Outline-Based Part Segmentation Using Intermediate-Level Symmetries

Guillaume-Alexandre Bilodeau and Robert Bergevin

Computer Vision and Systems Laboratory
Dept. of Electrical and Computer Engineering
Laval University, Sainte-Foy, Québec, Canada, G1K 7P4

bilodeau@gel.ulaval.ca, bergevin@gel.ulaval.ca

Keywords:
3D Parts, Segmentation, Outline, 2D contours, Symmetries, Complex images

Submitted to:
Vision Interface'99

Outline-Based Part Segmentation Using Intermediate-Level Symmetries

Abstract

A new part segmentation approach is presented which works on real images. These images contain 3D objects with textures and shadows on a complex background. The proposed approach relies on the outline of the object to guide the grouping of lines using symmetry and colinearity principles. This grouping of lines leads to simple shapes which can be modeled by 3D primitives such as geons or general cylinders. An algorithm implemented on the basis of this approach appears robust to noise and generic conditions. Besides, intermediate-level symmetries employed by the algorithm ensure a good robustness to internal textures and markings. The results obtained demonstrate the validity of the approach as a mean towards 3D generic object recognition from real 2D images.

1 Introduction

Despite many years of research, a great number of problems in computer vision still have no complete solutions. This is the case of generic object recognition. Indeed, many approaches have been proposed for recognition of specific objects in specific conditions, but no system has yet been able to successfully recognize objects in generic conditions. What is meant by generic conditions, is a recognition in an environment with variable lighting conditions, variable and complex backgrounds, variable visibility, a great number of different objects and variabilities in the object itself. These variabilities of the object are for instance textures and differences in shape which do not prevent humans to recognize an object as being part of a class of objects. The concept of class of objects was introduced in computer vision to mimic the human ability to label several different objects by the same name because of their common functions or their structural similarities.

A system capable of working in generic conditions is said to be a generic object recognition system. Object recognition systems are usually composed of three main steps. The first step is the extraction of low-level features (arcs, segments, junctions and corners) in a image. Then, the second step groups these features into parts and computes the relationship between the parts. Finally at the last step, the extracted parts and their relationships are compared with a database to identify the object.

To this day, the majority of works on 3D object recognition attempt to recognize objects without textures with a single color on a uniform background. The lighting conditions are also controlled and the objects are often very simple or synthetic. Despite the contributions of these works, it remains evident that more general methods are required to install an object recognition system on, say, a robot or on a computer controlled unit. This work addresses a new approach to the part segmentation step of an object recognition system. It is based on the outline and its symmetries and it works on a generalization of the usually processed objects. The proposed part segmentation system deals with textured objects, on nonuniform backgrounds with no need of a high contrast between background and object. The paper has the following structure. In Section 2, the recent works on 3D object part segmentation from 2D images are reviewed. Then, in Section 3, the proposed method is explained and in Section 4 some results obtained by this method are shown.

2 Past works on part segmentation

Let us look at why part segmentation is needed in generic object recognition. There exists two main approaches to object recognition, that is i) top-down recognition and ii) bottom-up recognition. The next two subsections will survey these two approaches.

2.1 Top-down recognition

Top-down recognition covers all the object recognition methods which are looking for or expecting something to be in the input image. These methods are used in specific applications where the goal is to find a particular object in an image. Different things can be searched for in the image by these methods depending on the target application. Some are feature-driven and others are part-driven. The former recognition scheme is based on a search of object specific features to identify an object in an image. The models database is composed of a set of features (corners, junctions, the length of a line, etc.) and their relative positions, if necessary, for each object to be recognized. The object recognition system then searches in the image for features of the expected object. The part-driven methods work in the same way except that higher-level structures (set of primitives grouped to form complete parts or shapes) are searched for instead of independent low-level features.

