Depth Computations from Polyhedral Images *

Gunnar Sparr

Dept. of Mathematics, Lund Institute of Technology,
Box 118, S-22100 Lund, Sweden

Abstract. A method is developed for the computation of depth maps, modulo scale, from one single image of a polyhedral scene. Only affine shape properties of the scene and image are used, hence no metrical information. Results from simple experiments show good performance, both what concerns exactness and robustness. It is also shown how the underlying theory may be used to single out and characterise certain singular situations that may occur in machine interpretation of line drawings.

1 Introduction

The topic of this paper is depth computation and scene reconstruction in the case when the scene is built up by planar surface patches, bounded by polygons. Having only this information, from one single image no quantitative information can be drawn. A common situation is that the scene contains patches which are parallelograms, often rectangles. It will be seen that under this rather weak assumption, without any knowledge about the sizes of these parallelograms, it is possible to compute a depth-map over the image, modulo a common scaling factor. The method may be used also for patches of other shapes. No camera calibration is needed.

The approach is inspired by the subjective experience that depth information seems to be contained in the shape of an image of e.g. a rectangle. In a series of papers, e.g. [8], [9], [10], [11], this hypothesis has been verified in quantitative terms. In the present paper emphasis will be laid on examples and experiments, rather than on the mathematical theory. For a thorough treatment of the latter, see [9].

The organization of the paper is as follows. In Sect. 2, the concept of 'shape' is described, with examples. In Sect. 3 the same will be done for 'depth', with some new theorems. In Sect. 4 is described a simple experiment, illustrating the applicability of the method for realistic data. Also the robustness properties are investigated. In Sect. 5, some degeneracies that may occur are treated. In Sect. 6, finally, the results and their relations to previous work are discussed.

Throughout the paper, it is assumed that the correspondence problem is solved beforehand, i.e. that a set of point matches between points in the image and in the scene is established.

2 Shape

Instead of working with individual points, we work with *m-point configurations*, by which is meant ordered sets of points, planar or non-planar,

$$\mathcal{X} = (X^1, \ldots, X^m) \ .$$

* The work has been supported by the Swedish National Board for Industrial and Technical Development, (NUTEK).

It turns out to be fruitful to work with a kind of duality (for motivations and proofs, see e.g. [9]), and consider the linear relations that exist between the points belonging to a particular configuration. It can be proved that the set (1) below is independent of coordinate representations. The following definition plays a crucial role in the sequel.

Definition 1. The *(affine) shape* of $\mathcal{X} = (X^1, \ldots, X^m)$ is the linear space

$$s(\mathcal{X}) = \{\xi = (\xi_1, \ldots, \xi_m) \mid \sum_1^m \xi_i X^i = 0, \quad \sum_1^m \xi_i = 0\} \ , \tag{1}$$

where X^i, $i = 1, \ldots, m$, stands for coordinates in an arbitrary affine coordinate system.

We use the notation

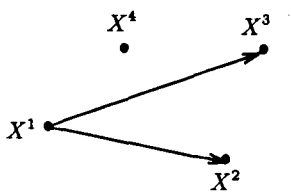
$$\mathcal{X}' \overset{s}{=} \mathcal{X}'' \iff \mathcal{X}' \text{ and } \mathcal{X}'' \text{ have equal shape} \ .$$

It can be shown that $\mathcal{X}' \overset{s}{=} \mathcal{X}''$ if and only if $\mathcal{X}'$ and $\mathcal{X}''$ can be mapped onto each other by an affine transformation.

