A Shape from Shading Approach for the Reconstruction of Polyhedral Objects using Genetic Algorithm

MANOJ KUMAR
RAMA BHARGAVA
R. BALASUBRAMANIAN

Indian Institute of Technology Roorkee, Roorkee
P.O. Box 247667, India
(mmathdma,rbharfma,balarfma)@iitr.ernet.in

Abstract. This paper presents a robust approach for the reconstruction of polyhedral objects from line drawings. Minimization technique has been used in this approach by exploiting the shading information contained in the image. The minimization energy function has been formulated between observed and calculated value of image intensity under the Lambertian reflectance model. Various constraints e.g. the surface normal coplanarity and edge length proportionality constraints are used to get the unique solution for the depth values. The minimization problem is nonlinear with nonlinear constraints, which is solved by using Genetic Algorithm. The proposed algorithm produces satisfactory results even in the case of slight error in computation of vertex positions caused in image processing. The algorithm has been tested on synthetic images and the results are shown.

Keywords: Polyhedra, superstrictness, Lambertian, albedo.

(Received March 6, 2007 / Accepted April 20, 2007)

1 Introduction

Reconstruction of 3D object shape and structure from 2D image(s) is the fundamental and most widely studied problem in Computer Vision. The techniques used for the reconstruction of 3D object shapes from 2D images are called shape - from - X techniques, where X stands for shading, stereo, texture, motion, etc. In these techniques, Shape from Shading (SFS) is the most robust because it reconstructs the 3D shape of an object from its single shaded image. The study of Shape from Shading was started by Horn and Brooks [9]. Since then enormous work has been done in this direction, in which most of the work deals with the reconstruction of smooth objects and very few work with the polyhedral objects [21], [18] and [17]. However, in the real world, there are many objects, especially man made objects which can be modeled or approximated by the polyhedral objects. Hence, the study of polyhedra is important and useful in many applications of computer vision. However very few elegant methods [20], [21], [18], [16], [17], [19] and [7] have been proposed for the reconstruction of polyhedral objects. The degree of freedom of a planar surface of a polyhedral object is three which can be reduced either by using more than one image or by imposing some extra constraints. For the unique reconstruction of polyhedral objects from their single 2D image the degree of freedom can be reduced by using different visual information such as light intensity [11], texture [2], [22], occluding boundaries [12], vanishing points [15], etc. In shape from shading techniques, light intensity cue is used as the visual information to reduce the degree of freedom of the planar surfaces of polyhedral objects.

Edge detection or line drawing is mainly used in extracting the vertices and edges in 2D images for the 3D reconstruction of the polyhedral shapes. The projection of an object on any plane is called a picture or line drawing of the object. Huffmann [10], categorized the picture plane in three categories on the basis of visibility of adjoining faces of the edges. If both the adjoining faces of the edge are visible, the edge is called convex edge and labeled as +, a concave line is called a concave edge and labeled as -, however, if exactly one of the adjoining face is visible, the edge is called occluding edge and labeled as $\rightarrow$. The main difficulty with the line drawing is the superstrictness problem i.e. the vertex position error caused by computation or inevitable noise in image processing. In figure (1), the situation of correct and incorrect line drawings have been shown in a truncated

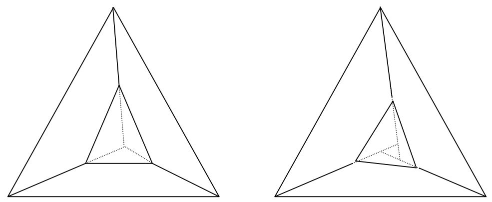


Figure 1: (a) Correct line drawing (b) Incorrect line drawing

pyramid. In figure 1(a) all the shortened edges meet at a common point, but in figure 1(b), the edges do not meet at a common point. Hence in figure 1(a), line drawings are correct while these are incorrect in figure 1(b). Due to vertex position error in line drawing, it is not always possible to reconstruct the 3D shape of a polyhedron. Few approaches have been proposed to reconstruct the 3D shape of a polyhedron in the presence of uncertainty in the vertex positions in line drawing.