2.2 Bottom-up recognition

In the case of bottom-up or data-driven recognition, each input image is processed as if its contents were initially unknown to the system. Only after a high-end modeling is obtained is the models database searched for to find a match. This approach has the advantage to be able to learn a model and then recognize new objects. Also, since nothing specific is searched for in the images, the approach is more general and applicable than the previous one. As in top-down methods, the object models can be described either by features or by parts. Parts describe objects with less ambiguity than low-level features since they group a number of low-level features into a higher-level structure. Part-based recognition is more generic and it offers better possibilities than low-level features[1]. This is why most recent object recognition approaches are data-driven and part-based.

2.3 Part segmentation methods

Many previous methods were reviewed to see if they could be used or modified to process the type of images used in this work. This review has shown that there are five principal types of methods. The first one is the symmetry axis based methods. A number of methods [2]-[6] rely almost uniquely on symmetry axes to extract parts. These methods compute the symmetry axes of the outline of the object, and the intersections of these axes give the clues necessary to extract the parts.

In these works, the symmetry axes are found in many different ways. Some use the standard skeleton [2], while others use more sophisticated methods, like annular operators [6]. Although these methods work fine on images without textures, their use on textured objects would give a large number of axes (most of them irrelevant), and it might be quite difficult to interpret the resulting data as parts. In another method [3][4], parts are found by computing distances and curvatures on the boundaries and shocks on the symmetry axes of the object to obtain neck-based and limb-based parts. This method is capable of dealing very well with occlusions. The parts obtained appear quite similar to what a human would find. Since this method also relies on symmetry axes, textures and noise may not permit an adequate extraction of parts. The calculations on boundaries would also be difficult because of the great possibility of incomplete or noisy outline. This first type of methods is thus not suited for the images intended to be processed in this work.

The second type of methods [7]-[9] is based on the minimum description length (MDL) criterion. Their goal is to group lines or pixels together if and only if they respect a MDL criterion. Again, with uniformly painted objects and not much noise, these methods offer a good alternative. However, with textured objects, shadows and noisy data, irrelevant parts are likely to be found since good lines and noisy lines could easily be grouped. Then, it would be difficult to separate the good parts from the irrelevant ones. Again, this method does not apply very well to real complex images.

The third type of part segmentation methods is exemplified by Bennamoun [10]. It finds parts by dividing the outline of an object into convex parts with the help of convex dominant points. This method has the disadvantage of not being independent of the viewpoint. The convex dominant points are likely to change with the viewpoint, which makes it difficult to have a consistent representation of a given object. Also, with noisy objects and textures, it is not clear whether the correct outline can always be found without mistakes. This type of method has been ruled out in this work.

The fourth type of methods, like the ones proposed by Jacobs [11] and Jacot-Descombes et all [12], consists of making closed cycles with lines. In the case of Jacot-Descombes, all possible cycles with the lines of the object are found and the pertinent ones (the ones that do not contain sub-cycles) are then labeled as parts. This approach is of great interest for the case where the gaps in and between the lines are short. Since, this cannot be guaranteed with real images, a number of parts would not be found. Indeed, in this method not all available structural informations, like the configuration of the lines, is used. The approach of Jacobs differs by the use of a proximity and a convexity criterion. The parts obtained are convex. They are determined by their saliency which takes into account gaps in lines. This method attains its goal, but the convex parts obtained are more and less consistent when the viewpoint is changed and when lines appear or disappear because of noise.

Finally, the last type of part segmentation methods groups lines into parts based on their configurations. PARVO [13] uses Biederman theory [14] to extract object parts. This system relies on symmetry, colinearity, corners and junctions to extract viewpoint invariant parts. It gives good performance on line drawings, but it has not been designed to deal directly with real images. Indeed, junctions and corners are not always easily interpretable on real objects, because of textures and noise. Further in the past, Brooks [15] proposed an object recognition system called ACRONYM in which the part segmentation step consisted of finding ribbons in the extracted edges. This idea of extracting ribbons is valuable, but Brooks's method was not robust enough to deal with generic conditions. As a result, parts were missing in the output structure. Also, because this method did not use local features (like corners and junctions) or the outline to guide the identification of ribbons, textures would generate noisy ribbons which, combined together, could lead to erroneous identifications of objects.

