DESCRIPTIVE GEOMETRY MEETS COMPUTER VISION — THE GEOMETRY OF MULTIPLE IMAGES

Hellmuth STACHEL
Vienna University of Technology, Austria

ABSTRACT: The geometry of multiple images has been a standard topic in Descriptive Geometry and Photogrammetry (Remote Sensing) for more than 100 years. During the last twenty years great progress has been made within the field of Computer Vision, a topic with the main goal to endow a computer with a sense of vision. The previously graphical or mechanical methods of reconstruction have been replaced by mathematical methods as offered by computer algebra systems. This paper will explain to geometers how to reconstruct two digital images of the same scene and how to recover metrical data of the depicted object — using standard software only. Not the presented results are new, but the way how they are deduced by geometric reasoning. The arguments are based on Linear Algebra and classical Descriptive Geometry results.

Section 1 deals with the difference between central perspectives and general linear images, i.e., between calibrated and uncalibrated images. Suppose a point $\mathbf{x}$ is imaged in two views, at $\mathbf{x}'$ in the first, and $\mathbf{x}''$ in the second. What is the relation between these corresponding image points $\mathbf{x}', \mathbf{x}''$? This will be explained in Section 2; the required relation called 'epipolar constraint' is based on the essential matrix. In Section 3 the problem of reconstruction is addressed, for the calibrated case as well as for uncalibrated images. For given epipolar constraint the reconstruction of the depicted scene is possible up to a collinear transform in the uncalibrated case and up to the scale for calibrated images. The related theorems are already 100 years old. Hence, the crucial point is the determination of the essential matrix. This problem, which is related to the classical Problem of Projectivity, is solved in Section 4. The paper ends with an algorithm which can be carried out with any computer algebra system like Maple.

Keywords: Descriptive Geometry, multiple images, two-views-system, essential matrix.

1. INTRODUCTION

Central projection:
The basic term in this paper is the *central projection* or linear perspective with *center* $\mathbf{z}$ and *image plane* π (see Fig. 1). This is the geometric idealization of the photographic mapping with $\mathbf{z}$ as the focal point or focal center of the lenses and π as the plane of the film or CCD sensor. The pedal point of $\mathbf{z}$ with respect to π is called *principal point* $\mathbf{h}$; the distance $d := \|\mathbf{z} - \mathbf{h}\|$ is the *focal length*.

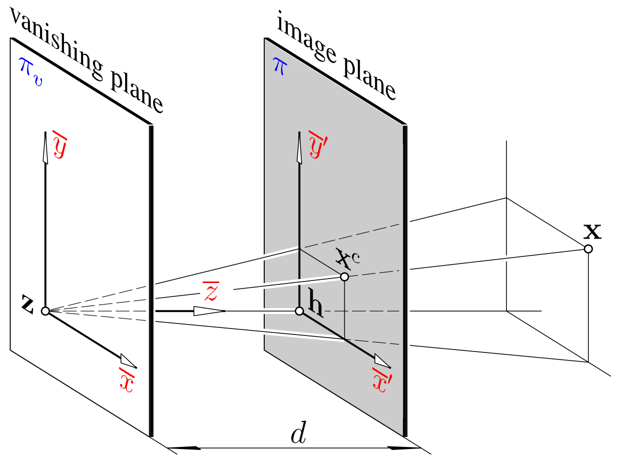

Figure 1: Central projection

Each central projection or photographic mapping defines a particular coordinate system in space, the *camera frame*. Its origin is placed at the center $\mathbf{z}$, the principal ray of the camera is the $\overline{z}$-axis. And the principal directions in the photosensitive plane serve as $\overline{x}$- and $\overline{y}$-axis. These coordinate axes span the vanishing plane π_v .

When at the same time the principal point $\mathbf{h}$ is the origin of 2D-coordines $(\overline{x}', \overline{y}')$ in the image plane, then the photographic mapping $\mathbf{x} \mapsto \mathbf{x}'$ obeys the matrix equation

$$\begin{pmatrix} \overline{x}' \\ \overline{y}' \end{pmatrix} = \frac{d}{\overline{z}} \begin{pmatrix} \overline{x} \\ \overline{y} \end{pmatrix}.$$

It is appropriate to introduce homogeneous 2D-coordinates $(\overline{x}'_0 : \overline{x}'_1 : \overline{x}'_2)$ by

$$\overline{x}' = \frac{\overline{x}'_1}{\overline{x}'_0}, \quad \overline{y}' = \frac{\overline{x}'_2}{\overline{x}'_0}.$$

In the same way we use homogeneous 3D-coordinates obeying

$$(\overline{x}_0 : \overline{x}_1 : \overline{x}_2 : \overline{x}_3) = (1 : \overline{x} : \overline{y} : \overline{z}).$$

