

# A Theory of Camera-Independent Correspondence

József Molnár<sup>1</sup>, Dmitry Chetverikov<sup>2,3</sup>, Zoltán Kató<sup>4</sup>, and Dániel Baráth<sup>2,3</sup>

<sup>1</sup> MTA SZBK, Szeged, Hungary

<sup>2</sup> MTA SZTAKI, Budapest, Hungary

csetverikov@sztaki.hu

<sup>3</sup> ELTE, Budapest, Hungary

<sup>4</sup> SZTE, Szeged, Hungary

**Abstract.** Projective geometry is a standard mathematical tool for image-based 3D reconstruction. Most reconstruction methods establish pointwise image correspondences using projective geometry. We present an alternative approach based on differential geometry of a surface observed by any camera, existing or potential, that satisfies very general conditions, namely, the differentiability of the surface and the bijective projection functions. Considering two views of the surface, we derive the pose equation that can be used to determine the relative pose of the two cameras. Then we discuss the generalized epipolar geometry and derive the generalized epipolar constraint along the epipolar curves. Applying the proposed theory to projective camera and assuming that affine mapping between small corresponding regions has been estimated, we obtain the minimal pose equation for the case when a fully calibrated camera is moved with its internal parameters unchanged. Equations for the projective epipolar constraint and the fundamental matrix are also derived. Then, the special cases of normalized coordinates and rectified image pair are discussed. Finally, we present test results for pose estimation showing that our solution is correct and operational.

## 1 Introduction

Most approaches to multi-view stereo reconstruction [15], [4], [5] use projective, affine or weak perspective camera models [6]. Solutions for central and non-central catadioptric cameras [17] [10] are also available. Many methods search for pointwise image correspondences, but attempts to avoid correspondence, e.g. [7], have also been made.

Despite the great variety of the methods, almost all of them rely on projective geometry as a basic tool to describe relations between scene points and image points or establish correspondence between points in different views. This mainstream research has led to the development of solutions providing impressive results in both sparse and dense reconstruction of scenes and objects with varying geometry and surface properties. Applications to vision-based SLAM [8] have also resulted in significant improvement in localization and mapping by mobile devices, autonomous robots and vehicles.

Differential properties of surfaces expressed by image gradients and affine distortions of local regions have been used in various areas related to 3D reconstruction. In particular, affine propagation of patch correspondences in wide-baseline stereo was proposed in [9]. A similar principle was successfully applied to multi-view stereo in [4].

The study [5] uses surface growing in multi-view reconstruction by image warping estimating the surface normal vector as a linear function of the camera matrix and the homography.

Affine-covariant regions and features [11] [18] [14] can be used to find image correspondences and estimate affine distortion of a surface patch between views. Alternatively, one can apply the correspondence-free approach [3] to register shapes and estimate local homography. In our study, we assume that such estimation has been done and the entries of the Jacobian describing the local mapping of two views are known.

Brightness and texture gradients reveal the surface geometry and are used in shape from shading and shape from texture, respectively [16]. These methods operate on single images and do not require correspondences.

In this paper, we consider a surface viewed by two cameras and derive relationships between local distortions of small corresponding regions, the parameters of the cameras and the local geometry of the surface in the regions. We present an alternative approach based on differential, rather than projective, geometry. In spirit, our theory is related to the work [2] that also relies on differential geometry. However, the study [2] considers only projective camera and uses a parameterization dependent, non-invariant representation, while we use a very general camera model and invariant representation.

The main contributions of this paper are as follows. The camera model we use is a mapping restricted only by the differentiability of the surface and the bijective projection functions. Projective, affine, weak-perspective and central and non-central catadioptric camera models are all special cases of our model. For this general model, we obtain the pose equation that can be used to calculate the relative pose of the cameras. Also, we derive the generalized epipolar constraint along the epipolar curves. For the special case of the widely applied projective camera model, the proposed theory results in the minimal pose equation that allows one to determine the new pose of a fully calibrated camera moved to another position with its internal parameters unchanged. Finally, we obtain equations for the projective epipolar constraint and the fundamental matrix.

The structure of the paper is as follows. Section 2 introduces notations and theoretical background. Then derivations and results for a surface observed by general camera are presented. Due to paper length limitations, we have to omit some technical details of lengthy derivations. The full version will be given in a forthcoming journal paper. In section 3, we apply the general theory to projective camera. Test results for pose estimation are shown and analyzed in section 4. Section 5 concludes the paper by discussion and outlook.

## 2 Theory for surface viewed by general camera

### 2.1 Basic equations

Consider an observed scene in the 3D space  $\mathbb{R}^3$ . The visible parts of the scene objects are treated as 2D surfaces embedded in  $\mathbb{R}^3$ . A standard basis in the space is defined by three orthonormal basis vectors  $\mathbf{i}$ ,  $\mathbf{j}$  and  $\mathbf{k}$ . For spatial coordinates, we use italic capital letters with superscripts:  $X^1, X^2, X^3$ ; for 3D vectors, we use bold capital letters,

while lowercase bold letters are used for 2D vectors. Homogeneous representations are marked with tilde to be distinguished from their inhomogeneous counterparts. Italic letters  $u^1, u^2$  are used for Gaussian point coordinates constrained to the embedded manifolds. Partial derivatives are denoted by subscripts.

Different images of a surface are distinguished with lower indices  $i, j$ ; only these two letters are used to identify the projection functions, any other letter in subscript means partial derivative. Scalar product between vectors is denoted by dot, vector product by cross. Triple scalar product of three vectors  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  is denoted by  $|\mathbf{abc}|$ .

Surfaces are parameterized using the general (Gaussian) coordinates:

$$\mathbf{S}(u^1, u^2) = X^1(u^1, u^2) \mathbf{i} + X^2(u^1, u^2) \mathbf{j} + X^3(u^1, u^2) \mathbf{k} \quad (1)$$

We assume that images of spatial points are projections given by two functions assigning two image coordinates  $(x^1, x^2)$  to spatial points. Spatial points lying on the surface are mapped onto the  $i$ -th image by composite functions

$$x_i^k = x_i^k(X^1(u^1, u^2), X^2(u^1, u^2), X^3(u^1, u^2)) = \hat{x}_i^k(u^1, u^2), \quad k = 1, 2 \quad (2)$$

To simplify notation, the hat in the right-hand side will be omitted. We suppose that the mappings in Eq. (2) are bijections in a small open disk around the point  $(u^1, u^2)$ . Assuming that both the projection functions and the surface are smooth, this is the condition for differentiability. The inverse functions  $u^1(x_i^1, x_i^2), u^2(x_i^1, x_i^2)$  of the bijective mappings also exist.

