Perceptual organization with image formation compatibilities

Josef Pauli *, Gerald Sommer

Christian-Albrechts-Universität zu Kiel, Institut für Informatik und Praktische Mathematik, Preußerstraße 1–9, D-24105 Kiel, Germany

Abstract

The work presents a methodology contributing to boundary extraction in images of approximate polyhedral objects. We make extensive use of basic principles underlying the process of image formation and thus reduce the role of object-specific knowledge. Simple configurations of line segments are extracted subject to geometric-photometric compatibilities. The perceptual organization into polygonal arrangements is based on geometric regularity compatibilities under projective transformation. The combination of several types of compatibilities yields a saliency function for extracting a list of most salient structures. Based on systematic measurements during an experimentation phase the adequacy and degrees of compatibilities are determined. The methodology is demonstrated for objects of various shapes located in cluttered scenes.

Key words: Perceptual organization, boundary extraction, geometric-photometric compatibility, geometric regularity compatibility, Hough transformation.

1 Introduction

Computer Vision procedures are based on expectations whose spectrum stretches from general assumptions, e.g. ramp profiles of gray-value edges, to specific models for object recognition, e.g. relational structures of geometric entities. A critical question is: *Which kind of expectations can be applied reasonably along the chain of processing and how are they acquired ?* If expectations are too general then the number of possible interpretations of image contents will

* Corresponding author. Tel. +49-431-560484; Fax. +49-431-560481.
Email address: jpa@ks.informatik.uni-kiel.de (Josef Pauli).
URL: www.ks.informatik.uni-kiel.de/∼jpa/ (Josef Pauli).

Article published in: **Pattern Recognition Letters 23 (2002) 803-817**

increase dramatically. Otherwise, if expectations are too specific and do not comply with the variability of possible situations then relevant structures can hardly be detected. We propose a methodology of treating this dilemma for the task of perceptual organization, e.g. for object boundary extraction in 2D images.

Early work on perceptual organization in Computer Vision dates back to the nineteen eighties, e.g. by Lowe (1985), and since then the principle of *nonaccidentialness* was postulated. It states that organized arrangements of image features are highly unlikely and hence their actual occurrence is significant for solving tasks such as object detection. Just a few *Gestalt principles*, e.g. proximity, similarity, closure, and continuation, which lead to simple organizations, can drastically reduce the search trees of recognition approaches. For example, Jacobs (1996) locates salient convex collections of line segments by an exclusive use of such general principles. Apart from Gestalt principles, Zisserman et al. (1995) also exploit more specific expectations, i.e. object shape models for extracting parallelepipeds and surfaces of revolution. Further principles of image formation are conceivable to reduce the role of object shape models.

The Gestalt principles are subject to real image formation, and therefore must involve inaccuracy and uncertainty. Relevant techniques have been developed, e.g. based on *fuzzy logic* by Walker et al. (1994) and by Ralescu et al. (1999), or based on *Dempster-Shafer theory* by Vasseur et al. (1999). However, the acquisition of fuzzy membership or probability assignment functions is hardly (or not at all) treated in the articles. A general learning-based mechanism would be necessary for automatically incorporating the complicated image formation effects which are caused by the specific camera and originate (among others) from the three-dimensionality of the world. Kim et al. (1999) apply artificial neural networks for automatical learning of conditional probabilities, which express relations between evidential image features and hypothetical organizations, and use these probabilities for decision making by *Bayes rules*. Only aerial images of buildings are treated and only 2D rectangles (roughly) are extracted.

The novelty of our work is fourfold. First, the theoretical concept of *invariance* (Wechsler , 1990, pp. 95-160) is relaxed into the practical concept of *compatibility*. The use of compatibilities reduces the amount of object-specific knowledge for medium-level vision tasks like attention control and *boundary extraction*. Second, we maximally exploit those kind of compatibilities which originate and are inherent in the *three-dimensional nature* of objects and in the *image formation principles*. Compatibilities between geometric and photometric features and between elementary and structured geometric entitities are considered. The related work in Zerroug et al. (1993) uses *geometric quasi-invariants* for curved objects, but doesn't treat the gap between geometry and photometry. Third, the compatibilities are determined on the basis of *statisti-*

cal measurements which must be taken during an *experimentation phase* prior to application (importance repeatedly stressed in Klette et al. (2000)). Systematic experiments are needed for *threshold setting* for procedures of line extraction and perceptual grouping. Fourth, we integrate a series of gestaltic cues spanning over *signal level*, *primitive level*, *structural level*, and *assembly level* (four-level classification proposed by Sarkar et al. (1993) and revisited by Boyer et al. (1999)). Evidences from the various levels are involved and integrated in the *quality assessment* of alternative organizations.

We present a catalogue of propositions each describing a certain type of compatibility (sections 2 and 3). The various types of compatibilities are integrated in procedures for extracting specific shapes of polygons and polygon arrangements (section 4). Finally, we demonstrate the usefulness of our methodology by analyzing five demanding images (section 5). The relevant compatibilities depend on thresholds δ_i which are learned in an experimentation phase.

2 Geometric-photometric compatibilities

The propositions in this section describe types of compatibilities between global geometric entities and local gray-value structures in the image.

2.1 Orientation compatibility between lines and edges

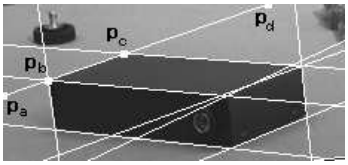
The *orientation-deviation* between orientation ϕ of an object boundary line in the image (assuming polar form representation) and the orientations $\mathcal{I}^O(p_i)$ of all gray-value edges along the points $(p_1, \cdots, p_N)$ of a segment $\mathcal{L}$ of the image line is defined by

$$D_{LE}(\phi, \mathcal{L}) := \frac{1}{N} \cdot \sum_{i=1}^{N} D_{OL}(\phi, \mathcal{I}^O(p_i)) \; ; \quad with \tag{1}$$

$$D_{OL}(\phi, \mathcal{I}^O(p_i)) := \frac{\min\{|h_i|, |h_i + 180°|, |h_i - 180°|\}}{90°} \; ; h_i := \phi - \mathcal{I}^O(p_i) \tag{2}$$

Proposition 1 *Given δ_1 as permissible orientation-deviation. The line/edge orientation compatibility holds subject to image formation if $D_{LE}(\phi, \mathcal{L}) \leq \delta_1$.*

Figure 1(a) shows a black box and candidate boundary lines which have been extracted by Hough transformation (Leavers , 1993). For one of them, going through points $\{p_a, p_b, p_c, p_d\}$, we show the series of edge orientations in Figure 1(b), which are the local gradient angles. In consensus with Proposition 1, just for the boundary segment $(p_b, \cdots, p_c)$ the series is close to the line orientation.

