Obtaining Function Models using Active Vision Strategies

C. P. Lam
School of Engineering,
Murdoch University,
South Street,
Murdoch,
W. Australia 6150,
peng@eng.murdoch.edu.au

G. A. W. West and S. Venkatesh
School of Compting,
Curtin University of Technology,
GPO Box U1987,
Perth 6001,
Western Australia,
{geoff, svetha}@cs.curtin.edu.au

Keywords: Active Vision, Functional Representation, 2D and 3D Scene Analysis

Abstract

This paper addresses the problem of determining which 3D shape is present, and more importantly, the dimensions of the shape in a scene. This is performed in an active vision system because it reduces the complexity of the problem through the use of gaze stabilisation, choice of foveation point and selective processing by adaptively processing regions of interest. In our case only a small number of equations and parameters are needed for each shape. For example, a container has width and height. These are incorporated into functional descriptions of the shapes.

1 Introduction

A major objective of computer vision has been, and still continues to be, 3D object recognition. We live in a 3D world and though many useful applications of computer vision can avoid 3D object recognition, it is still widely regarded that if it is solved, then many applications are solvable. 3D object recognition has been pursued using a small number of basic techniques. Ignoring the use of shape from shading, stereo or motion and concentrating on a single camera and monochrome images, there have been three main methods: *recognition by components* [3], 2D appearance based systems, and 3D model based systems. In the 2D approach, a number of 2D views of an object are learnt and compared with the unknown image. Hopefully, the most appropriate view of the correct object will match best in some sense. Such systems use viewspheres [8] and characteristic views [2] to determine the optimum set of views of each object and use well known 2D matching and registration methods. In the 3D model based approach, a 3D model is manipulated to match the 2D projection of the model to the unknown image, usually by some form of gradient descent.

Many proposed 3D object recognition systems use a combination of the two techniques with 2D views serving as an index into most probably views of 3D objects. 3D matching is then used to refine the matches and firm up the hypotheses. Ultimately only one view of one object should be determined as a match.

All the above techniques suffer greatly from complexity when many objects have to be considered. The database consists of many accurate models and is complex if articulation, deformation and variations in shape are allowed.

An alternative approach is to rethink exactly what are the tasks of a 3D object recognition system. One is to consider function. The shape of a table is immaterial as long as it can satisfy the current goal, say, to support some object. This invokes ideas such as a flat surface, a certain height above the ground, stability and support. A cup is something that can contain a liquid and is of a suitable size with shape not being that important.

In this paper a technique is proposed that can recognise various functions. As such it deals with various shapes of objects. It can be compared with other techniques that can recognise static and rigid objects, or dynamic and deformable models, usually sequentially i.e. one at a time. The technique we propose can deal with parameterised models of various shapes defined by function. Shape varies as each function is parameterised e.g. containers have width and height. All models are treated simultaneously in parallel and the best chosen based upon minimising the error between hypothesised models and the sequence of images in an active vision system. This paper describes (i) how the models related to the objects in the scene are initially determined and (ii) the strategy used to improve the selected hypotheses. The strong point of this strategy is that the required parametric models and their associated parameters are determined in a dynamic manner. As each image is captured, the information obtained from it will be used to update the parameter values within the parametric models. Thus as more images are captured, the stronger the belief that the parametric models and their associated parameter values are correct.

The estimation of the parameters in the parametric models involves moving the camera in a controlled manner. This is necessary as it is not possible to estimate the size of an object from an intensity image without any prior knowledge. The object could actually be quite small in 3D but if the camera is very near the object then this object would appear to be very large in the image. The reverse is also true.

An object is described in terms of its functionality (i.e. the suitability of the object to fulfil a particular task). When an object category is described using a functional representation scheme it is independent of any geometrical or structural properties thus avoiding the complexity issues associated with 3D object recognition systems which use models based on shape. In a function based approach, specific object models are not stored in a database. Rather, an object is defined by its potential to fulfil a particular function such as the ability to contain. Freeman and Newell [4] were the pioneers in work relating to form and function.

2 Methodology

A four stage methodology has been discussed in Lam *et al* [7] and this paper deals with the third stage: determination of the parametric models and the associated parameters.

The methodology relies on determining and tracking a foveation point (keeping it registered), determining the camera-to-foveation point distance and the related parametric models. Details regarding the choice of features to foveate on as well as the various parametric models have been discussed in Lam *et al* [7].