To get a geometric feeling for the notion of shape, consider the case of a planar 4-point configuration $\mathcal{X} = (X^1, \ldots, X^4)$. Suppose the configuration is non-degenerate, i.e. that it contains three non-collinear points, say X^1, X^2, X^3. Let ξ_2, ξ_3 be the coordinates of X^4 with respect to a coordinate system with origin in X^1 and basis vectors $\overrightarrow{X^1X^2}$ and $\overrightarrow{X^1X^3}$, cf. Fig. 1. The equation $\overrightarrow{X^1X^4} = \xi_2\overrightarrow{X^1X^2} + \xi_3\overrightarrow{X^1X^3}$ may then be written

$$\xi_1\overrightarrow{X^1X^1} + \xi_2\overrightarrow{X^1X^2} + \xi_3\overrightarrow{X^1X^3} + (-1)\overrightarrow{X^1X^4} = 0, \quad \xi_1 = 1 - \xi_2 - \xi_3 \ ,$$

where the coefficient ξ_1 of the null-vector $\overrightarrow{X^1X^1}$ is chosen so that the coefficient sum vanishes. This construction determines the shape-vector $(\xi_1, \xi_2, \xi_3, -1)$ in the case of 4-point configurations. Any multiple of this vector belongs to $s(\mathcal{X})$ too.



$$\overrightarrow{X^1X^4} = \xi_2\overrightarrow{X^1X^2} + \xi_3\overrightarrow{X^1X^3}$$

$$X^4 = \xi_1X^1 + \xi_2X^2 + \xi_3X^3, \quad \xi_1 + \xi_2 + \xi_3 = 1$$

Fig. 1. Affine coordinates.

Above the points X^1, X^2, X^3 form what is called an *affine basis* for the plane. The coordinates ξ_1, ξ_2, ξ_3, with $\sum \xi_i = 1$, are called the *affine* coordinates of X^4. An analogous construction can be done in space. Thus, if X^1, X^2, X^3, X^4 are vertices of a non-degenerate tetrahedron, they form an affine basis, and X^5 can be described by its affine coordinates $\xi_1, \ldots, \xi_4$, with $\sum \xi_i = 1$. Again $s(\mathcal{X})$ is a one-dimensional linear space.

For two-dimensional configurations with more than 4 points, and three-dimensional configurations with more than 5 points, $s(\mathcal{X})$ is a linear space of higher dimension. Generally, for the shape of an m-point configuration, the following can be said:

$$\begin{aligned} &\dim s(\mathcal{X}) = m - 2 \text{ if the points are collinear,} \\ &\dim s(\mathcal{X}) = m - 3 \text{ if the points are coplanar, but not collinear,} \\ &\dim s(\mathcal{X}) = m - 4 \text{ if the points are not coplanar.} \end{aligned} \tag{2}$$

Example 1. In Fig. 2 are shown two 2D and one 3D configurations. The dotted lines have no other meaning than to indicate relationship between the points.

In the left configuration, the point X^4 is the centroid of the triangle with vertices in X^1, X^2, X^3. In the middle one, two sides are parallel. The right configuration is a "joint" in 3D, consisting of two rectangular 4-point configurations. Bases for the shapes of these configurations are shown. They may be computed e.g. by means of the affine basis construction above. A natural way to select a basis for the joint is by means of the planar subconfigurations, as is done in the figure, but other choices are possible too.

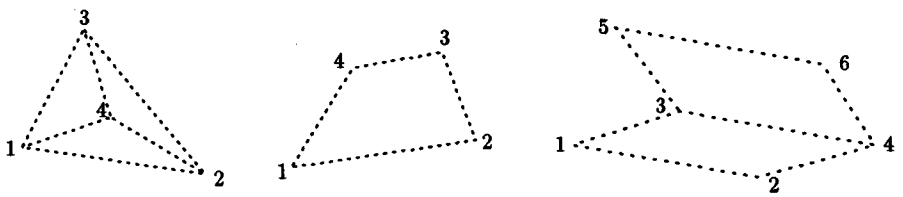


Fig. 2. Two 2D and one 3D configurations, and bases for their shapes.

To be of practical use, one needs algorithms for the *computation* of shape. One such algorithm suggests itself by the definition, namely the solving of a homogeneous system of linear equations. By means of e.g. a row echelon algorithm, a basis for $s(\mathcal{X})$ can be computed.