In the present paper, the vertex position in image is computed by line drawings or by edge detection viz. Canny edge detection [1] in the real images. The image intensity is used as the visualization cue and two geometrical constraints, plane normal and edge length proportionality are imposed to get the unique solution. Thus, we have formulated a nonlinear minimization problem with nonlinear constraints for the shape recovery of polyhedral object shape from its single 2D shaded image. The minimization of the formulated problem has been accomplished by using Genetic algorithm. The Lambertian reflectance map is most widely used in shape from shading due to its efficiency and simplicity in applications. Hence, we have also assumed that the scene is illuminated by Lambertian reflectance map.

The proposed work has following preferable properties over the existing works.

1. No prior information have been used in this work.
2. It deals with the orthographic as well as perspective images.
3. Only one shaded image is needed in this method.
4. Every planar surface of the object is treated individually and hence the algorithm is applicable in case of presence of slight error in vertex positions in image.
5. The number of variables in the optimization function are very small, the Genetic Algorithm has been used to get the global minima, so the initial guesses to the variables are not necessary.

Rest of the paper is organized as follows. In section 2, we have given a brief review of the related work. In section 3, the Lambertian reflectance map, energy function and imposing constraints are defined. The objective function is obtained and the nonlinear minimization is also described in this section. Results and discussions of experiment on the images are given in section 4. The conclusion of present work is given in section 5.

2 Review of the Related Work

Horn [8] and Kanade [13] are among the first, who have combined the polyhedral assumption with the light intensity information of a single image in order to get a unique solution of polyhedral objects. Most of the polyhedral object constraints were represented in the gradient space and the line representation of the pictures were taken. However, in gradient space, constraints can not be given completely and moreover the represented constraints are too strict. Therefore, in the gradient space, it is not always possible to reconstruct the proper shape of polyhedron from their single 2D image.

The another way of imposing polyhedral costraints is the algebraic representation [3]. Sugihara [20] has given a necessary and sufficient condition for the correct representation of polyhedral scene. Sugihara [21], has used the shading information contained in image as the visualization cue and has given an elegant method for the reconstruction of the polyhedral object. Sugihara has obtained the system of linear equations from the object planes. The line drawing is correct if and only if this system of equations, associated with the line drawing has the solutions. In this method, the new system of linear equations is obtained by detecting and deleting the redundant equations, which lead to superstrictness problem, and from this new system of equations, depth value is obtained by using a nonlinear optimization technique under linear constraints.

Shimshoni *et al.* [18] have taken vertex position error as a variable instead of deleting it as done by Sugihara [21]. The authors have used this variable in quadratic terms in the equations and reduced it to a nonlinear optimization problem under linear constraints.

Heyden [6], [7] and Gunnar [19] have detected the incorrect line drawings on the basis of rank of shape matrix formed from the image and corrected the incorrect line drawings by using singular value decomposition on the shape matrix.

Ros *et al.* [16] have given a method to correct the incorrect line drawing, in which all vertices are moved in their small neighborhood until a close correct line drawing is found. The problem is formulated as a nonlinear minimization problem.

Shimodaria [17] has proposed a SFS method to get the shape of polyhedral objects by using prior information, in which to handle with the superstrictness, gaps are permitted between the faces to avoid the vertex position error and joined them by minimizing the distance between them by using nonlinear cost function under nonlinear constraints.

Feng *et al.* [5] have given a Bayesian method for the reconstruction of 3D shapes from a single image. This method deals with the man-made and natural objects. The method consists of two probabilistic models for man-made and natural objects. The prior visual knowledge has been used as the constraints in the formulation of this method.

Marroqin *et al.* [14] have given a probabilistic solution for the ill-posed problems in Computer Vision. They have given stochastic approach to solve the inverse problems of Computer Vision.

3 Problem Formulation

In this method, shading information is used as the visual information. Shading (variation in image intensity) in the image depends on the condition under which the object is illuminated. The different reflectance maps have been given in order to map the scene radiance to image irradiance. If we assume that the viewer and the light sources are far from the object, then we can introduce the reflectance map, a means of specifying the dependence of brightness on surface orientation. If we elect to use the unit surface normal $\hat{n}$ as a way of specifying surface orientation, then the brightness can be computed as a function of orientation in the form $R(\hat{n})$. If we use p and q instead, we can use the form $R(p, q)$. The most widely used reflectance map in SFS is Lambertian reflectance map which is given as follows.

