at the same point. Figure 4.1(b) shows the hyperbolic mirror where all the reflected rays intersect at the focal point of the mirror F'. The focal point of the camera, called center of projection, coincides with the second mirror focal point F. Figure 4.1(c) shows the parabolic mirror. Reflected rays p'_1, p'_2 intersect themselves in the focal point F' however the second focal point F moves to infinity thus the *orthographic* projection has to be used.



(a) Spherical mirror. The reflected optical rays do not intersect in a unique point – spherical aberration.

(b) Hyperbolic mirror. The reflected optical rays intersect in the focal point of the hyperboloid.

(c) Parabolic mirror. The reflected optical rays intersect in the focal point of the paraboloid when orthographic projection is assumed.

Figure 4.1: Three mirrors.

Here we focus to the case when all reflected rays intersect at a single point, further called the center of projection. Panoramic cameras which posses this property shall be called *central panoramic cameras.* Among all panoramic cameras the central ones are important because:

1. The shape of the mirror is uniquely defined by the requirement that both rays p and p' intersect in the point F and F', respectively.

2. Standard epipolar constraint can be used, i.e. the translation vector $\mathbf{t}$ between F_1' and F_2', see Figure 4.3, is coplanar with the vectors X_{h1} and X_{h2}. The model of the panoramic camera can be in this case decomposed into a central projection from 3-D space onto a curved surface of the mirror and a central projection from the surface of the mirror into the image plane.

4.1.2 Model of a panoramic camera with a hyperbolic mirror

Figure 4.2 shows the composition of a perspective camera with a hyperbolic mirror. The camera center C coincides with the focal point of the mirror F. The perspective camera can be modeled by an *internal camera calibration matrix* K which relates pixel coordinates $\mathbf{q} = [q_u, q_v, 1]^T$ to the normalized image coordinates $\mathbf{u} = [u, v, 1]^T$:

$$\mathbf{u} = K^{-1}\mathbf{q}, \tag{4.1}$$

This matrix K encompasses first three columns of matrix $\tilde{K}$, see equation (3.10). See [6] for more information about camera calibration.

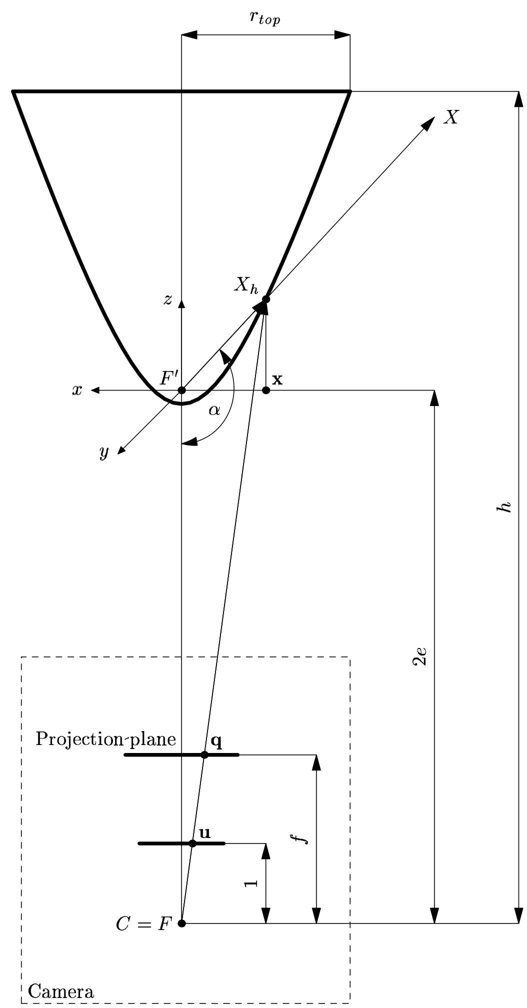


Figure 4.2: The geometry of the mirror and the camera.

