

# 3D Image Understanding in Virtual Environment

Patrick S. Wang, *IAPR Fellow*  
College of Computer Science  
Northeastern University  
Boston, MA 02115  
e-mail: pwang@ccs.neu.edu  
<http://www.ccs.neu.edu/home/pwang>

## Abstract

This paper discusses the state-of-the-art ideas in high level visualization, understanding and interpretation of line-drawings images. A new strategy using *parallel generalized linear combination* method is presented. It is simple and needs only very few learning samples. Further, it can strengthen advantages of current key methods while overcome their drawbacks. Also, it is able to distinguish objects with very similar patterns and is more accurate than other existing methods in the literature. In addition, it is readily applicable in virtual environments including internet, mulit-media, and WWW homepage implementations for image understanding, visulaization and interpretation.

## 1 Introduction

It is well known "*One picture is worth one thousand words.*" Indeed, it can be easily seen from mountains of documents, among other things, many of which include images, scenes, graphics, photos, or pictures. They are useful to illustrate certain ideas, to demonstrate some methods and algorithms, to describe a situation or scene, or to interpret a CAM/CAD engineering design. Therefore, how to understand these images attracts increasing attention among researchers, engineers, scientists, government administrators, and industrialists in recent years. [7,11,19,20,35,43,46]

Line drawing images, also called "noble class of images" by Sugihara [33] in which many important and essential phenomena of 3-D objects hold in the real world, provide a very effective and practical method to describe the 3d shape of an object. These drawings normally can be organized into two or more distinct views, known as characteristic views or aspects in which no one view by itself is sufficient to completely characterize the ob-

ject being represented.(Lysak and Kasturi)[21] One of the most challenging and difficult problems we will face in document analysis is that how can we visualize and interpret a 2d line drawing image as 3d object? How a 3d object in various rotations and topological scalings can be properly represented and recognized? Such a system for interpreting line drawings would have many practical applications in addition to theoretical interests. For instance, it can be used for interpreting engineering drawings and to convert from a paper-based environment to a CAD/CAM system or to efficiently store or transmit the information from a paper-based drawing.(Karima et)[17] It could also be used as a man-machine interface for a CAD/CAM system, providing an alternate methodology for a designer to describe the object shape. (Bobick and Bolles[5], Dori[11], Marti, Regincós, López-Krahe and Villanueva[23], Sakurai and Gossard[29], Sugihara[33])

There have been several developments in engineering drawing interpretation over the past twenty years. Perhaps, the earliest work was reported by Ejiri, Uno, Yoda, Goto and Takeyasu [12] in 1971. Later, Wesley and Markowsky [44] developed a bottom-up approach to interpretation, and Sakurai and Gossard [30] extended that work to include objects with curved surfaces. More recently, Nishihara and Ikeda [26] applied the bottom-up technique to interpret paper-based drawings of polyhedral objects. An alternative approach was presented by Whitaker and Huhns [45], who developed a rule-based method to produce a wire frame interpretation, however only for simple polyhedral objects. In 1988, Nagendra and Gujar did a comprehensive survey on various methodologies in the literature[24]. All of these methods appear to require a rigidly controlled drawing layout as there is no mechanism to determine relative viewing directions between the various views. In addition, therer is no provision

for any error in the alignment of the views or in the location of point features other than those that can be accommodated by a simple distance threshold. However, in 1986, Sugihara [33] developed a method for correcting misalignments in single-view drawings with a non-special view point; yet it does not appear to be extensible to multiple view drawings with special view points.

On the other hand, there have been various methods developed in computer vision and object recognition for visualizing and understanding 3d objects from 2d images. As pointed out by Negahdaripour and Jain [25], that *modeling*, *representation*, and *recognition* play important roles and have great potential for solving document analysis problems. An interesting approach was presented for recognizing 3d objects using a set of  $72 \times 72 = 5184$  sample views from learning by Thompson[34]. Further effort has reduced it to 100 views by Poggio [27].