In the existing methods, none has quite the strength we are looking for to handle textured 3D objects in real images with a background which is not necessarily uniform. In fact, in our literature review, we have not found any works attempting to deal with objects having the level of difficulty that this paper addresses. However, a combination of the different approaches may help produce a system capable of attaining our goal. For instance, a method grouping arcs and segments by symmetries could be well suited to handle 3D objects and parts. The next section explains the proposed approach and its basis.

3 The proposed approach

The proposed approach is based on Lowe's theory [16] which states that colinearity, proximity and symmetry are important clues used by humans to identify objects. Biederman [14] also shown that humans recognize objects by their component parts which are defined by the previously enumerated features. It has been deduced from these works, that parts can be obtained by looking for symmetry between lines, as in many previous works. However, with real images, symmetry must be used at a higher level as it is done, for instance, in PARVO. It means that symmetry axes cannot be computed on purely local basis. Instead, lines must be paired when they are globally symmetrical. This principle is used in our approach. Proximity and colinearity are also used to bridge gaps before pairing the lines in order to obtain the two main sides of the parts.

The originality of this approach is that the pairing of symmetrical lines is guided by the outline instead of junctions or corners. In PARVO, the lines are paired together if they can be interpreted as belonging to a single part at a high-level junction. In contrast, our new approach combines two lines by checking first that they are both on the outline and form a significant structure. It has been shown in many previous works that the outline conveys a lot of informations about the structure of an object, even a complex one. By guiding the pairing of part sides using the outline, less mistakes are likely to be made in noisy images. Besides, this approach does not require the outline to be perfect since it is only used to guide the line grouping process.

3.1 Outline extraction

In this work, low-level features are assumed to be provided. The required features are arcs and segments approximating contours. They are obtained using the SE2D system [17]. Starting with these features, the outline is extracted. This may be done in many ways. We have decided to develop an algorithm based on the following simple observation (Figure 1): when a contour line is found, it is possible to track the contour from it by making a clockwise cycle. A simple condition is imposed; that is, to go from a given line to the next at a multi-possibilities point (MPP, see Figure 1), one selects the line that makes the largest relative counterclockwise angle. The lines considered for possible continuation are those that are in a circular area centered on the endpoint of the current line. The pertinent endpoint is chosen such that the cycle is made clockwise. With this largest angle criterion, the algorithm is robust to distraction caused by internal texture lines.

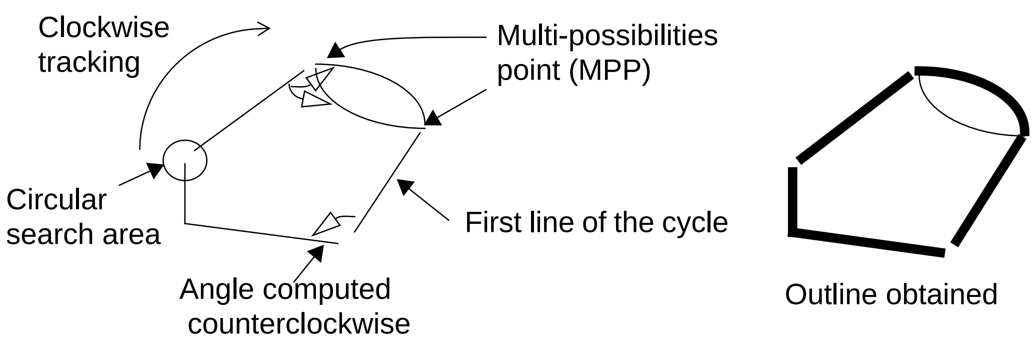

Figure 1: Tracking the outline of an object.