Consider a surface observed by two cameras that provide images  $i$  and  $j$ . A small shift on the surface results in small shifts  $\mathbf{dx}_i$  and  $\mathbf{dx}_j$  in the two images. As shown in [12], they are related as follows:

$$\mathbf{dx}_j = \mathbf{J}_{ij} \cdot \mathbf{dx}_i, \quad (3)$$

where the Jacobian of the image mapping  $i \rightarrow j$

$$\mathbf{J}_{ij} = \begin{bmatrix} \frac{\partial x_j^1}{\partial x_i^1} & \frac{\partial x_j^1}{\partial x_i^2} \\ \frac{\partial x_j^2}{\partial x_i^1} & \frac{\partial x_j^2}{\partial x_i^2} \end{bmatrix} = \begin{bmatrix} \frac{\partial x_j^1}{\partial u^1} & \frac{\partial x_j^1}{\partial u^2} \\ \frac{\partial x_j^2}{\partial u^1} & \frac{\partial x_j^2}{\partial u^2} \end{bmatrix} \begin{bmatrix} \frac{\partial x_i^1}{\partial u^1} & \frac{\partial x_i^1}{\partial u^2} \\ \frac{\partial x_i^2}{\partial u^1} & \frac{\partial x_i^2}{\partial u^2} \end{bmatrix}^{-1} \quad (4)$$

The equation is parameterized by  $(u^1, u^2)$ . We seek coordinate-independent, ‘invariant’ representation. The partial derivatives of any function  $f \in \{x_i^1, x_j^1, x_i^2, x_j^2\}$  can be written as

$$\frac{\partial f}{\partial u^k} = \frac{\partial X^1}{\partial u^k} \frac{\partial f}{\partial X^1} + \frac{\partial X^2}{\partial u^k} \frac{\partial f}{\partial X^2} + \frac{\partial X^3}{\partial u^k} \frac{\partial f}{\partial X^3} = \mathbf{S}_{u^k} \cdot \nabla f, \quad k = 1, 2, \quad (5)$$

where  $\mathbf{S}_{u^k}$  are the partial derivatives of the surface (1),  $\nabla f$  the spatial gradient of  $f$ . It has been shown in [12] that  $\mathbf{J}_{ij}$  can be expressed in invariant form as

$$\mathbf{J}_{ij} = \frac{1}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|} \begin{bmatrix} |\nabla x_j^1 \mathbf{n} \nabla x_i^2| & |\nabla x_i^1 \mathbf{n} \nabla x_j^1| \\ |\nabla x_j^2 \mathbf{n} \nabla x_i^2| & |\nabla x_i^1 \mathbf{n} \nabla x_j^2| \end{bmatrix}, \quad (6)$$

where  $|\nabla x_i^1 \mathbf{n} \nabla x_i^2|$  is the triple scalar product of the gradients and the normal unit vector  $\mathbf{n}$  of the surface.

## 2.2 Interpretation

Suppose the observed surface is parameterized by its image coordinates pushed forward to the surface. For example, image  $i$  induces the following parameterization:

$$\mathbf{S}(x_i^1, x_i^2) = X^1(x_i^1, x_i^2) \mathbf{i} + X^2(x_i^1, x_i^2) \mathbf{j} + X^3(x_i^1, x_i^2) \mathbf{k} \quad (7)$$

We wish the local basis  $\mathbf{S}_{1i} = \frac{\partial \mathbf{S}}{\partial x_i^1}$ ,  $\mathbf{S}_{2i} = \frac{\partial \mathbf{S}}{\partial x_i^2}$  to be expressed with invariants. (From now on, we will use the standard simplified notation  $\mathbf{S}_{1i} \equiv \mathbf{S}_{x_i^1}$ , etc.) Applying Eq. (5) to the coordinate functions  $x_i^1$  and  $x_i^2$  with  $u^1 = x_i^1$  and  $u^2 = x_i^2$ , we obtain

$$\mathbf{S}_p \cdot \nabla q = \delta_{pq}, p, q \in \{x_i^1, x_i^2\}, \quad (8)$$

where  $\delta_{pq}$  is the Kronecker delta. This fulfills the definition of the inverse basis for  $\nabla x_i^1, \nabla x_i^2$ . The inverse (contravariant) basis vectors will be denoted by  $\mathbf{S}_i^1, \mathbf{S}_i^2$ . Since they lie on the tangent plane of the surface, the following must hold:

$$\begin{aligned} \mathbf{S}_i^1 &= \nabla x_i^1|_T, \mathbf{S}_i^2 = \nabla x_i^2|_T \\ \nabla z|_T &= \nabla z \cdot (\mathbf{I} - \mathbf{nn}), z \in \{x_i^1, x_i^2\} \end{aligned} \quad (9)$$

Here  $\nabla z|_T$  is the projection of  $\nabla z$  to the tangent plane,  $\mathbf{I}$  the identity tensor,  $\mathbf{nn}$  the direct (dyadic) product. The cross-product of these contravariant vectors is perpendicular to the tangent plane, hence it is a surface normal with the length  $l_i = \mathbf{n} \cdot (\mathbf{S}_i^1 \times \mathbf{S}_i^2)$ . It can be easily shown that

$$l_i = |\nabla x_i^2 \mathbf{n} \nabla x_i^1|. \quad (10)$$

We observe that  $l_i$  equals the denominator in the Jacobian (6). Since the contravariant and covariant basis vectors are related as  $\mathbf{S}_{1i} = \frac{1}{l_i} (\mathbf{S}_i^2 \times \mathbf{n})$ ,  $\mathbf{S}_{2i} = \frac{1}{l_i} (\mathbf{n} \times \mathbf{S}_i^1)$ , we have

$$\begin{aligned} \mathbf{S}_{1i} &= \frac{1}{|\nabla x_i^2 \mathbf{n} \nabla x_i^1|} [\nabla x_i^2 - (\nabla x_i^2 \cdot \mathbf{n}) \mathbf{n}] \times \mathbf{n} = \frac{\mathbf{n} \times \nabla x_i^2}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|}, \\ \mathbf{S}_{2i} &= \frac{1}{|\nabla x_i^2 \mathbf{n} \nabla x_i^1|} \mathbf{n} \times [\nabla x_i^1 - (\nabla x_i^1 \cdot \mathbf{n}) \mathbf{n}] = \frac{\nabla x_i^1 \times \mathbf{n}}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|}. \end{aligned} \quad (11)$$