Then the linear perspective is expressed as a linear mapping

$$\begin{pmatrix} \overline{x}'_0 \\ \overline{x}'_1 \\ \overline{x}'_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & d & 0 & 0 \\ 0 & 0 & d & 0 \end{pmatrix} \cdot \begin{pmatrix} \overline{x}_0 \\ \vdots \\ \overline{x}_3 \end{pmatrix}.$$

Now we bring this in a more general form: We replace the camera frame by arbitrary world coordinates (x, y, z). And we admit that in the image plane π our particular frame is modified by a translation and by scalings to the system of (x', y')-coordinates. Then we come up with the general form of mapping equations:

$$\begin{pmatrix} x'_0 \\ x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ h'_x & d\,f_x & 0 \\ h'_y & 0 & d\,f_y \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ z_x & & & \\ z_y & & R & \\ z_z & & & \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ \vdots \\ x_3 \end{pmatrix}.$$

On the right hand side there is a triple product of matrices. The first matrix contains beside the focal distance d the new image coordinates (h'_x, h'_y) of the principal point $\mathbf{h}$ and the two scaling factors f_x, f_y which usually are set to 1 . These entries are called the *intrinsic calibration parameters* of the photo. An image where this parameters are known is called *calibrated*. In this case the image determines the bundle of rays $\mathbf{z} \vee \mathbf{x}$ up to a rigid spatial motion.

The last matrix in the equation above contains the orthogonal matrix R and the world coordinates (z_x, z_y, z_z) of the center $\mathbf{z}$. This defines the position of the camera frame with respect to the world coordinates; the involved entries are called *extrinsic calibration parameters*.

We can generalize the central projection by a *central axonometry*. It maps the 3-space by a (singular) collinear transformation into the image plane. Hence, collinearity of points remains invariant and cross ratios are preserved. In homogeneous coordinates a central axonometry can be expressed by a linear mapping, and therefore the images are called *linear images*. There are several results on how to characterize perspective views among linear images (see, e.g., [6, 13, 7, 11, 2, 12, 8]).

In the generic case linear images are *uncalibrated*. Such a linear image can, e.g., be obtained by taking a photo of a given photo. It can be proved that a linear image of a scene is always an *affine* transform of a central perspective of the same scene.

Singular value decomposition:
One technical tool from Linear Algebra, which will be used in the sequel, is the *singular value decomposition* of any matrix A. It expresses A as a matrix product

$$A = U \cdot \mathrm{diag}(\lambda_1, \dots, \lambda_r) \cdot V^T$$

with orthogonal U, V, i.e., $U^{-1} = U^T$ and $V^{-1} = V^T$. The non-zero entries $(\lambda_1, \dots, \lambda_r)$ in the main diagonal of the central matrix are called the *singular values* of A. They are uniquely determined as the positive square roots of the eigenvalues of the symmetric $A^T \cdot A$, which are non-negative.

There is an instructive geometric interpretation of this decomposition in dimension 2 (see Fig. 2) which can easily be generalized into the Euclidean n-space: Matrix A represents an affine transformation which maps any unit circle into an ellipse which might be degenerated. There are pairwise orthogonal diameters of the unit circle which are mapped onto the axes of symmetry of the corresponding ellipse. These particular frames define the directions of principal distortion for this affine map.

The singular values of A equal the semi-axes of the ellipse. Therefore the singular values are sometimes called the *principal distortions* of this affine map. The orthogonal matrices U and V^T represent the coordinate transformations between the given frames and that of the principal distortion directions.