3.1 Lambertian Reflectance Map

If $\hat{n}$ and $\hat{s}$ are the unit surface normal of the object and the unit illuminate vector respectively i.e.

$$\hat{n} = \frac{(-p,-q,1)}{\sqrt{1+p^2+q^2}}, \qquad \hat{s} = \frac{(-p_s,-q_s,1)}{\sqrt{1+p_s^2+q_s^2}} \qquad (1)$$

then the Lambertian reflectance map is given by their scalar product in the following form

$$R(p,q) = \rho\frac{1+pp_s+qq_s}{\sqrt{1+p^2+q^2}\sqrt{1+p_s^2+q_s^2}} \qquad (2)$$

where ρ is the albedo of the surface and $0 \leq \rho \leq 1$.

Let the equation of the i^{th} planar surface of the polyhedral object be given by

$$p_i x + q_i y + z + r_i = 0 \qquad (3)$$

then the normal vector to the i^{th} planar surface is $\hat{n}_i = (p_i, q_i, 1)$.

Let any three consecutive vertices of the planar surface in the image be $A(x_i, y_i, z_i)$, $B(x_{i+1}, y_{i+1}, z_{i+1})$ and $C(x_{i+2}, y_{i+2}, z_{i+2})$, in which x and y coordinates are known, while, the z coordinates are unknown, which are to be calculated.

Taking the vector product of the vectors $\overrightarrow{AB}$ and $\overrightarrow{AC}$, and making the third component of the product vector to unity, we get

$$p_i = \frac{(y_{i+1}-y_i)(z_{i+2}-z_i)-(y_{i+2}-y_i)(z_{i+1}-z_i)}{(x_{i+1}-x_i)(y_{i+2}-y_i)-(x_{i+2}-x_i)(y_{i+1}-y_i)} \qquad (4)$$

and

$$q_i = \frac{(x_{i+2}-x_i)(z_{i+2}-z_i)-(y_{i+2}-y_i)(x_{i+1}-x_i)}{(x_{i+1}-x_i)(y_{i+1}-y_i)-(x_{i+2}-x_i)(y_{i+1}-y_i)} \qquad (5)$$

If the object is illuminated under the Lambertian reflectance map, then the light intensity of the i^{th} planar surface is given by the following equation

$$I_i = \rho\frac{1+p_ip_s+q_iq_s}{\sqrt{1+p_i^2+q_i^2}\sqrt{1+p_s^2+q_s^2}} \qquad (6)$$

or

$$I_i - \rho\frac{1+p_ip_s+q_iq_s}{\sqrt{1+p_i^2+q_i^2}\sqrt{1+p_s^2+q_s^2}} = 0 \qquad (7)$$

3.2 Energy Function and Constraints

Due to the wide domain of applicability and almost fair approximation to the object surfaces, we have used the Lambertian reflectance map in the image irradiance equation of SFS technique. Theoretically, the left side of the Eqn. (7) should be 0, but experimentally, it is not fair to expect its value to be exactly 0, hence we are intended to minimize the square of the left term in the Eqn. (7) as the minimum of a square term is always 0. Thus the energy function is written as follows.

$$(I_i - \rho\frac{1+p_ip_s+q_iq_s}{\sqrt{1+p_i^2+q_i^2}\sqrt{1+p_s^2+q_s^2}})^2 \qquad (8)$$

On substituting the values of light source direction, albedo (which are known) and the values of variables p_i and q_i from Eqn. (4) and (5), in the above energy function, we get the energy function in which only depth values z_i, z_{i+1} and z_{i+2} are unknown. However, the above energy function is not sufficient to produce the unique depth values for the vertices of the polyhedron. Hence it is necessary to impose the extra constraints for getting the unique depth values. We have added two geometrical constraints [5] to get the unique solution.