Any vector  $\mathbf{v}$  in the tangential plane can be decomposed in two ways:

$$\mathbf{v} = (\mathbf{v} \cdot \mathbf{S}^1) \mathbf{S}_1 + (\mathbf{v} \cdot \mathbf{S}^2) \mathbf{S}_2 = (\mathbf{v} \cdot \mathbf{S}_1) \mathbf{S}^1 + (\mathbf{v} \cdot \mathbf{S}_2) \mathbf{S}^2, \quad (12)$$

where  $v^1 = \mathbf{v} \cdot \mathbf{S}^1$ ,  $v^2 = \mathbf{v} \cdot \mathbf{S}^2$  are the contravariant,  $v_1 = \mathbf{v} \cdot \mathbf{S}_1$ ,  $v_2 = \mathbf{v} \cdot \mathbf{S}_2$  the covariant vector coordinates. Applying such decomposition to Eq. (3), the components of  $d\mathbf{x}_i = \mathbf{S}_{1i} dx_i^1 + \mathbf{S}_{2i} dx_i^2$  in projection  $j$  can be expressed as

$$dx_j^k = \mathbf{S}_j^k \cdot (\mathbf{S}_{1i} dx_i^1 + \mathbf{S}_{2i} dx_i^2), \quad k = 1, 2 \quad (13)$$

Using (9) and (11), the Jacobian (4) can be written as

$$\mathbf{J}_{ij} = \begin{bmatrix} \nabla x_j^1|_T \cdot \frac{(\mathbf{n} \times \nabla x_i^2)}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|} & \nabla x_j^1|_T \cdot \frac{(\nabla x_i^1 \times \mathbf{n})}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|} \\ \nabla x_j^2|_T \cdot \frac{(\mathbf{n} \times \nabla x_i^2)}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|} & \nabla x_j^2|_T \cdot \frac{(\nabla x_i^1 \times \mathbf{n})}{|\nabla x_i^1 \mathbf{n} \nabla x_i^2|} \end{bmatrix} \doteq \begin{bmatrix} a_1^1 & a_2^1 \\ a_1^2 & a_2^2 \end{bmatrix}. \quad (14)$$

This form, which is equivalent to Eq. (6), expresses the image mapping  $i \rightarrow j$  by invariant first-order differential quantities, the projection gradients and the unit normal vector. The symbols  $a_1^1, a_2^1, \dots$  are introduced to simplify notation. The components of  $\mathbf{J}_{ij}$  can be estimated from image correspondences.

Applying the decomposition Eq. (12) to the tangential vectors  $\nabla x_j^1|_T, \nabla x_j^2|_T$ , we obtain

$$\nabla x_j^k|_T = (\nabla x_j^k|_T \cdot \mathbf{S}_{1i}) \nabla x_i^1|_T + (\nabla x_j^k|_T \cdot \mathbf{S}_{2i}) \nabla x_i^2|_T, \quad k = 1, 2 \quad (15)$$

The expressions in brackets are the components of  $\mathbf{J}_{ij}$ , hence Eq. (15) can be rewritten as

$$\begin{bmatrix} \nabla x_j^1|_T \\ \nabla x_j^2|_T \end{bmatrix} = \mathbf{J}_{ij} \cdot \begin{bmatrix} \nabla x_i^1|_T \\ \nabla x_i^2|_T \end{bmatrix}, \quad (16)$$

which means that contravariant basis vectors transform as coordinate differentials. We call this important relation the **pose equation** for the reason that will be explained later. The equation states that the same relationship exists between two images of a surface as between projection gradients constrained to the tangent plane.

Using Eq. (9), Eq. (16) can be re-written as

$$\nabla x_j^k \cdot (\mathbf{I} - \mathbf{nn}) = a_1^k \nabla x_i^1 \cdot (\mathbf{I} - \mathbf{nn}) + a_2^k \nabla x_i^2 \cdot (\mathbf{I} - \mathbf{nn}), \quad k = 1, 2 \quad (17)$$

Taking the dot product of both sides with  $(\nabla x_i^1 \times \nabla x_i^2)$ , we have

$$\frac{1}{l_i} \left[ \frac{|\nabla x_j^1 \nabla x_i^1 \nabla x_i^2|}{|\nabla x_j^2 \nabla x_i^1 \nabla x_i^2|} \right] = \left[ \frac{\nabla x_j^1|_{\mathbf{n}}}{\nabla x_j^2|_{\mathbf{n}}} \right] - \mathbf{J}_{ij} \cdot \left[ \frac{\nabla x_i^1|_{\mathbf{n}}}{\nabla x_i^2|_{\mathbf{n}}} \right]. \quad (18)$$

The right-hand side of Eq. (18) is the counterpart of Eq. (16) in the normal direction. Recall that  $l_i$  was introduced in Eq. (10), while  $\nabla z|_{\mathbf{n}} = (\nabla z \cdot \mathbf{n}) \mathbf{n}$  is the projection of  $\nabla z, z \in \{x_i^1, x_i^2, x_j^1, x_j^2\}$ , to the normal direction. The left-hand side is the basic expression for the **epipolar geometry** to be discussed below.

### 2.3 Epipolar geometry

Now we impose further restrictions on the projection functions (2). We assume that each image point has a dedicated ray associated with it. The rays may not intersect, that is, points in space may not have same image coordinates, except for the case when they have common projection center. We emphasize that this does not necessarily mean central projection, since each image point may have its own origin denoted by  $\mathbf{C} = \mathbf{C}(x^1, x^2)$ . We only assume that origins and rays vary smoothly keeping all differentiability criteria valid.

A back-projected ray  $\mathbf{X}(t), t \in (0, \infty], \mathbf{X}(0) = \mathbf{C}$ , is characterized by constant image coordinates  $x^1(\mathbf{X}(t)) = (x^1)_0, x^2(\mathbf{X}(t)) = (x^2)_0$  for any ray parameter  $t$ . The derivative wrt  $t$  is  $\nabla x^k \cdot \dot{\mathbf{X}} = 0, k = 1, 2$ , where  $\dot{\mathbf{X}}(t) = \frac{d\mathbf{X}}{dt}$  is the direction of the ray. That is,  $\dot{\mathbf{X}}(t)$  is perpendicular to both gradients and

$$\dot{\mathbf{X}}(t) = c (\nabla x^1 \times \nabla x^2) \quad (19)$$

for any real constant  $c$ , which can be selected freely. Since the ray direction  $\frac{\dot{\mathbf{X}}(t)}{|\dot{\mathbf{X}}(t)|}$  is independent of  $t$ , the unit vector  $\frac{\nabla x^1 \times \nabla x^2}{|\nabla x^1 \times \nabla x^2|}$  depends only on the image coordinates  $(x^1)_0, (x^2)_0$ . Integrating this normalized version of Eq. (19), we obtain the **equation for back-projected ray**:

$$\mathbf{X}(t) = \mathbf{C} + \frac{\nabla x^1 \times \nabla x^2}{|\nabla x^1 \times \nabla x^2|} t = \mathbf{C} + \frac{\nabla x^1 \times \nabla x^2}{r} t, \quad r \doteq |\nabla x^1 \times \nabla x^2|, \quad (20)$$

where the constant vector  $\mathbf{C} = X(0)$  is the origin of the ray, the ‘projection center’ associated with the image coordinates  $(x^1)_0, (x^2)_0$ .