More recently, a powerful new computational method called linear combinations (LC) using alignment techniques was developed by Ullman and Basri[2,36,37]. It involves easy computations and needs few learning sample views, as few as *two*, which coincides with the finding independently by Poggio [28], in contrast to other methods that required a huge set of learning samples even for describing and recognizing only one single object by Weng and Huang [43], Poggio and Edelman [27], Thompson and Mundy [33], and Wong [47], and some other methods which require more complicated primitive structures such as ellipsoid and bent cylinder etc. by Dickinson, Pentland and Rosenfeld [10], and Koenderink and Doorn [19], which are normally not easy to extract from input images. Yet, it has many limitations and may misrecognize invalid objects as well as reject valid ones. Such difficulties were overcome by another approach using parallel graph matching by Wang [41,43] with several explicit rules for learning and recognition of concave rigid objects with very few sample views needed in learning. Other recent developments and surveys can be found in Besl and Jain [4], Chakravarty and Freeman [8], and Chin and Dyer [9]. All these efforts were in principle aimed at finding a way that can interpret, understand and recognize a 3d object from 2d image with : (1)higher accuracy rate, (2)faster speed, (3)simpler computations, (4)smaller space, and (5)fewer learning samples.

In this research project we intend to find a new avenue for tackling a more general problem involving more complicated objects such as *articulated* objects, which can change shapes in part in the sense that some portion of the objects can change its rela-

tive relations (such as angles, distance and lengths) to the other portions, but each portion itself is rigid. There are many interesting objects that fall into this category, such as boxes, closets, cabinets, refrigerators and pairs of scissors etc. Clearly, the patterns and characteristics of such articulated objects have even more variations and are more complicated than rigid ones. This class of objects is of special interests and importance since it includes most of the industrial robots and man made factory tools. For example, in designing a one-door closet, its engineering line drawings will contain different views of the closet and various status or range of how wide the door can be open. How can such process be automated by using computer to visualize, understand and interpret such line drawings?

Though interesting and important, yet recognizing such objects by computers is more difficult and challenging, and most of the work done so far concentrated on rigid objects only. Up to date very little work was done on articulated object analysis, interpretation and understanding. The first attempt to tackle such problem was perhaps by Brooks [6] using symbolic reasoning in his famous ACRONYM system. Grimson extended the interpretation tree approach to deal with 2-d objects with rotating subparts [16]. Goldberg and Lowe [15] extended Lowe's system to deal with 3-d articulated objects such as staplers. In [3], Beinglass and Wolfson reviewed different aspects in this area and generalized the generalized Hough Transform (GHT) to recognize single joint articulated objects such as pairs of scissors. Yet, it remains to be seen how to overcome the added limitations of the GHT method, as usually happens when an existing object recognition technique is extended to handle articulated objects (Beinglass and Wolfson[3], Engelbracht and Wahl [13]). The method proposed by Wang in [39,42] uses very few learning samples and can handle articulated objects well, but it only works for wire-frame objects, and needs 3d coordinates of the objects, which are normally not easily obtainable.

Our main goal is to propose a new strategy for analyzing, visualizing, understanding, and interpreting complicated objects including concave, partially self-occluded and articulated objects from 2d line drawings. Such approach will be capable of overcoming above mentioned disadvantages and difficulties, while maintaining or enhancing advantages of many existing key methods. We'll concentrate on non-perspective view (equivalent to long range view) line-drawing objects, or also known as *noble class* of objects in which many important and interesting phenomena hold for 2d images of 3d objects

(Baird et al [1], Dori [11], Girod and Scherock [14], Kasturi et al [18], Marill [22], Shepard and Metzler [50], Stark and Bowyer [31], Waltz [38], Wang [42], Wong [47]).

## 2 Notations and Definitions

We focus on non-perspective views of polyhedral line drawings, in which many important and interesting phenomena of scenes and objects hold in the realistic world. We observe (1) in 2d geometric line drawings of a 3d object even though the angles and lengths may vary sharply when rotated, yet the property of parallelism is *preserved* under any circumstances. (2) there is a range between  $0^\circ$  and  $\pi$  for concave angles, and between  $\pi$  and  $2\pi$  for convex angles of 2d orthographic images of 3d objects.

Notice that for an articulated object, even though all views are observed from the same angle, they look quite different because of various articulated status. Also, the *dangling nodes* play a very important role. [42,43] Further, it is observed that even though there are *infinite* variations, yet they correspond to only a *finite* number of *categories*, each of which can be represented by a graph by *thinned lines* [32,39,40,42,43].