The hyperbolic mirror is defined in the "mirror coordinate system", which is centered at the focal point F', by the equation

$$\frac{(z + \sqrt{a^2 + b^2})^2}{a^2} - \frac{x^2 + y^2}{b^2} = 1, \tag{4.2}$$

where a, b are parameters of the mirror. The image formation can be expressed as a composition of the coordinate transformations and projections. We want to find relationship between point X in world coordinates and the camera point $\mathbf{q}$ in pixels. The derivation of the image formation is omitted here, details can be found in [18]. Complete model can be concisely written as

$$\mathbf{q} \simeq KR_C\Big(\mathcal{F}(R_M(X-\mathbf{t}_M))R_M(X-\mathbf{t}_M)-\mathbf{t}_C\Big), \tag{4.3}$$

where $\simeq$ denotes equality up to similarity, $R_C, \mathbf{t}_C$ characterize the transformation between the mirror coordinate system and the camera frame, $R_M, \mathbf{t}_M$ denote transformation between the world and the mirror frame and where $\mathcal{F}(R_M(X-\mathbf{t}_M))$ is given by the following nonlinear function of a vector $\mathbf{v} = [v_1, v_2, v_3]^T$:

$$\mathcal{F}(\mathbf{v}) = \frac{b^2(ev_3 + a||\mathbf{v}||)}{b^2v_3^2 - a^2v_1^2 - a^2v_2^2}. \tag{4.4}$$

There are 6 external calibration parameters (3 for $\mathbf{t}_M$ and 3 for R_M) and 9 internal parameters (two for the mirror, two for the rotation R_C, and 5 for K). The matrix R_C has only two free parameter as it is used to model the angle between the image plane and the axis of the mirror.

In order to establish the equations of epipolar geometry, it is necessary to find the vector X_h for each image point $\mathbf{q}$. The formula for computing X_h from pixel coordinates $\mathbf{q}$ reads as:

$$X_h = \mathcal{F}(R_C^T K^{-1}\mathbf{q})R_C^T K^{-1}\mathbf{q} + \mathbf{t}_C, \tag{4.5}$$

where $\mathcal{F}(R_C^T K^{-1}\mathbf{q})$ is given by equation (4.4).

4.1.3 Design of a useful hyperbolic mirror

The shape of the real mirror has to be designed carefully since there are two main requirements: (1) The whole camera–mirror system has to be as compact as possible since it is to be used on a mobile robot, (2) the spatial angle of view has to be very close to whole sphere. This angle α is defined by

$$\alpha = \frac{\pi}{2} + \mathrm{atan}\Big(\frac{h-2e}{r_{top}}\Big), \tag{4.6}$$

where h is the distance between camera center C and the top of the mirror and r_{top} is the radius of the mirror rim, Figure 4.2, value of which has to be determined by a designer. Using mirror equation (4.2), h can be computed as

$$h = e + a\sqrt{1 + \frac{r_{top}^2}{b^2}}. \tag{4.7}$$

It is obvious that α is a function of a, b. In order to achieve good resolution, the projected mirror rim has to occupy the whole image. This requirement determines projected radius r_u in normalized image coordinates. A designer draws up values r_{top} and r_u. Then the height h from equations (4.6)(4.7) can be computed as

$$h = \frac{r_{top}}{r_u}. \tag{4.8}$$

Knowing h, the ratio a/b can be determined from equation (4.7). The maximum of angle α increases with increasing a/b. However the size (height) of the mirror increases as well. A compromise is necessary. The ground for the constraint $2 < a/b < 3$ can be found in [18].

4.2 Epipolar geometry for panoramic cameras

The epipolar geometry of central panoramic cameras, see Figure 4.3, related by the translation $\mathbf{t}$ pointing from F_1' to F_2' and the rotation R relating the coordinate systems of the mirrors, assigns to each point X_{h1} a curve on the second mirror. Points X_{h1}, X_{h2} and vector $\mathbf{t}$ are coplanar. We can rewrite

