Copyright © 1998, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

VISION THEORY IN SPACES OF CONSTANT CURVATURE

by

Yi Ma and Shankar Sastry

Memorandum No. UCB/ERL M98/36

20 June 1998

VISION THEORY IN SPACES OF CONSTANT CURVATURE

by

Yi Ma and Shankar Sastry

Memorandum No. UCB/ERL M98/36

20 June 1998

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Vision Theory in Spaces of Constant Curvature *

Yi Ma Shankar Sastry

Electronics Research Laboratory
University of California at Berkeley
Berkeley, CA 94720-1774
{mayi, sastry}@robotics.eecs.berkeley.edu

June 20, 1998

Abstract

In this paper, vision theory for Euclidean, spherical and hyperbolic spaces is studied in a uniform framework using differential geometry in spaces of constant curvature. It is shown that the epipolar geometry for Euclidean space can be naturally generalized to the spaces of constant curvature. In particular, it is shown that, in the general case, the bilinear epipolar constraint is exactly the same as in the Euclidean case; also, there are only bilinear, trilinear and quadrilinear constraints associated with multiple images of a point. Differential (continuous) case is also studied. For the structure from motion problem, 3D structure can only be determined up to a universal scale, the same as the Euclidean case. Approaches are proposed to reconstruct 3D structure with respect to a normalized curvature.

Key words: space of constant curvature, spherical space, hyperbolic space, computer vision, epipolar constraint, motion recovery, structure from motion, absolute geometry.

1 Spaces of Constant Curvature

Spaces of constant curvature are Riemannian manifolds with constant sectional curvature. In differential geometry, they are also referred to as **space forms**. A Riemannian manifold of constant curvature is said to be **spherical**, **hyperbolic** or **flat** (or **locally Euclidean**) according as the sectional curvature is positive, negative or zero. Geometry about spaces of constant curvature is also called **absolute geometry** [1], due to one of the co-founders non-Euclidean geometry: Janos Bolyai.

Not until Einstein's general relativity theory, non-Euclidean geometry, or Riemannian geometry in general, is just a pure mathematical creation rather than geometry of physical spaces. In general relativity theory, the physical space is typically described as a (3 dimensional) Riemannian manifold (with possibly non-zero curvature). In such a space, **light travels the geodesics** of the manifold

*Computer vision research report. This work is supported by ARO under the MURI grant DAAH04-96-1-0341. The authors also would like to thank professor Alan Weinstein from Berkeley Mathematics Department for invaluable discussions during the preparation of this manuscript.

(corresponding to straight lines in the Euclidean case). Locally, the curvature of a Riemannian manifold is approximately constant. Thus the study of vision theory in spaces of constant curvature will help understand vision problems in general Riemannian manifolds.

In this paper, we study vision theory in 3 dimensional spaces of constant curvature, as a natural generalization of the vision theory in 3 dimensional Euclidean space. In particular, we study vision in spherical and hyperbolic spaces since the Euclidean case has been well studied. On the other hand, the Euclidean case will always show up in discussion as a special (limit) case of the general theory.

Geometric properties of n dimensional space of constant curvatures have been well studied in differential geometry [2, 3, 9] (as an important case of **symmetric spaces**). In the rest of this section, we *review* some of the main results which are important for studying vision.

1.1 Characteristics of Spaces of Constant Curvature

In this section, we characterize 3 dimensional spaces of constant curvature. In fact, most of the results directly follow from general results about n dimensional spaces of constant curvature, in Kobayashi [2, 3] and Wolf [9].

The next theorem following from Kobayashi [2] (Theorem 3.1 Chapter V) characterizes the 3 dimensional space of constant curvatures:

Theorem 1 (Three Dimensional Spaces of Constant Curvature) *Let (x_1, x_2, x_3, x_4) be the coordinate system of $\mathbb{R}^4$ and M be the hyper-surface of $\mathbb{R}^4$ defined by:*

$$x_1^2 + x_2^2 + x_3^2 + r x_4^2 = r \quad (r\text{: a nonzero constant}). \tag{1}$$

Let g be the Riemannian metric of M obtained by restricting the following form to M:

$$dx_1^2 + dx_2^2 + dx_3^2 + r\, dx_4^2.$$

Then

1. *M is a 3 dimensional space of constant curvature with sectional curvature $1/r$.*
2. *The group G of linear transformations of $\mathbb{R}^4$ leaving the quadratic form $x_1^2 + x_2^2 + x_3^2 + r x_4^2$ invariant acts transitively on M as the group of isometries of M.*
3. *If $r > 0$, then M is isometric to a sphere of a radius $\sqrt{r}$. If $r < 0$, then M consists of two mutually isometric connected manifolds each of which is diffeomorphic with $\mathbb{R}^3$.*

Let Q be the 4×4 matrix associated to the quadratic form defining M:

$$Q = \begin{pmatrix} I_3 & 0 \\ 0 & r \end{pmatrix}.$$

The **isometry group** G of M is then given as a subgroup of $GL(4, \mathbb{R})$:

$$G = \left\{ g \in \mathbb{R}^{4\times 4} \,\middle|\, g^T Q g = Q \right\}. \tag{2}$$

For an element $g \in G$, it has the form:

$$g = \begin{pmatrix} W & y \\ z^T & w \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

with $W \in \mathbb{R}^{3\times 3}, y \in \mathbb{R}^3, z \in \mathbb{R}^3, w \in \mathbb{R}$ and the conditions:

$$W^T W + r \cdot zz^T = I_3, \quad W^T y + r \cdot wz = 0, \quad y^T y + r \cdot w^2 = r. \tag{3}$$

It follows that the **Lie algebra** $\mathfrak{g}$ of the group G (as a Lie group) is the set of the matrices of the form:

$$\xi = \begin{pmatrix} A & b \\ c^T & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4} \tag{4}$$

where $A \in \mathbb{R}^{3\times 3}, b \in \mathbb{R}^3$ and $c \in \mathbb{R}^3$ satisfy the conditions:

$$A^T + A = 0, \quad b + r \cdot c = 0. \tag{5}$$

The **isotropy group** H of G which leaves the point $o = (0,0,0,1)^T \in M$ fixed is isomorphic to $O(3)$:

$$H = \begin{pmatrix} O(3) & 0 \\ 0 & 1 \end{pmatrix}. \tag{6}$$

As a result, the manifold M is identified with the **homogeneous space** G/H. In fact, the orthonormal frame bundle of M is isomorphic to G as a principle H bundle, Kobayashi [2].

