IAPR Workshop on CV - Special Hardware and Industrial Applications OCT.12-14, 1988, Tokyo

RECONSTRUCTION OF CONSISTENT SHAPE FROM INCONSISTENT DATA: OPTIMIZATION OF 2½D SKETCHES

Ken-ichi Kanatani

Department of Computer Science, Gunma University
Kiryu, Gunma 376, Japan

ABSTRACT

Although the 3D orientations of edges and surfaces are theoretically sufficient for reconstructing the 3D object shape, this does not mean that the 3D object shape can actually be reconstructed: Inconsistency may result if image data contain errors. We propose a scheme of *optimization* to construct a consistent object shape from inconsistent data. Our optimization is achieved by solving a set of *linear* equations. This technique is first applied to the problem of *shape from motion* and then to the 3D recovery based on the *rectangularity hypothesis* and the *parallelism hypothesis*.

1. Constraints on 2½D Sketches

In the past, various 3D shape recovery techniques called *shape from ...* have been proposed (*shape from motion, shape from shading, shape from texture*, etc.). Now, we must ask the following question: Do these techniques really enable us to recover the 3D object shape? The shape from ... paradigms usually present us with object images equipped with the following types of 3D information:

(i) The surface gradient (p, q), or equivalently the unit surface normal $\boldsymbol{n}$, is *densely* estimated (Fig. 1(a)).

(ii) The region corresponding to the object surface is segmented into planar patches, and the surface gradient (p, q), or equivalently the unit surface normal $\boldsymbol{n}$, is estimated for each patch (Fig. 1(b)).

(iii) The region corresponding to the object surface is segmented into planar patches, and 3D edge orientations are estimated (Fig. 1(c)).

We call an image equipped with such 3D information a *2½D sketch*.[5]

We consider cases (ii) and (iii), and assume that the object surface is approximated by a polyhedron. We say that a vertex is *incident* to a face if the vertex is on the boundary of the face. Let $V=\{V_1, ..., V_n\}$ be the set of its vertices, and $F=\{F_1, ..., F_m\}$ be the set of its faces. The *incidence structure* is specified by a set R of *incidence pairs* (F_α, V_i) meaning that vertex V_i is incident to face F_α. Let l be the number of such incident pairs.

Let (X_i, Y_i, Z_i) be the scene coordinates of vertex V_i, $i=1, ..., n$, and let $Z=p_\alpha X+q_\alpha Y+r_\alpha$ be the equation of face F_α, $\alpha=1, ..., m$. The pair (p_α, q_α) indicates its surface gradient. Let us call $p_\alpha, q_\alpha, r_\alpha$ the *surface parameters* of face F_α. The incidence pair (F_α, V_i) states that vertex V_i is incident to face F_α:

$$Z_i=p_\alpha X_i+q_\alpha Y_i+r_\alpha. \tag{1.1}$$

In this paper, we use a coordinate system fixed to the camera in such a way that the Z-axis coincides with the optical axis and point $(0, 0, -f)$ (which we henceforth call the *viewpoint*) coincides with the center of the lens. Then, we can think of the XY-plane as the image plane (Fig. 2).

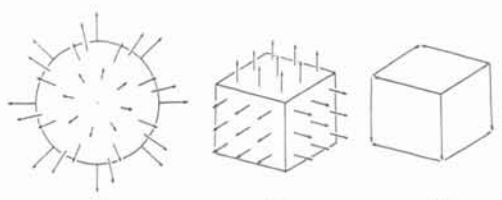

Fig. 1 *2½D sketch.* (a) The surface gradient is densely estimated. (b) The surface gradient is estimated for each planar patch. (c) The 3D orientation is estimated for each edge.

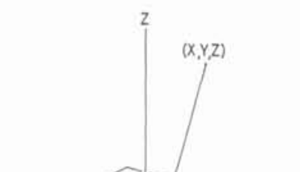

Fig. 2 Perspective projection.