Observing by camera  $j$  a back-projected ray of camera  $i$ , we have the following correspondence equation:

$$x_j^k(t) = x_j^k \left( \mathbf{C}_i + \frac{1}{r_i} (\nabla x_i^1 \times \nabla x_i^2) t \right), \quad k = 1, 2 \quad (21)$$

Since the normalized cross product  $\frac{1}{r_i} (\nabla x_i^1 \times \nabla x_i^2)$  is independent of  $t$ ,

$$\frac{dx_j^k}{dt} = \nabla x_j^k \cdot \frac{\nabla x_i^1 \times \nabla x_i^2}{r_i}, \quad k = 1, 2 \quad (22)$$

From this, we obtain the first-order ordinary differential equation

$$\frac{dx_j^2}{dx_j^1} = \frac{|\nabla x_j^2 \nabla x_i^1 \nabla x_i^2|}{|\nabla x_j^1 \nabla x_i^1 \nabla x_i^2|} \quad (23)$$

expressed as a ratio of triple scalar products that contains neither  $t$  nor  $r_i$ . The initial condition is given by the ‘epipoles’  $x_j^2(x_j^1((\mathbf{C}_i))) = x_j^2(\mathbf{C}_i)$ , and solution associating possible image coordinate pairs  $(x_j^1, x_j^2(x_j^1))$  to the image point  $(x_i^1, x_i^2)$  is uniquely defined.

According to Eq. (18), the differential equation compatible with Eq. (6) can be expressed via image gradients and the entries of  $\mathbf{J}_{ij}$ :

$$\frac{dx_j^2}{dx_j^1} = \frac{\mathbf{n} \cdot (\nabla x_j^2 - a_1^2 \nabla x_i^1 - a_2^2 \nabla x_i^2)}{\mathbf{n} \cdot (\nabla x_j^1 - a_1^1 \nabla x_i^1 - a_2^1 \nabla x_i^2)} \quad (24)$$

Eq. (24) can be considered as **generalized epipolar constraint** since it provides equations for the components of  $\mathbf{J}_{ij}$ , i.e., the components of  $\mathbf{J}_{ij}$  are not independent along the **epipolar curves**. Examples will be given in section 3.

In the case of **central projection** with constant  $\mathbf{C}_i$  and  $\mathbf{C}_j$ , the vector  $(\mathbf{C}_i - \mathbf{C}_j)$  and the two rays  $(\nabla x_i^1 \times \nabla x_i^2), (\nabla x_j^1 \times \nabla x_j^2)$  define the **epipolar plane**. Its images are the above mentioned epipolar curves. With an epipolar plane given, the two associated epipolar curves are defined by

$$\frac{dx_i^2}{dx_i^1} = \frac{|\nabla x_i^2 \nabla x_j^1 \nabla x_j^2|}{|\nabla x_i^1 \nabla x_j^1 \nabla x_j^2|}, \quad x_i^2(x_i^1(\mathbf{C}_j)) = x_i^2(\mathbf{C}_j), \quad (25)$$

and similarly for  $j$ , with  $i$  and  $j$  swapped. Any observed object point on an epipolar plane has two projected points on its associated epipolar curves. Searching a point along the corresponding epipolar curves means searching an object point on the epipolar plane.

### 3 Application to projective camera

As long as the differentiability criteria are valid, the presented theory does not assume any particular camera model. Below, we apply the theory to finite projective CCD camera because of its practical importance. In this case, the projection matrix  $\mathbf{P} = \mathbf{K} \cdot [\mathbf{R}, \mathbf{t}]$  where  $\mathbf{K}$  is an upper-triangular matrix,  $\mathbf{R}$  the rotation matrix,  $\mathbf{t}$  the translation vector. In homogeneous coordinates, a spatial point  $\mathbf{X}$  is projected onto image point  $\mathbf{x}$  as

$$\tilde{\mathbf{x}} = \mathbf{K}^{-1} \cdot \mathbf{P} \cdot \tilde{\mathbf{X}}, \quad (26)$$

where  $\tilde{\mathbf{X}} = [X^1 \ X^2 \ X^3 \ 1]^T$  and  $\tilde{\mathbf{x}} = s [x^1 \ x^2 \ 1]^T$  with unknown scale factor  $s$ . In practice, the skew-free (CCD) camera model is widely used. In this case  $\mathbf{K}$  and  $\mathbf{K}^{-1}$  take simple form

$$\mathbf{K} = \begin{bmatrix} \alpha & 0 & u^1 \\ 0 & \beta & u^2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}^{-1} = \begin{bmatrix} \frac{1}{\alpha} & 0 & -\frac{u^1}{\alpha} \\ 0 & \frac{1}{\beta} & -\frac{u^2}{\beta} \\ 0 & 0 & 1 \end{bmatrix}. \quad (27)$$

Introduce  $\boldsymbol{\rho}^k = [r_1^k \ r_2^k \ r_3^k]$  for the  $k$ -th row of the rotation matrix. Then the projection function becomes

$$\begin{aligned} x^k &= \frac{1}{s} [(\beta \boldsymbol{\rho}^k + u^k \boldsymbol{\rho}^3) \cdot \mathbf{X} + p_4^2], \quad k = 1, 2 \\ s &= \boldsymbol{\rho}^3 \cdot \mathbf{X} + p_4^3 \end{aligned} \quad (28)$$

with  $\mathbf{X} = [X^1 \ X^2 \ X^3]^T$  and  $\mathbf{K} \cdot \mathbf{t} = [p_4^1 \ p_4^2 \ p_4^3]^T$ . The gradient components are

$$\begin{aligned} \nabla x^1 &\rightarrow \frac{\partial x^1}{\partial X^k} = \frac{1}{s} [\alpha r_k^1 - (x^1 - u^1) r_k^3], \\ \nabla x^2 &\rightarrow \frac{\partial x^2}{\partial X^k} = \frac{1}{s} [\beta r_k^2 - (x^2 - u^2) r_k^3], \quad k = 1, 2, 3. \end{aligned} \quad (29)$$

The following problems can be addressed using the proposed theory: 1. **Reprojection.** For a calibrated camera system and an approximately reconstructed surface, transformation between images can be estimated to evaluate similarity and refine the surface. This problem is considered in [12]. 2. **Normal vector calculation.** For a calibrated camera system and estimated Jacobian (14), the surface normal vector can be computed, enabling reconstruction from sparse correspondences. The Jacobian is the local affine transformation with the two origins aligned, which can be estimated by different means [11], [18], [3]. 3. **Pose estimation.** For one fully calibrated camera and another one with only internal parameters known, the pose of the second camera can be calculated given the Jacobian. Below, we address the third problem assuming that the Jacobian components  $a_1^1, a_2^1, \dots$  have been estimated.