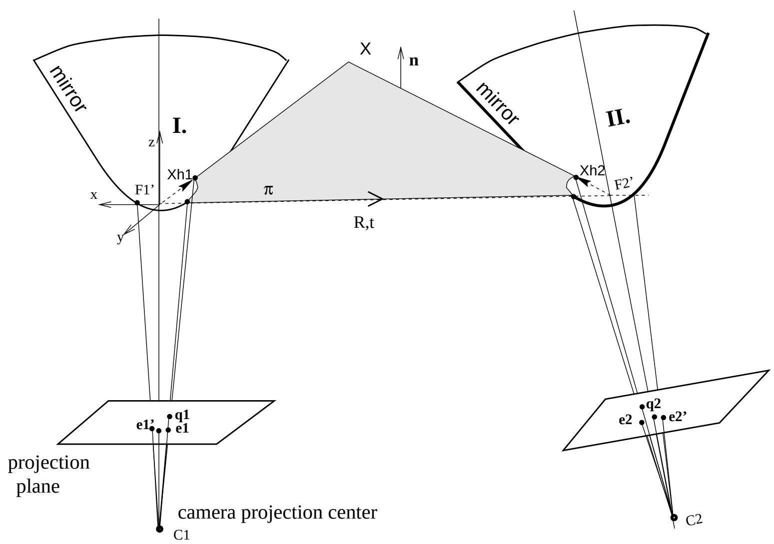


Figure 4.3: The epipolar geometry of two panoramic cameras with hyperbolic mirrors.

this coplanarity in the coordinate system of the second camera as

$$X_{h2} R(\mathbf{t} \wedge X_{h1}) = 0. \tag{4.9}$$

Introducing the antisymmetric matrix S containing elements of $\mathbf{t}$, see equation (3.17), we can rewrite the coplanarity constraint (4.9) in matrix form as

$$X_{h2}^T E X_{h1} = 0. \tag{4.10}$$

Matrix $E = RS$ stands for the *essential matrix*.

The epipolar curves are conics since they are intersections between planes and the quadratic surface (hyperboloid in our case). Each epipolar plane intersects the mirror in a planar conic. This conic is then projected into another conic to the image plane by a central (perspective) projection. To a point $\mathbf{q}_1$ in the first image a conic in the second image

$$\mathbf{q}_2^T A_2(E, \mathbf{q}_1)\ \mathbf{q}_2 = 0 \tag{4.11}$$

is assigned. The matrix $A_2(E, \mathbf{q}_1)$ is in general case a nonlinear function of the essential matrix E, point $\mathbf{q}_1$, and the calibration parameters of the panoramic cameras and the mirrors.

All epipolar conics pass through two points which are images of the intersection of mirrors with the line $F_1'F_2'$. Therefore there are usually two epipoles, denoted e_1 and e_1' resp. e_2 and e_2' in Figure 4.3. The epipoles can degenerate into one double epipole if the camera is translated along the axis of the mirror.

After intensive derivation, see [18], the equation (4.11) can be rewritten as

$$\mathbf{q}_2^T K^{-T} R_C B R_C^T K^{-1} \mathbf{q}_2 = 0, \tag{4.12}$$

leaving us with $A_2 = K^{-T} R_C B R_C^T K^{-1}$.

$$B = \begin{bmatrix} -4s^2a^2e^2 + p^2b^4 & pqb^4 & psb^2(-2e^2+b^2) \\ pqb^4 & -4s^2a^2e^2 + q^2b^4 & qsb^2(-2e^2+b^2) \\ psb^2(-2e^2+b^2) & qsb^2(-2e^2+b^2) & s^2b^4 \end{bmatrix} \tag{4.13}$$

is a nonlinear function of a, b, and

$$[p, q, s]^T = E(\mathcal{F}(R_C^T K^{-1} \mathbf{q}_1) R_C^T K^{-1} \mathbf{q}_1 + \mathbf{t}_C), \tag{4.14}$$

where $\mathcal{F}(R_C^T K^{-1} \mathbf{q}_1)$ is defined by equation (4.4). The equation (4.12) defines the curve on which the projected corresponding point has to lie and it is indeed an equation of a conic as alleged by equation (4.11).

Using the epipolar geometry The epipolar geometry presented above can be used similarly as using perspective cameras. Finding correspondences is the one reason to use it. Once the epipolar geometry between two panoramic images is established the search for correspondences is nicely reduced to 1 degree of freedom problem. At least 8 correspondences are needed to solve essential matrix linearly (4.10), using method [11], for instance. Knowing essential matrix E we can compute epipolar conic for each point of interest in the first image, on which conic the corresponding point has to lie. Then we can employ an iterative algorithm to establish epipolar geometry and to find correspondences more robustly [24].

Let the essential matrix E be robustly estimated, from equation (4.10). Recall that $E = RS$, where R is rotation matrix and matrix S contains elements of the translation vector $\mathbf{t}$. The *motion parameters* R and $\mathbf{t}$ can be recovered using the approach described in [10], for instance.