It is easy to see that the image coordinates (x_i, y_i) of vertex V_i are related to its scene coordinates (X_i, Y_i, Z_i) by the projection equations

$$x_i=\frac{fX_i}{f+Z_i}, \quad y_i=\frac{fY_i}{f+Z_i}. \tag{1.2}$$

Now, we introduce a new quantity

$$z_i\equiv\frac{fZ}{f+Z} \tag{1.3}$$

for each vertex V_i and call it the *reduced depth* of the vertex. Then, we can express the reduced depth z_i of vertex V_i in terms of its image coordinates (x_i, y_i):

$$z_i=\frac{fp_\alpha}{f+r_\alpha}x_i+\frac{fq_\alpha}{f+r_\alpha}y_i+\frac{fr_\alpha}{f+r_\alpha}. \tag{1.4}$$

We define new parameters

$$P_\alpha\equiv\frac{fp_\alpha}{f+r_\alpha}, \quad Q_\alpha\equiv\frac{fq_\alpha}{f+r_\alpha}, \quad R_\alpha\equiv\frac{fr_\alpha}{f+r_\alpha}, \tag{1.5}$$

and call these the *reduced surface parameters*. In terms of the reduced surface parameters $P_\alpha, Q_\alpha, R_\alpha$, the reduced depth z_i is written as

$$z_i=P_\alpha x_i+Q_\alpha y_i+R_\alpha. \tag{1.6}$$

Since the image coordinates (x_i, y_i) of vertex V_i are known, the 3D position of vertex V_i is determined if its reduced depth z_i is known. Consequently, the reduced depths z_i, $i=1, ..., n$, can be taken as unknowns for the 3D vertex positions instead of the original depths Z_i, $i=1, ..., n$. Similarly, the reduced surface parameters P_α, Q_α, R_α, $\alpha=1, ..., m$, can serve as unknowns for the surface shape instead of the original surface parameters p_α, q_α, r_α, $\alpha=1, ..., m$. Thus, we obtain l (= the number of incidence pairs) equations

$$x_iP_\alpha+y_iQ_\alpha+R_\alpha-z_i=0,\quad (F_\alpha, V_i)\in R. \tag{1.7}$$

2. Optimization of a 2½D Sketch

2.1 Surface Gradients Estimated

Suppose we are given a 2½D sketch. Let $R=\{(F_\alpha, V_i)\}$ be its incidence structure. Let $(\hat p_\alpha, \hat q_\alpha)$ be the estimate of the surface gradient of face F_α. Let (p_α, q_α) be the true surface gradient of face F_α. Here, we seek, from among the infinitely many consistent polyhedron solutions which are exactly projected onto the observed image, the one whose surface gradients are the closest to the given estimates on the average (Fig. 3). Specifically, let us consider the least square method to minimize

$$J=\frac{1}{2}\sum_{\alpha=1}^{m}W_\alpha[(p_\alpha-\hat p_\alpha)^2+(q_\alpha-\hat q_\alpha)^2], \tag{2.1}$$

where W_α is the weight for face F_α. If eqns (1.7) are substituted, eqn (2.1) is rewritten in terms of the reduced surface parameters as follows.

$$J=\frac{1}{2}\sum_{\alpha=1}^{m}W_\alpha\left(\frac{f+r_\alpha}{f}\right)^2[(P_\alpha+\frac{\hat p_\alpha}{f}R_\alpha-\hat p_\alpha)^2+(Q_\alpha+\frac{\hat q_\alpha}{f}R_\alpha-\hat q_\alpha)^2]. \tag{2.2}$$

In order to make our analysis easy, we replace r_α in the above equation by its *estimate* $\hat r_\alpha$, assuming that it is somehow available. Then, we put

$$W_\alpha\left(\frac{f+\hat r_\alpha}{f}\right)^2=\frac{1}{\zeta_\alpha}. \tag{2.3}$$

Now, ζ_α is a constant assigned to face F_α whose value is yet to be determined.

The problem now reduces to the minimization of J under the constraints (1.7) for all incidence pairs $(F_\alpha, V_i)\in R$. As long as we use surface gradient cues, we must give the depth Z, or equivalently the reduced depth z, to one vertex. Let that vertex be V_n. Since J is quadratic and the constraints (1.7) are linear in the unknowns, the minimization is achieved by solving a set of linear equations. If we introduce Lagrangian multipliers $\Lambda_{\alpha i}$ to all the incidence pairs $(F_\alpha, V_i)\in R$, the final result becomes as follows:

