STATIC ANALYSIS OF MOVING JOINTED OBJECTS

Jon A. Webb

Department of Computer Science
University of Texas at Austin

ABSTRACT

The problem of interpreting images of moving jointed objects is considered. Assuming the existence of a connectedness model, an algorithm is presented for calculating the rigid part lengths and motion of the jointed objects just from the positions of the joints and some depth information. The algorithm is proved correct.

I INTRODUCTION

Vision research has only recently begun considering the three-dimensional motion of jointed objects, but progress has been relatively rapid. This paper presents a method for using a very general model to discover the motion and structure of a jointed object. The method is proved correct under reasonable conditions, which are stated precisely. These conditions are found to be satified in most normal observation of normal jointed object movement.

Jointed objects are important because they include most of the significant moving objects in this world, e.g. rigid objects, humans, and animals. The method to be described allows the recovery of a wealth of information by a single monocular observer of a moving jointed object. This information could aid recognition from a distance.

This paper, like most other research in three-dimensional motion ([1-4]), adopts the feature point model. In this model only the positions of points rigidly attached to the object are recorded. This method makes the mathematical analysis more direct. Moreover, psychological research has shown that humans can readily interpret movies of people where only the joints can be seen [5-7]. It is therefore reasonable to try to construct a program that could interpret such images.

II THE MODEL

A. Introduction

This paper assumes the existence of a connectedness model. This model could be constructed by the methods of [3] or by other methods under development. The jointed object model for a jointed object consists of three parts: joints, rigid parts, and feature points. The feature points are fixed on the rigid parts, which are connected by the joints. In this paper, it will be assumed that the jointed object forms a tree (i.e., that is has no cycles) and that the feature points coincide with the joints. The rigid parts are not allowed to bend or stretch. The lengths of the rigid parts are unknown, but are calculated by the algorithm through observation of the jointed object.

A connectedness model for a humanoid figure is shown in figure 1. Feature points are indicated by letters.

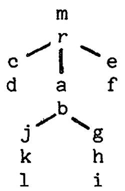

Figure 1.

B. Input Description

The analysis proposed in this paper applies equally well whether the central projection or parallel projection model of vision is used, but central projection will be assumed as it most accurately describes the way cameras work. The camera will be assumed to be at the origin, with the focal plane at (0,0,f). Figure 2 shows this model.

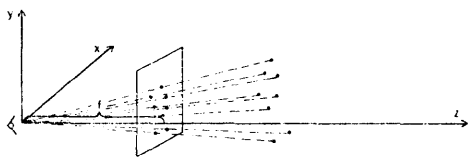

Figure 2.

The correspondence between model and image feature points must be established. The correspondence problem for moving objects has been considered in [2-4]. These correspondence algorithms are based on nearest neighbor, and work

---

* This research supported by the Air Force Office of Scientific Research under grant number AFOSR 77-3190.

well ([3] reports 98% accuracy) for frames with small time intervals between them.

The algorithm to be described requires a z coordinate for some feature point in every frame. This point will be called the reference point. For simplicity, it will be assumed that the reference point is the same in every frame. The z coordinate of the reference point can be obtained by several means, including the support assumption (used in [1] for this purpose and proposed for psychological reasons in [9]) but no method is entirely satisfactory. This will be discussed briefly in section IV.

III THE ALGORITHM

A. Introduction

The algorithm treats the model as a tree with the root being the reference point. Figure 3 shows this tree for the humanoid model. The starting point of a rigid part is its joint nearest the reference point (in this tree); its ending point is the joint farthest from the reference point. A first rigid part is said to be above a second if it lies on a path from the second to the reference point. Similarly, the second is said to be below the first.

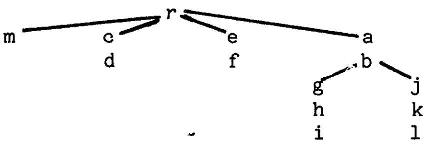

Figure 3.

The algorithm works by calculating the lengths and the positions of the ending points of the topmost rigid parts (these ending points are m, c, e, and a in figure 4). Next, rigid part lengths and ending point positions immediately below these rigid parts are calculated. The process continues until the positions of all the joints and the lengths of all the rigid parts have been calculated.

The calculation of the lengths of rigid parts is done using known lower bounds on their lengths. These lower bounds are obtained from previous frames. (In the first frame a lower bound of zero is used). If the lower bound is too small to account for the observed positions of the joints, the smallest rigid part length that will work is calculated and a new lower bound is established.

B. Formal Statement of the Algorithm

For each frame, do the following for each rigid part in the tree, going from top to bottom:

1. Let the position of the starting point of this rigid part be (x,y,z), the observed coordinates of the ending point be (u,v), and the lower bound on the rigid part length be r. If the rigid part length is exactly r, then the ending point lies on a sphere of radius r with center at (x,y,z). At the same time, the ending point lies on a line through the origin and (u,v,f), where f is the focal length. This situation is shown in figure 4. The coordinates of the ending point under these assumptions can easily be calculated using the quadratic formula.

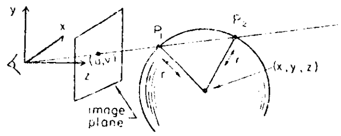

Figure 4.