Let $\mathfrak{m}$ be the linear subspace of the Lie algebra $\mathfrak{g}$ of G consisting of matrices of the form:

$$\begin{pmatrix} 0 & b \\ c^T & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4} \tag{7}$$

with $b, c \in \mathbb{R}^3$ and $b + rc = 0$. Let $\mathfrak{h}$ be the Lie algebra of H as a subspace of $\mathfrak{g}$ consisting of matrices of the form:

$$\begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4} \tag{8}$$

with $A \in \mathbb{R}^{3\times 3}$ and $A^T + A = 0$. Then we have a **canonical decomposition:**

$$\mathfrak{g} = \mathfrak{h} + \mathfrak{m}. \tag{9}$$

It is direct to check the following relations between the subspaces hold:

$$[\mathfrak{h}, \mathfrak{h}] \subset \mathfrak{h}, \quad [\mathfrak{h}, \mathfrak{m}] \subset \mathfrak{m}, \quad [\mathfrak{m}, \mathfrak{m}] \subset \mathfrak{h} \tag{10}$$

where $[\cdot,\cdot]$ stands for **Lie bracket**. Let $\mathfrak{h}$ be the **vertical tangent subspace** of G and $\mathfrak{m}$ be the **horizontal tangent subspace.** Then this decomposition gives a **canonical connection** on the principle bundle $G(G/H, H)$ (Theorem 11.1 of Chapter II, Kobayashi [2]) which induces constant sectional curvature $1/r$ on $G/H = M$.

The space M is a symmetric space with the symmetry s_o of M at the point $o = (0,0,0,1)^T$ given by:

$$\begin{array}{rcl} s_o : M & \to & M \\ (x_1, x_2, x_3, x_4)^T & \mapsto & (-x_1, -x_2, -x_3, x_4)^T. \end{array}$$

Obviously, $s_o^2 = Id(M)$. Due to Kobayashi [3] (Theorem 1.5 of Chapter XI), this induces a (symmetric) automorphism σ on G such that H lies between G_σ (subgroup of G fixed under σ) and the identity component of G_σ.

Denote the projection from G to G/H as π and Let $\exp(\cdot)$ be the exponential map from $\mathfrak{g}$ to G. Then according to Kobayashi [3] (Theorem 3.2 of Chapter XI), we have:

Theorem 2 (Geodesics in 3D Spaces of Constant Curvature) *Consider the 3 dimensional space of constant curvature $M = G/H$ as above. For each $X \in \mathfrak{m}$, $\pi(\exp tX) = (\exp tX) \cdot o$ is a geodesic starting from o and, conversely, every geodesic from o is of this form.*

As we will soon see, this theorem is very important for modeling and studying vision in the spaces of constant curvature.

Let T be the subset of G consisting of all the matrices of the form $\exp(X)$ with $X \in \mathfrak{m}$. Then T corresponds to **transvection** on M (see Kobayashi [3]), an analogy to the translation in the Euclidean space. Notice that in general T is not a subgroup of G (although it is in the Euclidean case) and its representation depends on the base point. Naturally, the subgroup H of G corresponds to **rotation** on M. As in the Euclidean case, for a "rigid body motion" on M, it is natural to consider the rotation is in the special orthogonal group $SO(3)$ instead of the full group $O(3)$. One of the reasons for only considering $SO(3)$ is that it preserves the orientation of the space.

1.2 Euclidean Space as a Space of Constant Curvature

Theorem 1 requires the curvature parameter $r \in \mathbb{R} \setminus \{0\}$ hence only the spherical and hyperbolic spaces were considered. However, the Euclidean case can be regarded as the limit case when r goes to infinite, *i.e.* the curvature $1/r$ goes to zero.

When $r = \infty$, a point in $\mathbb{R}^4$ which satisfies the quadratic form (1) always has the form $(x_1, x_2, x_3, 1)^T \in \mathbb{R}^4$. This is just the homogeneous representation of the 3 dimensional Euclidean space $\mathbb{R}^3$, see Murray [7]. From (3), we have $w^2 = 1$, $z = 0$, $W^T W = I_3$ and $y \in \mathbb{R}^3$. Thus the group G is just the Euclidean group $E(3)$. In particular, the special Euclidean group $SE(3)$ with elements:

$$g = \begin{pmatrix} R & p \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{4\times4} \tag{11}$$

with $R \in SO(3)$ and $p \in \mathbb{R}^3$ is a subgroup of $G = E(3)$. $SE(3)$ then represents the **rigid body motion** in $M = \mathbb{R}^3$.

When $r = \infty$, the Lie algebra $\mathfrak{se}(3)$ of $SE(3)$ or $\mathfrak{e}(3)$ of $E(3)$ then has the form given in (4) with the condition $c = 0$. In robotics literature [7], an element this Lie algebra is usually represented as:

$$\xi = \begin{pmatrix} \hat{\omega} & v \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times4} \tag{12}$$

where $\omega, v \in \mathbb{R}^3$ and $\hat{\omega}$ is the skew-symmetric matrix associated with $\omega = (\omega_1, \omega_2, \omega_3)^T$:

$$\hat{\omega} = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times3}. \tag{13}$$

According to Theorem 2, the geodesics in $\mathbb{R}^3$ are given in the form:

$$\exp t \begin{pmatrix} 0 & v \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} I_3 & vt \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{4\times4}. \tag{14}$$

This is exactly the straight line in $\mathbb{R}^3$ in the direction of v.