At each MPP, the lines are recorded in a stack. If at a given point, a continuation line cannot be found, the algorithm returns to the previous MPP and retries using the second largest angle line. This continues until either a closed cycle is obtained or all the lines of the MPPs have been tried. Only in the latter case would the algorithm fails to recover a closed outline.

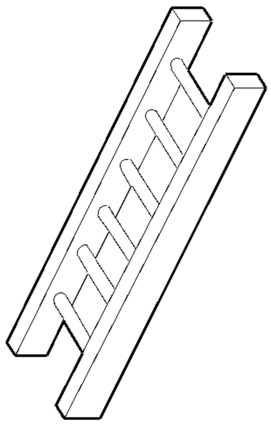

Figure 2: The outline of a ladder.

In order to increase the algorithm performance, the lines added to the cycle are monitored. If a series of consecutive randomly oriented short lines is added, they are erased and the algorithm returns to the last MPP. This way, the algorithm avoids building cycles with a large number of irrelevant random background textures. The performance of the algorithm is adjusted with parameters that control the extent of the circular area. One for the lower bound and one for the higher bound. These bounds permit to increase the circular area and recover from a failure of the algorithm when it is unable to make up a closed cycle. Figure 2 shows the outline found for a synthetic ladder object.

3.2 Part segmentation

To avoid making early decisions that could lead to mistakes, the algorithm is divided in two steps. The first step makes sure that only the most obvious parts are extracted. Then at step 2, all the remaining possible parts are extracted and then interpreted. Previously to these two steps, the lines (arcs and segments) positioned outside of the closed outline are removed. The remaining lines are assumed to be on the object. Also, before the execution of these two steps, lines are grouped by a cocurvilinearity criterion, whenever possible, in order to bridge noise gaps in lines.

3.2.1 Parts extracted at step one

Beginning with the longest outline line (arc or segment), the lines which make a symmetrical or parallel pair [18] with it are searched for (Figure 3). Using a quality factor based on proximity and similarity, the best match is determined. This match must involve another line which is completely of partially on the outline. To validate this match, the best possible symmetrical or parallel pair (including both outline and interior lines) made with each of the two previously matched lines is found (Figure 4). If the two previously matched lines are matched with each other again or with the

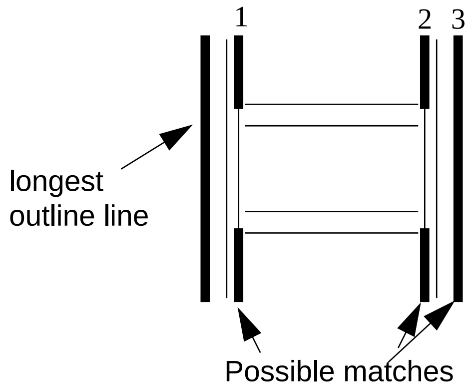

Figure 3: The extraction of part sides by pairing symmetrical or parallel lines. Line 1 is the best match because of proximity. It will form the basis of a part with the longest outline line.

same line, the match is validated and the pair is accepted as the basis of a part to be later completed (See Section 3.2.3). If not, the match is discarded and other lines are tried to avoid any risk of ambiguity. This process is run until no part can be extracted without ambiguities. Figure 6a) shows the step one parts obtained for the ladder.

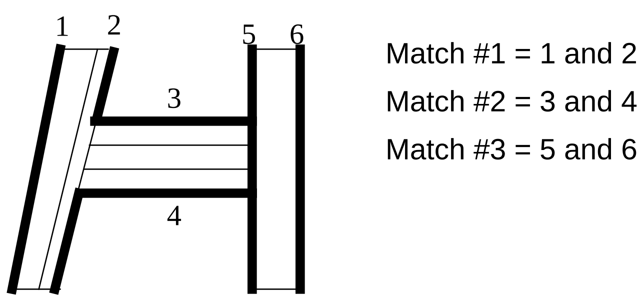

Figure 4: The validation of a match. Match #1 is validated because the two matched lines are parallel with the one in the middle. Match #2 is not validated, because both lines form better parallel pairs with the other lines between. Match #3 is validated since the two lines are again grouped together.