## 3 Limitations and Drawbacks of Conventional Method

There are some main limitations and drawbacks of current computational LC method:

(1) It only works for rigid objects with fixed shapes, but does not work for *articulated objects*, which are widely used in realistic world and appear frequently in various documents.

(2) The key matching points for recognition are not defined, instead, they are just randomly selected. Also, it may mis-recognize due to inappropriate bases used [41,42].

(3) It can mis-recognize due to inappropriate angles of view known as ill-posed or ambiguous phenomenon [41,42]. To remedy above drawbacks, the following rules are used:

- Rule 1: Every characteristic view of an object in learning should include all visible sides from that angle and should cover all such sides' maximum area as possible as it can.
- Rule 2: Every dangling line should not be seen through in recognition, i.e. should not cross the touching boundary line(s).

- Rule 3: Incomplete lines or dangling lines with no touching boundary lines are not acceptable, unless they are matched with the corresponding lines and nodes in learning samples, and
- Rule 4: Select nodes, including dangling nodes in the representing graph to be correspondence points for matching by linear combinations (LC).
- Rule 5: Should avoid views that involve a side just barely visible or a near unstable view that will change dramatically with a tiny bit rotation. (else, it will be ambiguous or ill-posed)

## 4 Generalized Parallel Linear Combination Method (GPLCM)

Let  $O$  be a rigid object.  $P_1$  is a 2d image of  $O$  and  $P_2$  the image of  $O$  following a rotation by a  $3 \times 3$  matrix  $R$ . Let  $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$  be the three rows of  $R$ , and  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  the three rows of the identity matrix. For a given 3d point  $p$  in  $O$ , its coordinates  $(x_1, y_1)$  in the first image view are  $x_1 = \mathbf{e}_1 p, y_1 = \mathbf{e}_2 p$ . Its coordinates  $(x_2, y_2)$  in the second view are given by :  $x_2 = \mathbf{r}_1 p, y_2 = \mathbf{r}_2 p$ . Any other view obtained by applying another  $3 \times 3$  matrix  $U$  to the points of  $O$  will have coordinates  $\hat{x}, \hat{y}$  of  $p$  as :  $\hat{x} = \mathbf{u}_1 p, \hat{y} = \mathbf{u}_2 p$ , where  $\mathbf{u}_1, \mathbf{u}_2$  are the first and second rows of  $U$ , respectively. Assuming that  $\mathbf{e}_1, \mathbf{e}_2$  and  $\mathbf{r}_1$  span  $\mathcal{R}^3$ , then :  $\mathbf{u}_1 = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{r}_1$  for some scalars  $a_1, a_2, a_3$ . Therefore :

$$\hat{x} = \mathbf{u}_1 p = (a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + a_3 \mathbf{r}_1) p = a_1 x_1 + a_2 y_1 + a_3 x_2$$

It holds for every point  $p$  in  $O$ . Let  $\mathbf{x}_1$  be the vector of  $x$ -coordinates of the points in the first view,  $\mathbf{x}_2$  in the second,  $\hat{\mathbf{x}}$  in the third, and  $\mathbf{y}_1$  the vector of  $y$ -coordinates in the first view. Then :

$$\hat{\mathbf{x}} = a_1 \mathbf{x}_1 + a_2 \mathbf{y}_1 + a_3 \mathbf{x}_2$$

For any other image of the same object, its vector  $\hat{\mathbf{x}}$  of  $x$ -coordinates must be a linear combination of these basis  $\mathbf{x}_1, \mathbf{y}_1$  and  $\mathbf{x}_2$ . Similarly for the  $y$ -coordinates :

$$\hat{\mathbf{y}} = b_1 \mathbf{x}_1 + b_2 \mathbf{y}_1 + b_3 \mathbf{x}_2$$

where coefficients  $a_1, a_2, a_3$ , and  $b_1, b_2, b_3$  can be got from solving simultaneous equations [2,45].