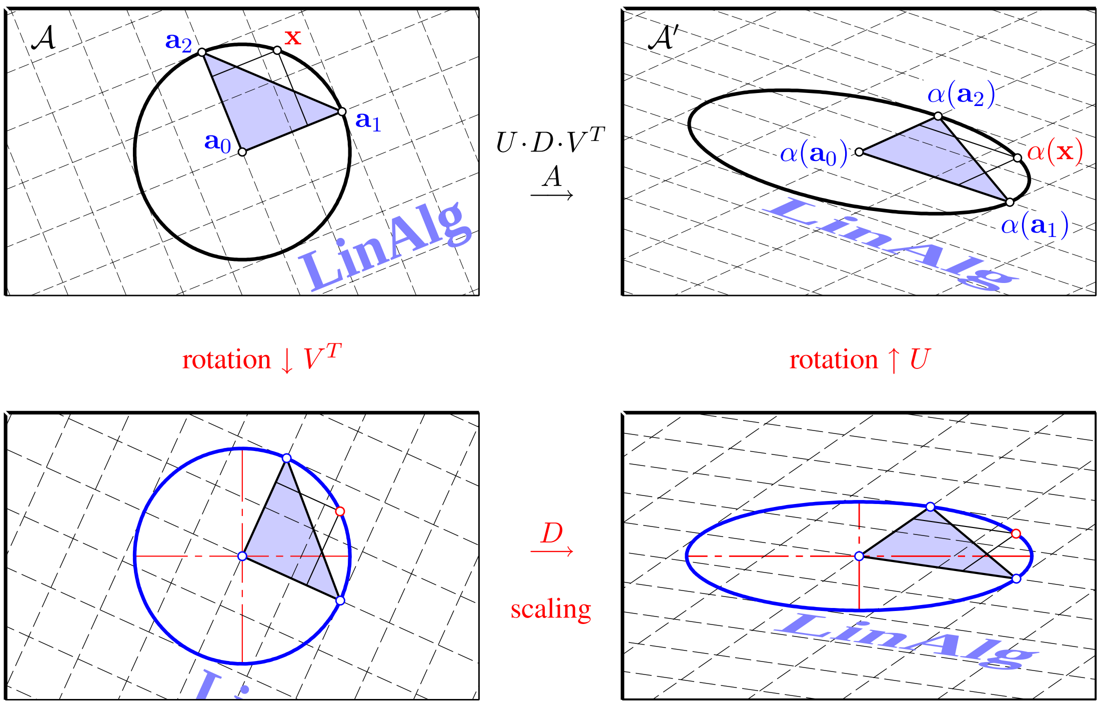

Figure 2: Singular value decomposition

2. THE GEOMETRY OF IMAGE PAIRS

The geometry of pairs of central views has been a classical subject of Descriptive Geometry. Important results are, e.g., due to S. FINSTERWALDER, E. KRUPPA [9], J. KRAMES, W. WUNDERLICH, H. BRAUNER [1].

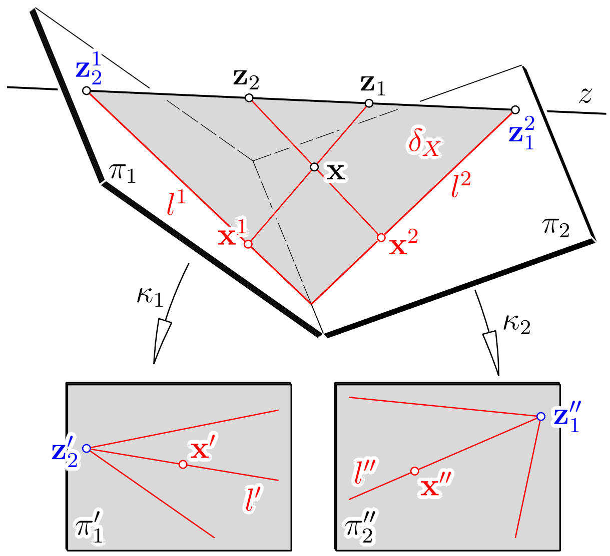

Figure 3: Epipolar constraints in a two-views-system

Uncalibrated case:
Let two central projections be given with centers $\mathbf{z}_i$ and image plane π_i for $i = 1, 2$. This refers to the *viewing situation* in 3-space as displayed in Fig. 3. In addition, let κ_1, κ_2 be collinear transformations which map the images into π_1' and π_2'', respectively. In this way we have defined a *general two-views-system* consisting of two linear images. Any space point $\mathbf{x}$ different from the two centers is represented by its views $\mathbf{x}', \mathbf{x}''$.

The basic geometric property of two-views-systems results from the fact that for space points $\mathbf{x}$ which are not aligned with the two centers, the two rays of sight $\mathbf{z}_1 \vee \mathbf{x}$ and $\mathbf{z}_2 \vee \mathbf{x}$ are coplanar (see Fig. 3). They are located in a plane $\delta_\mathbf{x}$ which in both linear images appears in an edge view. In the viewing situation the images of the pencil of planes $\delta_\mathbf{x}$ constitute two *perspective* line pencils. After applying the collinear transformations κ_1, κ_2 there remain to *projective* line pencils, the socalled *epipolar lines*. The centers $\mathbf{z}_2'$ and $\mathbf{z}_1''$ of these pencils are called *epipoles*. As expressed in the notation, each epipole is the image of one center under the other projection. The projectivity is called *epipolar constraint*. We summarize:

Theorem 1: *1) For any two linear images of a scene there is a projectivity between two particular line pencils*

$$\mathbf{z}_2'(\delta_\mathbf{x}') \;\overline{\wedge}\; \mathbf{z}_1''(\delta_\mathbf{x}'')$$

such that two points $\mathbf{x}', \mathbf{x}''$ are corresponding, i.e., images of the same space point, if and only if they are located on corresponding epipolar lines.

2) Using homogeneous coordinates, there is a matrix $B = (b_{ij})$ of rank 2 such that two points $\mathbf{x}' = (x_0' : x_1' : x_2')$ and $\mathbf{x}'' = (x_0'' : x_1'' : x_2'')$ are corresponding if and only if

$$\sum_{i,j=0}^{2} b_{ij}\, x_i'\, x_j'' = \mathbf{x}'^T \cdot B \cdot \mathbf{x}'' = 0\,. \qquad (1)$$

The vanishing bilinear form in (1) defines a correlation which is singular because of the rank deficiency of the socalled *essential matrix* B.