From the above discussion, the Euclidean space can be treated as a limit case of general spaces of constant curvature given in Theorem 1. Because of this, the vision theory for Euclidean space should also be a limit case of vision theory for general spaces of constant curvature.

2 Camera Motion and Projection Model

Based upon the mathematical facts given in the preceding section, we are ready to study vision in the spaces of constant curvature. Similar to the Euclidean case, we first need to specify the (valid) motion of the camera and the projection model of the camera, *i.e.* how the 2 dimensional image is formulated in spaces of constant curvature.

First notice that, as in the Euclidean case, the transvection set T of the isometry group G acts transitively on a space M of constant curvature. Then for any $g \in G$, there exists $g_t \in T$ such that $g_t^{-1}(g(o)) = o$, *i.e.* $g_t^{-1}g$ fixs the origin o. So $g_t^{-1}g = g_h \in H$, the isotropy group of o. It then follows that the group G is equal to $G = TH$. This is the so-called **Cartan decomposition.** By **rigid body motion** in spaces of constant curvature, we mean the connected subgroup of G which preserve the orientation of the space M. That is, the rotation group H is just $SO(3)$ (the subgroup of $O(3)$ which is connected to the identity element). We still use G to denote the group of rigid body motion:

$$G = TH \quad \text{with } H \in SO(3).$$

Assumption 1 *The motion of a camera in spaces of constant curvature is the group G of rigid body motion.*

A point q, in the space M of constant curvature, can be represented in **homogeneous coordinates** as $q = (q_1, q_2, q_3, q_4)^T \in \mathbb{R}^4$ which satisfies the quadratic form:

$$q_1^2 + q_2^2 + q_3^2 + rq_4^2 = r$$

with $1/r$ the sectional curvature of M. Then under the motion $g(t) \in G, t \in [t_0, t_f] \subset \mathbb{R}$ of the camera, the homogeneous coordinates of the point q (with respect to the camera frame) satisfy the transformation:

$$q(t) = g(t)q(t_0). \tag{15}$$

Notice that, with this representation, the point $o = (0,0,0,1)^T \in \mathbb{R}^4$ is always in M. We then call the point o the **origin** in the homogeneous representation of M. It is natural to assume that

Assumption 2 *The optical center of the camera is the same as the origin o in the homogeneous representation of M.*

According to Theorem 2, any geodesic connecting a point $q = (q_1, q_2, q_3, q_4)^T \in M$ to the origin o has the form: $q = (\exp tX) \cdot o$ for some $t \in \mathbb{R}, X \in \mathfrak{m}$. Without loss of generality, we may assume X has the form:

$$X = \begin{pmatrix} 0 & b \\ -b^T/r & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

for some unit vector $b \in \mathbb{R}^3$, $\|b\| = 1$. It is then direct to check that:

$$q = (\exp tX) \cdot o = \begin{pmatrix} f(r,t)bb^T & h_1(r,t)b \\ h_2(r,t)b^T & g(r,t) \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} h_1(r,t)b \\ g(r,t) \end{pmatrix} \in \mathbb{R}^4$$

for some real scalar functions $f(r,t), g(r,t), h_1(r,t)$ and $h_2(r,t)$ of r and t (the explicit expressions of these functions are given in the next section). Thus the unit vector b is equal to:

$$b = \frac{(q_1, q_2, q_3)^T}{\sqrt{q_1^2 + q_2^2 + q_3^2}}. \tag{16}$$

This is exactly the unit tangent vector of M at the origin o. In this way, we may identify the tangent space $T_o(M)$ of M at o to the subspace $\mathfrak{m}$ by:

$$\begin{aligned} \phi:\ T_o(M) &\rightarrow \mathfrak{m} \\ b \in T_o(M) &\mapsto \begin{pmatrix} 0 & b \\ -b^T/r & 0 \end{pmatrix} \in \mathfrak{m}. \end{aligned}$$

Under this identification, the exponential map $\exp : T_o(M) \rightarrow M$ is given by:

$$\exp b = (\exp \phi(b)) \cdot o, \quad b \in T_o(M).$$

In general relativity, it is assumed that

Assumption 3 *In a Riemannian manifold M, light always travels the geodesics with constant speed.*

Then from previous discussion, the light from $q = (q_1, q_2, q_3, q_4) \in M$ to the origin o has the direction $b \in T_o(M)$ given by (16). In homogeneous coordinate, the vector b can be represented as

$$b = (q_1, q_2, q_3)^T \in \mathbb{R}^3$$

which only keeps the information of the direction of the light from q. We may assume that

Assumption 4 *In a Riemannian manifold M with the optical center at o, the image of a point $q \in M$ is the direction of the tangent vector at $T_o(M)$ which corresponds to the geodesic connecting q and the optical center o.*

Then in the case of the space M of constant curvature, if the space M is represented by the homogeneous coordinates as above, the image of a point $q = (q_1, q_2, q_3, q_4)^T \in M$ is simply given by $\mathbf{x} = \lambda^{-1}(q_1, q_2, q_3)^T \in \mathbb{R}^3$ where $\lambda \in \mathbb{R}^+$ and $\mathbf{x} \in \mathbb{R}^3$. Define the **projection matrix** to be:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3\times 4}.$$

We then have the relation:

$$\lambda \mathbf{x} = Pq. \tag{17}$$

We call the scalar λ the **scale** of the point q with respect to the image $\mathbf{x}$. The scale λ then encodes the depth information of the point q in the scene.