#### 3.1 Pose estimation

Assume a camera had been calibrated, then moved with the internal parameters unchanged. Without loss of generality, we can suppose that camera  $i$  has been calibrated

to the origin of the tangent plane  $\mathbf{n} = \mathbf{k}$  ( $Z = 0$ ). Then the pose equation (16) becomes

$$\nabla x_j^k|_T = a_1^k \nabla x_i^1|_T + a_2^k \nabla x_i^2|_T, \quad k = 1, 2. \quad (30)$$

The right-hand side has known entries, the parameters of the completely calibrated camera and the estimated Jacobian components. The left hand side has 7 unknowns, 6 components of the rotation matrix and  $s$ . The number of equations available is also 7: 4 independent equations (30) written for the tangential ( $k = 1, 2$ ) components of (29), and the constraint on the rotational matrix properties, i.e., the norms of the columns are 1 and their dot product is zero. Equations (30) can therefore be considered as **minimal pose equations**.

Since all unknowns are in camera  $j$ , in the equations below we omit this index. Introduce  $\mathbf{r}_k = [r_k^1 \ r_k^2 \ r_k^3]^T$ ,  $k = 1, 2, 3$ , for the  $k$ -th column of  $\mathbf{R}$  in the decomposition  $\mathbf{P} = \mathbf{K} \cdot [\mathbf{R}, \mathbf{t}]$ . The right-hand side of Eq. (30) can be given in the standard basis. Denote these components by  $A_l^k$ ,  $k, l = 1, 2$ :

$$a_1^k \nabla x_i^1|_T + a_2^k \nabla x_i^2|_T \doteq A_1^k \mathbf{i} + A_2^k \mathbf{j}, \quad k = 1, 2, \quad (31)$$

where  $A_l^k$  are known. Using properties of  $\mathbf{R}$  and (27), (31), one can derive

$$\begin{aligned} (B_1^1 s + C_1^1 r_1^3)^2 + (B_1^2 s + C_1^2 r_1^3)^2 + (r_1^3)^2 &= 1, \\ (B_2^1 s + C_2^1 r_2^3)^2 + (B_2^2 s + C_2^2 r_2^3)^2 + (r_2^3)^2 &= 1, \\ (B_1^1 s + C_1^1 r_1^3)(B_2^1 s + C_2^1 r_2^3) + (B_1^2 s + C_1^2 r_1^3)(B_2^2 s + C_2^2 r_2^3) + r_1^3 r_2^3 &= 0. \end{aligned} \quad (32)$$

Here we introduced notations  $B_k^1 \doteq \frac{1}{\alpha} A_k^1$ ,  $B_k^2 \doteq \frac{1}{\beta} A_k^2$ ,  $k = 1, 2$ ,  $C^1 \doteq \frac{1}{\alpha} (x^1 - u^1)$ ,  $C^2 \doteq \frac{1}{\alpha} (x^2 - u^2)$ .  $r_k^i$  is the element of  $\mathbf{R}$  in  $i$ -th row and  $k$ -th column.

The first two equations in (32) can be parametrically solved for  $r_1^3$  and  $r_2^3$  as functions of  $s$ , then the absolute value of the left-hand side in the third equation can be used as error function for  $s$ . Fixed-length iteration can be used. The maximum value for  $s$  is estimated as the lower bound of the two discriminants of the first two equations (32). Finally, 4 solutions are available for positive  $s$ , from which the unique solution can be chosen by reprojection.

### 3.2 Epipolar lines

For projective camera, the gradients are

$$s \nabla x^l = \mathbf{p}^l - x^l \mathbf{p}^3, \quad l = 1, 2 \quad s = \mathbf{p}^3 \cdot \mathbf{X} + p_4^3, \quad (33)$$

where  $(\mathbf{p}^T)^k = [p_1^k \ p_2^k \ p_3^k]$ ,  $k = 1, 2, 3$ , is the  $k$ -th row of the left  $3 \times 3$  submatrix of  $\mathbf{P}$ . In Eq. (23), the scale factors  $s_i, s_j$  are eliminated:

$$\frac{dx_j^2}{dx_j^1} = \frac{|\nabla x_j^2 \nabla x_i^1 \nabla x_i^2|}{|\nabla x_j^1 \nabla x_i^1 \nabla x_i^2|} = \frac{(\mathbf{p}_j^2 - x_j^2 \mathbf{p}_j^3) \cdot [(\mathbf{p}_i^1 - x_i^1 \mathbf{p}_i^3) \times (\mathbf{p}_i^2 - x_i^2 \mathbf{p}_i^3)]}{(\mathbf{p}_j^1 - x_j^1 \mathbf{p}_j^3) \cdot [(\mathbf{p}_i^1 - x_i^1 \mathbf{p}_i^3) \times (\mathbf{p}_i^2 - x_i^2 \mathbf{p}_i^3)]}. \quad (34)$$

This can be re-arranged as

$$\frac{x_j^2 - \frac{(x_i^1 D_{23}^2 - x_i^2 D_{13}^2 + D_{12}^2)}{(x_i^1 D_{23}^3 - x_i^2 D_{13}^3 + D_{12}^3)}}{x_j^1 - \frac{(x_i^1 D_{23}^1 - x_i^2 D_{13}^1 + D_{12}^1)}{(x_i^1 D_{23}^3 - x_i^2 D_{13}^3 + D_{12}^3)}} = \frac{x_j^2 - d_3^2}{x_j^1 - d_3^1}, \quad (35)$$

where

$$d_3^k \doteq \frac{(x_i^1 D_{23}^k - x_i^2 D_{13}^k + D_{12}^k)}{(x_i^1 D_{23}^3 - x_i^2 D_{13}^3 + D_{12}^3)}, \quad k = 1, 2.$$

Here the notation  $D_{mn}^l = |\mathbf{p}_j^l \mathbf{p}_i^m \mathbf{p}_i^n|$ ,  $l, m, n \in \{1, 2, 3\}$  was introduced for triple scalar products with the first vector from camera  $j$  and two vectors from camera  $i$ . For a fixed image point  $(x_i^1, x_i^2)$  whose corresponding epipolar line is sought in image  $j$ , the expression (35) is a function of  $(x_j^1, x_j^2)$  and

$$\frac{dx_j^2}{dx_j^1} = \frac{x_j^2 - d_3^2}{x_j^1 - d_3^1}, \quad (36)$$

with the point  $(d_3^1, d_3^2)$  lying on the epipolar line.