$$x_iP_\alpha+y_iQ_\alpha+R_\alpha-z_i=0,\quad (F_\alpha, V_i)\in R, \tag{2.4}$$

$$P_\alpha+\frac{\hat p_\alpha}{f}R_\alpha+\zeta_\alpha\sum_{i:(F_\alpha,V_i)\in R}x_i\Lambda_{\alpha i}=\hat p_\alpha,\quad \alpha=1, ..., m, \tag{2.5}$$

$$Q_\alpha+\frac{\hat q_\alpha}{f}R_\alpha+\zeta_\alpha\sum_{i:(F_\alpha,V_i)\in R}y_i\Lambda_{\alpha i}=\hat q_\alpha,\quad \alpha=1, ..., m, \tag{2.6}$$

$$\sum_{i:(F_\alpha,V_i)\in R}\left(\frac{\hat p_\alpha x_i+\hat q_\alpha y_i}{f}-1\right)\Lambda_{\alpha i}=0,\quad \alpha=1, ..., m, \tag{2.7}$$

$$\sum_{\alpha:(F_\alpha,V_i)\in R}\Lambda_{\alpha i}=0,\quad i=1, ..., n-1. \tag{2.8}$$

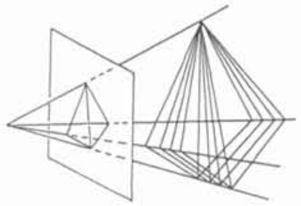

Fig. 3 *Optimization.* From among infinitely many consistent polyhedron solutions which are exactly projected onto the observed image, the one whose surface gradients are the closest to the given estimates on the average is chosen.

2.2 Edge Orientations Estimated

Next, consider case (iii) - the 3D edge orientations estimated. Let $E_\alpha=\{e_1, ..., e_{N_\alpha}\}$ be the set of edges constituting the boundary of face F_α, and let $\hat e_k=(\hat e_{k(1)}, \hat e_{k(2)}, \hat e_{k(3)})$ be the unit vector indicating the estimated 3D orientation of edge $\hat e_k$. If (p_α, q_α) is the surface gradient of face F_α, the unit surface normal $n_\alpha=(n_{\alpha(1)}, n_{\alpha(2)}, n_{\alpha(3)})$ to face F_α is given by

$$n_\alpha=\left(\frac{p_\alpha}{\sqrt{p_\alpha^2+q_\alpha^2+1}}, \frac{q_\alpha}{\sqrt{p_\alpha^2+q_\alpha^2+1}}, \frac{-1}{\sqrt{p_\alpha^2+q_\alpha^2+1}}\right). \tag{2.9}$$

The vectors $\hat e_k$ for $\hat e_k\in E_\alpha$ are supposed to be all orthogonal to n_α, but this is not necessarily guaranteed in the presence of noise. Hence, it is reasonable to estimate the surface normal n_α by the least square method which minimizes

$$\frac{1}{2}\sum_{e_k\in E_\alpha}(\hat e_k, n_\alpha)=\frac{1}{2}\sum_{e_k\in E_\alpha}(\hat e_{k(1)}n_{\alpha(1)}+\hat e_{k(2)}n_{\alpha(2)}+\hat e_{k(3)}n_{\alpha(3)})^2. \tag{2.10}$$

However, we need not do this minimization for each face separately. The surface normals of all the faces are estimated by minimizing

$$J=\frac{1}{2}\sum_{\alpha=1}^{m}W_\alpha\sum_{e_k\in E_\alpha}(\hat e_{k(1)}n_{\alpha(1)}+\hat e_{k(2)}n_{\alpha(2)}+\hat e_{k(3)}n_{\alpha(3)})^2$$

$$=\frac{1}{2}\sum_{\alpha=1}^{m}\frac{W_\alpha}{p_\alpha^2+q_\alpha^2+1}\sum_{e_k\in E_\alpha}(\hat e_{k(1)}p_\alpha+\hat e_{k(2)}q_\alpha-\hat e_{k(3)})^2, \tag{2.11}$$

where W_α is the weight for face F_α. Following the procedure shown earlier, we finally obtain the following result.