### Figure 4.1 here

Note that in the *learning* of Figure 4.1, we need *two* line images to decide if it is an *articulated object*. Further, these two images must be in the same *characteristic view* (or *aspect view*) with different status, as outlined in Section 3. Once this is done,

the compiled library can be partitioned into two parts, namely: (i) graph representations (with measurements) for rigid, non-articulated objects (single joint), and (ii) those for articulated objects. For the category (ii), each representative graph has two portions, i.e. (a) main body and (ii) *articulated portion*. In this paper, we begin with *articulated objects* with two portions only. Here the *main body* and the *articulated portion* are only relative terms. If we consider the *articulated portion* the *main body*, then the rest part is the *articulated portion*, and vice versa. For example, in Figure 3.3 of previous section, one can consider the *door* the *main body*, leaving the rest part *articulated portion*.

In the *recognition* phase, when an line drawing image is scanned and extracted from a sample document page, our strategy will first compute its graph representation, using the maximum neighborhood heuristic method, with its measurements [39,41]. Then, it will be compared with those in the library, level by level, in *parallel*, i.e. comparing all nodes at the same level *simultaneously*. If they match one in the rigid category, the algorithm then computes their measurements to determine if the input image is a *linear combination* of two of the image in the same *characteristic views*, subject to the rules outlined in this section. If it matches one in the *articulated portion* category, then the algorithm will invoke the *linear combination* subtask *twice*, namely, one for the *main body*, and one for the *articulated portion*. When both match, the input image will be accepted.

## 5 Illustrative Examples and Applications

Figure 5.1 shows 12 line images, which are from three types of closets learned from Figures 2.5-6. Note that the dotted lines of several images show how the parallel line invariance can be used to determine the angles of the door open status. For example, in image (A),  $\angle cde = 30^\circ$  because  $de$  is parallel to  $ac$ , and  $\angle cde = \angle acd$  and  $\angle acd = 30^\circ$  from *learning* (Figure 2.5). Similarly, in image (F),  $\angle cde = 150^\circ$  because  $\angle cdb = 180^\circ$  and  $\angle cde = 30^\circ$  (from top view of learning in Figure 2.5). Figure 5.2 is a table summarizing recognition results. Figure 5.3 shows 4 partitions of type I closet from learning (Figures 2.5 and 2.6 in Section 2).

### Figure 5.1-5.4 here

The objects in Fig. 5.4 are kinds of closets and their doors can be rotated along their hinges. The three model images in (i), (ii) and (iii) are set up by

rotating both main part and articulated part separately. Fig 5.7(iv) shows the different view of the same object and the following data verify that it matches the model images.

The match procedures on the main part are described as follows:

The feature points on the main parts of the model images are as follows:

Model1: (-1.90,0.40), (1.00,-0.20), (1.70,-1.10), (0.70,-2.10)

Model2: (-2.00,0.60), (0.70,-0.50), (1.70,-1.20), (0.30,-2.20)

Model3: (-2.00,0.70), (0.40,-0.70), (1.70,-1.30), (-0.00,-2.40)

and the selected feature points on the main part of this unknown image are

(-2.0,0.8), (0.3,-0.8), (1.7,-1.3), (-0.2,-2.4), and (-2.3,-0.8).

After the equations are solved, the constrain values are as follows:

$$\alpha_x = -1.12, \beta_x = 1.78, \gamma_x = 0.31, \delta_x = 0.05$$

$$\alpha_y = -0.32, \beta_y = 0.09, \gamma_y = 1.21, \delta_y = 0.03$$

The predicted x and y coordinate values are -2.4030 and -0.7106 and the absolute differences between the predicted value and the original one are 0.1030 and 0.0894. Let the threshold be 0.6, both the differences are less than threshold, so far the main part of this unknown image is not accepted.

The match procedures on the articulated portion are described as follows:

The feature points on the articulated portion of the model images are as follows:

Model1: (-1.90,0.40), (1.00,0.30), (2.00,1.20), (-0.80,1.30)

Model2: (-2.00,0.60), (0.70,0.60), (2.10,1.40), (-0.40,1.40)

Model3: (-2.00,0.70), (0.40,0.80), (2.20,1.70), (-0.10,1.60)

and the feature points on the articulated portion of this unknown image are

(-2.0,0.8), (0.1,1.2), (2.2,2.1), (0.0,1.7), and (-0.95,1.0).