O.d.e. (36) is separable in its variables, and its general solution

$$x_j^2 = \kappa x_j^1 + (d_3^2 - \kappa d_3^1) \quad (37)$$

is a one-parameter family of straight lines with the slope  $\kappa$ . For a particular solution we need an initial value condition to be satisfied. Denote the epipole coordinates by  $e_j^1, e_j^2$ . Then the initial condition is  $e_j^2 = \kappa e_j^1 + (d_3^2 - \kappa d_3^1)$ ,  $\kappa = \frac{e_j^2 - d_3^2}{e_j^1 - d_3^1}$  and Eq. (37) transforms to

$$(e_j^1 - d_3^1) x_j^2 - (e_j^2 - d_3^2) x_j^1 + (e_j^2 d_3^1 - e_j^1 d_3^2) = 0. \quad (38)$$

Any of the following ratios expresses the same property, the slope  $\kappa$  of the epipolar line:

$$\frac{e_j^2 - d_3^2}{e_j^1 - d_3^1} = \frac{x_j^2 - d_3^2}{x_j^1 - d_3^1} = \frac{e_j^2 - x_j^2}{e_j^1 - x_j^1} \quad (39)$$

All of them lead to the same solution (38).

Eq. (38) is related to the **fundamental matrix**. It can be written in the form expressing that three points are on a same line:

$$\det \begin{bmatrix} x_j^1 & x_j^2 & 1 \\ e_j^1 & e_j^2 & 1 \\ d_3^1 & d_3^2 & 1 \end{bmatrix} = 0, \quad (40)$$

or, equivalently, using the notation of Eq. (35)

$$\tilde{\mathbf{x}}_j \cdot [\tilde{\mathbf{e}}_j]_{\times} \cdot \begin{bmatrix} D_{23}^1 & -D_{13}^1 & D_{12}^1 \\ D_{23}^2 & -D_{13}^2 & D_{12}^2 \\ D_{23}^3 & -D_{13}^3 & D_{12}^3 \end{bmatrix} \cdot \tilde{\mathbf{x}}_i = 0 \rightarrow \tilde{\mathbf{x}}_j \cdot \mathbf{F} \cdot \tilde{\mathbf{x}}_i = 0 \quad (41)$$

Here the fundamental matrix appears in the factorized form  $\mathbf{F} = [\mathbf{e}]_{\times} \cdot \mathbf{H}$  with the homography  $\mathbf{H}$ . The properties  $\text{rank}(\mathbf{F}) = 2$  and  $\mathbf{e}_j \cdot \mathbf{F} = 0$  are obvious.

Applying Eq. (24) to Eq. (33), we obtain

$$\mathbf{n} \cdot (\kappa \nabla x_j^1 - \nabla x_j^2 + a_1^2 \nabla x_i^1 + a_2^2 \nabla x_i^2 - \kappa a_1^1 \nabla x_i^1 - \kappa a_2^1 \nabla x_i^2) = 0. \quad (42)$$

Substituting (33) and (39), we have

$$\begin{aligned} s_j \mathbf{n} \cdot [(a_1^2 - \kappa a_1^1) (\mathbf{p}_i^1 - x_i^1 \mathbf{p}_i^3) + (a_2^2 - \kappa a_2^1) (\mathbf{p}_i^2 - x_i^2 \mathbf{p}_i^3)] = \\ s_i \mathbf{n} \cdot [\mathbf{p}_j^2 - \kappa \mathbf{p}_j^1 + (\kappa e_j^1 - e_j^2) \mathbf{p}_j^3] \end{aligned} \quad (43)$$

$s_i, s_j$  are the homogeneous scale factors (projective depths) for cameras  $i$  and  $j$ . Since the equation must hold for any normal unit vector, including  $\mathbf{n} = \mathbf{i}, \mathbf{j}, \mathbf{k}$ , we have three equations from which two independent ratios can be used to eliminate the projective depths. These two equations represent the **epipolar constraint** on the components of  $\mathbf{J}_{ij}$ , reducing its DOF to two.

For normalized coordinates, however,  $s_i = d_i, s_j = d_j$  become ‘real’ Euclidean depths, and their ratio has a well-defined meaning. We consider two special cases of the epipolar constraint, for normalized coordinates and for rectified image pair.

For calibrated cameras, we can **normalize** image coordinates and projection matrix:

$$\begin{aligned} \bar{\mathbf{x}} &= (\mathbf{K}^{-1} \cdot \mathbf{P}) \cdot \bar{\mathbf{X}}, \\ \bar{\mathbf{P}} &= \mathbf{K}^{-1} \cdot \mathbf{P} = [\mathbf{R}, -\mathbf{RC}], \end{aligned} \quad (44)$$

where  $\bar{a}$  denotes normalization. Note that any  $\lambda \bar{\mathbf{P}}, \lambda \neq 0$ , is a possible choice for the normalized projection matrix, but the specific representation can easily be chosen forcing the determinant of the  $3 \times 3$  left submatrix of  $\bar{\mathbf{P}}$  to be 1. Denote the coordinates for this special case by  $\bar{\mathbf{X}} = [X \ Y \ Z]^T$  and  $\bar{\mathbf{x}}_i = [x_i \ y_i]^T$ ,  $\bar{\mathbf{x}}_j = [x_j \ y_j]^T$ . Using notation similar to Eq. (33), we have

$$s \nabla x = \rho^1 - x \rho^3, \quad s \nabla y = \rho^2 - y \rho^3, \quad s = \rho^3 \cdot \bar{\mathbf{X}} + \rho_4^3, \quad (45)$$

where  $\rho^k$  is the  $k$ -th row of  $\mathbf{R}$ . The following properties hold:

$$\begin{aligned} \det \mathbf{R} &= |\rho^1 \rho^2 \rho^3| = 1 \\ \rho^1 \times \rho^2 &= \rho^3, \quad \rho^2 \times \rho^3 = \rho^1, \quad \rho^3 \times \rho^1 = \rho^2, \quad \rho^l \cdot \rho^k = \delta_{lk} \\ s &= d \end{aligned} \quad (46)$$