(a) 

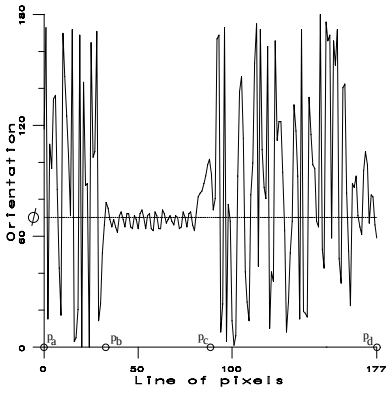
(b) 

Fig. 1. (a) Black box, boundary lines; (b) Edge orientations along the line with indices.

2.2 Junction compatibility between pencils and corners

A *pencil* is a simple configuration of M line segments meeting at one common *pencil point* (Faugeras , 1993, pp. 8,17). At the gray-level corner located nearest to a pencil point the two-dimensional gray-value structure will be considered. The *junction-deviation* between a pencil at pencil point p_l with line orientations $\mathcal{A} := (\alpha_1, \cdots, \alpha_M)$ and a collection of edge sequences meeting at corner point p_c with local orientations $\mathcal{B} := (\beta_1, \cdots, \beta_M)$ is defined by

$$D_{PC}(p_l, p_c, \mathcal{A}, \mathcal{B}) := \omega_1 \cdot D_{JP}(p_l, p_c) + \omega_2 \cdot D_{JO}(\mathcal{A}, \mathcal{B}) \tag{3}$$

$$D_{JP}(p_l, p_c) := \frac{\|p_l - p_c\|}{I_d} \tag{4}$$

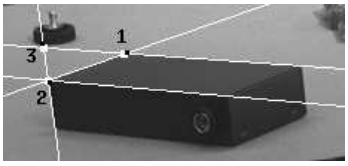
$$D_{JO}(\mathcal{A}, \mathcal{B}) := \frac{\sum_{i=1}^{M} \min\{|\alpha_i - \beta_i|, |\alpha_i - \beta_i + 360^\circ|, |\alpha_i - \beta_i - 360^\circ|\}}{180^\circ \cdot M} \tag{5}$$

It is a weighted summation of two components, i.e. the Euclidean distance between pencil point and corner point (normalized by the constant diagonal I_d of a standard image size, e.g. 512×512 pixel), and the deviation between the orientation of a pencil line and of a corresponding edge sequence (averaged over all such pairs).

Proposition 2 *Given δ_2 as permissible junction-deviation. The pencil/corner junction compatibility holds subject to image formation if $D_{PC}(p_l, p_c, \mathcal{A}, \mathcal{B}) \leq \delta_2$.*

Figure 2(a) shows a subset of four boundary lines, three pencil points (white dots) with indices $1, 2, 3$, and a subset of three nearest gray-value corner points (black dots). The latter are extracted by the SUSAN operator (Smith et al. , 1997). For example, the pencil/corner junction compatibility holds for point

2, where we have a pencil of three lines. The diagram in Figure 2(b) shows the characterization of the local gray-value structure, i.e. orientation-dependent significance measurement for the occurrence of edge sequences, which is computed by a steerable wedge filter (Simoncelli et al. , 1996). The three peaks, which indicate the occurrence of three edge sequences for certain orientations, are close to three vertical diagram lines, which indicate the orentations of the pencil lines. This kind of compatibility holds as well for point 1 but not for point 3 (not shown).

(a) 

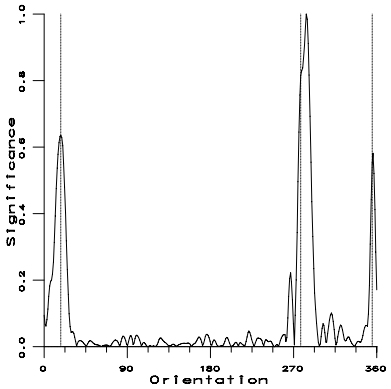
(b) 

Fig. 2. (a) Subset of boundary lines, pencil points (white dots), corner points (black dots); (b) Orientation-dependent significance measurements for edge sequences at point 2.

2.3 Phase compatibility between parallels and ramps

The *local phase* characterizes the type of a gray-value edge, i.e. ascending or descending ramps, and top or bottom directed roofs (Granlund et al. , 1995, pp. 258-278). Being a one-dimensional concept, we show the series of local phases exemplary by scanning the virtual, vertical line in Figure 3(a) from top to bottom. At the first and second intersection points with the object boundary the ramps are descending, and at the third point the ramp is ascending. This behavior is in consensus with the quantitative series of the polar angle (representing the local phases), as shown in Figure 3(b). In particular, the sign of the local phase at the first boundary line is converse to the sign at the opposite boundary line of the object. Generally, this is true if all gray values of the object are lower or higher than the gray values of the local background. Based on this observation and assumption, a criterion for the detection of opposite boundary lines of an object is proposed.

Let $\mathcal{L}_1$ and $\mathcal{L}_2$ be two approximate parallel line segments. The two mean values of the local phases along these segments (computed orthogonal to the line orientations) are denoted by $f^{ph}(\mathcal{L}_1)$ and $f^{ph}(\mathcal{L}_2)$. We define the *phase-similarity* between the two mean phases such that the similarity between equal

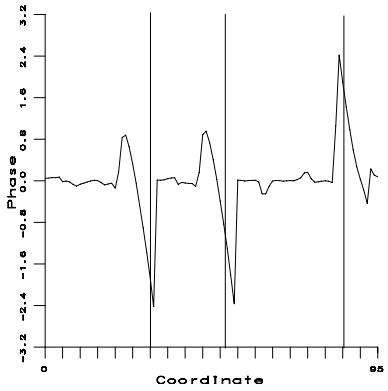
(a)  (b) 

Fig. 3. (a) Virtual straight line over the object surface; (b) Local phases along the line.

phases is 1 and the similarity between phases with converse signs is 0.

$$D_{PR}(\mathcal{L}_1, \mathcal{L}_2) := \left| 1 - \frac{|f^{ph}(\mathcal{L}_1) - f^{ph}(\mathcal{L}_2)|}{\pi} \right| \tag{6}$$

Proposition 3 *Given δ_3 as permissible deviation from 0. The parallel/ramp phase compatibility holds subject to image formation if $D_{PR}(\mathcal{L}_1, \mathcal{L}_2) \leq \delta_3$.*

The presented geometric-photometric compatibilities are the foundation for applying the following list of pure geometric compatibilities.