$$x_iP_\alpha+y_iQ_\alpha+R_\alpha-z_i=0,\quad (F_\alpha,V_i)\in R, \tag{2.12}$$

$$A_{11}P_\alpha+A_{12}Q_\alpha+\frac{1}{f}A_{13}R_\alpha+\zeta_\alpha\sum_{i:(F_\alpha,V_i)\in R}x_i\Lambda_{\alpha i}=A_{13},\quad \alpha=1, ..., m, \tag{2.13}$$

$$A_{21}P_\alpha+A_{22}Q_\alpha+\frac{1}{f}A_{23}R_\alpha+\zeta_\alpha\sum_{i:(F_\alpha,V_i)\in R}y_i\Lambda_{\alpha i}=A_{23},\quad \alpha=1, ..., m, \tag{2.14}$$

$$A_{31}P_\alpha+A_{32}Q_\alpha+\frac{1}{f}A_{33}R_\alpha+\zeta_\alpha\sum_{i:(F_\alpha,V_i)\in R}\Lambda_{\alpha i}=A_{33},\quad \alpha=1, ..., m, \tag{2.15}$$

$$\sum_{\alpha:(F_\alpha,V_i)\in R}\Lambda_{\alpha i}=0,\quad i=1, ..., n-1, \tag{2.16}$$

where

$$A_{ij}\equiv\sum_{e_k\in E_\alpha}\hat e_{k(i)}\hat e_{k(j)},\quad i,j=1,2,3. \tag{2.17}$$

3. Optimization for Shape from Motion

Suppose we are given a sequence of images of a polyhedron moving in a scene. Let us assume that the point-to-point correspondence has already been detected. Consider a face which has four or more corners. If the image velocities are

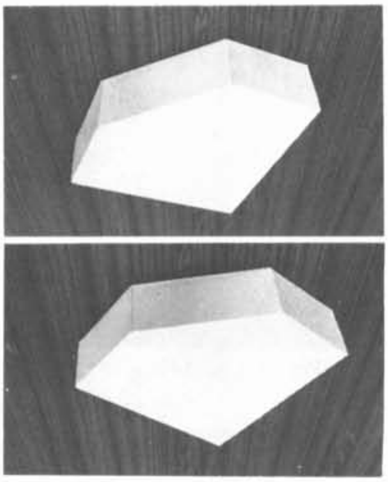

Fig. 4 Two object images.

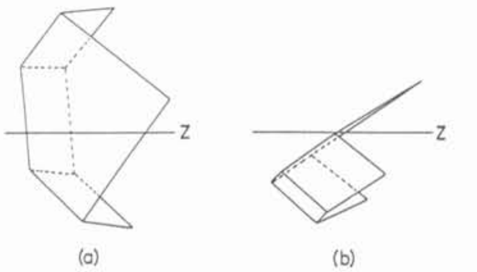

Fig. 5 The 3D shape reconstructed from Fig. 4: (a) The top view. (b) The side view.

observed at at least four vertices, we can compute the rotation velocity $(\omega_1, \omega_2, \omega_3)$ and the surface gradient (p, q) of this face by the analytical formula of Kanatani[3]. Thus, we can estimate the surface gradient (p, q) for the faces which have four or more corners. Then, the optimization technique is applied to the resulting 2½D sketch.

Example 1. Consider the two images of Fig. 4. Regarding the displacements as instantaneous velocities, and applying the procedure described above, the 3D shape shown in Fig. 5 is obtained. Fig. 5(a) shows the top view (orthographic projection onto the YZ-plane), while Fig. 5(b) shows the side view (orthographic projection onto the ZX-plane).

4. Optimization of Rectangularity Heuristics

If one corner is known to be rectangular and has three visible edges, we can compute the 3D orientations of the three edges[1,2,4,6] Here, we use the formulation of Kanatani[4]. Each edge orientation indicates the surface normal to the face defined by the other two edges. Hence, we can determine the surface gradient of the three faces. Then, the optimization technique is applied to the resulting a 2½D sketch.

Thus, the remaining question is how to find rectangular corners. Our algorithm is based on the following two considerations.

- **Rectangularity test**. Corner images which cannot be projections of rectangular corners are removed.[4]

- **Compatibility test**. Choose two corners which share at least one face, and compute the 3D edge orientations and the surface gradients, assuming that both are rectangular corners. If this assumption is correct, the computation must predict an identical 3D orientation for the connecting edge (if the two corners are connected) and identical surface gradients for the common faces (within some fixed tolerance).

- **Maximal compatible sets**. Form maximal compatible sets of corners in such a way that as many corners are included as possible unless incompatible pairs arise among them. Assuming that the corners belonging to each set are all rectangular, we end up with as many 2½D sketches as these maximal compatible sets.