In the definition of shape, the coordinate invariancy is very important. It makes it possible to compute the shape of an image configuration from intrinsic measurements in the image plane, in terms of an arbitrary coordinate representation. The same holds for the object configuration. These shapes can thus be computed independently, without reference to the imaging process.

In the rest of this paper, point configurations defined by the vertices of polyhedral objects will be considered. (Here the word 'polyhedral' is used in a wide sense for configurations, not necessarily solid, built up by planar polygonal patches.) Besides being a point configuration of its own,

$$\mathcal{X} = (X^1, \ldots, X^m) \ ,$$

such a configuration has a lot of additional structure. In fact, each of the f polygonal faces of the object contributes with a sub-configuration, defined by the vertices of the polygon

$$\mathcal{X}_i = (X_i^1, \ldots, X_i^{m_i}), \quad i = 1, \ldots, f \ .$$

The whole configuration may be considered as an ordered set of these sub-configurations

$$\mathcal{O}^{\mathcal{X}} = (\mathcal{X}_1, \ldots, \mathcal{X}_f) \ .$$

We will work in parallel with both these representations $\mathcal{X}$ and $\mathcal{O}^{\mathcal{X}}$, and by abuse of notation write $\mathcal{O}^{\mathcal{X}}$ in both instances.

3 Depth

In this section will be examined how the shape, as defined above, transforms under projective transformations. The results are fundamental for the applications below.

By a *perspectivity* with *center* Z in 3-space is meant a mapping with the property that every point on a line through Z is mapped onto the intersection of the line with some plane π, the image plane, where $Z \notin \pi$. For a perspectivity between two m-point configurations, $\mathcal{X} \longrightarrow \mathcal{Y}$, there exist α_i, $i = 1, \ldots, m$, such that

$$\overline{ZX^i} = \alpha_i \overline{ZY^i}, \quad i = 1, \ldots, m \ .$$

Here α_i is called the *depth* of X_i with respect to Y_i, $i = 1, \ldots, m$, and the vector $\alpha = (\alpha_1, \ldots, \alpha_m)$ is called the depth of $\mathcal{X}$ with respect to $\mathcal{Y}$.

By a *projectivity* is meant a composition of perspectivities. The product of the depths of these perspectivities defines the depth of the projectivity. (It can be shown, cf. [9], that this product is independent of decomposition.)

The following theorem shows how the knowledge of the shapes of two configurations $\mathcal{X}$ and $\mathcal{Y}$ makes it possible to characterise all projectivities P such that $\mathcal{Y} \overset{s}{=} P(\mathcal{X})$. In particular, pose information is attained about the location of $\mathcal{X}$ relative to $\mathcal{Y}$.

Theorem 2. *The following conditions are equivalent:*

- *There exists a projectivity P such that $P(\mathcal{X}) \overset{s}{=} \mathcal{Y}$, where $\mathcal{X}$ has depth α with respect to $P(\mathcal{X})$,*
- *$diag(\alpha)s(\mathcal{X}) \subset s(\mathcal{Y})$, with equality if $\mathcal{X}$ is planar.*

Note that here α is determined up to proportionality only. In accordance with this, the depth will be considered as a homogeneous vector.

The meaning of Theorem 2 is illustrated in Fig. 3, drawn for 4-point configurations. Let there be given two planes, containing the point configurations $\mathcal{X}$ and $\mathcal{Y}$ respectively. Suppose it is possible to move around these planes in space. Take an arbitrary point Z, and form a pencil of rays connecting Z with the points of $\mathcal{Y}$. The theorem says that whenever an $\mathcal{X}$-configuration fits on this pencil, the depth-values α are given by the formula of the theorem, independently of the location of Z and $\mathcal{Y}$. This is even true when $\mathcal{X}$ and $\mathcal{Y}$ are replaced by configurations having the same respective shapes. Conversely, if instead α and $\mathcal{Y}$ are given, and by means of them an $\mathcal{X}$-configuration is constructed, then the shape of the latter is given by the theorem.