3 Geometric compatibilities for perceptual organization

The propositions describe compatibilities between elementary and structured geometric entitites which are subject to the process of image formation, i.e. approximate perspective transformation.

3.1 Patterns of Hough peaks for approximate-parallel lines

Based on polar parameters r and ϕ of straight lines we apply *Hough transformation* for line extraction. The horizontal and vertical axes of the Hough image are taken correspondingly. The Hough transformation of parallel image lines (having identical value ϕ) yields a horizontal sequence of peaks in the Hough image. Under projective transformation, two parallel lines in 3D remain almost parallel in the image, i.e. there is a small *angle-deviation* $D_{OL}(\phi_1, \phi_2)$.

Proposition 4 *Given δ_4 as permissible angle-deviation. The parallelism compatibility of two lines holds subject to image formation if $D_{OL}(\phi_1, \phi_2) \leq \delta_4$.*

Considering Proposition 4, parallel 3D lines occur as peaks in the Hough image located within a horizontal stripe of height δ_4. Figure 4(b) shows the Hough image when applying Hough transformation to the image in Figure 4(a). We restricted the process to a quadrangle image window around the black box and selected 12 peaks which are organized in four stripes of three peaks, respectively. For example, three *approximate parallel lines* are shown on the left, which are specified by the peaks in the third stripe of the Hough image.

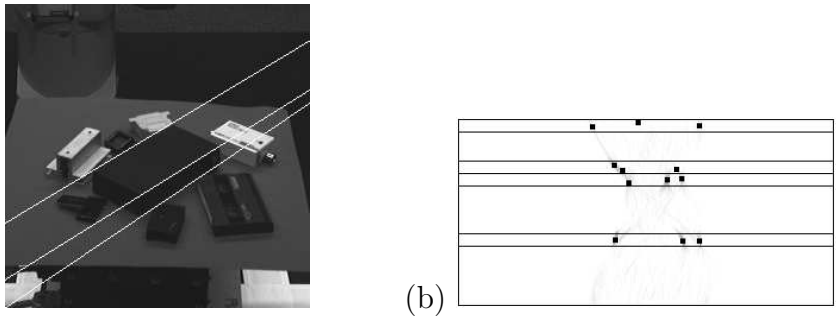


Fig. 4. (a) Subset of three approximate parallel boundary lines for the black box object; (b) Hough image and peaks marked by black dots.

3.2 Regularity compatibilities for polygons

Approximate parallel line segments may occur in *approximate regular polygons*. The basic component for describing polygon regularities is a *polyline*. We specify a polygon as the union of two non-overlapping polylines $\mathcal{G}_1$ and $\mathcal{G}_2$, possibly including single line segments located at the end of each polyline, respectively. Figure 5 shows in the two pictures a regular polygon, respectively on the left side. The one in Figure 5(a) contains a pair of reflected polylines (reflection-symmetric), the one in Figure 5(b) contains a pair of parallel polylines (parallel-symmetric). The angle-deviation $D_{OP}(\mathcal{G}_1, \mathcal{G}_2)$ between two approximate parallel polylines is defined as the mean value of angle-deviations between the constituting approximate parallel line segments.

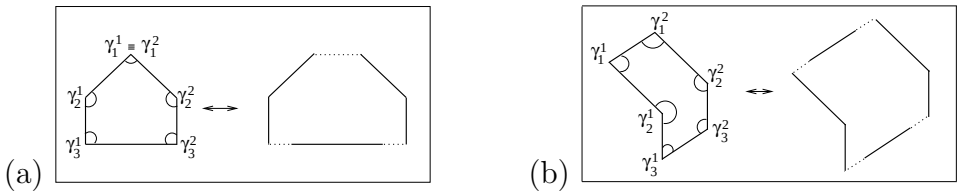


Fig. 5. Examples of regular polygons: (a) Reflection-symmetric polylines; (b) Parallel-symmetric polylines.

Proposition 5 *Given δ_5 as permissible angle-deviation. The parallel-symmetry compatibility of two polylines holds subject to image formation if $D_{OP}(\mathcal{G}_1, \mathcal{G}_2) \leq \delta_5$.*

Similar propositions can be formulated for *reflection-symmetry compatibility* and *right-angle compatibility*. They are based on permissible deviations from exact reflections or exact right-angles.

3.3 Vanishing-point compatibility of boundary lines

The projective transformation of parallel boundary lines yields image lines whose extensions should meet in one vanishing-point (see Figure 6). This imposes certain qualitative constraints on the series of Hough peaks within a horizontal stripe, which we summarize as the *vanishing-point compatibility*. A similar constraint was formulated by Stahs et al. (1992) but under slope/intercept parameterization of lines.

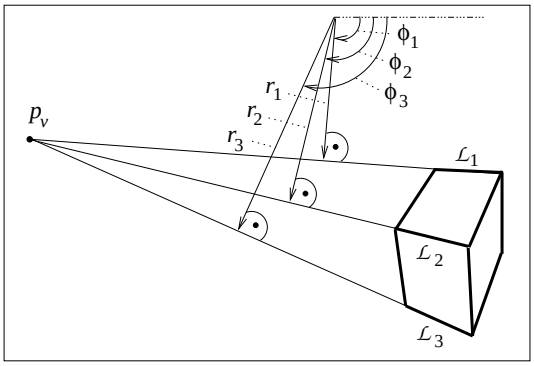


Fig. 6. Projected parallelepiped, vanishing point, monotonous relationship between distance parameter and angle parameter.

Proposition 6 *Let* $\{\mathcal{L}_1, \cdots, \mathcal{L}_V\}$ *be a set of approximate parallel line segments in the image, which originate from projective transformation of parallel line segments of the 3D object boundary. The extensions of the image line segments meet at a common vanishing point* p_v *and can be sorted according to the strong monotony* $r_1 < \cdots < r_i < \cdots < r_V$ *of the parameter* r*. For this arrangement there is a weak monotony of the angle parameter,*

$$\phi_1 \geq \cdots \phi_i \geq \cdots \geq \phi_V \quad or \quad \phi_1 \leq \cdots \phi_i \leq \cdots \leq \phi_V \tag{7}$$

The vanishing-point compatibility can be examined for the Hough image in Figure 4(b). Proposition 6 holds for the third and fourth stripe but not for the first and second stripe. To make it completely valid, we must apply a strategy which slightly modifies parameters r and ϕ of the relevant image lines.