3.2.2 Parts extracted at step two

At this point, the step one parts have been extracted and removed from the data to speedup the processing. All possible matches of two lines having a good quality factor are then extracted and assumed as parts until they are further analyzed to remove overlapping parts. Parts found at this

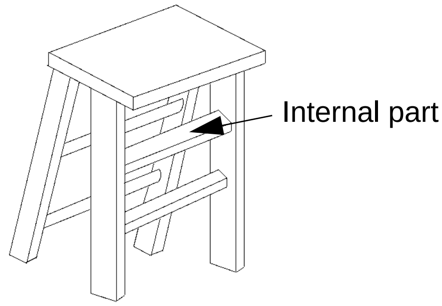

Figure 5: Example of an internal part.

stage are usually internal parts that are not in contact with the outline of the object (Figure 5). The best step 2 parts kept after analysis are usually actual parts of the object because of its inherent structure and the previous removal of step one parts. Figure 6b) shows the step two parts obtained for the ladder.

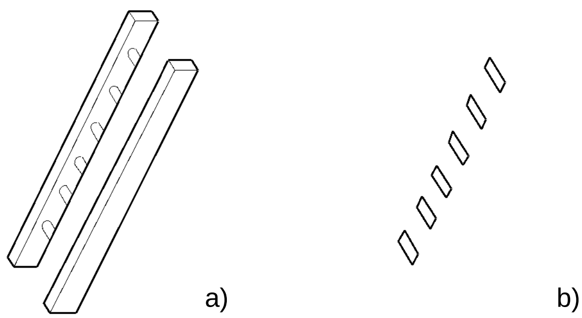

Figure 6: The completed parts extracted from a synthetic ladder. a) Step one parts. b) Step two parts. Left extremities of the rungs are enclosed in step one parts because lines are split at junction points in pre-processing.

3.2.3 Parts completion

Parts are completed by making closed paths from matched lines. Two paths are formed to connect the matched lines endpoints. These paths are formed in a way similar to the outline extraction and, as such, they must respect a number of criteria. First, the paths must not intersect themselves. Also, the lines making up the paths must remain in the area enclosed by the two matched lines (sides) to ensure that the shape of the extracted part is mainly determined by its two sides. If paths cannot be completed, straight lines complete the parts by joining the endpoints. The part description is then stored, including its internal lines which are in the area enclosed by the extracted boundaries.

4 Experimental Results

In this section, the results of processing two real images are presented and commented. The goal is