Proof: Using homogeneous line coordinates, the projectivity between the line pencils can be expressed as

$$(\mathbf{u}_1'\lambda_1 + \mathbf{u}_2'\lambda_2)\mathbb{R} \;\mapsto\; (\mathbf{u}_1''\lambda_1 + \mathbf{u}_2''\lambda_2)\mathbb{R}$$

for $(\lambda_1, \lambda_2) \in \mathbb{R}^2 \setminus \{(0,0)\}$. $\mathbf{x}'$ and $\mathbf{x}''$ are corresponding iff there is a nontrivial pair (λ_1, λ_2) such that

$$\begin{aligned} (\mathbf{u}_1'\lambda_1 + \mathbf{u}_2'\lambda_2)\cdot \mathbf{x}' &= 0 \\ (\mathbf{u}_1''\lambda_1 + \mathbf{u}_2''\lambda_2)\cdot \mathbf{x}'' &= 0\,. \end{aligned}$$

These two linear homogeneous equations in the unknowns (λ_1, λ_2) have a nontrivial solution if and only if the determinant vanishes. This gives the stated bilinear form

$$(\mathbf{u}_1'\cdot\mathbf{x}')(\mathbf{u}_2''\cdot\mathbf{x}'') - (\mathbf{u}_2'\cdot\mathbf{x}')(\mathbf{u}_1''\cdot\mathbf{x}'') = 0\,.$$

There are singular points of this correspondance: $\mathbf{z}_2'$ corresponds to all $\mathbf{x}''$, and vice versa all points $\mathbf{x}'$ correspond to $\mathbf{z}_1''$. Therefore $\mathrm{rk}(b_{ij}) = 2$. □

Calibrated case:
In the calibrated case we can express the essential matrix B in a particular form. For this purpose it is necessary to specify the homogeneous coordinates used in the bilinear relation (1): For each image point we take its 3D coordinates with respect to the camera frame as homogeneous 2D coordinates (see Fig. 4).

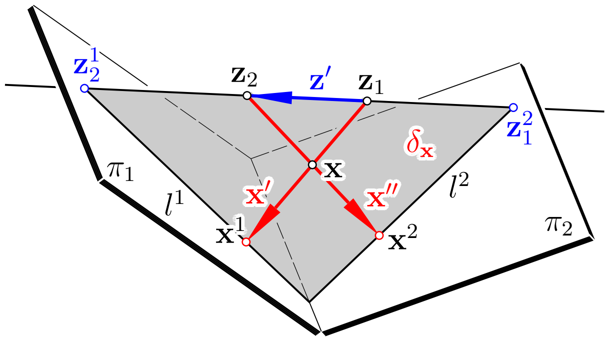

Figure 4: Epipolar constraints for calibrated images

Theorem 2: *In the calibrated case the essential matrix B is the product of a skew symmetric matrix and an orthogonal one, i.e.,*

$$B = S \cdot R \text{ with } S^T = -S \text{ and } R^{-1} = R^T. \quad (2)$$

Therefore the two singular values of B are equal.

Proof: According to Fig. 4 the three vectors

$$\mathbf{z}' := \mathbf{z}_2 - \mathbf{z}_1, \quad \mathbf{x}' \text{ and } \mathbf{x}''$$

are coplanar. Therefore their triple product vanishes. However, we have to pay attention to the fact that $\mathbf{x}'$ and $\mathbf{x}''$ are given in two different camera frames. Let

$$\mathbf{x}_1 = \mathbf{z}' + R \cdot \mathbf{x}_2 \quad (3)$$

be the conversion of the second camera frame into the first one with an orthogonal R. Now the complanarity is equivalent to

$$0 = \det(\mathbf{x}', \mathbf{z}', R \cdot \mathbf{x}'') = \mathbf{x}' \cdot (\mathbf{z}' \times R \cdot \mathbf{x}'').$$

We may replace the cross product by the product of $\mathbf{x}''$ with a skew-symmetric matrix, i.e.,

$$\mathbf{z}' \times R \cdot \mathbf{x}'' = S \cdot R \cdot \mathbf{x}''$$

with

$$\begin{pmatrix} 0 & -z_z' & z_y' \\ z_z' & 0 & -z_x' \\ -z_y' & z_x' & 0 \end{pmatrix}, \quad (4)$$

provided (z_x', z_y', z_z') are the coordinates of $\mathbf{z}'$ with respect to the first camera frame. It is important to notice that according to (3) the two factors S and R define the relative position between the two camera frames uniquely.