3.4 Pencil compatibility of meeting boundary lines

The most prominent corner type of man-made objects is a pencil of three lines. For a subset of the object corners all three boundary lines are visible. In the case that these image lines are extracted, we can impose the following *pencil compatibility*.

Proposition 7 *Let us assume a 3D pencil and the pencil point of three meeting boundary lines of an approximate polyhedral object. The projective transformation of the 3D pencil must yield just one pencil in the image plane, i.e. just one 2D pencil point.*

For example, Figure 2(a) shows three lines intersecting in a close neighborhood at point 2. According to Proposition 7 just one intersection point is accepted which must be considered in the process of line extraction.

4 Generic procedures for perceptual organization

The compatibilities are the foundation of procedures for perceptual organization. Prior to the working phase, the thresholds δ_i must be determined in an experimentation phase. We present generic procedures for the extraction of quadrangles, polygons, and arrangements of polygons. These procedures are generic in the sense that weighted combinations of compatibilities can be configured, and specific types of polygonal or polyhedra shapes can be treated.

4.1 Procedure for the extraction of specific quadrangles

The following procedure extracts a set of most salient quadrangles of a certain type, e.g. rectangles, parallelograms, or rhombuses. These are special cases of right-angled, parallel-symmetric or relection-symmetric polygons.

1. In the Hough image determine a certain number of peaks which must be organized in a certain number of stripes.
2. For each pair of stripes, and for each pair of peaks in the first stripe, and for each pair of peaks in the second stripe, do:
   - Intersect the lines specified by the four Hough peaks and construct line quadrangles.
   - Compute the mean line/edge orientation-deviation based on function D_{LE}.
   - Optionally compute the mean pencil/corner junction-deviation based on function D_{PC}.
   - Optionally compute the mean parallel/ramp phase-deviation based on function D_{PR}.
   - Erase those quadrangles which don't meet the geometric-photometric compatibility.
   - Compute the deviation from a certain quadrangle type based on functions D_{OL} and D_{OP}, i.e. regularity compatibilities.
3. Compute the saliency values for the quadrangles by combining the above measurements, i.e. the geometric-photometric compatibilities and the regularity compatibilities.

4.2 Procedure for the extraction of specific polygons

The following procedure is an extension of the previous one. It extracts a set of most salient polygons of a certain type, e.g. right-angled, parallel-symmetric or relection-symmetric polygons.

1. In the Hough image determine a certain number of peaks which must be organized in a certain number of stripes.
2. For each combination of three Hough peaks, subject to the constraint that first and third peak don't belong to the same stripe as the second peak, do:
   - Extract the line segment by intersecting first and third line with the second one.
   - Compute the line/edge orientation-deviation.
   - Erase the line segment if the geometric-photometric compatibility doesn't hold.
3. Compute a graph which represents the neighborhood among the remaining line segments, i.e. create a knot for each meeting point and an arc for each line segment.
4. Compute the set of minimal, planar cycles in the graph, i.e. minimal numbers of knots and no arc in the graph is intersecting the cycles, which results in a candidate set of polygons.
5. For each polygon compute the mean line/edge orientation-deviation, and optionally the mean pencil/corner junction-deviation.
6. Erase those polygons which don't fulfill the geometric-photometric compatibility.
7. For the remaining polygons compute the deviations from a certain polygon type.
8. Compute the saliency values for the polygons by combining the above measurements, i.e. the geometric-photometric compatibilities and the regularity compatibilities.

The following procedure treats polygons in combination in order to extract arrangements of polygons which may approximate the complete boundary (i.e. the visible part) of a 3D object. In this article we consider only the case of parallelepipeds whose visible surface consists of three parallelograms (in general), respectively.

1. Detect a quadrangle image window which contains an object of approximate parallelepiped shape.
2. Extract the 2D boundary of the object silhouette in the image which is assumed to be the most salient hexagon consisting of three pairs of approximate parallel line segments.
3. Propagate the silhouette lines (outer boundary lines) to the interior of the silhouette in order to extract the inner lines. This is done by considering the parallelism compatibility, the vanishing-point compatibility, and the pencil compatibility.
4. Compute for the set of line segments the mean line/edge orientation-deviation, and optionally the mean pencil/corner junction-deviation and the mean parallel/ramp phase-deviation.
5. Compute the saliency values for the parallelepipeds by combining geometric-photometric compatibilities and regularity compatibilities.

This procedure is very specific, but other more sophisticated procedures can be implemented for treating other polyhedral shapes as well. The interested reader is refered to Pauli (2001, pp. 78-81).

5 Experiments on perceptual organization

Systematic measurements are performed during an experimentation phase to determine thresholds for the degrees of compatibilities, i.e. statistically reasonable values for threshold parameters $\delta_1, \cdots, \delta_5$. These verified image formation compatibilities prove to be useful for extracting boundaries of objects in complex environments. We show exemplary results of the extraction of polygons

and polygon arrangements.

5.1 Learning the degree of parallelism compatibility

Generally, the experiments produce series or distributions from which to determine extreme or variance values, and these in turn are used to specify reasonable values for various parameters. For example, in numerous applications the distribution can be approximated by a Gaussian function respectively, and the Gaussian support can be used to specify the degree of compatibility. In this article we only present (exemplary) experimental measurements for determining the degree of parallelism compatibility, i.e. parameter δ_4 as introduced in Subsection 3.1. This value will be used for analyzing Hough images, i.e. specifying the height of horizontal stripes within which to search for peaks.

We mount a task-relevant objective, put the camera at a task-relevant place, and take images from a training object under varying rotation angle. An elongated, rectangular paper is used as training object with the color in clear contrast to the background. From the images of the rotating object we extract a certain pair of object boundary lines, i.e. the pair of approximate parallel lines which are the longest. In the experiment, a camera objective with focal length $6mm$ has been used, the distance between camera and object was about $300mm$, and rectangular paper has been rotated in discrete steps of $10°$ from approximately $0°$ to $180°$. Figure 7(a) shows a subset of four images at the rotation angles $10°$, $50°$, $100°$, $140°$, and therein the extracted pairs of approximate parallel lines. For the whole series of discrete object rotations the respective difference between the polar angles of the lines is shown in the diagram of Figure 7(b). This difference varies between $1°$ and $7°$, and the maximum is reached when the principal object axis is collinear with the direction of the optical axis. The maximum value is used for specifying δ_4 (including a small multiplicative uncertainty factor).