Assumptions 1, 2, 3 and 4 specify the camera motion and projection model for a vision system in spaces of constant curvature. They give a natural generalization of the model of **perspective** and **spherical** projections in the Euclidean case (see [4] for a comparison).

3 Epipolar Geometry in the Spaces of Constant Curvature

In this section, we study the relation between the images of a point $q \in M$ before and after a rigid body motion of the camera. We know that the motion of the camera can be expressed in the form:

$$g = g_t \cdot g_h, \quad g_t \in T, g_h \in H.$$

The transvection part g_t and rotation part g_h respectively have the forms:

$$g_t = \exp X = \begin{pmatrix} W & y \\ z^T & w \end{pmatrix}, \quad g_h = \begin{pmatrix} R & 0 \\ 0 & 1 \end{pmatrix}, \quad X \in \mathfrak{m}, R \in SO(3). \tag{18}$$

We will later give the expressions of $W \in \mathbb{R}^{3\times 3}, y \in \mathbb{R}^3, z \in \mathbb{R}^3$ and $w \in \mathbb{R}$ in terms of X.

Denote the images of $q = (q_1, q_2, q_3, q_4)^T$ before and after the transformation g are $\mathbf{x}_1 \in \mathbb{R}^3$ and $\mathbf{x}_2 \in \mathbb{R}^3$, respectively. Then according to (15) and (17) we have:

$$\lambda_1 \mathbf{x}_1 = Pq, \quad \lambda_2 \mathbf{x}_2 = Pgq.$$

It yields:

$$\lambda_2 \mathbf{x}_2 = WR \cdot \lambda_1 \mathbf{x}_1 + q_4 y \quad \Rightarrow \quad y \times \lambda_2 \mathbf{x}_2 = y \times (WR \cdot \lambda_1 \mathbf{x}_1) \quad \Rightarrow \quad \mathbf{x}_1^T R^T W^T \hat{y} \mathbf{x}_2 = 0. \tag{19}$$

In the Euclidean case, (19) would exactly give the well-known bilinear epipolar constraint. In the case of spaces of constant curvature, the role of essential matrix is replaced by $R^T W^T \hat{y}$. We need to study the structure of such matrices.

Any matrix $X \in \mathfrak{m}$ has the form:

$$\begin{pmatrix} 0 & b \\ -b^T/r & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

with vector $b \in \mathbb{R}^3$. To simply the notation, denote $\gamma = \|b\|$ and $p = b/\gamma$. We consider $\sin(\cdot)$ and $\cos(\cdot)$ as the complex functions:

$$\begin{aligned}\sin(u) &= \frac{1}{2i}(e^{iu} - e^{-iu}), \quad u \in \mathbb{C} \\ \cos(u) &= \frac{1}{2}(e^{iu} + e^{-iu}), \quad u \in \mathbb{C}.\end{aligned}$$

Also define $\rho = \sqrt{1/r} \in \mathbb{C}$. Then through direct calculation we get:

$$\exp X = \begin{pmatrix} W & y \\ z^T & w \end{pmatrix} = \begin{pmatrix} I_3 + (\cos(\gamma\rho) - 1)pp^T & \rho^{-1}\sin(\gamma\rho)p \\ \rho\sin(\gamma\rho)p^T & \cos(\gamma\rho) \end{pmatrix}. \tag{20}$$

Notice that we always have $pp^T\hat{p} = 0$. Then suppose $\sin(\gamma\rho) \neq 0$, (19) yields:

$$\mathbf{x}_1^T R^T W^T \hat{p}\mathbf{x}_2 = 0 \quad \Leftrightarrow \quad \mathbf{x}_1^T R^T (I_3 + (\cos(\gamma\rho) - 1)pp^T)\hat{p}\mathbf{x}_2 \quad \Leftrightarrow \quad \mathbf{x}_1^T R^T \hat{p}\mathbf{x}_2 = 0. \tag{21}$$

This is exactly the well-known bilinear **epipolar constraint** in Euclidean space (for a comparison see [5]). Here we see that this constraint also holds in spaces of constant curvature. Notice that in the Euclidean case the matrix $E = R^T\hat{p}$ is called (normalized) **essential matrix.**

Comments 1 *The condition* $\sin(\gamma\rho) \neq 0$ *is equivalent to the condition the translation* $p \neq 0$ *in the Euclidean case. The reason is when* $\sin(\gamma\rho) = 0$*, we have* $\exp X = I_4$*, i.e. the motion is equivalent to the identity transformation on* M*. In spaces of constant curvature, we may have* $\sin(\gamma\rho) = 0$ *without* $p = 0$*. This occurs only when the curvature* r *is positive, i.e. the space is spherical. If so, let* $\gamma = 2k\pi\sqrt{r} \in \mathbb{R}, k = 1, 2, \ldots,$ *we then have* $\sin(\gamma\rho) = \sin(2k\pi) = 0$*. This implies that translation with distance* $2\pi\sqrt{r}$ *along the geodesics (big circles) in a spherical space of radius* $\sqrt{r}$ *is equivalent to the identity transformation (back to the initial position). One can simply check this phenomenon on the 2 dimensional sphere* S^2*.*

As a summary of the above discussion, we have the following theorem:

Theorem 3 (Epipolar Constraint) *Consider a rigid body motion of a camera in a space* M *of constant curvature. If* $p \in \mathbb{R}^3$ *is the vector associated to the direction of the translation and* $R \in SO(3)$ *the rotation, then the images* $\mathbf{x}_1 \in \mathbb{R}^3$ *and* $\mathbf{x}_2 \in \mathbb{R}^3$ *of a point* $q \in M$ *before and after the motion satisfy the epipolar constraint:*

$$\mathbf{x}_1^T R^T \hat{p}\mathbf{x}_2 = 0. \tag{22}$$

As in the Euclidean case, the normalized essential matrix E can be estimated from more than eight image correspondences $\{(\mathbf{x}_1^j, \mathbf{x}_2^j)\}_{j=1}^n, n \geq 8$ in general positions using linear or nonlinear estimation schemes [8, 6]. The rotation matrix R and the translation vector p can then be recovered from the essential matrix E [5].