After calculation, the constrain values are as follows:  $\alpha_x = 0.62$ ,  $\beta_x = -2.00$ ,  $\gamma_x = 2.38$ ,  $\delta_x = -0.07$   
 $\alpha_y = 0.33$ ,  $\beta_y = 0.67$ ,  $\gamma_y = 1.00$ ,  $\delta_y = 1.60$

The predicted x and y coordinate values are -1.4835 and 1.2656 and the absolute differences between the predicted value and the original one are 0.5335 and 0.2656. If the threshold is 0.6, both the differences are less than threshold, so far the articulated portion of this unknown image also matches those of the model images. Therefore, this unknown image is accepted.

Figure 5.7 (v) is a closet which has the same main part as that of model object, however, its door is different from that of the model one: it has a various size. Therefore, when this input image is recognized, the main part should match the models and the articulated portion should not. Hence, the entire image will not match model images. The followings are the recognition data which demonstrate the above statement.

Because the main part of this unknown image is the same as that in (iv) and from the above calculation, we know that it matches the main parts of the model images.

Next, let us focus on the articulated portion recognition.

The feature points on the articulated portion of this unknown image are

(-2.0,0.8), (-1.0,1.0), (1.1,1.9), (0.0,1.7), and (-1.5,0.9).

After calculation, the constrain values are as follows:

$\alpha_x = 0.26$ ,  $\beta_x = -1.68$ ,  $\gamma_x = 2.62$ ,  $\delta_x = -0.61$   
 $\alpha_y = 0.50$ ,  $\beta_y = -0.92$ ,  $\gamma_y = 0.67$ ,  $\delta_y = 0.40$

The predicted x value is -0.0273 and the absolute difference between the predicted value and the original one is 1.4727. If the threshold is 0.6, both the difference is much greater than threshold, the articulated portion of this unknown image does not match that of the model images, therefore this unknown image does not match the model images.

## 6 Conclusion

We have presented a new methodology for visualizing, understanding and interpreting 3d objects. It is simple and robust, yet needs very few learning

samples and can distinguish similar objects. More illustrations are given in the Appendix.

## 7 References

1. L.Baird and P.S.Wang, "3D Object Perception Using Gradient Descent", *Int. Journal of Mathematical Imaging and Vision (IJMIV)*, 5, 111-117, 1995
2. R. Basri, "Viewer-centered representations in object recognition - a computational approach", *Handbook of Pattern Recog. and Comp. Vision* (eds C.Chen,L.Pau and P.Wang), WSP, (1993) 863-882
3. A. Beinglass and H. Wolfson, "Articulated object recognition, or how to generalize the generalized Hough transform", *Proc. ICCVPR*, 461-466, 1991
4. P. Besl and R.Jain, "3-d object recognition", *ACM Computing Survey*, 17(1): 75-154, 1985
5. A.F.Bobick and R.C.Bolles, "The representation space paradigm of concurrent evolving object descriptions", *IEEE-PAMI*, v14 n2 (1992) 146-156
6. R. Brooks, "Symbolic reasoning around 3-d models and 2-d images", *Arti. Int.*, 17, 285-348, 1981
7. E. Brown, P.S.Wang, "3D Object Recovery from 2D Images: A New Approach", *SPIE'96*, v. 2904 *Robotics and Computer Vision*, Boston, November 1996, 138-145
8. I. Chakravarty and H. Freeman, "Characteristic views as a basis for 3-d object recognition", *SPIE: Robot Vision*, 336, 37-45, 1982
9. R.T.Chin and C.R.Dyer, "Model-based recognition in robot vision", *ACM Computing Survey*, 18(1), 67-108, 1986
10. S.I.Dickinson, A.P. Pentland and A. Rosenfeld, "3-D Shape Recovery Using Distributed Aspect Matching", *IEEE-PAMI*, v14, n2, 174-197 (1992)
11. D.Dori, "Self-structured syntax-directed pattern recognition of dimensioning components in engineering drawing", *Structured Document Analysis* (ed. H.Baird, H.Bunke, K.Yamamoto), Springer Verlag, (1992) 359-384
12. M.Ejiri, T.Uno, H.Yoda, T.Goto and K.Takeyasu, "An intelligent robot with cognition and decision-making ability", *Proc. 2nd IJCAI*, 350-358, 1971
13. J.R.Engelbracht and F.M.Wahl, "Polyhedral object recognition using hough-space features", *PR*, v21 n2 (1988) 155-167
14. B. Girod and S. Scherock, "Depth from defocus of structured light", *em SPIE Proc. Optics, Illumination and Image Sensing for Machine Vision IV* (1989), v1194, 129-146
15. R.Goldberg and D.Lowe, "Verification of 3-d parametric models in 2-d image data", *Proc. IEEE Workshop on Computer Vision*, 255-267, 1987