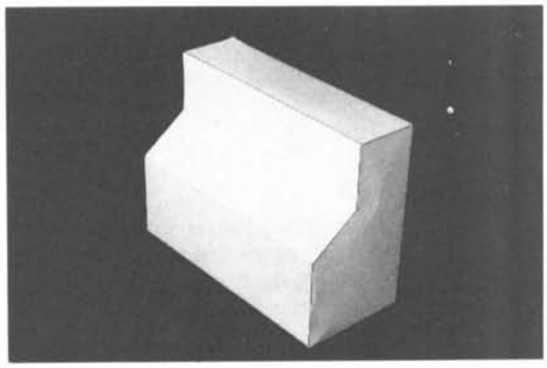

Fig. 6 A polyhedron image.

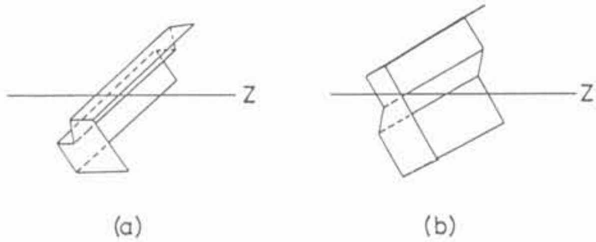

Fig. 7 One 3D shape reconstructed from Fig. 6: (a) The top view. (b) The side view.

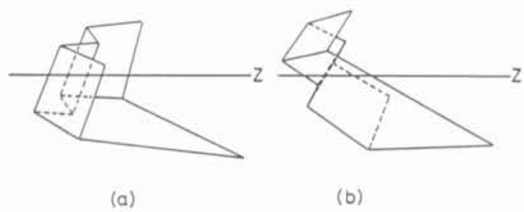

Fig. 8 Another 3D shape reconstructed from Fig. 6: (a) The top view. (b) The side view.

Example 2. Fig. 6 is a real image of a polyhedron. The rectangularity test cannot reject any vertices as definitely nonrectangular. The compatibility test tells us that these vertices are split into two compatible groups. Hence, we obtain two solutions. Applying the optimization technique, we can reconstruct the two 3D shapes shown in Figs. 7 and 8. In both of them, (a) shows the top view (orthographic projection onto the YZ-plane) while (b) shows the side view (orthographic projection onto the ZX-plane).

6. Optimization of Parallelism Heuristics

If two lines in the image are interpreted to be projections of parallel lines in the scene, they define a *vanishing point* on the image plane, which determines the 3D orientation of these lines. Hence, if we can find a set of edges which are parallel in the scene, their 3D orientation is computed from their vanishing

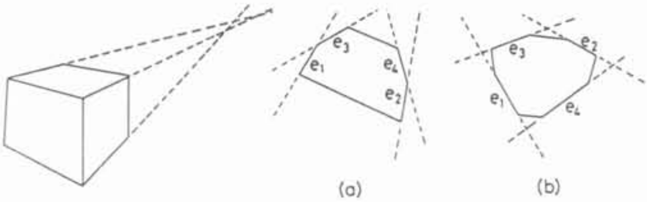

Fig. 9 *Concurrency test*: Parallel edges in the scene are projected onto concurrent lines on the image plane (within some tolerance). *Parallelogram test*: If edges e_1, e_2 are parallel, edges e_3, e_4 cannot be parallel in (a) but can be parallel in (b).

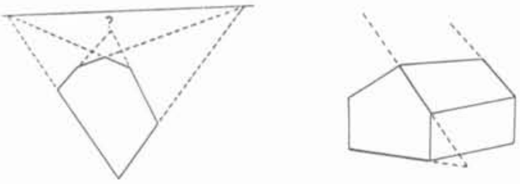

Fig. 10 *Collinearity test*: The vanishing points of parallel edges belonging to the same face must be collinear (within some tolerance). *Vanishing point heuristic*: Two edges are more likely to be parallel in the scene if their intersection is farther away from the image origin O.