Comments 2 *Notice that the epipolar constraint is independent of the scale* λ *of the point* q*, the scale* γ *of the translational motion* b *and the curvature* $1/r$ *of the space* M*. The motion recovery is then decoupled from the 3D structure, as in the Euclidean case.*

It is already known that in the Euclidean case, m images of a point satisfy more general multilinear constraints besides the bilinear epipolar constraint. Similar constraints exist in the case of spaces of constant curvature. Suppose $\mathbf{x}_i \in \mathbb{R}^3, i = 1, 2, \ldots, m$ are m images of the same point q with the camera at m different position. Suppose the relative motion between the i^{th} and $(i-1)^{th}$ positions is $g_i \in G, i = 1, 2, \ldots, m$. Without loss of generality, we may always assume $g_1 = I$. Let $\lambda_i \in \mathbb{R}^+, i = 1, 2, \ldots, m$ be the scales of $\mathbf{x}_i, i = 1, 2, \ldots, m$ with respect to q. Then we have the following equation:

$$\begin{pmatrix} \lambda_1 \mathbf{x}_1 \\ \lambda_2 \mathbf{x}_2 \\ \vdots \\ \lambda_m \mathbf{x}_m \end{pmatrix} = \begin{pmatrix} P g_1 \\ P g_2 g_1 \\ \vdots \\ P g_m \cdots g_1 \end{pmatrix} q.$$

Now define the **motion matrix** $A \in \mathbb{R}^{3m \times 4}$ to be:

$$A = \begin{pmatrix} P g_1 \\ P g_2 g_1 \\ \vdots \\ P g_m \cdots g_1 \end{pmatrix} \in \mathbb{R}^{3m \times 4}$$

and the four columns of A are denoted by a_1, a_2, a_3, a_4 respectively. Define the vector $\tilde{\mathbf{x}}_i \in \mathbb{R}^{3m}$ associated to the i^{th} image $\mathbf{x}_i$ as:

$$\tilde{\mathbf{x}}_i = (0, \ldots, 0, \mathbf{x}_i^T, 0, \ldots, 0)^T \in \mathbb{R}^{3m}, \quad 1 \leq i \leq m.$$

Similar to the Euclidean case [4], in spaces of constant curvature, we also have:

Theorem 4 (Projective Constraint) *Consider m images $\{\mathbf{x}_i\}_{i=1}^m \in \mathbb{R}^3$ of a point q in a space M of constant curvature, and the motion matrix is $A = (a_1, a_2, a_3, a_4) \in \mathbb{R}^{3m \times 4}$ as defined above. Then the associated vectors $\{\tilde{\mathbf{x}}_i\}_{i=1}^m \in \mathbb{R}^{3m}$ satisfy the following wedge product equation:*

$$a_1 \wedge a_2 \wedge a_3 \wedge a_4 \wedge \tilde{\mathbf{x}}_1 \wedge \ldots \wedge \tilde{\mathbf{x}}_m = 0. \tag{23}$$

The proof is essentially the same as in the Euclidean case [4]. The reason that this wedge product constraint is called **projective constraint** is because it is invariant under projective transformation (see [4]). For the same reasons as in Euclidean case, the non-trivial constraints given by the wedge product equation are either **bilinear, trilinear** or **quadrilinear** [4]. One may use these constraints to design more delicate motion estimation schemes.

4 Structure From Motion

Knowing the motion, the next problem is how to reconstruct the scale information from images. The scale information includes the depth λ of the point q with respect to its image $\mathbf{x}$, the scale of the translational motion p and if possible the constant curvature $1/r$ of the space M (but we will soon see, the curvature cannot be recovered from vision). Although our formulation allows to study reconstruction from multiple image frames, we here only study the case of two image frames. To generalize to the case of multiple image frames, one may refer to [4].

To simplify the notation, in this section, we assume the image $\mathbf{x}$ of a point q is always normalized, *i.e.* $\|\mathbf{x}\| = 1$ (in the Euclidean case, this corresponds to the spherical projection). Suppose the distance from q to the optical center o is $\eta \in \mathbb{R}^+$. Recall that $\phi(\cdot)$ is the map from $T_o(M)$ to $\mathfrak{m}$. Then the homogeneous coordinate of q is given in terms of $\mathbf{x}$ and η by:

$$q = (\exp \eta\phi(\mathbf{x})) \cdot o = \begin{pmatrix} \rho^{-1}\sin(\eta\rho)\mathbf{x} \\ \cos(\eta\rho) \end{pmatrix} \in \mathbb{R}^4.$$

Consequently, the scale λ of q with respect to $\mathbf{x}$ is given by $\lambda = \rho^{-1}\sin(\eta\rho)$. To differentiate from the scale λ, the distance quantity η will be called the **depth** of the point q with respect to the image $\mathbf{x}$.