Generally, in all experiments we must take care of a sufficient and appropriate set of images in order to obtain adequate extreme or variance values. The training sitations have to be representative for the working phase in the sense that similar conditions of image formation are provided, e.g. same camera objective. As the experimentation-based approach catches the process of image formation generically, it is possible to make use of the learned compatibilities in all images taken under these conditions. More experiments on parameter learning are reported in Pauli (2001, pp. 85-96). In the following two subsections we demonstrate our methodology of perceptual organization.

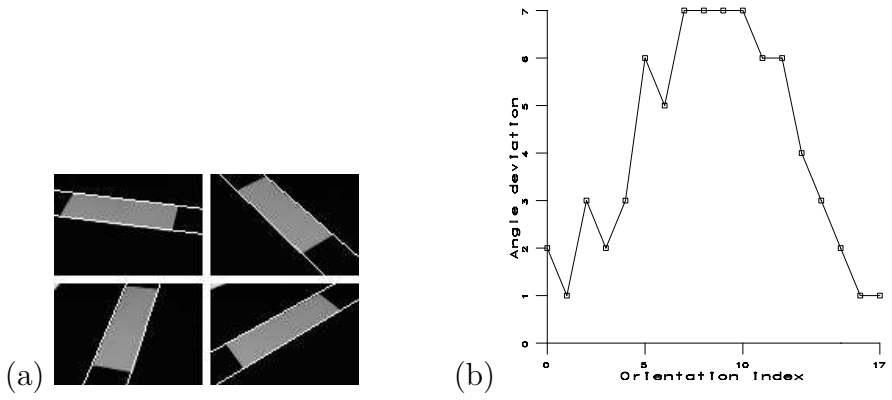


Fig. 7. (a) Images taken under focal length $6mm$, rectangle object in four orientations from a series of 18, and extracted pairs of approximate parallel lines; (b) Series of deviations from exact parallelism for the rectangle object under rotation.

5.2 Exemplary results of the extraction of polygons

The first series of experiments consists of three applications which treat natural scenes of buildings and of electrical scrap. We concentrate on two-dimensional (more or less) object components. They are located within complex environments, and the gray value structure of the surfaces may also be complicated (e.g. for electronic boards).

In the first application the image depicts a factory plant including two buildings (see Figure 8(a) [1]). The perceptual task is to organize line segments into approximate parallelograms which may serve as an early step in the process of detecting the buildings.

The generic procedure in Subsection 4.1 will be used which searches salient quadrangles of a certain shape by taking just compatibilities into account and applies no other object-specific knowledge. We configured the procedure to consider the line/edge orientation compatibility, the parallelism compatibility, the parallel-symmetry compatibility, and to search a subset of most salient approximate parallelograms. In the Hough image we detected a set of 30 most maximal peaks organized in 5 stripes which yield the lines in Figure 8(b). A subset of 11 most salient parallelograms, extracted from the lines, is shown in Figure 8(c). Among the parallelograms, a subset of 2 most salient rhombuses is shown in Figure 8(d), which includes the roof silhouette of a building.

In the second application the image shows the interior of a computer including an electronic board which is of approximate rectangular shape (see Figure 9(a)). We configured the procedure in Subsection 4.1 to consider the line/edge orientation compatibility, the parallelism compatibility, the parallel-

[1] It is taken from Sarkar's web site http://marathon.csee.usf.edu/∼sarkar/ .

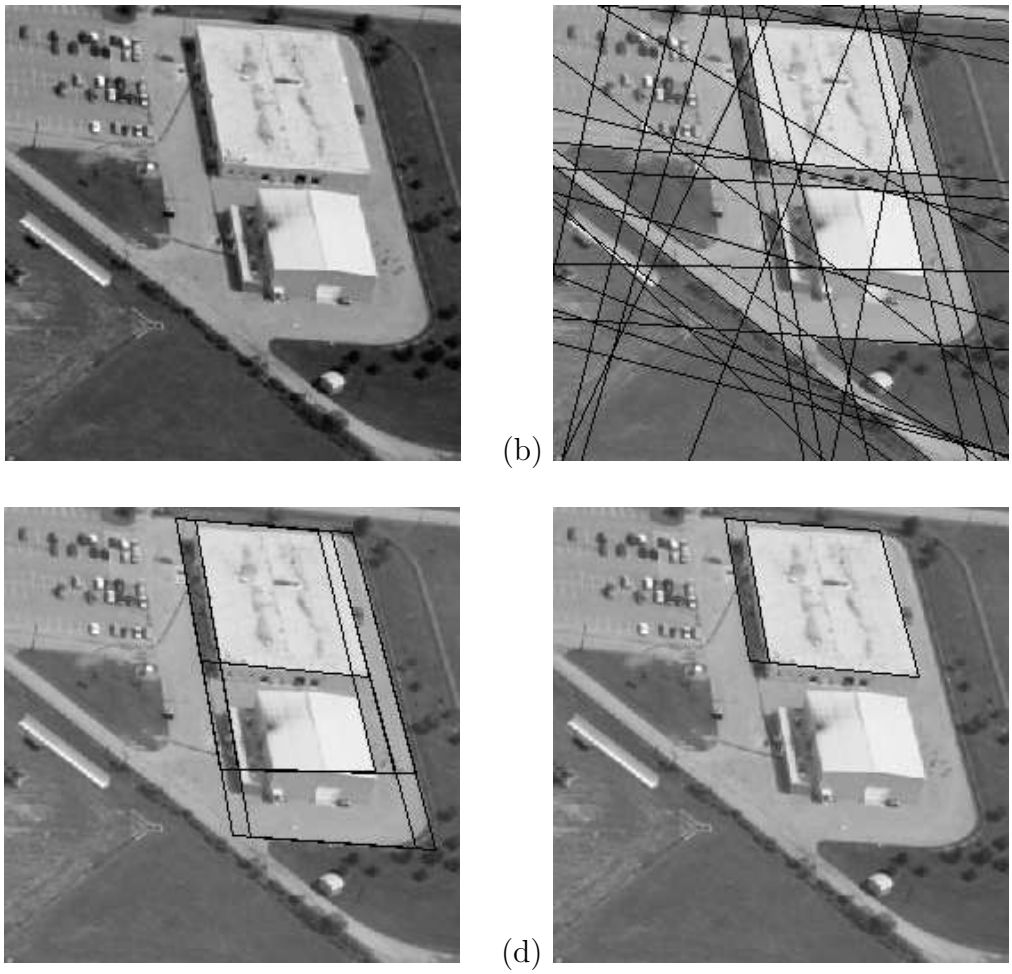


Fig. 8. Extraction of salient quadrangles: (a) Factory plant including two buildings; (b) Subset of 30 most salient lines; (c) Subset of 11 most salient parallelograms; (d) Subset of two most salient rhombuses.

symmetry compatibility, and to search a subset of most salient approximate rectangles. In the Hough image we detected a set of 10 most maximal peaks organized in 2 stripes which yield the lines in Figure 9(b). A subset of 4 most salient rectangles, extracted from the lines, is shown in Figure 9(c), and the most salient one in Figure 9(d).