The first constraint is imposed as the normal plane constraint. The vector product of each face of each pair of concurrent 3D edges in the same planar face gives the same normal vector to the face

$$\frac{(AB \times AC) \cdot (AC \times BC)}{\|AB \times AC\| \|AC \times BC\|} = 1 \tag{9}$$

which can be written in the minimization form as follows

$$\{1 - \frac{(AB \times AC) \cdot (AC \times BC)}{\|AB \times AC\| \|AC \times BC\|}\}^2. \tag{10}$$

The second constraint is the edge length proportionality constraint. Because the length of the edges of the object should be proportional to the length of the corresponding edges in its 2D projected image. It is assumed that the nature of the projection is known to us.

If e_i, $i = 1, 2, ..., |E|$ are the edges in a planar surface in 3D space and $e_{i'}$, $i = 1, 2, ..., |E|$ are the corresponding edges in the projected image then the minimization can be written as follows:

$$\{\frac{1}{|E|} \frac{\|e_i\|}{\|e_{i'}\|} - \overline{r}\}^2 \tag{11}$$

where $\overline{r} = \frac{1}{|E|} \sum \frac{\|e_i\|}{\|e_{i'}\|}$. The edge length is calculated as the Euclidean length.

Hence the complete minimization functional becomes

$$\begin{aligned}\{I_i - \rho \frac{1 + p_i p_s + q_i q_s}{\sqrt{1 + p_i^2 + q_i^2}\sqrt{1 + p_s^2 + q_s^2}}\}^2 + \\ \lambda\{1 - \frac{(AB \times AC) \cdot (AC \times BC)}{\|AB \times AC\| \|AC \times BC\|}\}^2 + \\ \mu\{\frac{1}{|E|} \frac{\|e_i\|}{\|e_{i'}\|} - \overline{r}\}^2\end{aligned} \tag{12}$$

where λ and μ are the weight factors whose values are given to control the minimization functional. In the above minimization function all the terms are written in the form of coordinate of vertices, so the function has only z_i, z_{i+1} and z_{i+2} unknowns, which are calculated by the minimization of the function using Genetic Algorithm. Since, we are dealing with each planar face individually, there is no place left for the superstrictness problem and the method works well in the case of small error in vertex positions.

3.3 Nonlinear Optimization

In order to search a global minimum of the objective function in Eqn. (12), we have used the existing Genetic Algorithm tool box in MATLAB. The value of control parameter λ in objective function is taken as the area of individual planar surface in the picture because the area of the planar surface in the projected image is proportional to the area of the planar surface in the scene. The value of μ can be taken any scalar quantity as the reconstructed shape is proportional to the original shape and not necessarily to be the original shape. We have taken the value of μ as unity due to the simplification in the calculation. Generally, there is no criteria for the initial guess of the variables, which are to be computed by optimization algorithms. The Genetic Algorithm [4] is preferable in this sense because in Genetic Algorithm, the optimization is possible without giving the initial value to the variables. The objective function in Eqn. (12) has been given as the fitness function with two variables. The minimization has been done at least up to the order of 10^{-8} accuracy. The minimization process has been done repeatedly for each planar surfaces of the polyhedral object. We assign the value of z as 0 to one of the vertices of the initial (starting) plane surface. By minimizing the above function, the z values are calculated of the three vertices of each planar face and the z values for the rest vertices are calculated by using the equation of the planar surface.

4 Experiments and Results

For the construction of the scene of truncated pyramid, the coordinates of the vertices of the truncated pyramid are taken as by Sugihara [21] and given in table 1. The constructed scene is projected orthographically on xy -plane.

There are six vertices and five visible planes in the projected image, which are numbered in the parenthesis. First plane in background is assumed as xy-plane. The light source direction to illuminate the object is considered as $(1, -1, 4)$ and the albedo (ρ) as 1. Using Eqn. (6), the image intensity is calculated for each visible planar face of the polyhedral scene and a shaded image is generated as shown in figure 2(b).

Since the image is synthetic, the vertex positions are corrected upto digitization error. For showing the applicability of the presented algorithm, the incorrect

Table 1: Co-ordinates of Truncated Pyramid Scene.