Let η_1 and η_2 be the depths of the point q with respect its two images $\mathbf{x}_1$ and $\mathbf{x}_2$ taken by the camera at two positions, respectively. Suppose the camera motion $g \in G$ is specified by the rotation $R \in SO(3)$, the translation direction $p \in S^2$ and the scale of translation γ (as in the preceding section). Then the first equation in (19) yields:

$$\rho^{-1}\sin(\eta_2\rho)\mathbf{x}_2 = \left(I_3 + (\cos(\gamma\rho) - 1)pp^T\right) R \cdot \rho^{-1}\sin(\eta_1\rho)\mathbf{x}_1 + \cos(\eta_1\rho)\rho^{-1}\sin(\gamma\rho)p. \tag{24}$$

This is the **coordinate transformation formula in spaces of constant curvature.** It looks kind of complicated. However, it is no more than a natural generalization of the Euclidean coordinate transformation formula which people are with. Notice when the curvature $1/r$ goes to zero, so does ρ. Since

$$\lim_{\rho\to 0}\cos(x\rho) = 1, \quad \lim_{\rho\to 0}\rho^{-1}\sin(x\rho) = x, \quad x \in \mathbb{R},$$

then when the curvature of the space goes to zero, we have:

$$\lambda_i = \lim_{\rho\to 0}\rho^{-1}\sin(\eta_i\rho) = \eta_i, \quad i = 1, 2,$$

and (24) simply becomes:

$$\lambda_2\mathbf{x}_2 = R\lambda_1\mathbf{x}_1 + \gamma p. \tag{25}$$

That it, in the limit case, the scale λ and the depth η are the same; and the equation (24) gives the Euclidean coordinate transformation formula. The Euclidean transformation (25) is extensively used for reconstructing Euclidean structure [4]. Naturally, to reconstruct structure in spaces of constant curvature, the equation (24) has to be exploited.

Notice equation (24) is homogeneous in the scale of ρ. Since the quantities η_1, η_2 and γ are all multiplied with ρ, they can only be determined with respect to an arbitrary scale of ρ. In Euclidean case, this corresponds to the fact that the Euclidean structure can only be reconstructed up to a universal scale [4]. Thus in the case of spaces of constant curvature, we may normalize everything with respect to the scale of the curvature: if $r > 0$, let $\rho = 1$; if $r < 0$, let $\rho = i = \sqrt{-1}$. That is, now the space M has constant sectional curvature of either $+1$ or -1. Then (24) becomes:

$$\begin{aligned} \sin(\eta_2)\mathbf{x}_2 &= \left(I_3 + (\cos(\gamma) - 1)pp^T\right) R \cdot \sin(\eta_1)\mathbf{x}_1 + \cos(\eta_1)\sin(\gamma)p, \quad \rho = 1; && (26) \\ \sinh(\eta_2)\mathbf{x}_2 &= \left(I_3 + (\cosh(\gamma) - 1)pp^T\right) R \cdot \sinh(\eta_1)\mathbf{x}_1 + \cosh(\eta_1)\sinh(\gamma)p, \quad \rho = i. && (27) \end{aligned}$$

These two equations correspond to coordinate transformations in (normalized) spherical and hyperbolic spaces, respectively.

From the preceding section, we know R and p can be estimated from epipolar constraints. The problem left is to reconstruct η_1, η_2 and γ. In computer vision, this problem is usually referred to as **structure from motion** (this name is used by some authors for the problem of reconstructing both motion and structure from images, but we shall maintain the distinction). One may directly use the above coordinate transformation formula to formulate objective function for estimating scales η_1, η_2 and γ. In the Euclidean case, such objective functions are linear in the scales [4]. However, in the Non-Euclidean case, such objective functions are usually nonlinear.

In stead of directly using the coordinate transformation formula, one may use some well-known constraints in spaces of constant curvature, *i.e.* **Bolyai's law of sine** and **law of cosine** (for absolute geometry), which have been well summarized by Hsiang in [1]. Define functions:

$$\alpha(x) = \begin{cases} \sin(x), & \rho = 1, \\ \sinh(x), & \rho = i, \end{cases} \qquad \beta(x) = \begin{cases} \cos(x), & \rho = 1, \\ \cosh(x), & \rho = i. \end{cases}$$

The next theorem follows from Hsiang [1] as a special case:

Theorem 5 (Laws of Absolute Trigonometry) *Consider a geodesic triangle $\triangle ABC$ in a space M of constant curvature ± 1, and let a, b, c be the lengths of the opposite sides of angles A, B, C respectively. Then we have:*

$$\frac{\sin(A)}{\alpha(a)} = \frac{\sin(B)}{\alpha(b)} = \frac{\sin(C)}{\alpha(c)}, \qquad \textit{Bolyai's sine law.} \tag{28}$$

and

$$\begin{aligned} \alpha(a)\alpha(b)\cos(C) &= \beta(c) - \beta(a)\beta(b), \\ \alpha(b)\alpha(c)\cos(A) &= \beta(a) - \beta(b)\beta(c), \qquad \textit{law of cosine} \\ \alpha(c)\alpha(a)\cos(B) &= \beta(b) - \beta(c)\beta(a). \end{aligned} \tag{29}$$

Suppose the two optical centers of the camera are o_1 and o_2. A geodesic triangle is formed by the three points (o_1, o_2, q), see Figure 1. The angle A is given by the angle between the two

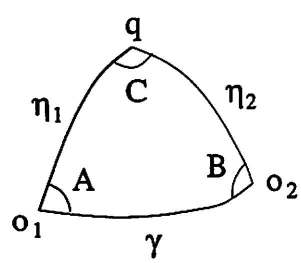

Figure 1: Geodesic triangle formed by two optical centers o_1, o_2 and a point q in the scene.

vectors $R\mathbf{x}_1$ and $-p$; B is given by the angle between $\mathbf{x}_2$ and p; C is given by the angle between $R\mathbf{x}_1$ and $\mathbf{x}_2$. The quantities $\sin(A), \sin(B), \sin(C), \cos(A), \cos(B), \cos(C)$ can be directly calculated from those vectors.