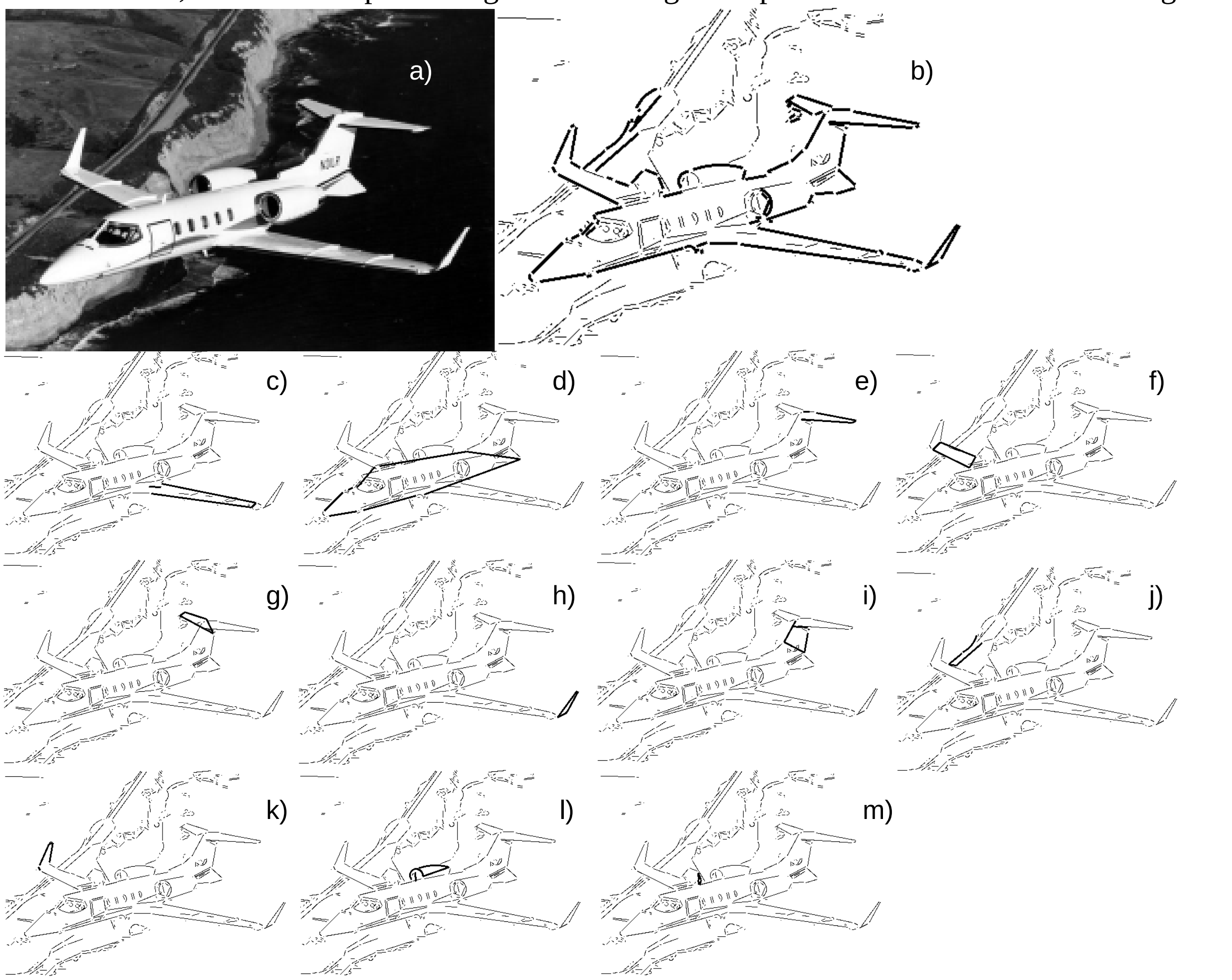

Figure 7: The proposed approach applied to an airplane image. a) The original image. b) The outline obtained from our algorithm. The sequence of images from c)-k) are step one parts, and images l) and m) are step two parts.

to show the capabilities of the proposed algorithm and its limitations. Figure 7 presents the results obtained on a real image of an airplane on a nonuniform background. It can be seen on Figure 7b) that the outline obtained captures well the shape of the airplane. However, it also follows a section of the background road. This is not surprising, as our algorithm is designed to be very robust to internal object textures, but not as much for background textures. Nevertheless, the monitoring of small random lines permits us to not continuously follow the background textures while constructing the outline. Since a section of the road is retained, it can be expected that some spurious parts are to be found. In all experiments, the first outline line used for tracking is found by intersecting rays issued from eight uniformly distributed locations on the image border. The best closed contour is then chosen by considering the aspect ratio computed from the perimeter and the area enclosed by the cycle.

Figure 7 c)- m) show the parts obtained. It can be seen that only two parts are irrelevant (j and m). One is caused by the road section included in the outline. The other is also caused by a pattern in the background. All the other parts correspond to sub-structures of an airplane. The part in Figure 7c) appears not to be closed, but in fact the two sides are joined at each extremity by closed cycles. When making cycles, gaps are permitted in order to be robust to noisy images.