No.	X	Y	Z
1	0.7	3.8	0
2	10	10	0
3	7.6	7.5	0
4	6.0	4.9	8.0
5	3.02	3.9	5.27
6	0.7	3.8	0
7	6.6	2.6	6.5
8	8.3	0.5	0
9	0	0	0
10	10	0	0

Table 2: Co-ordinates of Cube Scene.

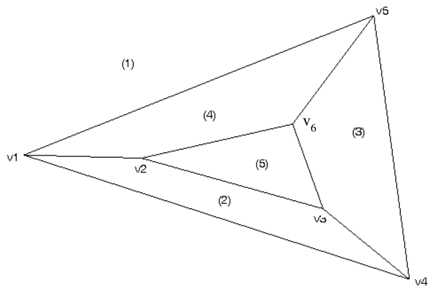
No.	X	Y	Z
1	4	1	1
2	5	4	3
3	7	5	8
4	6	2	6
5	0	0	0
6	1	3	2
7	3	4	7
8	2	1	5

Table 3: Co-ordinates of Book Scene.

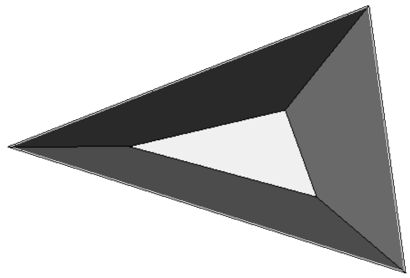
No.	X	Y	Z
1	1	3	0
2	6	3	0
3	2	5	3
4	7	5	3
5	1	7	0
6	6	7	0

Table 4: Co-ordinates of Truncated Cube Scene.

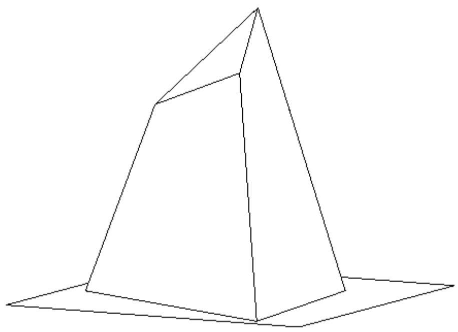
No.	X	Y	Z
1	2	1	2
2	3	1	2
3	3	1.5	2
4	2.5	2	2
5	2	2	2
6	2	1	1
7	3	1	1
8	3	2	1
9	2	2	1
8	3	2	1.5



(a)



(b)

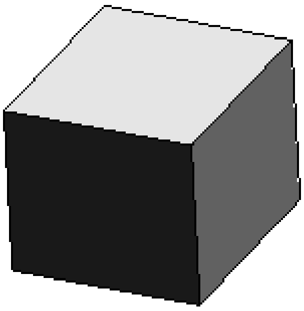


(c)

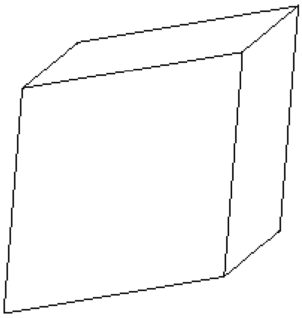
Figure 2: (a) Structure of the truncated pyramid (b) Shaded image of truncated pyramid (c) Recovered shape of truncated pyramid

line drawing was generated by adding random perturbation ± 0.1 in the image coordinate values of each vertex in the correct line drawing. The random perturbation were generated by using the uniform distributed numbers within the above range.