In the third application the image shows again the interior of a computer, and the electronic board is of approximate right-angled hexagonal shape (see Figure 10(a)). The generic procedure in Subsection 4.2 will be used which searches salient polygons of a certain shape by taking just compatibilities into account. We configured the procedure to consider the line/edge orientation compatibility. In the Hough image we detected a set of 50 most maximal peaks which yield the lines in Figure 10(b). A subset of 10 most salient hexagons, extracted from the lines, is shown in Figure 10(c). Furthermore, if the parallelism compatibility and the right-angle compatibility are considered additionally, we obtain among the hexagons the most salient right-angled hexagon as shown

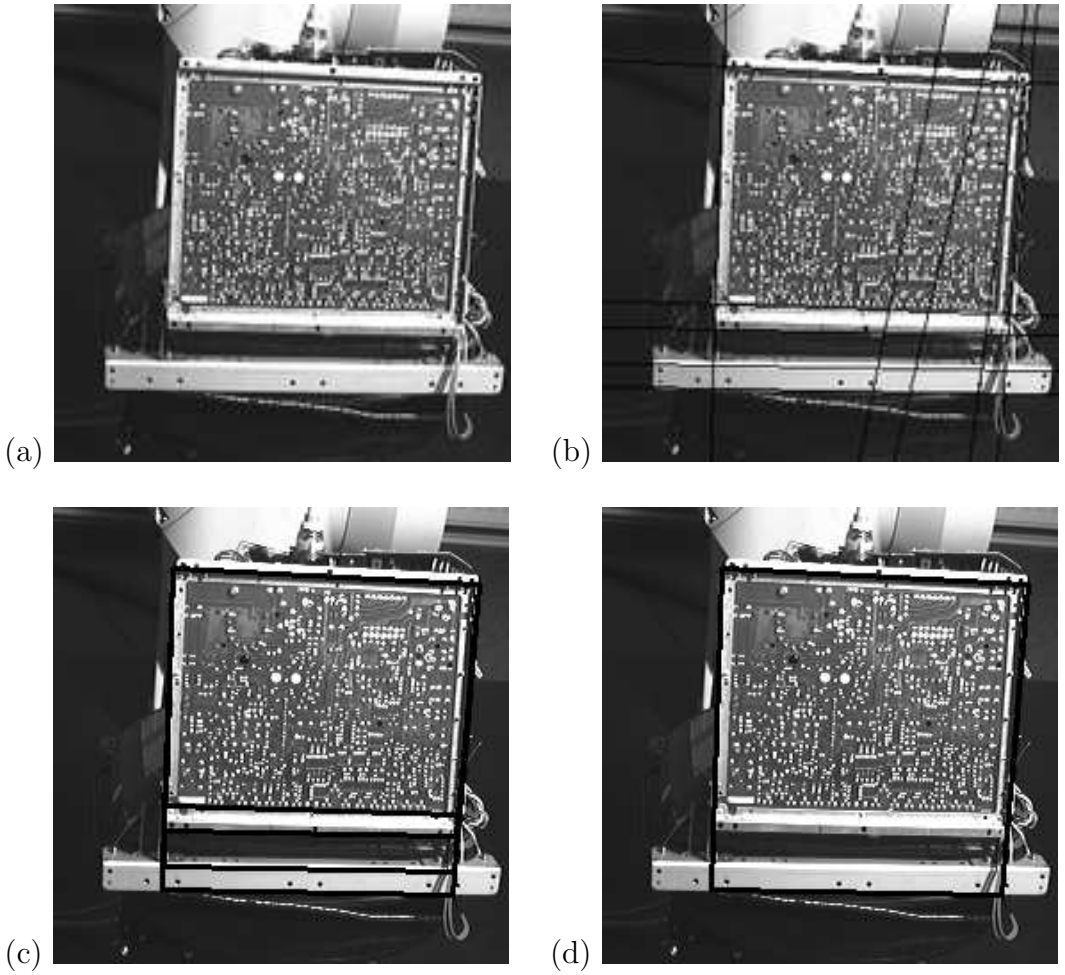


Fig. 9. Extraction of salient rectangles: (a) Computer interior with rectangular board; (b) Subset of 10 most salient lines; (c) Subset of four most salient rectangles; (d) Most salient rectangle.

in Figure 10(d).

5.3 Exemplary results of the extraction of polygon arrangments

The second series of experiments consists of two more applications (number 4 and 5) which treat natural scenes of electrical scrap whose objects are of three-dimensional shape.

In the fourth application the image shows a set of 3D objects including a large black box, which is of approximate right-angled, parallelepiped shape (see Figure 11(a)). The object is located in a complex background, and the gray-value structure of the object surface is nearly homogeneous with low gray-value contrast between neighboring faces. The generic procedure in Subsection 4.3 will be used which searches salient parallelepipeds by taking just