The singular values of $B = S \cdot R$ can either be computed straight forward as the positive square-roots of eigenvalues of $B^T \cdot B$, i.e., of $S^T \cdot S = -S \cdot S$. But we can also proceed in a more geometric way:

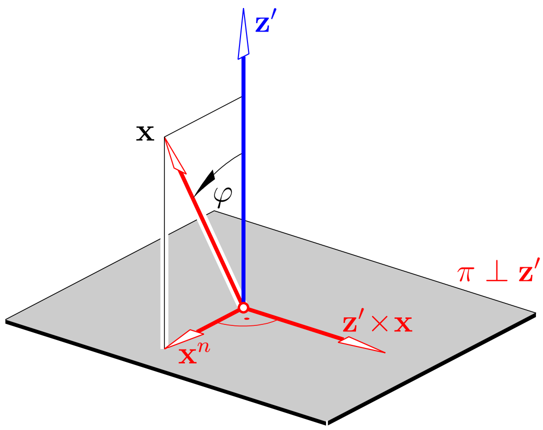

Figure 5: $\mathbf{x} \mapsto S \cdot \mathbf{x} = \mathbf{z}' \times \mathbf{x}$ is the product of an orthogonal projection, a $90°$-rotation, and a scaling with factor $\|\mathbf{z}'\|$

The cross product $\mathbf{z}' \times \mathbf{x}$ is orthogonal to the plane spanned by $\mathbf{z}'$ and $\mathbf{x}$, and it has the length

$$\|\mathbf{z}' \times \mathbf{x}\| = \|\mathbf{z}'\| \, \|\mathbf{x}\| \sin\varphi = \|\mathbf{z}'\| \, \|\mathbf{x}^n\|$$

where $\mathbf{x}^n$ is the orthogonal projection of $\mathbf{x}$ in direction of $\mathbf{z}'$ (see Fig. 5). So, the mapping $\mathbf{x} \mapsto S \cdot \mathbf{x}$ is the composition of an orthogonal projection, of a $90°$-rotation, and a scaling with factor $\|\mathbf{z}'\|$ which in the sense of Fig. 2 reveals the above-mentioned singular values of S. □

3. THE FUNDAMENTAL THEOREMS

What means *'reconstruction'* from two images ? The photos have been taken in a particular viewing situation. But afterwards we have only the two images, and we know nothing about how the camera frames where mutually placed in 3-space. Hence, reconstruction means both, recovering the viewing situation and recovering the depicted scene.

The problem of recovering a scene from two or more images is a basic problem in Computer Vision (see, e.g., [3, 4, 15, 5]). It is remarkable, that sometimes in the cited books the authors really refer to results which have already been achieved in Descriptive Geometry (note, e.g., the high estimation of E. KRUPPA's results [9] in [15]). However, Computer Vision focuses on numerical solutions, and the use of computers brought new insight and progress in this problem. Since measuring pixels in any image can be carried out with standard software, it has become possible to recover an object with high precision from two digital images just by using a laptop.

Theorem 3: *From two uncalibrated images with given projectivity between epipolar lines the depicted object can be reconstructed up to a collinear transformation.*

Sketch of the proof: The two images can be placed in space such that pairs of epipolar lines are intersecting:

For this purpose we start with a position where the two images are coplanar and two corresponding lines are aligned. Then the two pencils of epipolar lines are perspective with respect to an axis a. Now we rotate one of the image planes about this axis a. The corresponding epipolar lines are still intersecting on a. Then we specify arbitrary centers $\mathbf{z}_1, \mathbf{z}_2$ on the baseline z which connects the two epipoles. This gives rise to a reconstructed 3D object.

Any other choice of the viewing situation gives a collinear transform of the previously recovered 3D object. □

Theorem 4: (S. Finsterwalder, 1899)
From two calibrated *images with given projectivity between epipolar lines the depicted object can be reconstructed up to a similarity.*

Sketch of the proof: In the corresponding bundles of rays the pencils of epipolar planes $\delta_{\mathbf{x}}$ for both projections need to be *congruent*. There is a rigid motion of one camera frame such that any two corresponding epipolar planes are coincident. For any choice of $\mathbf{z}_2$ relative to $\mathbf{z}_1$ on the carrier line z of the unified pencil of planes there exists a reconstructed 3D object. Any other choice of $\mathbf{z}_2$ gives a similar 3D object. □

In this sense the problem of recovering a scene is reduced to the determination of epipoles. This problem is equivalent to a classical problem of Projective Geometry, the Problem of Projectivity (see Fig. 6):

Given: 7 pairs of corresponding points $(\mathbf{x}'_1, \mathbf{x}''_1)$, $\ldots, (\mathbf{x}'_7, \mathbf{x}''_7)$.