The shaded image in figure 2(b) is the input image for the 3D reconstruction. The image intensities of each planar face $I_i, i = 1, 2, ..., 5$ are known in case of synthetic images and are calculated from the image in case of real images. The (x, y) coordinates of each vertex are known in case of synthetic image and are calculated by using any appropriate edge detection technique (Canny

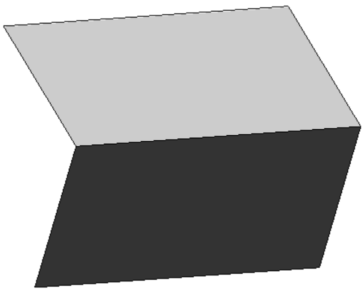


(a)

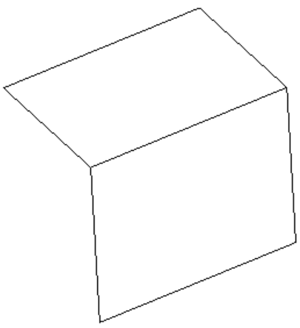


(b)

Figure 3: (a) Shaded image of cube (b) Reconstructed Shape of cube

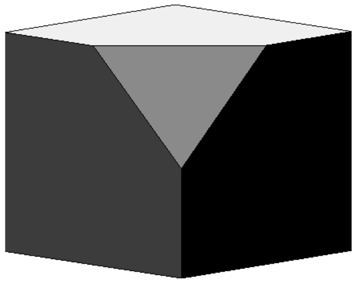


(a)

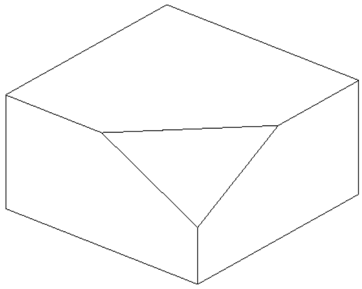


(b)

Figure 4: (a) Shaded image of Open Book (b) Reconstructed Shape of Open Book



(a)



(b)

Figure 5: (a) Shaded image of Truncated Cube (b) Reconstructed Shape of Truncated Cube

edge detection) [1], in case of real images. The z coordinate values of any three vertices in a plane are calculated by minimizing Eqn. (12) as discussed in section 3.3 and the z values of other vertices in the same plane are calculated by using the equation of plane. The reconstructed shape of the truncated pyramid is shown in figure 2(c). The Synthetic image of cube is generated by using the coordinates of the vertices given in table 2. The orthographic shaded image of cube is generated by using the Lambertian reflectance map and shown in figure 3(a). The shape of cube is reconstructed by using the proposed methodology and shown in figure 3(b). The third test image is the synthetic image of the open book shaped structure. The shape of open book is made by using the coordinates given in table 3. The shaded image of open book is shown in figure 4(a). The shape of open book is reconstructed by using the proposed method and shown in figure 4(b). The final test is the synthetic image of a truncated cube with four visible faces. First, the 3D scene of truncated cube is generated by using the coordinates shown in table 4 and then the projected scene is projected orthographically xy plane. The intensity of each visible face is calculated by using the Lambertian reflectance map and a shaded image of truncated cube is obtained as shown in figure 5(a). The proposed method is applied on this obtained image and a 3D shape of truncated pyramid is reconstructed as shown in figure 5(b). In all experiments, the recovered shapes are similar to the shapes of actual objects.

We briefly compare the proposed method with some well known existing methods [21], [18] and [17].

In Sugihara's method [21], all the faces are considered together, thus the fundamental equations for the line drawing are too strict. The system of equations have no solution in case of incorrect line drawing. For getting the solution in case of incorrect line drawing, the redundant equations responsible for superstrictness are deleted. However, in our method, each planar face of the polyhedral object is dealt separately and the depth values at each vertices in the same planar surface are calculated. Hence, no place left for the superstrictness problem.

In Shimshoni's method [18], uncertainty at the vertices due to the error in vertex positions is taken as variable and the fundamental equations are formulated in non-linear terms. The non-linear terms are linearized by introducing new variables. The number of variables in the proposed method are relatively too small than the Sugihara and Shimshoni's method.

In Shimodaria's method, planar faces are treated separately and joined by Sequentially - Face - Positioning method. The prior information are used in this method as the constraints for the reconstruction. In the proposed method, no prior information and the number of variables are used in this method are very small as compared to Shimodaria's method. The Genetic algorithm is used for the global minimum search which is better than the optimization used in [21], [18] and [17].

5 Conclusions

The proposed method is simple and robust, because it has only two variables in objective function and the method is also applicable in case of slight error in the vertex positions. In this method, no prior information is needed and the Genetic algorithm is used for the minimization of energy function. Hence, the optimization is possible without any initial guess of the variables. The method is tested on synthetic images and satisfactory results are obtained.