Figure 8 shows the results obtained on a stool image. It can be seen from Figure 8b) that the outline is obtained without any major mistake. The cycle did not follow the shadow region between the

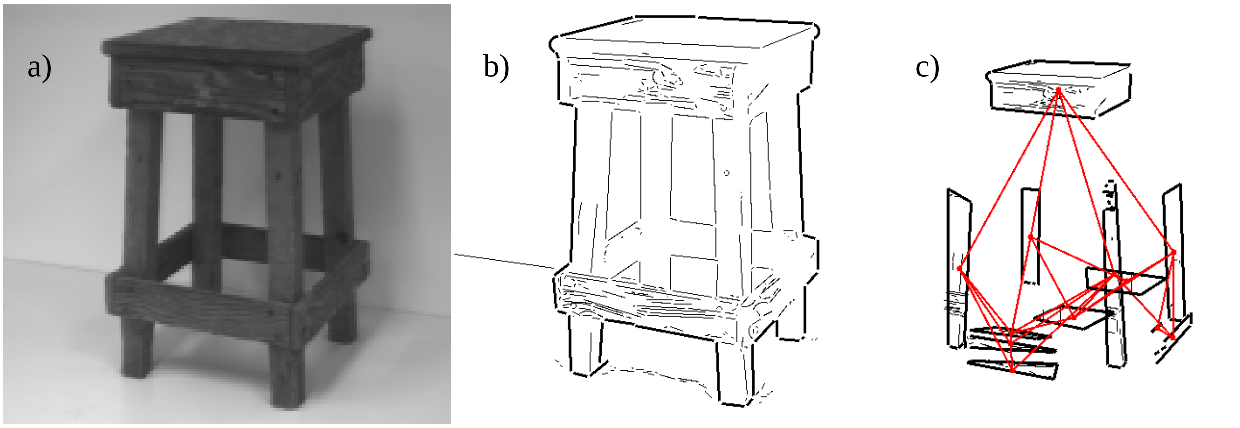

Figure 8: Our algorithm applied to an object with self-occlusion. a) The original image. b) The extracted outline. c) The parts obtained and their links.

two front legs. Figure 8c) shows that the top part and the legs of the stool have been found successfully. It can also be seen that the back rung, which is occluded, has been found. However, this figure also shows that highly textured objects may cause over-segmentation. For example, the front rung of the stool is decomposed into three parts. A spurious part has been found which corresponds to an internal hole in the object. This is due to the fact that the rungs are found as step two parts and the outline information is not used at this step. Merging of oversegmented parts is to be considered in future works.

5 Conclusion

In this paper, an original approach to 3D part segmentation has been presented. It has been implemented and tested on 2D images obtained in generic conditions. The main difference with the majority of existing systems is that our method relies mainly on the outline of the object to guide the high-level line grouping process. Informations contained in the outline give important structural clues on the object. It is thus possible, using this simple knowledge and a global symmetry principle, to decompose the object into its component parts. This approach allows one to segment into parts 3D objects of a complexity greater than many existing systems. It produce good results for 2D images with shadows, textures, a moderately complex background, and object self-occlusions.

The outlines obtained with the contour extraction algorithm have demonstrated the ability of the algorithm to not be distracted by internal object textures and, in most cases, by shadows. The part segmentation algorithm in itself gives results that, even though hard to precisely quantify at this stage, appear consistent with the high-level object structure. If one consider parts visible from a particular viewpoint, it is not necessarily the exact shape of the parts found that is the most important. Indeed, the ability of a part segmentation algorithm to give a unique and consistent description of an object in each viewpoint, might be sufficient to realize a good match with a generic model in a database. The results obtained indicate that such a unique and consistent description is attainable using our algorithm.

The results obtained confirm the validity of the approach. Many future works can be envisaged. For instance, the performance of the part segmentation algorithm could be improved by the knowledge of internal contours. These could be obtained by 3D data or, to stay in the 2D domain, by using the changes in the structural informations caused by the extraction of parts. Indeed, when a part is found, its outline lines can be removed. This removal causes the outline to become incomplete. Reconstructing the outline would give new outline lines that are internal to the object. Also, to obtain a complete object recognition system, the parts extracted could be modeled as 3D volumetric primitives. In fact, the part segmentation algorithm has been designed with that goal in mind. Most parts obtained could easily be modeled by generalized cylinders or geons.