Applying Bolyai's sine law (28) to the geodesic triangle, $\alpha(\eta_1), \alpha(\eta_2)$ and $\alpha(\gamma)$ are determined up to a unknown scalar $k \in \mathbb{R}$ by linear equations:

$$\sin(A)\alpha(\eta_1) = \sin(B)\alpha(\eta_2), \quad \sin(C)\alpha(\eta_2) = \sin(A)\alpha(\gamma). \tag{30}$$

The scalar k can be then determined by using one of the cosine law (29). Suppose

$$(s_1, s_2, s_3)^T = (k\alpha(\eta_1), k\alpha(\eta_2), k\alpha(\gamma))^T \in \mathbb{R}^3$$

is a solution of (30). In the hyperbolic case, from the first equation of (29), the scalar k satisfies:

$$s_1 s_2 \cos(C) = k\sqrt{s_3^2 - k^2} - \sqrt{(s_1^2 - k^2) \cdot (s_2^2 - k^2)}. \tag{31}$$

In the spherical case, we may assume $0 \leq \eta_1, \eta_2, \gamma \leq \pi/2$ (*i.e.* comparing to the size of the whole space, the structure we consider is relatively small). Then the first equation of (29) yields:

$$s_1 s_2 \cos(C) = k\sqrt{k^2 - s_3^2} - \sqrt{(k^2 - s_1^2) \cdot (k^2 - s_2^2)}. \tag{32}$$

In order to calculate k, the above equations can be easily reduced to algebraic equations in k^2 of degree 4. Since there is a general formula for roots of algebraic equations of degree 4, k **has a closed-form solution.** Knowing k, $\alpha(\eta_1), \alpha(\eta_2)$ and $\alpha(\gamma)$ can be calculated hence η_1, η_2 and γ. This approach is obviously easier than directly optimizing the (multi-variable) objective functions associated with the coordinate transformation formula.

5 Differential Case

Suppose the motion of the camera is given as $g(t), t \in [t_0, t_f]$, a smooth curve in the isometry group G. Without loss of generality we may assume $g(t_0) = I$. Then according to (15) and (17), for the image $\mathbf{x} \in \mathbb{R}^3$ of a point $q = (q_1, q_2, q_3, q_4)^T \in M$:

$$\lambda(t)\mathbf{x}(t) = Pg(t)q, \quad t_0 \leq t \leq t_f. \tag{33}$$

Since $g(t_0) = I$, the derivative $\dot{g}(t_0)$ is an element in the Lie algebra $\mathfrak{g}$ hence it has the form:

$$\dot{g}(t_0) = \begin{pmatrix} \hat{\omega} & v \\ -v^T/r & 0 \end{pmatrix} \in \mathbb{R}^{4\times 4}$$

where $\omega, v \in \mathbb{R}^3$. The vector ω then corresponds to the angular velocity of the camera and v the linear velocity. Now differentiate (33) at time $t = t_0$ then we have:

$$\dot{\lambda}\mathbf{x} + \lambda\dot{\mathbf{x}} = \hat{\omega}\lambda\mathbf{x} + q_4 v \quad \Rightarrow \quad v \times \dot{\lambda}\mathbf{x} + v \times \lambda\dot{\mathbf{x}} = v \times \hat{\omega}\lambda\mathbf{x} \quad \Rightarrow \quad \dot{\mathbf{x}}^T\hat{v}\mathbf{x} + \mathbf{x}^T\hat{v}\hat{\omega}\mathbf{x} = 0. \tag{34}$$

This is exactly the differential version of the bilinear epipolar constraint in the Euclidean case. It also holds in spaces of constant curvature.

Theorem 6 (Differential Epipolar Constraint) *Consider a moving camera in a space M of constant curvature. If $v \in \mathbb{R}^3$ is the linear velocity and $\omega \in \mathbb{R}^3$ the angular velocity, then the image $\mathbf{x} \in \mathbb{R}^3$ of a point $q \in M$ and its optical flow $\dot{\mathbf{x}} \in \mathbb{R}^3$ satisfy the differential epipolar constraint:*

$$\dot{\mathbf{x}}^T\hat{v}\mathbf{x} + \mathbf{x}^T\hat{v}\hat{\omega}\mathbf{x} = 0. \tag{35}$$

Given more than eight optical flow measurements $\{(\mathbf{x}^j, \dot{\mathbf{x}}^j)\}_{j=1}^n, n \geq 8$ in general position, the velocities ω and v are recoverable with v determined up to a scale using the same linear or nonlinear estimation schemes designed for Euclidean case [5, 6].

As a generalization of the bilinear epipolar constraint, we give the multilinear constraints satisfied by higher order optical flow. At time $t = t_0$, differentiating the equation (33) $(m-1)$ times, we obtains:

$$\begin{pmatrix} \mathbf{x} & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \dot{\mathbf{x}} & \mathbf{x} & 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \cdots & \cdots & \vdots \\ \mathbf{x}^{(i)} & \vdots & c_j^i \mathbf{x}^{(i-j)} & \ddots & \ddots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \cdots & \ddots & \ddots & \vdots \\ \mathbf{x}^{(m-2)} & \cdots & \cdots & \cdots & \cdots & \mathbf{x} & 0 \\ \mathbf{x}^{(m-1)} & \cdots & \cdots & \cdots & \cdots & \cdots & \mathbf{x} \end{pmatrix} \begin{pmatrix} \lambda \\ \dot{\lambda} \\ \vdots \\ \lambda^{(i)} \\ \vdots \\ \lambda^{(m-2)} \\ \lambda^{(m-1)} \end{pmatrix} = \begin{pmatrix} Pg \\ P\dot{g} \\ \vdots \\ Pg^{(i)} \\ \vdots \\ Pg^{(m-2)} \\ Pg^{(m-1)} \end{pmatrix} q.$$