This method gives two values for the position of the end of the rigid part. These two values represent two reflections of the rigid part that could account for the observed position of the ending point. For the algorithm to work, and calculate the correct rigid part lengths, the correct reflection must be chosen. It is assumed that the correct reflection is always chosen by some process. While deciding which of the reflections is correct might be a hard problem (see section IV), once the correct reflection is chosen it can be tracked fairly easily since the two reflections normally differ greatly in the z coordinate, and in the angle they make at the starting point.

2. If the quadratic formula yields no value for the position of the end of the rigid part this means that the rigid part length must be longer than r. Calculate a new lower bound on the rigid part length by the formula

(1) $r = \mathrm{SQRT}[(x-pu)^2+(y-pv)^2+(z-pf)^2]$

where

(2) $p = \dfrac{(ux+vy+fz)}{u^2+v^2+f^2}$

The coordinates of the ending point are (pu,pv,pf). The situation giving rise to this formula is shown in figure 5.

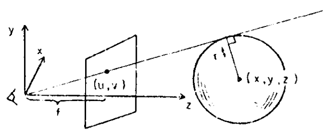

Figure 5.

Whenever a rigid part length is changed, the previously calculated lower bounds on rigid part lengths below the changed rigid part become invalid, so they must be set to zero. This action introduces an order dependence into the algorithm; for the algorithm to work correctly, the proper view of a rigid part must be seen after the proper views of rigid parts above it are seen. This restriction will be discussed in greater detail later.

C. Experimental Results

An experiment was run using three hand-drawn humanoid figures and the algorithm given above. The figures were drawn with specific rigid part lengths in mind. The rigid part lengths were recovered by the algorithm to within an average relative error of about 10-15%.

D. Proof of the Algorithm

It will now be shown that the algorithm will eventually calculate the correct rigid part lengths and three-dimensional joint positions. In order to show this, these assumptions are necessary:

1. The correct reflections of the joints must be known.

2. Each rigid part must be seen at some time in a position that satisfies figure 6. That is, the angle between the origin, the endpoint, and the starting point of the rigid part must be a right angle.

3. If rigid part A is above rigid part B, condition 2 must be satisfied for B after it is satisfied for A.

Theorem. Under the above conditions, the given algorithm will correctly calculate the length and endpoint position for every rigid part.

Proof. Let R be a rigid part. The proof will be by induction by the number of rigid parts above R. If there are no rigid parts above R then R is attached to the reference point. As soon as condition (2) is satisfied for R formula (1) will correctly calculate R's length and R's endpoint will be correct.

If there are any rigid parts above R then their correct lengths and endpoint positions will eventually be found. Once this has happened, conditions (3) guarantees that condition (2) will be satisfied for R, at which time formula (1) will be used to correctly calculate R's length. This completes the proof.

IV EXTENSIONS TO THE ALGORITHM

There are several restrictions placed on the data available to the system that are undesirable in the sense that humans cannot make them in their observation of jointed objects. The most serious restrictions are the necessity of a connectedness model for the jointed object, needing a z-coordinate for the reference point in every frame, the necessity of knowing the correct reflections of the rigid parts, and the order dependence in rigid part views. These restrictions are necessary because the analysis of the moving object is only static, and does not take into account invariants in the object's motion. Dynamic analysis of the moving object is under active investigation and is yielding quite encouraging results that suggest that most, and perhaps all, of these restrictions can be removed.

V SUMMARY

A mathematical approach to the problem of jointed object observation has been presented. Given a connectedness model of the jointed object to be observed, the actual three-dimensional motion and rigid part lengths of the jointed object can be discovered by observation of the jointed object. This is done by constantly making minimizing assumptions about the object.

Further research must take into account the actual motion of the object in a more sophisticated way. In order to overcome the deficiencies of the currently proposed method it is necessary to have a more complete understanding of how objects can be expected to move.

ACKNOWLEDGEMENTS

Fruitful discussions with J. K. Aggarwal, Larry Davis, Worthy Martin, and John Roach are gratefully acknowledged.

REFERENCES

1. Roberts, L G, "Machine perception of three-dimensional solids." in Optical and electro-optical information processing, J.T. Tippett, et al., Eds., 159-197. 1965.

2. Ullman, S, The interpretation of visual motion. The MIT Press, Cambridge, MA. 1979.

3. Rashid, R F, "Lights: A study in motion." In Proc of the ARPA image understanding workshop, Los Angeles, CA. 57-68. November 1979.

4. Roach, J W, Determining the three-dimensional motion and model of objects from a sequence of images. Ph.D. dissertation, University of Texas at Austin, Department of Computer Science. May 1980.

5. Johansson, G, "Visual perception of biological motion and a model for its analysis." Perception and Psychophysics, 14, (2), 201-211. 1973.

6. Kozlowski, L T and J E Cutting, "Recognizing the sex of a walker from dynamic point-light displays." Perception and Psychophysics, 21, (6), 575-580. 1977.

7. Johansson, G, "Spatio-temporal differentiation and integration in visual motion perception." Psychological Research, 38, 379-383. 1976.

8. Johansson, G and G Jansson, "Perceived rotary motion from changes in a straight line." Perception and Psychophysics, 4, (3), 165-170. 1968.

9. Gibson, J J, The perception of the visual world. Houghton Mifflin Co., Boston, 1950.