6 References

[1] A. Pentland, Recognition by Parts, Technical Note No. 406, Artificial Intelligence Center, SRI International, Standford, CA, 1986.

[2] A. Held and K. Abe, On the Decomposition of Binary Shapes into Meaningful Parts, Pattern Recognition, Vol. 27, No.5, 1994, pp. 637-647.

[3] H. Tek, P.A. Stoll and B.Kimia, Shocks From Images: Propagation of Orientation Elements, in Proceedings of IEEE Comp. Vision and Patt. Recog. (CVPR ‘97), Puerto Rico, June 1997, pp. 839-845.

[4] K. Siddiqi and B.J. Kimia, Parts of Visual Form: Computational Aspects, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 3, March 1995, pp. 239-251.

[5] T.Liu, D. Geiger, A.L.Yuille, Segmenting by Seeking the Symmetry Axis, in Proceedings of 14th IAPR Int’l Conf. on Pat. Recog. (ICPR’98), Brisbane, Australia, August 1998, pp. 994-998.

[6] M.F. Kelly and M. D. Levine, Region-Based Grouping Operations for Locating and Describing Objects, in Proceedings of CIPPRS Vision Interface ’95. Quebec City, Canada, June 15-19, 1995, pp. 212-219.

[7] C. Huang, O.I. Camps and T. Kanungo, Object Recognition Using Appearance-Based Parts and Relations, in Proceedings of IEEE Comp. Vision and Patt. Recog. (CVPR ‘97), Puerto Rico, June 1997, pp. 877-883.

[8] M. Pilu, R.B. Fisher, Part Segmentation from 2D Edge Images by the Minimum Description Length Criterion , British Machine Vision Conference, Edinburgh, Sept 1996.

[9] M. Pilu, R.B. Fisher , Model-Driven Grouping and Recognition of Generic Object Parts from Single Images , 4th International Symposium on Intelligent Robotic Systems, Lisbon, July 1996.

[10] M. Bennamoun and B.Boashash, A Structural-Description-Based Vision System for Automatic Object Recognition, IEEE Trans. on Sys., Man and Cybernetics - Part B: Cybernetics, Vol.27, No.6, December 1997, pp. 893-906.

[11] D. Jacobs, Robust and Efficient Detection of Convex Groups, in Proceeding of IEEE Conf. on Comp. Vision and Patt. Recog. (CVPR ‘93), 1993, pp.770-771.

[12] A. Jacot-Descombes and T. Pun, Asynchronous Perceptual Grouping: From Contours to Relevant 2-D Structures, Computer Vision and Image Understanding, Vol.66, No.1, April 1997, pp. 1-24.

[13] R. Bergevin and M.D. Levine, Generic Object Recognition: Building and Matching Coarse Descriptions from Line Drawings, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 15, No. 1, January 1993, pp. 19-36.

[14] I. Biederman, Human Image Understanding: Recent Research and a Theory, Computer Vision Graphics Image Processing, Vol. 32, 1985, pp. 29-73.

[15] R. Brooks, Model-Based Three-Dimensional Interpretations of Two-Dimensional Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. Pami-5, No. 2, March 1983, pp. 140-149.

[16] D.G. Lowe, Perceptual Organization and Visual Recognition, Kluwer Academic Publishers, 1985.

[17] J.L. Arsenault, R. Bergevin and D. Laurendeau, Extraction of 2D groupings for 3D Object Recognition, SPIE Orlando ‘94, Visual Information Processing III, Orlando, Florida, April 1994.

[18] R. Bergevin and M.D. Levine, Extraction of Line Drawing Features for Object Recognition, Pattern Recognition, Vol. 25, No.3, 1992, pp. 319-334.