4.3 Experiments

We present a simulated experiment with a synthetic scene rendered by POV-ray software package[1] as our hyperbolic mirror has been manufactured just at December 1997. We did tests with the "bar" scene taken from SCED package[2]. A pair of panoramic images is shown in Figure 4.4. We can see

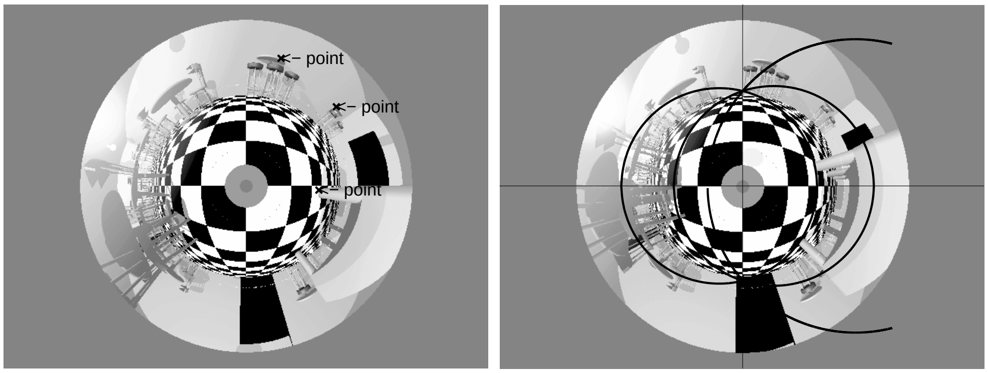


Figure 4.4: Left: A pair of panoramic images with three marked points. Right: Epipolar conics pass through the corresponding points.

three points of interest in the top image and the epipolar conics passing through the corresponding points in the bottom image. Epipolar conics intersect themselves on the y-axis, since the normalized displacement is $\mathbf{t} = [0,1,0]^T$ for this pair.

4.4 Summary

This chapter has presented the foundations of epipolar geometry for panoramic cameras and the approach to design a useful hyperbolic mirror. We have shown that panoramic cameras using convex hyperbolic or parabolic mirrors, so called *central panoramic cameras*, can be decomposed into two central projections and therefore allow the same epipolar geometry as perspective cameras. We have defined the model of image formation by a central panoramic camera. It has been shown that epipolar curves are conics and their equation was derived. The theory has been demonstrated by a simulated experiment.

[1] `http://www.povray.org`
[2] `http://http.cs.berkeley.edu/~schenney/sced/sced.html`

Chapter 5

Thesis plan

My thesis will deal with mobile vehicle position estimation from panoramic camera. The plan is to propose a compact theory of the problem and verify it on real experiments.

There are many opened questions. Namely, the influence of changing resolution in the field of view has to be studied. It is a question how lower resolution will affect the quality of motion estimation. The epipolar alignment of central panoramic images can be further worked out so that the corresponding epipolar curves have the same equations in both images.

The experiments with real data have to be carried out. Special care needs to be paid to finding out an automatic calibration and adjustment method for real mirrors and perspective cameras.

To prepare my PhD thesis I will further work on

1. the alignment (rectification) of the panoramic images,
2. the development of useful adjustment and calibration method,
3. the analysis of the reliability of using central panoramic cameras instead of perspective ones.

Bibliography

[1] J. K. Aggarwal and W. Martin. *Motion Understanding.* Kluwer Academic Publishers, Boston, Ma, 1988.

[2] Simon Baker and Shree K. Nayar. A theory of catadioptric image formation. *Submitted to IJCV*, 1997. Also to appear in ICCV98, and exists as TR CUCS-015-97, and it was presented in Darpa IUW 1997.

[3] R. Benosman, T. Maiere, and J. Devars. Multidirectional stereovision sensor, calibration and scenes reconstruction. In *International Conference on Pattern Recognition 1996, Vienna, Austria*, pages 161–165. IEEE Computer Society Press, September 1996.

[4] T. Brodský, C. Fernmüller, and Y. Aloimonos. Directions of motion fields are hardly ever ambiguous. In Buxton. B. and R. Cipola, editors, *4th European Conference on Computer Vision, Cambridge UK*, volume 2, pages 119–128, 1996.

[5] K. Daniilidis and H.-H. Nagel. The coupling of rotation and translation in motion estimation of planar surfaces. In *IEEE Conf. on Computer Vision and Pattern Recognition*, pages 188–193, New York, NY, June 1993.