Wanted: A pair of points $(\mathbf{s}', \mathbf{s}'')$ (= epipoles) such that the connecting lines with $\mathbf{x}'_i$ and $\mathbf{x}''_i$, respectively, are included in a projectivity, i.e.,

$$\mathbf{s}'(\mathbf{s}' \vee \mathbf{x}'_i) \;\barwedge\; \mathbf{s}''(\mathbf{s}'' \vee \mathbf{x}''_i), \quad i = 1, \ldots, 7\,.$$

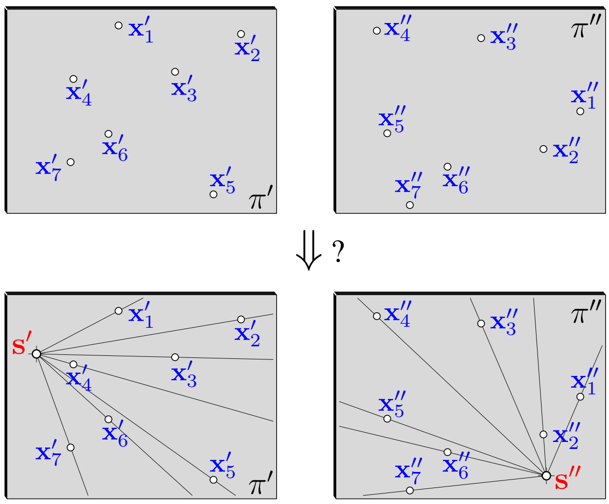

Figure 6: Problem of Projectivity

The Problem of Projectivity is a *cubic* problem. This follows from the following reasoning:

Due to (1) the 7 given pairs of corresponding points give 7 linear homogeneous equations

$$\mathbf{x}_i^T \cdot B \cdot \mathbf{x}_i'' = 0, \quad i = 1, \ldots, 7\,, \tag{5}$$

for the 9 entries in the essential (3×3)-matrix $B = (b_{ij})$. The condition $\mathrm{rk}(B) = 2$ gives the additional cubic equation $\det B = 0$ which fixes all b_{ij} up to a common factor.

4. COMPUTING THE ESSENTIAL MATRIX

For noisy image points it is recommended to use more than 7 points and to apply methods of least squares approximation for obtaining the 'best fitting' matrix B:

Let A denote the coefficient matrix in the linear system (5) of homogeneous equations for the entries of B. Then the 'least square fit' $\widetilde{B}$ is an eigenvector for the smallest eigenvalue of the symmetric matrix $A^T \cdot A$ which minimizes

$$\mathbf{y}^T \cdot A^T \cdot A \cdot \mathbf{y} = \|A \cdot \mathbf{y}\|^2$$

under the side condition $\|\mathbf{y}\| = 1$.

Any essential matrix has the rank 2, and in particular in the calibrated case the two singular values must be equal. In order to obtain such a 'best fitting' essential matrix B for our obtained $\widetilde{B}$, we use what sometimes is called the *'projection into the essential space':*

This is based on the singular value decomposition of $\widetilde{B}$, which has been presented in Section 1. It factorizes $\widetilde{B}$ as a matrix product

$$\widetilde{B} = U \cdot D \cdot V^T, \quad D = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3),$$

with orthogonal U, V. For the singular values of $\widetilde{B}$ we suppose $\lambda_1 \geq \lambda_2 \geq \lambda_3$.

Then in the uncalibrated case the best fitting essential matrix reads

$$B = U \cdot \mathrm{diag}(\lambda_1, \lambda_2, 0) \cdot V^T. \tag{6}$$

In the calibrated case

$$B = U \cdot \mathrm{diag}(\lambda, \lambda, 0) \cdot V^T \text{ with } \lambda = \frac{\lambda_1 + \lambda_2}{2} \tag{7}$$

is optimal in the sense of the Frobenius norm $\|A\|_f$ for square matrices A (see, e.g., [10, 15]). $\|A\|_f^2$ equals the trace of $A^T \cdot A$ and therefore the square sum of the singular values of A.

In the uncalibrated case the solution B of (6) gives

$$\|\widetilde{B} - B\|_f = \lambda_3$$

which is minimal among all rank 2 matrices. In the calibrated case the solution B presented in (7) yields the error

$$\|\widetilde{B} - B\|_f = \sqrt{(\lambda - \lambda_1)^2 + (\lambda - \lambda_2)^2 + \lambda_3^2}\,,$$

which is minimal among all possible essential matrices.

The factorization $B = S \cdot R$ according to Theorem 2 reveals already the relative position of the two camera frames. Therefore we need

Theorem 5: *The factorization of the essential matrix $B = U \cdot D \cdot V^T$, $D = \mathrm{diag}(\lambda, \lambda, 0)$, into the skew symmetric matrix S and the orthogonal matrix R reads:*

$$S = \pm U \cdot R_+ \cdot D \cdot U^T, \quad R = \pm U \cdot R_+^T \cdot V^T$$
$$\text{where } R_+ = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{8}$$

Proof: It is sufficient to factorize the product of the first two matrices by

$$U \cdot D = S \cdot R',$$

because this implies immediately

$$B = S \cdot (R' \cdot V^T), \text{ i.e., } R = R' \cdot V^T.$$

We focus on the affine 3D transformations which are represented by the involved matrices:

- $U \cdot D$ is composed from the orthogonal projection parallel to the $\overline{z}$-axis, the scaling with factor λ and the rotation U which transforms the $\overline{z}$-axis into the kernel of $U \cdot B$.
- On the other hand, the skew symmetric matrix S represents the orthogonal projection parallel $\mathbf{z}'$ composed with a $90°$-rotation about $\mathbf{z}'$ and a scaling with factor $\|\mathbf{z}'\|$ (see Fig. 5).