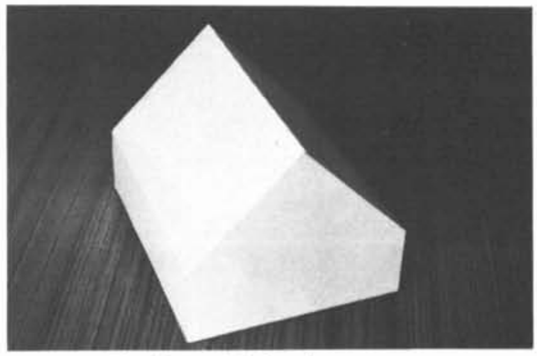

Fig. 11 A polyhedron image.

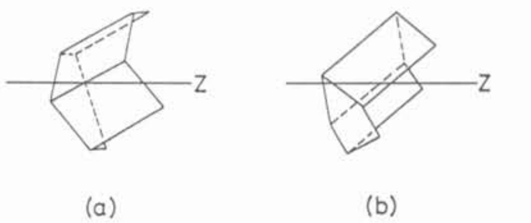

Fig. 12 The 3D shape reconstructed from Fig. 11: (a) The top view. (b) The side view.

points, and we can obtain a 2½D sketch with estimated 3D edge orientations. Then, we apply the optimization technique.

Thus, it remains to construct an algorithm for finding parallel edges. We apply the following procedures:[8]

- **Concurrency test**. A set of edges concurrent on the image plane are judged to be parallel in the scene (within some tolerance) (Fig. 9).

- **Common incidence heuristic**. The concurrency test is incapable of detecting a pair of parallel edges if no other edges are parallel to them. It is reasonable to check only those pairs which share a common face but no common vertex, because it is highly unlikely that two edges belonging to different faces are parallel, yet no other edges are parallel to them.

- **Parallelogram test**. If two pairs of parallel lines lying on the same plane are projected onto half-lines starting from their respective vanishing points, they must intersect with each other at exactly four points on the image plane (Fig. 9).

- **Collinearity test**. If three or more sets of parallel lines belong to the same face, their vanishing points must be collinear (within some tolerance) (Fig. 10).

- **Vanishing point heuristic**. Those edge pairs which have passed these two tests are assumed to be parallel in the scene. However, if one edge appears in different pairs of parallel edges, we choose the pair whose intersection is located farthest away from the image origin (Fig. 10).

Example 3. Fig. 11 is a real image of a polyhedron. Applying the concurrency test, we detect two sets of parallel edges. From among the remaining edges, the pairs which share common faces but no common vertices are the next candidates for parallel pairs. All of these pairs pass both the parallelogram test and the collinearity test. Since some edges belong to multiple pairs, we invoke the vanishing point heuristic. Applying the optimization technique, we can recover the 3D shape up to a single scale factor, and obtain the 3D shape shown in Fig. 12. Fig. 12(a) shows the top view (orthographic projection onto the YZ-plane), while Fig. 12(b) shows the side view (orthographic projection onto the ZX-plane).

Acknowledgment: This work was supported in part by Casio Science Promotion Foundation, Yazaki Memorial Foundation for Science and Technology, Inamori Foundation, and the Japanese Ministry of Education, Science, and Culture under Grant in Aid for Scientific Research (No. 63550268).

REFERENCES

1. S. T. Barnard, Choosing a basis for perceptual space, *Comput. Vision Graphics Image Process.*, **29** (1985), 87 - 99.
2. K. Kanatani, The constraints on images of rectangular polyhedra, *IEEE Trans. Pattern Anal. Machine Intell.*, **PAMI-8** (1986), 456 - 463.
3. K. Kanatani, Structure and motion from optical flow under perspective projection, *Comput. Vision Graphics Image Process.*, **38** (1987), 122 - 146.
4. K. Kanatani, Constraints on length and angle, *Comput. Vision Graphics Image Process.*, **41** (1988), 28 - 42.
5. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, W. H. Freeman, San Francisco, 1982.
6. T. Shakunaga and H. Kaneko, Perspective angle transformation and its application to 3-D configuration recovery, *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, June 1986, Miami Beach, FL, pp. 594 - 601.
7. K. Sugihara, *Machine Interpretation of Line Drawings*, MIT Press, Cambridge, MA, 1986.
8. T. Tanaka, K. Kawashima and K. Kanatani, *3D Recovery of Polyhedra by Parallelism Heuristics*, Technical Report CS-87-6, Department of Computer Science, Gunma University, December 1987.