16. W.E.L.Grimson and T. Lozano-Perez, "Localizing overlapping parts by searching the interpretation tree", *IEEE-PAMI*, 9(4), 469-482, 1987
17. M.Karima, K.S.Sadhal and T.O.McNeil, "From paper drawing to computer aided design", *IEEE Trans. on Computer Graphics and Appl.*, 27-39, Feb., 1985
18. R.Kasturi, S.T.Bow, W. El-Masri, J.R. Gattiker, U.B. Mokate, "A system for interpretation of line drawings", *IEEE-PAMI*, PAMI-12, 10, October 1990, 978-992
19. J.J.Koenderink and A.J.van Doorn, "The internal representation of solid shape with respect to vision", *Biol. Cybernetics*, 32, 1979, 211-216
20. Y.Leclerc and M.Fishler, "An optimization-based approach to the interpretation of simple line drawings as 3d wire frames", *IJCV*, v9-2, 113-136, 1992
21. B. Liu and P.S. Wang, "3D Articulated Object Recognition - A Case Study", *SPIE'96*, v. 2904 *Robotics and Computer Vision*, Boston, November 1996, 14-24
22. T. Marill. "Emulating the human interp. of line-drawings as 3d objects", *IJCV*, v6-2, 1991, 147-161
23. E.Marti, J.Regincós, J. López-Krahe and J.Villanueva, "A system for interpretation of hand line drawings as 3-d scene for CAD input", *Proc. ICDAR'91*, Saint-Malo, France, September 1991, 472-481
24. I.V.Nagendra and U.G. Gujar, "3-D objects from 2-D orthographic views - a survey", *Computers and Graphics*, v12, no 1, 111-114, 1988
25. S. Negahdaripour and A.K. Jain, "Challenges in computer vision: future research directions", *IEEE-CVPR'92*, 189-198
26. S. Nishihara and K. Ikeda, "Interpreting engineering drawings of polyhedrons", *Proc. 9th ICPR*, Rome, Nov, 1988, 869-871
27. T.Poggio and S. Edelman, "A network that learns to recognize 3D objects", *Nature*, (1990), 343: 263-266
28. T.Poggio, "3d object recognition: on a result by Basri and Ullman", TR 9005-03, IRST, Italy, 1990
29. H.Sakurai and D.C.Gossard, "Solid model input through orthographic views", *Computer Graphics*, v 17, no 3, 243-252, 1983
30. R. N. Shepard and J.Metzler, "Mental rotation: effects of dimensionality of objects and type of task", *I. Exp. Psychol. : Human Perception and Performance*, 14 : 3-11 (1988)
31. L.Stark and K.W.Bowyer, "Achieving generalized object recognition through reasoning about association of function to structure", *IEEE-PAMI*, v13 (1991) 1097-1104
32. C.Y. Suen and P.S.P. Wang (eds), *Advances of Thinning Methodologies*, WSP, 1993
33. K. Sugihara, *Machine interpretation of line drawings*, MIT Press, Cambridge (1986)
34. D.W. Thompson and J.L. Mundy, "3D model matching from an unconstrained viewpoint", *Proc. IEEE Int. Conf. on robotics and automation*, Raleigh, NC, (1987) 208-220
35. Y.Y.Tang, C.D.Yan, M.Cheriet and C.Y.Suen, "Automatic analysis and understanding of documents", *Handbook of Pattern Recog. and Comp. Vis.*, (C.H.Chen, L.Pau and P.S.P.Wang eds), WSP, 1993, 625-654
36. S. Ullman and R. Basri, "Recognition by Linear Combinations of Models", *IEEE-PAMI*, v13, no 10, (1991) 992-1006
37. S. Ullman, "Aligning pictorial descriptions: an approach to object recognition", *Cognition*, 32(3):193-254, 1989
38. D. Waltz, "Understanding line drawings of scenes with shadows", in *The Psychology of Computer Vision* (ed. by P. Winston), McGraw-Hill, 1975, 19-92
39. P.S.P.Wang, "3D Line Image Analysis - A Heuristic Parallel Approach", *Int. J. of Information Sciences*, 81/3-4, 155-175, 1994
40. P.S.P.Wang and Y.Y.Zhang, "A fast and flexible thinning algorithm", *IEEE-Computers*, v 38, no 5, (1989) 741-745
41. P.S.P.Wang, "Parallel Matching of 3D Articulated Object Recognition", *IJPRAI*, to appear 1998
42. P.S.P.Wang, "Analysis, Learning, Recognition and Understanding of Articulated Line-drawing Images", *IJCSIM*, to appear 1998
43. J.J.Weng and T.Huang, "3d motion for image sequence using point correspondence", *Handbook of PRCV*, (C.H.Chen, L.Pau and P.S.P.Wang eds), WSP, 395-442, 1993
44. M.A.Wesley and G.Markowsky, "Fleshing out projections", *IBM J. Res. Deve.* v25, n6, 934-954, 1981
45. E.T. Whitaker and M.N. Huhns, "Rule-based geometric reasoning for the interpretation of line drawings", *Appl. of A.I. III, Proc. SPIE*, v635, 621-627, 1986
46. J.Wieser and A.Pinz, "Layout and Analysis: Finding text, files, and photos in digital images of newspaper pages", *Proc. IEEE-ICDAR'93*, Japan, 1993, 774-777
47. E.K.Wong, "Model matching in robot vision by subgraph isomorphism", *P.R.*, v25 no3, 287-303, 1992