The projective depth now becomes the distance  $d$  to the principal plane of the camera. The specific form of Eq. (43) is

$$\begin{aligned} \frac{d_j}{d_i} \mathbf{n} \cdot [(a_1^2 - \kappa a_1^1) (\rho_i^1 - x_i \rho_i^3) + (a_2^2 - \kappa a_2^1) (\rho_i^2 - y_i \rho_i^3)] = \\ \mathbf{n} \cdot [\rho_j^2 - \kappa \rho_j^1 + (\kappa e_j^1 - e_j^2) \rho_j^3]. \end{aligned} \quad (47)$$

To simplify Eq. (47), we can choose the world coordinate system to coincide with that of camera  $i$ :  $\rho_i^1 = \mathbf{i}$ ,  $\rho_i^2 = \mathbf{j}$ ,  $\rho_i^3 = \mathbf{k}$ . Then

$$\begin{aligned} \frac{d_j}{d_i} \mathbf{n} \cdot [(a_1^2 - \kappa a_1^1) (\mathbf{i} - x_i \mathbf{k}) + (a_2^2 - \kappa a_2^1) (\mathbf{j} - y_i \mathbf{k})] = \\ \mathbf{n} \cdot [\rho_j^2 - \kappa \rho_j^1 + (\kappa e_j^1 - e_j^2) \rho_j^3]. \end{aligned} \quad (48)$$

Component-wise, applying the normals  $\mathbf{n} = \mathbf{i}, \mathbf{j}, \mathbf{k}$ , we have

$$\begin{aligned} \frac{d_j}{d_i} (a_1^2 - \kappa a_1^1) &= r_{1j}^2 - \kappa r_{1j}^1 + (\kappa e_j^1 - e_j^2) r_{1j}^3, \\ \frac{d_j}{d_i} (a_2^2 - \kappa a_2^1) &= r_{2j}^2 - \kappa r_{2j}^1 + (\kappa e_j^1 - e_j^2) r_{2j}^3, \\ \frac{d_j}{d_i} [-x_i (a_1^2 - \kappa a_1^1) - y_i (a_2^2 - \kappa a_2^1)] &= r_{3j}^2 - \kappa r_{3j}^1 + (\kappa e_j^1 - e_j^2) r_{3j}^3. \end{aligned} \quad (49)$$

$\rho_j^1, \rho_j^2, \rho_j^3$  are the rows of the relative rotation matrix  $\mathbf{R}_j = [r_{kj}^i], i, k = 1, 2, 3$ .

For known camera poses and selected (fixed) image point  $(x_i, y_i)$ , Eq. (48) provides three equations (49). One of them can be solved for  $\frac{d_j}{d_i}$ . Eliminating this parameter, we have two equations for the four entries of the Jacobian. They can be parameterized by the two components of the unit normal vector.

**Rectified image pair** can be characterized by two special camera matrices and image coordinate system with origin in the optical center:

$$\mathbf{P}_i = \mathbf{K} [\mathbf{I}, \mathbf{0}] = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{P}_j = \mathbf{K} [\mathbf{I}, -d\mathbf{i}] = \begin{bmatrix} \alpha & 0 & 0 & -\alpha d \\ 0 & \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \quad (50)$$

Using the finite CCD model (27), we have  $\mathbf{p}_j^1 = \mathbf{p}_i^1 = [\alpha \ 0 \ 0]$ ,  $\mathbf{p}_j^2 = \mathbf{p}_i^2 = [0 \ \beta \ 0]$ ,  $\mathbf{p}_j^3 = \mathbf{p}_i^3 = [0 \ 0 \ 1]$ .

Two trivial observations can be made for any imaged spatial point, namely,  $x_j^2 = x_i^2$  and  $s_j = s_i$ . The slope parameter  $\kappa$  given by Eq. (35) is zero:  $\kappa = 0$ . Since  $\mathbf{p}_j^r = \mathbf{p}_i^r$ , Eq. (43) becomes

$$\mathbf{n} \cdot [a_1^2 (\mathbf{p}_i^1 - x_i^1 \mathbf{p}_i^3) + (a_2^2 - 1) (\mathbf{p}_i^2 - x_i^2 \mathbf{p}_i^3)] = 0. \quad (51)$$

In the directions  $\mathbf{i}, \mathbf{j}, \mathbf{k}$  this yields, respectively,

$$\begin{aligned} a_1^2 \alpha &= 0 \Rightarrow a_1^2 = 0, \\ (a_2^2 - 1) \beta &= 0 \Rightarrow a_2^2 = 1, \\ a_1^2 (u^1 - x_i^1) + (a_2^2 - 1) (u^2 - x_i^2) &= 0. \end{aligned} \quad (52)$$

Note that the third condition is satisfied by the solutions of the first two, expressing the fact that the depth parameters are identical:  $s_j = s_i$ . The epipolar constraint-compatible Jacobian is therefore written as

$$\mathbf{J}_{ij} = \begin{bmatrix} a_1^1 & a_2^1 \\ 0 & 1 \end{bmatrix}. \quad (53)$$

It has two degrees of freedom. Note that this result can be obtained directly from the correspondence equation (6). In this case, the epipolar constraint and the correspondence equation are identical. The correspondence equation can also be used to translate parameterization (53) into parameterization with components of the unit normal vector. This has been done by purely geometric considerations in [9].

## 4 Tests

This paper is essentially theoretical. We propose a novel theoretical framework providing an alternative to the mainstream approach. The sole purpose of the initial tests presented in this section is to demonstrate that our theory is technically correct and operational. We use synthetic data and projective camera model to test the minimal pose equation (30) applying the solution (32). A fully calibrated virtual camera views a virtual, elliptical surface patch from a randomly generated position on a plane. Then the camera is randomly moved to another position on the plane preserving the visibility of the patch. A lower and an upper limit on the distance between the two positions were introduced to avoid too close and too far views. The precise Jacobian components  $a_1^1, a_2^1, \dots$  were calculated based on the known geometry of the stereo pair and the patch.

To simulate the imprecision of the Jacobian estimation, random noise was added to the patch contour points in the second view. Then the normalized DLT algorithm [6] for planar homography estimation was applied between the two views. For each noise level, 100 sets of perturbed Jacobians were obtained. For each set, the camera generation procedure was repeated 100 times resulting in 100 camera pairs viewing the patch. In each trial, the relative pose of the second camera was calculated as proposed and compared to the ground truth.