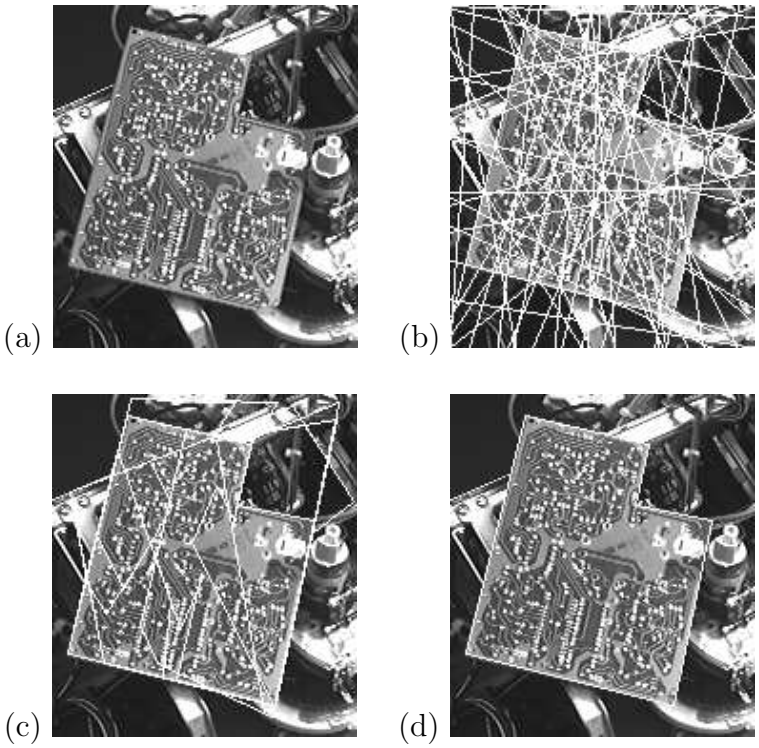


Fig. 10. Extraction of salient hexagons: (a) Computer interior with hexagonal board; (b) Subset of 50 most salient lines; (c) Subset of 10 most salient hexagons; (d) Most salient right-angled hexagon.

compatibilities into account. We configured the procedure to consider the line/edge orientation compatibility, the parallel/ramp phase compatibility, the parallel-symmetry compatibility, the vanishing-point compatibility, and the pencil compatibility. The process of perceptual organization is restricted to a quadrangle image window around the black box (see Figure 11(b)). This quadrangle has been extracted (among others) using the generic procedure in Subsection 4.1 and selected by considering appearance-based features based on the gray-level distribution within the quadrangles. Based on the windowed Hough transformation we detected in the Hough image a set of 12 most maximal peaks organized in 4 stripes, and these peaks represent the lines shown in Figure 11(c). The most salient parallelepiped is shown in Figure 11(d).

In the fifth application the image shows only one 3D object, i.e. a transceiver box (see Figure 12(a)). For robotic tasks like object grasping we are interested in extracting the coarse shape, e.g. a parallelepiped. The difficulty in perceptual organization is that the shape is only roughly a parallelepiped, i.e. including protrusions and indentations, and the gray-level structure of the object surface contains natural inscription. The generic procedure in Subsection 4.3 will be applied which is configured to consider the line/edge orientation compatibility, the parallel-symmetry compatibility, the vanishing-point compatibility, and the pencil compatibility. In the Hough image we detected a set of 15 most maximal peaks organized in 3 stripes which yield the lines in

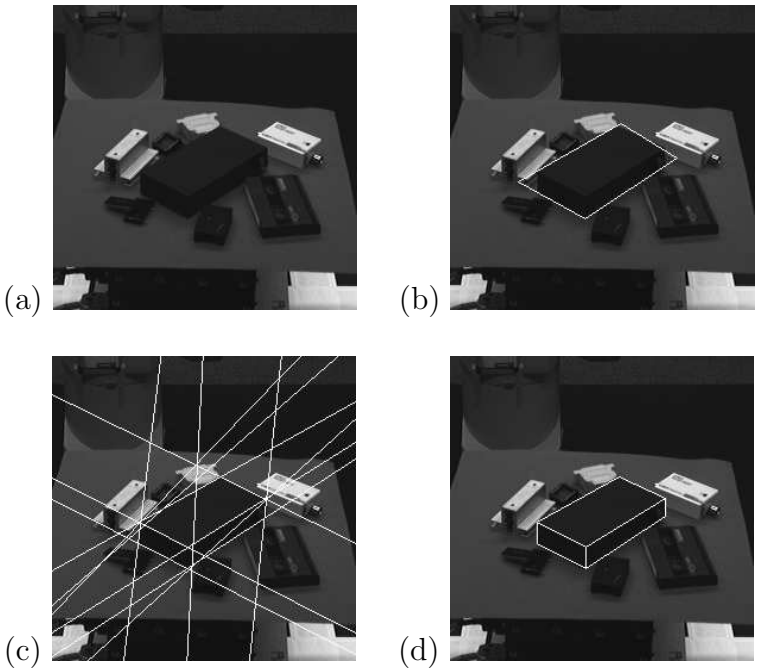


Fig. 11. Extraction of salient parallelepipeds: (a) Central black box and other objects; (b) Black box localized in quadrangle image window; (c) Subset of 12 most salient lines; (d) Most salient parallelepiped.

Figure 12(b). The two most salient parallelepipeds are shown in Figures 12(c) and 12(d). Furthermore, if we apply also the pencil/corner junction compatibility then the parallelepiped in Figure 12(d) is favoured clearly, because some of the line intersections in Figure 12(b) are less compatible with the underlying local image structures.

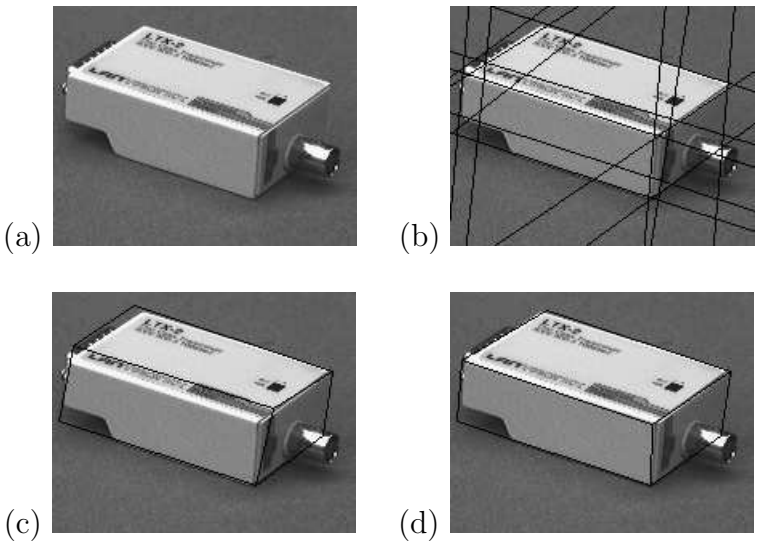


Fig. 12. Extraction of salient parallelepipeds: (a) Transceiver box; (b) Subset of 15 most salient lines; (c) and (d) Two most salient parallelepipeds.

We presented five applications of perceptual organization in demanding images. The results serve respectively as an intermediate step in the process of boundary extraction and/or detection of objects. Useful sets of polygons or arrangements of polygons have been extracted in spite of complex background, shape deviations from standard forms, low gray-value contrast between neighboring faces of the object surface, and natural inscriptions on object surfaces.