Figure 7: Given photos: Historical ‘Stadtbahn’ station Karlsplatz
in Vienna (Otto Wagner, 1897)

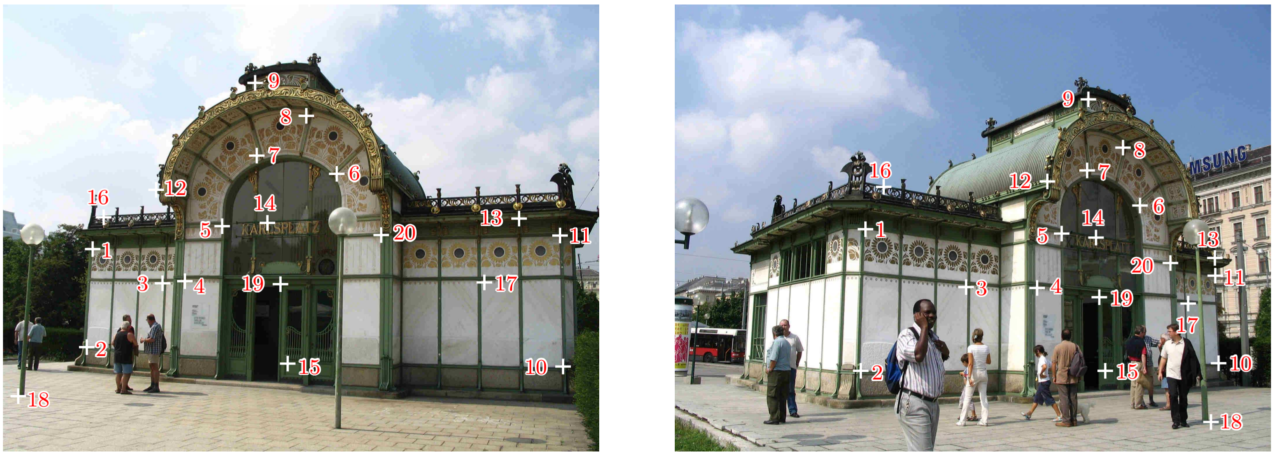

Figure 8: Identifying 20 reference points

Figure 9: Epipolar lines

Let R_+ denote the matrix representing the $90°$-rotation about the $\overline{z}$-axis. Then R_+ is of the form stated in Theorem 5. The product $R_+ \cdot D = D \cdot R_+$ is skew-symmetric and thus we obtain the following two solutions:

$$S = \pm U \cdot R_+ \cdot D \cdot U^T \text{ and } R' = \pm U \cdot R_+^T .$$

For the following reason, these are the only two possible factorizations of the required type:
As matrix B represents an orthogonal axonometry, the column vectors are the images of an orthonormal frame. We know from Descriptive Geometry that apart from translations parallel to the rays of sight there are exactly two different triples of pairwise orthogonal axes with images in direction of the given column vectors. The two triples are mirror images from each other. So, we can't expect more than two factorizations. □

There are *critical configurations* where the specified reference points are not sufficient to determine the epipoles uniquely. This is, e.g., the case when only coplanar 3D points are chosen as reference points. But there are also other cases related to quadrics. For details see, e.g., [14, 15, 5]).

5. THE ALGORITHM

We summarize: The numerical reconstruction of two calibrated images with the aid of any computer algebra system (e.g., Maple) consists of the following five steps:

1) Specify $n > 7$ pairs $(\mathbf{x}'_i, \mathbf{x}''_i)$, $i = 1, \ldots, n$, of corresponding points under avoidance of critical configurations.

2) Set up the homogeneous linear system of equations $\mathbf{x}'^T_i \cdot B \cdot \mathbf{x}''_i = 0$ for the unknown fundamental matrix B. The optimal solution $\widetilde{B}$ is an eigenvector of the smallest eigenvalue of $A^T \cdot A$ with A as the coefficient matrix of this system.

