

A SYSTEM FOR STEREO COMPUTER VISION WITH GEOMETRIC MODELS*

D. J. Burr and R. T. Chien
Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
Urbana, Illinois 60801

The goal of this work is the specification of a system which extracts 3-D features from images by stereo comparison, and which recognizes objects by comparing these features to geometric models of objects. Models are encoded from hand measurements and consist of piecewise-linear wire frame structures. The work is in progress and some partial results are shown.

Bulk correlation methods [4] are used with three images of a scene to insure accurate triangulation for smooth directed edges. The following pictures are obtained: 1. center. 2. scene rotated 20° east. 3. scene rotated 10° north. The program compares center and east pictures for vertical edges and center and north pictures for horizontal edges.

Edge orientation is computed after preprocessing the center picture to find edge chains. An operator developed by Burr [1,2] finds local gradient extrema and tracks them by searching near neighbors (Fig. 1a). These chains are approximated by line segments [3] in Fig. 1b. The edge direction at a node, or line end, is taken as the average of the line direction on each side weighted by its length.

Stereo correlation is implemented as a mean-square difference of image intensity over 9 x 9 pixel windows. Search for the minimum in the corresponding view is restricted to a single line segment [4], and further, to locations where the intensity gradient exceeds a threshold. Triangulation determines a 3-D location for each node of Fig. 1b. A separate program attempts to match this 3-D structure (Fig. 2a) to a model.

Model matching proceeds in two steps as follows: (1) A search is made to find a model edge and perceived edge whose lengths agree. If found, each defines the z-axis of a cylindrical coordinate system (r, θ, z). The rotational, or θ -ambiguity, is relieved by finding an additional edge pair which have nearly equal lengths, center positions (r_c, z_c), and ($\hat{r}, \hat{\theta}, \hat{z}$) direction cosines. The position of this feature relative to each z-axis now defines the y-axis direction of a cartesian reference frame for both model and perceived structure. The x-axes are just $\hat{y} \times \hat{z}$.

(2) There may be further implied edge matches due to similarity of x-y-z edge coordinates and directions relative to each reference frame. An attempt is made to measure the confidence of this coordinate transformation by finding these implied edge matches. The criterion for an edge match is -- perceived edge length $\approx$ model edge length,

perceived edge center lies within a tolerance cylinder about the model edge, and direction cosines agree. The confidence is thus the total length of all matched edges divided by the total length of all scene edges. If this confidence exceeds the last computed value then it becomes the new threshold, and subsequent proposals must either exceed it or be rejected. The match of highest confidence is taken as the correct one, and the model is rotated and projected into the center view for observation (Fig. 2b).

Further work is being done on efficient tracking of 3-D edges and piecewise circular representations for 3-D edges. The ultimate goal is automatic identification of an object in an occluded scene from a model data base.

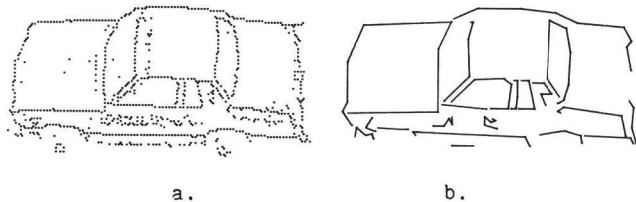


Figure 1. a. Thresholded gradient extrema of car image. b. Linear approximation of edges in 1a.

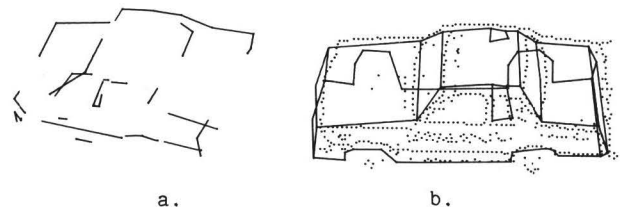


Figure 2. a. Rotated view of Fig. 1b after 3-D computation and removal of highly sloped depth edges. b. Best match of car model to 3-D scene (confidence=0.713, PDP-10 run time=13 sec.).

References

1. Burr, D. J. and R. T. Chien, "The Minimal Spanning Tree in Visual Data Segmentation," Proc. 3rd IJCPR, San Diego, California, Nov., 1976.
2. Chien, R. T. and L. J. Peterson, "Image Compression by Feature Extraction and Reconstruction," Proc. Workshop on Picture Data Description and Management, Chicago, Ill., April 21-22, 1977.
3. Duda, R. O. and P. E. Hart, Pattern Classification and Scene Analysis, p. 338, John Wiley and Sons, New York, 1973.
4. Hannah, M. J., "Computer Matching of Areas in Stereo Images," Stanford AI Memo AIM-239, Stanford University, July, 1974.

*This work was supported by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DAAB-07-72-C-0259.