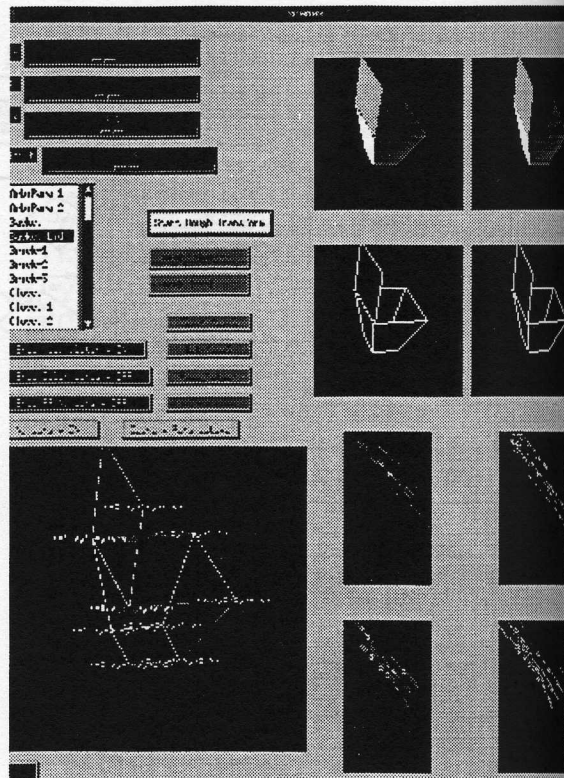


# Appendix

More illustrations of images in multi-media xwindow environment. The two screens shows two different versions of the same articulated object - a basket with lid, with different views and the lid status. Note that the object can be rotated around Ax,Ay and Az axis. The light intensity from x, y, z directions can be adjusted in the "Light-X,Y or Z" scroll bars, respectively.

There are about 40 3D objects (represented in graphics format) in the left blue-colored sub-window with scroll bar. Single-joint articulated objects (e.g. basket with lid, folders, dryer, pairs of scissors etc) use the "Joint" scroll bar, while double-joint ones (e.g. the two tank models, Tank1 and Tank2) use the additional "Joint2" scroll bar for rotating the "articulated" portion(s).

The two images shown on the upper right portion of the xwindow "scenes" are of the same object, one from the "left" eye, the other from the "right" eye, equivalent to about 12.5 degrees rotation difference. The line drawing images below them are the "hough transform" (edge detection) results, respectively.



More examples can be found in:  
[www.ccs.neu.edu/home/pwang/ObjectRecognition/](http://www.ccs.neu.edu/home/pwang/ObjectRecognition/)  
and  
[www.ccs.neu.edu/home/pwang/ObjectRecognition/scenes](http://www.ccs.neu.edu/home/pwang/ObjectRecognition/scenes)