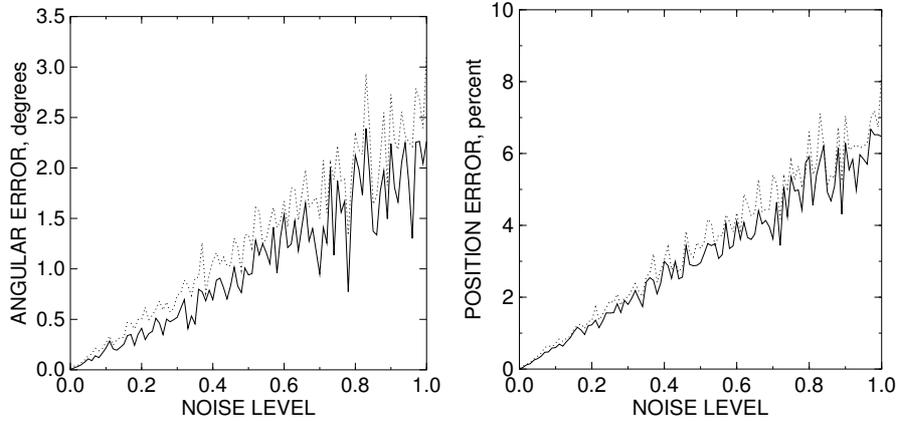
Recall that the Eq. (32) has four solutions, and the solution with the smallest reprojection error is selected. By setting an error threshold, we excluded the cases when the smallest reprojection error is still too large. In such cases, which were rare (less than 5%), the proposed method may not provide an acceptable solution. A major source of the potential failures is a poor estimate of the homography, which is not a part of the proposed theory.

The mean and the median errors of the 100 trials for each noise level were obtained. Both values were averaged over the 100 different camera pairs. Fig. 1 shows the plots of the angular and position errors for varying noise level which is the variance of the Gaussian noise, in pixels. The continuous line is the averaged median, the dotted line the averaged mean. The position error of the second camera is measured as the percentage of the distance between the patch and the camera. The angular error was obtained as follows. Given the ideal rotation matrix  $\mathbf{R}^{\text{id}}$  and the estimated matrix  $\mathbf{R}^{\text{es}}$ , we calculated the correction matrix  $\mathbf{R}^{\text{cr}}$  that relates the ideal and the estimated matrices:  $\mathbf{R}^{\text{id}} = \mathbf{R}^{\text{es}}\mathbf{R}^{\text{cr}}$ . Then the angle of the axis-angle representation [1] was obtained as

$$\theta = \arccos \frac{\text{trace } \mathbf{R}^{\text{cr}} - 1}{2}.$$

The absolute value of this angle was used as the angular error.

Analyzing Fig. 1, we observe that in the noise-free case the errors are zero, that is, the estimates are precise demonstrating that the proposed theory is technically correct.



**Fig. 1.** Plots of angular (left) and position (right) errors. Continuous line: averaged median. Dotted line: averaged mean.

The small difference between the averaged median and the averaged mean indicates that imposing an upper limit on the smallest reprojection error efficiently filters out the rare cases when the proposed method may become unreliable.

## 5 Discussion and conclusion

Traditional approaches to image correspondence are based on projective geometry that operates with points and lines to obtain the fundamental matrix or the trifocal tensor. The proposed alternative approach uses differential geometry and operates with two-dimensional entities, small surface patches. The correspondence equation (6) is valid when the surface is close to the tangent plane, and the derivatives of the projection functions are approximately constants. However, for projective camera viewing a planar patch, the Jacobian can be *exactly* determined from homography. This means that for flat surfaces the proposed theory provides exact solution to the surface normal and camera pose estimation problems.

Recently, we have applied the general theory to different kinds of camera models. Results for 3D reconstruction of planar patches viewed by omnidirectional cameras appeared in our study [13]. A promising direction of research could be the development of a second-order theory of image correspondence along the lines proposed in [12]. The first-order theory allows for camera pose estimation. Additive second-order entries could possibly bring additional information allowing for planar autocalibration with less images than the current approaches. A complete reconstruction pipeline could be built based exclusively on the proposed theory and its second-order extension.

## References

1. H.M. Choset. *Principles of robot motion: theory, algorithms, and implementation*. MIT Press, 2005.
2. F. Devernay and O. Faugeras. Computing differential properties of 3-D shapes from stereoscopic images without 3-D models. In *Conf. on Computer Vision and Pattern Recognition*, pages 208–213. IEEE, 1994.
3. C. Domokos, J. Nemeth, and Z. Kato. Nonlinear shape registration without correspondences. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 34(5):943–958, 2012.
4. Y. Furukawa and J. Ponce. Accurate, dense, and robust multiview stereopsis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 32(8):1362–1376, 2010.
5. M. Habbecke and L. Kobbelt. A surface-growing approach to multi-view stereo reconstruction. In *Conf. on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
6. R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2005.
7. K.N. Kutulakos and S.M. Seitz. A theory of shape by space carving. In *Proc. International Conf. on Computer Vision*, volume 1, pages 307–314, 1999.
8. T. Lemaire, C. Berger, I.-K. Jung, and S. Lacroix. Vision-based SLAM: Stereo and monocular approaches. *International Journal of Computer Vision*, 74(3):343–364, 2007.
9. Z. Megyesi, G. Kós, and D. Chetverikov. Dense 3D reconstruction from images by normal aided matching. *Machine Graphics & Vision*, 15:3–28, 2006.
10. B. Micusik and T. Pajdla. Autocalibration and 3D reconstruction with non-central catadioptric cameras. In *Conf. on Computer Vision and Pattern Recognition*, volume 1, pages I–58, 2004.
11. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool. A comparison of affine region detectors. *International Journal of Computer Vision*, 65:43–72, 2005.
12. J. Molnár and D. Chetverikov. Quadratic transformation for planar mapping of implicit surfaces. *Journal of Mathematical Imaging and Vision*, 48:176–184, 2014.
13. J. Molnár, R. Frohlich, D. Chetverikov, and Z. Kató. 3D Reconstruction of Planar Patches Seen by Omnidirectional Cameras. In *Proc. International Conf. on Digital Image Computing: Techniques and Applications*, 2014. Accepted for publication.
14. Oxford University, Katholieke Universiteit Leuven, INRIA, Center for Machine Perception. Affine Covariant Features. [www.robots.ox.ac.uk/~vgg/research/affine/](http://www.robots.ox.ac.uk/~vgg/research/affine/), 2007.
15. S.M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. In *Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 519–528, 2006.
16. M. Sonka, V. Hlavac, and R. Boyle. *Image Processing, Analysis, and Machine Vision*. Thomson, 2008.
17. T. Svoboda and T. Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 49(1):23–37, 2002.
18. T. Tuytelaars and K. Mikolajczyk. Local invariant feature detectors: a survey. *Foundations and Trends in Computer Graphics and Vision*, 3(3):177–280, 2008.