However, this method has some drawbacks. This method deals only with the visible surface of the object but not any hidden surface. The Lambertian reflectance map has been used in this method, however, all the surfaces specially, shiny objects can not be approximated by the Lambertian reflectance map.

Acknowledgements

One of the authors, Manoj Kumar, gratefully acknowledges the financial support of MHRD to carry out this research work.

References

[1] Canny, J. A computational approach to edge detection. *EEE Trans. Pattern Analysis and Machine Intelligence*, 8(6):679–698, Nov 1986.

[2] Coleman, E., Jr. and Jain, R. Obtaining 3-dimensional shape of textured and specular surface using four-source photometry. *Jones and Bartlett Publishers, Inc. , USA*, pages 180–199, Jan 1992.

[3] Falk, G. Interpretation of imperfect line data as a three dimensional scene. *Artificial intelligence*, 3:101–144, 1972.

[4] Goldberg, D. Genetic algorithms in search, optimization and machine learning. *Addison-Wesley, New York*, 1989.

[5] Han, F. and Zhu, S.-C. Bayesian reconstruction of 3d shapes and scenes from a single image. *1st IEEE Int. workshop on Higher-Level Knowledge in 3D Modeling and Motion Analysis*, 2003.

[6] Heyden, A. A new metod for correction of images of origami/polyhedral objects. *Swedish Symposium on Image Analysis*, 1993.

[7] Heyden, A. Consistency and correction of line-drawings, obtained by projection of piecewise planar objects. *LNCS, Springer Berlin*, 1(800):411–419, June 1994.

[8] Horn, B. Understanding image intensities. *Artificial intelligence*, 8:201–231, 1977.

[9] Horn B. K. P., M. J. B. Shape from shading. *The MIT Press, Cambridge*, 1989.

[10] Huffman, D. A. Realizable configurations of lines in pictures of polyhedra. *Machine Intelligence*, 8:493–509, 1977.

[11] Ikeuchi, K. Determinining surface orientation of speculare surfaces by using the photometric streo method. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 3:635–641, 1981.

[12] Ikeuchi, K. and Horn, B. Numerical shape from shading and occluding boundaries. *Mit Press Series Of Artificial Intelligence Series, Shape from shading*, pages 245–299, June 1989.

[13] Kanade, T. Recovery of the three-dimensional shape of an object from a single view. *Artificial intelligence*, pages 409–460, 1981.

[14] Marraoquin J., P. T., Mitter S. Probablistic solution of ill-posed in computational vision. *Journal of American Statistical Asssociation*, 3(82):76–89, 1987.

[15] Nakatani, H. and Kitahashi, T. Inferring 3-d shape from line drawing using vanishing point. *1st Int. Conf. Computer Applications, Peking*, pages 411–419, 1984.

[16] Ros, L. and Thomas, F. Overcoming superstrictness in line drawing interpretation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(4):456–466, Apr 2002.

[17] Shimodaria, H. A shape-from-shading method of polyhedral objects using prior information. *EEE Trans. Pattern Analysis and Machine Intelligence*, 28(4):612–624, Apr 2006.

[18] Shimshoni, I. and Ponce, J. Recovering the shape of polyhedra using line-drawing analysis and complex reflectance models. *Computer Vision and Image Understanding*, 65(2):296–310, Feb 1997.

[19] Sparr, G. Depth computations from polyhedral images. *ECCV,Proceedings of the Second European Conference on Computer Vision,Springer-Verlag*, 588:378–386, May 1992.

[20] Sugihara, K. Mathematical structure of line drawings of polyhedrons: Toward man-machine communication by means of line drawings. *IEEE Trans. pattern Analysis and Machine Intelligence*, 4(5):458–469, 1982.

[21] Sugihara, K. An algebraic approach to shape-from-image problems. *Artificial intelligence*, 23(1):59–95, May 1984.

[22] Witkin, A. Recovering surface shape and orientation from texture. *Artificial intelligence*, pages 17–45, 1981.