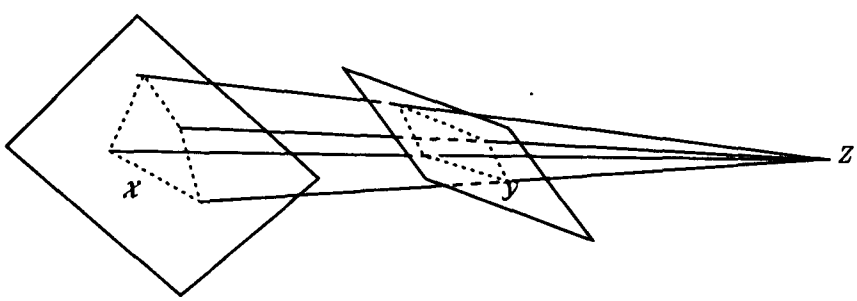


Fig. 3. Perspective mapping of point configurations.

Example 2. The right hand configuration of Fig. 2 was said to illustrate a three-dimensional figure, a joint, with rectangular faces. Looking upon the figure as it is printed on the paper, i.e. as a two-dimensional perspective image of the joint, measurements in the image give that its two parts have the following shapes:

$$s((Y^1, Y^2, Y^3, Y^4)) = (\eta_1, \eta_2, \eta_3, \eta_4) = (1.1, -1, -1.2, 1.1) \ ,$$
$$s((Y^3, Y^4, Y^5, Y^6)) = (\eta_3, \eta_4, \eta_5, \eta_6) = (-1.2, 1.1, 1.1, -1) \ .$$

The result of applying Theorem 2 to the two parts separately may be summarised in a matrix equation

$$\operatorname{diag}(\alpha) \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ -1 & 1 \\ 1 & -1 \\ 0 & -1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1.1 & 0 \\ -1.0 & 0 \\ -1.2 & -1.2 \\ 1.1 & 1.1 \\ 0 & 1.1 \\ 0 & -1.0 \end{bmatrix} \operatorname{diag}(1, -1) \ . \tag{3}$$

Here the diagonal matrix on the right hand side is needed to adjust for the arbitrariness in the choice of the columns of the 6×2-matrices. The system has the depth solution

$$\alpha^T = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4 \ \alpha_5 \ \alpha_6] = [1.1 \ 1.0 \ 1.2 \ 1.1 \ 1.1 \ 1.0]$$

(together with all multiples of this vector).

In this example, a claim on *depth-consistency* is met, formulated in terms of the matrix-equation (3). The general situation is covered by the following theorem.

Theorem 3. *Let $\mathcal{O}^{\mathcal{X}} = (\mathcal{X}_1, \ldots, \mathcal{X}_f)$ be a polyhedral point configuration, and let $\mathcal{O}^{\mathcal{Y}} = (\mathcal{Y}_1, \ldots, \mathcal{Y}_f)$ be its image. Let*

$$S^{\mathcal{X}} = [S^{\mathcal{X}_1}, \ldots, S^{\mathcal{X}_f}]$$

be a matrix with sub-matrices $S^{\mathcal{X}_i}$, having columns that form a basis for $s(\mathcal{X}_i)$, and let $S^{\mathcal{Y}_i}$ be defined in the corresponding way, $i = 1, \ldots, f$. Then holds, for some α and c,

$$diag(\alpha) S^{\mathcal{X}} = S^{\mathcal{Y}} diag(c) \ . \tag{4}$$

For noisy data, the equation (4) can't be expected to be satisfied exactly. Moreover, the system in α is in general overdetermined. In the next section it will be solved in the least square sense.

For projective mappings from one plane to another, the following theorem gives an analytic expression for the depth function.