[6] Olivier Faugeras. *3-D Computer Vision, A Geometric Viewpoint.* MIT Press, 1993.

[7] Margaret M. Fleck. Perpective projection: The wrong imaging model. Research report 95-01, Department of Computer Science, University of Iowa, Iowa City, USA, 1995.

[8] Francis Hamit. New video and still cameras provide a global roaming viewpoint. *Advance Imaging*, pages 50–52, March 1997.

[9] R. M. Haralick and L. G. Shapiro. *Computer and Robot Vision, Volume II.* Addison Wesley, Reading, Ma., 1993.

[10] Richard I. Hartley. Euclidean reconstruction from uncalibrated views. In *Workshop on Invariants in Computer Vision*, pages 187 – 202, October 1993.

[11] Richard I. Hartley. In defence of the 8-point algorithm. In *Fifth International Conference on Computer Vision, MIT Cambridge Massachusetts*, pages 1064–1070. IEEE Copmuter Society Press, 1995.

[12] Hiroshi Ishiguro, Masashi Yamamoto, and Saburo Tsuji. Omni–directional stereo. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, 14(2):257–262, February 1992.

[13] H.C. Longuett-Higgins. A computer algorithm for reconstruction a scene from two projections. *Nature*, 293:133–135, 1981.

[14] Roger Mohr. *Hanbook of Pattern Recognition and Computer Vision*, chapter 2.4, Projective Geometry and Computer Vision, pages 369–393. World Scientific Publishing Company, 1993.

[15] Shree K. Nayar. Catadioptric omnidirectional camera. In *International Conference on Computer Vision and Pattern Recognition 1997, Puerto Rico USA*, pages 482–488. IEEE Computer Society Press, June 1997.

[16] David Southwell, Anup Basu, Mark Fiala, and Jereome Reyda. Panoramic stereo. In *International Conference on Pattern Recognition 1996, Vienna, Austria*, volume A, pages 378–382. IEEE Computer Society Press, August 1996.

[17] Tomáš Svoboda and Tomáš Pajdla. Eficient motion analysis. Research report K335/95/95, Czech Technical University, Faculty of Electrical Engineering, October 1995. 29 pages. Available at `ftp://cmp.felk.cvut.cz/pub/cvl/articles/svoboda`.

[18] Tomáš Svoboda, Tomáš Pajdla, and Václav Hlaváč. Central panoramic cameras: Geometry and design. Research report K335/97/147, Czech Technical University, Faculty of Electrical Engineering, Center for Machine Perception, December 1997. Available at ftp://cmp.felk.cvut.cz/pub/cmp/articles/svoboda/TR-K335-97-147.ps.gz.

[19] Tomáš Svoboda, Tomáš Pajdla, and Václav Hlaváč. Epipolar geometry for panoramic cameras. Submitted to ECCV'98, November 1997.

[20] Tomáš Svoboda and Peter Sturm. A badly calibrated camera in ego-motion estimation, propagation of uncertainty. In Gerald Sommer, Kostas Daniilidis, and Josef Pauli, editors, *7-th International Conference Computer Analysis of Images and Patterns, Kiel, Germany*, pages 183–190. Springer Verlag in the LNCS series, September 1997.

[21] Juyang Weng, Thomas S. Huang, and Ahuja Narendra. Motion and structure from two perspective views: Algorithms, error analysis, and

error estimation. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, 11(5):451–476, May 1989.

[22] Yasushi Yagi, Yoshimitsu Nishizawa, and Masahiko Yachida. Map-based navigation for a mobile robot with omnidirectional images sensor COPIS. *IEEE Transaction on Robotics and Automation*, 11(5):634–648, October 1995.

[23] Yasushi Yagi, Kazuma Yamazawa, and Masahiko Yachida. Rolling motion estimation for mobile robot by using omnidirectional image sensor hyperomnivision. In *International Conference on Pattern Recognition 1996, Vienna, Austria*, pages 946–950. IEEE Computer Society Press, September 1996.

[24] Zhengyou Zhang, Rachid Deriche, Olivier Faugeras, and Quang-Tuang Luog. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Inteligence*, 78:87–119, October 1995. Also Research Report No.2273, INRIA Sophia-Antipolis.