where $c_j^i = \begin{pmatrix} i \\ j \end{pmatrix} \in \mathbb{Z}^+$ for $0 \leq j \leq i \leq (m-1)$. The quantities $\mathbf{x}^{(i)}, 0 \leq i \leq (m-1)$ are the i^{th} order derivatives of the image point. If we define $c_j^i = 0$ for $i < j$, the $(i,j)^{th}$ entry (in fact a tuple) of the first matrix in the above equation has the unified form $c_j^i \mathbf{x}^{(i-j)}$, $0 \leq i, j \leq (m-1)$. We may define matrices $\mathbf{U} \in \mathbb{R}^{3m \times m}$ and $B \in \mathbb{R}^{3m \times 4}$:

$$\mathbf{U} = (c_j^i \mathbf{x}^{(i-j)}), \quad B = (Pg^{(i)}), \quad 0 \leq i, j \leq (m-1). \tag{36}$$

Let $\tilde{\mathbf{u}}_i \in \mathbb{R}^{3m}, 1 \leq i \leq m$ be the i^{th} column of the matrix $\mathbf{U}$ and $b_1, b_2, b_3, b_4 \in \mathbb{R}^{3m}$ be the four columns of the matrix B. We then have the differential version of the Theorem 4.

Theorem 7 *Consider the image $\mathbf{x}(t) \in \mathbb{R}^3$ of a point q under the camera motion $g(t) \in SE(3)$. Then for the matrices $\mathbf{U} \in \mathbb{R}^{3m \times m}$ and $B \in \mathbb{R}^{3m \times 4}$ defined in (36), the column vectors $\{\tilde{\mathbf{u}}_i\}_{i=1}^m \in \mathbb{R}^{3m}$ of the matrix $\mathbf{U}$ and the column vectors $b_1, b_2, b_3, b_4 \in \mathbb{R}^{3m}$ of the matrix B satisfy the following wedge product equation:*

$$b_1 \wedge b_2 \wedge b_3 \wedge b_4 \wedge \tilde{\mathbf{u}}_1 \wedge \ldots \wedge \tilde{\mathbf{u}}_m = 0. \tag{37}$$

The proof is the same as the Euclidean case. See [4] for the proof and for more detailed discussion of this wedge product equation.

Remark 1 *Similarly, all the results about hybrid cases (when both image correspondences and optical flows are available), which we have discussed in [4], can be similarly generalized to the case of spaces of constant curvature. The structure from motion problem can also be generalized to the differential case and hybrid cases in a similar fashion. We here do not discuss them in detail.*

6 Discussions and Conclusions

In this paper, we have generalized basic vision theorems in Euclidean space to spaces of constant curvature. A uniform treatment is possible because a unified homogeneous representation of these spaces exists and the isometry groups of these spaces have similar structures. As we have seen, the Euclidean vision theory can always be viewed as a limit case of the general one.

One may have noticed that the epipolar geometry in spaces of constant curvature is remarkably similar to that of Euclidean space. Especially, the bilinear epipolar constraint is exactly the same. As in the Euclidean case, the motion is nicely decoupled from structure by the epipolar constraint. This allows us to use most of the motion recovery algorithms which were previously developed only for Euclidean space to spherical and hyperbolic spaces, without any modification. In the differential case, the epipolar geometry also remains to be the same as in Euclidean case.

As for the structure from motion problem, the three dimensional structure can only be reconstructed up to a universal scale, the same as the Euclidean case [4]. In a space of non-zero curvature, the curvature of the space can not be recovered from vision. However, the three dimensional structure of objects can be determined with respect to the curvature. In this paper, we normalize the curvature with absolute value 1. Although the structure from motion is a linear problem in the Euclidean case, it is no longer linear in spherical and hyperbolic spaces. It is shown in the paper that using sine and cosine laws for Absolute Geometry there is a closed-form solution for the structure from motion problem.

Although any Riemannian manifold locally can be approximated by spaces of constant curvature, it is still interesting to know if the results of epipolar geometry hold for more general classes of Riemannian manifolds (for example, symmetric spaces); and how the structure from motion problem needs to be changed in general. These will be our research topics in the future.

7 Acknowledgment

We would like to thank professor Alan Weinstein (Berkeley Mathematics Department) for invaluable discussions during the preparation of this manuscript.

References

[1] Wu-Yi Hsiang. Absolute geometry revisited, Center for Pure and Applied Mathematics, University of California at Berkeley. *PAM-628*, 1995.

[2] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume I. John Wiley & Sons, Inc., 1996.

[3] S. Kobayashi and K. Nomizu. *Foundations of Differential Geometry*, volume II. John Wiley & Sons, Inc., 1996.

[4] Yi Ma, Jana Kosecka, and Shankar Sastry. Euclidean structure and motion from image sequences. *UC Berkeley, ERL*, 1998.

[5] Yi Ma, Jana Kosecka, and Shankar Sastry. Motion recovery from image sequences: Discrete viewpoint vs. differential viewpoint. In *UC Berkeley Memorandum No. UCB/ERL M98/11*, 1998.

[6] Yi Ma, Jana Kosecka, and Shankar Sastry. Optimal motion from image sequences: A Riemannian viewpoint. In *submitted to IEEE Conference on Decision and Control*, 1998.

[7] Richard M. Murray, Zexiang Li, and Shankar Sastry. *A Mathematical Introduction to Robotic Manipulation.* CRC Press, 1993.

[8] G. Toscani and O. D. Faugeras. Structure and motion from two noisy perspective images. *Proceedings of IEEE Conference on Robotics and Automation*, pages 221–227, 1986.

[9] Joseph A. Wolf. *Spaces of Constant Curvature.* Publish or Perish, Inc., 5th edition, 1984.