A major subset of parameters include the thresholds for the degrees of compatibilities, i.e. $\delta_1, \cdots, \delta_5$, which can be learned in the experimentation phase. As is usual in all learning-based approaches the training situations must be similar to the testing situations. In our case the situation refers to the condition of image formation and therefore is of general nature, e.g. in the experimentation and in the application phase the images must be taken with the same camera. With statistical learning the parameters proved to be appropriate without the need for readjustment.

Further parameters must be specified to search for an appropriate set of peaks in the Hough image, i.e. the number stripes and the number peaks in the stripes. No boundary line of a relevant object should be missing and therefore a superset of lines is requested from which to select appropriate subsets. The relevant parameters are critical, because a large number of peaks may be required to include (represent) all object-relevant image lines, but it would reduce the efficiency of subsequent processing steps. This problem becomes serious in case of complex background and inhomogeneous object surfaces. For this reason we propose a layered approach which extracts in the first step the object silhouette and in the second step the object boundary within the relevant image window (see fourth application and Figure 11).

Among the three types of geometric-photometric compatibilities we realized the line/edge orientation compatibility as the most important one, which has been applied in all applications. The parallel/ramp phase compatibility was useful in case of homogeneous surfaces, e.g. fourth application, and the pencil/corner junction compatibility proved to be useful under homogeneous background, e.g. fifth application. Among the various types of pure geometric compatibilities we successfully applied the parallelism compatibility, the parallel-symmetry compatibility, and the right-angle compatibility in our first series of applications (2D object components), and additionally the vanishing-point compatibility and the pencil compatibility in the second series of applications (3D objects).

So far, the learning-based procedures for perceptual organization are not optimized concerning the execution time. Various improvements can be considered,

and parallel processing would be the most important one. This is relevant in treating the original image, e.g. extracting gray-level edges and estimating their orientations and local phases, and extracting junctions and estimating junction angles. It is also relevant in performing Hough transformation, in analyzing the Hough image, and in successive steps, e.g. extracting local peaks and treating combinations between them in parallel. Certainly, with these improvements the processing of the five images would take less than a second, respectively, using a SUN Enterprise with four ULTRASPARC processors.

6 Summary

In the introductory section we raised the question concerning the *kind* of expectations to include in Computer Vision approaches. The answer to this question and the novelty of this work is that we maximally apply compatibilities of image formation for extracting necessary information from images. Compatibilities are degradations of invariants and are based on the actual effects of image formation. The first attempt of relaxing invariants has been undertaken by Binford et al. (1993), who introduced *quasi-invariance* under transformations of geometric features. Our compatibility concept is a more general one, because more general transformations are considered, maybe with different types of features prior and after the mapping. For boundary extraction it is reasonable to consider compatibilities between global geometric entities and local gray-value features, as well as compatibilities between elementary and structured geometric entitites.

An appended question has been concerning the *methodology* of acquiring expectations. We proposed a learning-based approach. Based on systematic measurements during an experimentation phase one *approximates* the quality of certain procedures statistically and determines *degrees of compatibilities* thereof. However, it is an approximation which may serve only as initial configuration in the working phase. In many applications robust adaptation mechanisms would be necessary for reliable perceptual organization and object detection in the working phase.

For characterizing the performance of this learning-based approach it is necessary to extend the set of measures, e.g. those presented in Borra et al. (1997), by also evaluating the underlying learning mechanism.

Acknowledgements

We wish to thank the anonymous reviewers for their useful comments.

References

Binford, T., Levitt, T., 1993. Quasi-invariants – Theory and exploitation. Image Understanding Workshop, pp. 819-829.

Borra, S., Sarkar, S., 1997. A framework for performance characterization of intermediate-level grouping modules. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 1306-1312.

Boyer, K., Sarkar, S., 1999. Perceptual organization in Computer Vision – Status, challenges, and potential. Computer Vision and Image Understanding 76, 1-5.

Faugeras, O., 1993. 3-Dimensional Computer Vision. MIT Press, Massachusetts.

Granlund, G., Knutsson, H., 1995. Signal Processing for Computer Vision. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Jacobs, D., 1996. Robust and efficient detection of salient convex groups. IEEE Transactions on Pattern Analysis and Machine Intelligence 18, 23-37.

Kim, Z., Nevatia, R., 1999. Uncertain reasoning and learning for feature grouping. Computer Vision and Image Understanding 76, 278-288.

Klette, R., Stiehl, S., Viergever, M., Vincken, V., editors, 2000. Performance Evaluation of Computer Vision Algorithms. Kluwer Academic Publishers, Amsterdam, The Netherlands.

Leavers, L., 1993. Survey – Which Hough transform ? Computer Vision and Image Understanding 58, 250-264.

Lowe, D., 1985. Perceptual Organization and Visual Recognition. Kluwer Academic Publishers, Boston, USA.

Pauli, J., 2001. Learning-Based Robot Vision. Springer Verlag, Berlin, Germany. Lecture Notes in Computer Science 2048.

Ralescu, A., Shanahan, J., 1999. Perceptual organization for inferring object boundaries in an image. Pattern Recognition 32, 1923-1933.

Sarkar, S., Boyer, K., 1993. Perceptual organization in Computer Vision – A review and a proposal for a classificatory structure. IEEE Transactions on Systems, Man, and Cybernetics 23, 382-399.

Simoncelli, E., Farid, H., 1996. Steerable wedge filters for local orientation analysis. IEEE Transactions on Image Processing 5, 1377-1382.

Smith, S., Brady, J., 1997. A new approach to low level image processing. International Journal of Computer Vision 23, 45-78.

Stahs, T., Wahl, F., 1992. Recognition of polyhedral objects under perspective views. Computers and Artificial Intelligence 11, 155-172.

Vasseur, P., Pegard, C., Mouaddib, E., Delahoche, L., 1999. Perceptual organization approach based on Dempster-Shaver theory. Pattern Recognition 32, 1449-1462.

Walker, E., Kang, H.-B., 1994. Fuzzy measures of uncertainty in perceptual grouping. IEEE International Conference on Fuzzy Systems, pp. 2020-2024.

Wechsler, H., 1990. Computational Vision. Academic Press, San Diego, USA.

Zerroug, M., Nevatia, R., 1993. Quasi-invariant properties and 3D shape re-

covery of non-straight, non-constant generalized cylinders. Image Understanding Workshop, pp. 725-735.

Zisserman, A., Mundy, J., Forsyth, D., Liu, J., 1995. Class-based grouping in perspective images. International Conference on Computer Vision, pp. 183-188.