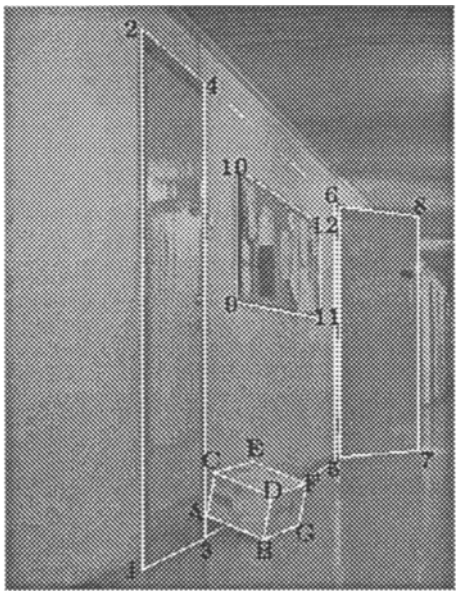
Theorem 4. *Let the planar configuration $\mathcal{X} = (X^1, X^2, X^3, X)$ be mapped onto $\mathcal{Y} = (Y^1, Y^2, Y^3, Y)$ by a perspectivity, so that Y^1, Y^2, Y^3 have depths $\alpha_1, \alpha_2, \alpha_3$ with respect to X^1, X^2, X^3. If $s(\mathcal{Y}) = (\eta_1, \eta_2, \eta_3, \eta)$, then the depth α of Y with respect to X is given by*

$$\frac{\eta_1}{\alpha_1} + \frac{\eta_2}{\alpha_2} + \frac{\eta_3}{\alpha_3} + \frac{\eta}{\alpha} = 0 \ .$$

If $\eta \in s(\mathcal{Y})$ is so normed that $\eta = -1$, i.e. $\eta_1 + \eta_2 + \eta_3 = 1$, then η_1, η_2, η_3 are the affine coordinates of Y with respect to the affine basis Y^1, Y^2, Y^3. The formula says that the depth α of Y is the weighted harmonic mean of $\alpha^1, \alpha^2, \alpha^3$, with the affine coordinates as weights.

4 An Experiment

A simple experiment will illustrate how the theory may be used. In Fig. 4 is shown an image of a corridor scene, containing a number of rectangular objects: two doors, part of a wall, a board and the faces of a box. The doors are at approximate distances 6 m and 12 m from the camera. The dimensions of the box are $35 \times 35 \times 20$ cm. The size of the image is about 300×400 pixels, where the box occupies about 50×50 pixels.



	Wall+doors							
	1	2	3	4	5	6	7	8
comp	1.00	1.01	1.19	1.20	2.08	2.09	2.23	2.23
meas	1.00	1.02	1.20	1.22	2.07	2.08	2.21	2.22

	Board			
	9	10	11	12
comp	1.35	1.35	1.92	1.92
meas	1.35	1.36	1.90	1.91

	Box						
	A	B	C	D	E	F	G
comp	1.02	1.01	1.02	1.00	1.07	1.05	1.05
meas	1.03	1.01	1.03	1.00	1.07	1.05	1.05

Fig. 4. A corridor scene. Measured and computed depth values.

The purpose is to investigate the capacity of the method, both what concerns exactness and robustness. The pixel coordinates of the points of interest were picked out by hand from the image. The wall and the two doors form a configuration similar to the joint of Example 1, and may be treated as in Example 2, by means of Theorem 3. In the same way, the depth values for the box are computed. Theorem 4 then gives the depth values for the board. For the wall, the doors and the board, the depth values were normalised against Point 1, while for the box, the normalisation was done against Point D. The so normalised computed depth values are shown in the first lines of the respective tables in Fig. 4.

For comparison, the distances from the camera to the marked scene points were measured by a measuring tape. The values were normalised as described above. Since these distance ratios are not exactly the same as depth, they have to be corrected by means of coefficients, which depend on the angle of sight relative to the focal line. The so obtained measured depth values are printed in the second lines of the respective tables in Fig. 4. Since the distance to Point 1 is about 6 m, a deviation of 0.01 in depth corresponds to about 6 cm in distance. This is also the estimated uncertainty in the measurements. As can be seen, the computed depth values show good agreement with the measured ones.