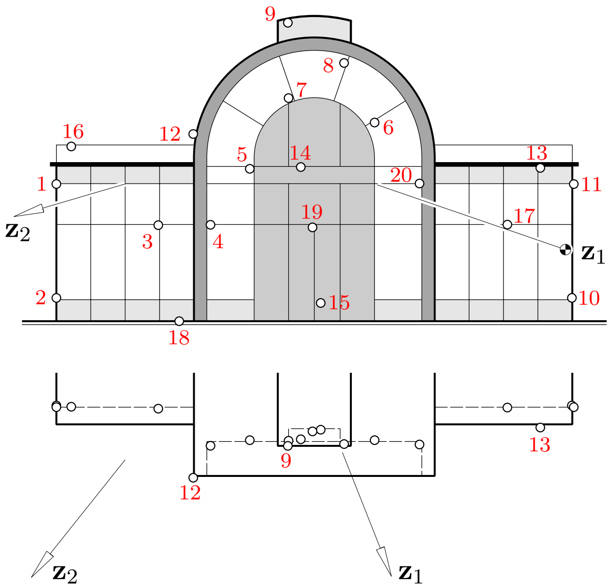

Figure 10: The result of the reconstruction

3) Based on the singular value decomposition of $\widetilde{B}$ compute the closest rank 2 matrix B with two equal singular values.

4) Factorize $B = S \cdot R$ as a product of a skew symmetric matrix S and an orthogonal R according to Eq. (8). This fixes the relative position between the two camera frames.

5) In one of the camera frames compute the approximate point of intersection between corresponding rays $\mathbf{z}_1 \vee \mathbf{x}'_i$ and $\mathbf{z}_2 \vee \mathbf{x}''_i$, $i = 1, 2, \ldots$

6) Transform the reconstructed coordinates of points of the scene into any world coordinate system.

Figs. 7, 8 and 9 show on example with the determination of epipolar lines and epipoles. The solution is displayed in Fig. 10.

References

[1] Brauner, H., Lineare Abbildungen aus euklidischen Räumen, *Beitr. Algebra Geom. 21,* (1986), 5–26.

[2] Dür, A., An Algebraic Equation for the Central Projection, *J. Geometry Graphics 7*, (2003), 137–143.

[3] Faugeras, O., *Three-Dimensional Computer Vision. A Geometric Viewpoint,* MIT Press, Cambridge, Mass., 1996.

[4] Faugeras, O., and Q.-T. Luong, *The Geometry of Multiple Images,* MIT Press, Cambridge, Mass., 2001.

[5] Hartley, R., and Zisserman, A., *Multiple View Geometry in Computer Vision.* Cambridge University Press, 2000.

[6] Havel, V., O rozkladu singulárních lineárních transformací, *Časopis Pést. Mat. 85,* (1960), 439–447.

[7] Havlicek, H., On the Matrices of Central Linear Mappings, *Math. Bohem. 121,* (1996), 151–156.

[8] Hoffmann, M., and Yiu, P., Moving Central Axonometric Reference Systems, *J. Geometry Graphics 9*, (2005), 133–140.

[9] Kruppa, E., Zur achsonometrischen Methode der darstellenden Geometrie, *Sitzungsber., Abt. II, österr. Akad. Wiss., Math.-Naturw. Kl. 119,* (1910), 487–506.

[10] Lawson, C.L., and Hanson, R.J., *Solving Least Squares Problems*, siam, 1995, 1st ed.: Prentice-Hall, Inc., Englewood Cliffs, NJ, 1974.

[11] Stachel, H., Zur Kennzeichnung der Zentralprojektionen nach H. Havlicek, *Sitzungsber., Abt. II, österr. Akad. Wiss., Math.-Naturw. Kl. 204,* (1995), 33–46.

[12] Stachel, H., On Arne Dür's Equation Concerning Central Axonometries, *J. Geometry Graphics 8*, (2004), 215–224.

[13] Szabó, J., Stachel, H. and Vogel, H., Ein Satz über die Zentralaxonometrie, *Sitzungsber., Abt. II, österr. Akad. Wiss., Math.-Naturw. Kl. 203,* (1994), 3–11.

[14] Tschupik, J., and Hohenberg, F., *Die geometrische Grundlagen der Photogrammetrie,* in Jordan, Eggert, Kneissl (eds.): *Handbuch der Vermessungskunde III a/3,* 10. Aufl., Metzlersche Verlagsbuchhandlung, Stuttart 1972, 2235–2295.

[15] Yi Ma, Soatto, St., Kosecka, J., and Sastry, S. Sh., *An Invitation to 3-D Vision.* Springer-Verlag, New York 2004.

ABOUT THE AUTHOR

Hellmuth Stachel, Ph.D, is Professor of Geometry at the Institute of Discrete Mathematics and Geometry, Vienna University of Technology, and editor in chief of the "Journal for Geometry and Graphics". His research interests are in Higher Geometry, Kinematics and Computer Graphics. He can be reached by e-mail: stachel@dmg.tuwien.ac.at, by fax: (+431)-58801-11399, or through the postal address:
Institut für Diskrete Mathematik und Geometrie / Technische Universität Wien / Wiedner Hauptstr. 8-10/104 / A 1040 Wien / Austria, Europe.