To get a feeling for the robustness, rectangularly distributed noise of ± 1 pixels was added to each coordinate of the points of interest above. New depth-values were then computed. In order to compare homogeneous depth-vectors, the normalised differences

$$\alpha_{\text{ref}} - \frac{|\alpha_{\text{ref}}|}{|\alpha_{\text{new}}|}\alpha_{\text{new}}$$

were computed, where α_{ref} stands for the computed depth values of Fig. 4. Table 1 shows the outcome of 10 random simulations. For the wall+doors, the results indicate good robustness properties, while for the box the results are not equally favourable. The latter isn't surprising, since the object is small and so distant that the perspective effects are small.

Table 1. Effects of noise.

Wall+doors							
1	2	3	4	5	6	7	8
-0.01	0.00	0.00	-0.01	0.02	-0.01	0.00	-0.01
0.00	0.00	0.02	0.01	0.03	-0.02	0.00	-0.02
0.01	0.00	-0.01	0.00	-0.01	0.00	0.00	0.01
-0.01	0.00	0.00	-0.01	0.02	-0.01	0.02	-0.01
0.02	-0.02	0.01	-0.01	-0.01	0.01	0.00	0.01
0.01	0.00	0.00	0.01	0.00	-0.02	0.02	-0.02
0.00	0.00	0.00	0.01	0.00	0.02	-0.01	-0.01
-0.01	0.01	-0.01	0.01	0.00	0.04	-0.03	-0.01
-0.02	0.00	0.00	0.01	0.00	0.00	0.00	0.00
0.00	-0.01	0.01	0.00	0.03	-0.03	0.01	-0.02

Box						
A	B	C	D	E	F	G
0.00	0.00	0.00	-0.06	0.01	-0.03	0.08
-0.04	0.02	-0.01	0.00	-0.01	0.00	0.04
-0.02	-0.06	0.02	0.02	0.02	0.05	-0.03
0.03	0.01	0.05	0.05	-0.05	-0.02	-0.05
-0.02	-0.04	0.03	-0.01	0.06	0.00	-0.03
0.00	-0.01	0.00	0.01	0.04	0.01	-0.04
0.01	0.01	-0.04	0.01	0.09	0.02	-0.09
-0.02	0.01	-0.01	0.04	-0.08	0.02	0.05
0.00	0.03	-0.06	-0.04	0.00	0.01	0.06
-0.03	0.01	-0.01	0.02	-0.02	0.02	0.02

To summarise, using only that the scene contains parallelograms, nothing about their sizes, it has been possible to compute depths from the image. No knowledge at all is used about the camera, only that the imaging process is projective. In fact, the camera used in the experiment turned out to give about 5 % deficiency in the width and height scales. However, since the method only uses affine properties, it is robust also to such errors, as long as they can be modeled by affine transformations in the image plane. They need not even be compensated for. The results of the experiment are accurate and seem to have good robustness properties.

5 Degeneration

When analyzing an image of a polyhedral scene, difficulties with the interpretation may occur, even for human observers. Some possible explanations are the occlusions and accidental alignments that may occur in the image, without having a counterpart in the scene. A computational algorithm, based on linear programming, for the interpretation of line drawings is given in [12].

In this section, it will be indicated how the concepts of shape and depth can be used to give a simple criterion for correctness. First some definitions. Let $\mathcal{O}^{\mathcal{Y}} = (\mathcal{Y}_1, \ldots, \mathcal{Y}_f)$ be a planar configuration, built up by a number of polygonal configurations. Then $\mathcal{O}^{\mathcal{Y}}$ is called an *impossible picture* if

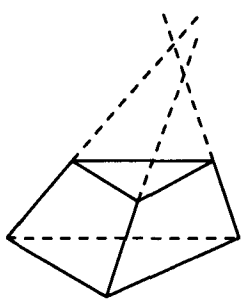
$$P : \mathcal{O}^{\mathcal{X}} \longrightarrow \mathcal{O}^{\mathcal{Y}},\ P \text{ a projectivity} \implies \mathcal{O}^{\mathcal{X}} \text{ is a planar configuration.}$$

By construction, cf. Theorem 3, all columns of $S^{\mathcal{Y}}$ belong to the shape $s(\mathcal{O}^{\mathcal{Y}})$. The same holds for $S^{\mathcal{X}}$, $s(\mathcal{O}^{\mathcal{X}})$. From the depth consistency (4) it follows that $S^{\mathcal{Y}}$ and $S^{\mathcal{X}}$ have the same ranks. Hence $\dim s(\mathcal{O}^{\mathcal{X}}) \geq \operatorname{rank} S^{\mathcal{X}} = \operatorname{rank} S^{\mathcal{Y}}$. From (2) it is known that the m–point configuration $\mathcal{O}^{\mathcal{X}}$ is non-planar if and only if $\dim s(\mathcal{O}^{\mathcal{X}}) = m - 4$. Combining these facts and definitions, we have proved the sufficiency part of the following theorem. The necessity is omitted here.

Theorem 5. *$\mathcal{O}^{\mathcal{Y}}$ is an impossible picture if and only if $\operatorname{rank} S^{\mathcal{Y}} \geq m - 3$.*

Having an image of a true three-dimensional polyhedral scene, this rank condition must be fulfilled. If it is violated because of noise, it may be possible to "deform" $\mathcal{O}^{\mathcal{Y}}$ to fulfill the condition. In doing this, any deformation can't be allowed. Let us say that a deformation is *admissible* if it doesn't change the topological and shape properties of the configuration, where the latter claim may be formulated Two configurations $\mathcal{O}^{\mathcal{Y}}$ and $\mathcal{O}^{\tilde{\mathcal{Y}}}$ are *topologically shape-equivalent* iff for every choice of matrix $S^{\mathcal{Y}}$ in Theorem 3, there exists a corresponding matrix $S^{\tilde{\mathcal{Y}}}$ for $\mathcal{O}^{\tilde{\mathcal{Y}}}$, such that their non-vanishing elements have the same distributions of signs. This gives a constructive criterion, possible to use in testing. As a final definition, we say that $\mathcal{O}^{\mathcal{Y}}$ is a *correctable impossible picture* if there exists an admissible deformation which makes $\operatorname{rank} S^{\mathcal{Y}} \leq m - 4$. An example of a configuration with this property is given in Fig. 5. For a method to find admissible deformations, see [7].

$$\begin{bmatrix} 0.65 & -0.60 & 0.00 \\ -0.65 & 0.00 & -0.61 \\ 0.00 & 0.60 & 0.54 \\ -1.00 & 1.00 & 0.00 \\ 1.00 & 0.00 & 1.06 \\ 0.00 & -1.00 & -1.00 \end{bmatrix}$$



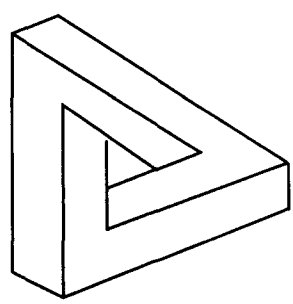


Fig. 5. A truncated tetrahedron and its shape. The Reutersvärd–Penrose tribar.

A more severe situation is met when it is impossible to correct the picture by means of admissible deformations. We then talk about an *absolutely impossible picture*. When dealing with such an image, one knows that the topology of the object isn't what it seems to be in the image. Accidental alignments or occlusions have occurred, and must be discovered and loosened.

A celebrated example of an "impossible picture" in the human sense is the tribar of Fig. 5. It is alternately called the "Reutersvärd tribar" or the "Penrose tribar", after two independent discoveries (1934 and 1958 respectively.) For historical facts, see the article of Ernst in [1]. For this configuration it can be proved that there exists no admissible deformation which makes the tribar fulfill the rank condition of Theorem 5. For more details, see [10], [11]. In terms of the concepts introduced above, the discussion of this section may be summarised:

- The truncated tetrahedron is a correctable impossible picture.
- The Reutersvrd–Penrose tribar is an absolutely impossible picture.

6 Discussion

Above a method for the computation of depth, modulo scale, from one single image of a polyhedral scene has been presented, under the assumption of known point correspondences between scene and image. Only affine information about the scene is used, e.g. that the objects contain parallelogram patches, nothing about their sizes. Other affine shapes may be used as well. In the image, no absolute measurements are needed, only relative (affine) ones. The image formation is supposed to be projective, but the method is insensitive to affine deformations in the image plane. No camera parameters are needed. The problem considered may be called an "affine calibration problem", with a solution in terms of relative depth values. The weak assumptions give them good robustness properties. All computations are linear.

The relative depth values may be combined with metrical information to solve the full (metrical) calibration problem (cf. [8], [9]). That problem is usually solved by methods that make extensive use of distances and angles, cf. [3] for an overview.

Relative depth information is also of interest in its own. For instance, in the case of rectangular patches in the scene, the relative depth values may be interpreted as the "motion" of the camera relative a location from which the patch looks like a rectangle. Looked upon in this way, our approach belongs to the same family as [2], [4], [6].

Crucial for the approach is the use of affine invariants (the 'shape'). In this respect the work is related to methods for recognition and correspondence, cf. [5].

In the last part of the paper is sketched an approach to the line drawing interpretation problem. Its relations to other methods, notably the one of [12], need further investigations.

References

1. Coxeter, H.M.S., Emmer, M., Penrose, R., Teuber, M.L.: M.C. Escher: Art and Science. Elsevier, Amsterdam (1986)
2. Faugeras, O.D.: What can be seen in three dimensions with an uncalibrated stereo rig? Proc. ECCV92 (1992) (to appear)
3. Horn, B.K.P.: Robot Vision. MIT Press, Cambridge, MA. (1986)
4. Koenderink, J.J., van Doorn, A.J.: Affine Structure from Motion. J. of the Opt. Soc. of America (1992) (to appear)
5. Lamdan, Y., Schwartz, J.T., Wolfson, H.J.: Affine Invariant Model-Based Object Recognition. IEEE Trans. Robotics and Automation **6** (1990) 578-589
6. Mohr, R., Morin, L., Grosso, E.: Relative positioning with poorly calibrated cameras. In Proc. DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision (1991)
7. Persson, A.: A method for correction of images of origami/polyhedral objects. Proc. Swedish Society for Automated Image Analysis. Uppsala, Sweden. (1992) (to appear)
8. Sparr, G., Nielsen, L.: Shape and mutual cross-ratios with applications to exterior, interior and relative orientation. Proc. Computer Vision – ECCV90. Springer Verlag, Lect. Notes in Computer Science (1990) 607-609
9. Sparr, G.: Projective invariants for affine shapes of point configurations. In Proc. DARPA-ESPRIT Workshop on Applications of Invariance in Computer Vision (1991)
10. Sparr, G.: Depth computations from polyhedral images, or: Why is my computer so distressed about the Penrose triangle. CODEN:LUFTD2(TFMA-91)/7004, Lund (1991)
11. Sparr, G.: On the "reconstruction" of impossible objects. Proc. Swedish Society for Automated Image Analysis. Uppsala, Sweden. (1992) (to appear)
12. Sugihara, K.: Mathematical Structures of Line Drawings of Polyhedrons – Toward Man-Machine Communication by Means of Line Drawings. IEEE Trans. Pattern Anal. Machine Intell. **4** (1982) 458-469

This article was processed using the LaTeX macro package with ECCV92 style