An issue to consider is the camera motion and defined camera motions are used in this strategy [7]. In this paper, the camera is moved in a circular trajectory in a vertical plane about the foveation point (see Figure 1). This reduces the complexity of analysis as the CFP distance is then kept constant.

2.1 Parametric Models

Given that we know the focal length and orientation of the camera, the features of the object can be tracked with a model using the distance of the object to the camera and the relevant spatial information between the foveation point and the feature point. This requires estimates for both the CFP (camera-to-foveation point) distance and the associated parameters. As the camera changes its viewpoint, the relationship between the foveation point and the feature point currently being tracked in the image plane changes as a function of the camera's position. The camera moves in a known controlled manner and thus this relationship can be computed and predicted. According to the theory explored by Ballard *et al* [1], the representation of the products of early vision is in an object-centred coordinate frame of reference with the foveation point at the origin of this frame. In addition, knowledge of self motion and the foveation geometry will allow an active vision system to compute spatial information about surfaces relative to the foveation point. The computation becomes less complex as the coding of this geometric information will be in terms of the object-centred coordinate frame of reference.

We have previously described [7] the parametric models for three canonical cases which can be applied to describe all situations. Only the first canonical case is described in detail in this paper: both the foveation point and the feature point being tracked are stationary points. However results for other cases are presented in the results in Section 3. Figure 2 shows two points, the camera is foveating on T and tracking B. The camera moves along a circular trajectory with respect to T and is currently at point F. The points T and B are projected onto the image plane at T' and B' respectively with $FT'T$ forming the optical axis. The camera now moves to F'. Now the points T and B are projected onto the image plane at T'' and B'' respectively. The CFP distance is $FT = F'T = R$ and the focal length is $FT' = F'T''$. The value of d, the distance between the projections of T and B on the image plane at the different camera positions is given by

$$d = f \tan(\theta - \Psi) \tag{1}$$

where $\theta = \tan^{-1}(\frac{y_3 - y_1}{x_3 - x_1})$ and $\Psi = \tan^{-1}(\frac{y_3 - y_2}{x_3 - x_2})$. Hence the distance d can be computed from Equation 1 for both of the camera positions. The deviation of the measured value d_e from this theoretical value d for each camera position i is given by

$$\varepsilon^i = d_e^i - d^i \tag{2}$$

If the model used is correct then the predicted value of d^i should match the measured value i.e. $\varepsilon^i = 0$. If this is not the case, then the model is wrong or the parameters of the model have been incorrectly estimated.

The parametric equation is given by:

$$d = f \tan\left(\theta - \arctan\left(\frac{R\sin\theta + H}{R\cos\theta + D}\right)\right) \tag{3}$$

where f is the focal length, R is the camera-to-foveation point distance, H is the vertical height difference between the foveation point and the current point tracked, D is the horizontal distance between the foveation point and the current point tracked, and θ is the angle between the line joining the viewpoint and the foveation point and the horizontal plane.

Depending on the value of H, three situations can occur:

1. The points T and B are on the same horizontal plane.

2. The point T is above point B.

3. The point T is below point B.

If H=0 then the first case occurs, if $H < 0$ then the third case occurs, and lastly if $H > 0$ then the second case occurs. The case for $H = 0$ enables the verification that the foveation point and point being tracked are on the same horizontal plane. This allows for the determination of the function *flatness*. Similarly, if $H \neq 0$ the conclusion is that the foveation point and the point being tracked are not on the same horizontal plane and thus fail to provide the function *flatness*.

2.2 Parameters of Models

The function *flatness* requires only one parametric model and in this primitive there is only one parameter (i.e. D) that is required to be estimated. Similarly for the functions *vertical-ness*, *spherical-ness*, *cylindrical-ness* and *rollability*, only one parametric model is involved and within the parametric model only one unknown parameter needs to be determined (discussed in Lam *et al* [7]). However, tracking one feature may be insufficient to confirm that the object in question possesses a specific function such as the function *flatness*. The more features found on the object which can be fitted to the general function primitive (i.e. in the case of *flatness*, $H = 0$ and $D \neg 0$) the stronger the belief that the object fulfils the function.

For a complex function such as *containment*, two parametric models are required. In the first parametric model, one unknown parameter (D) is to be determined and in the second parametric model there are two unknown parameters (D and H) involved. If the recognition task involves a complex object, then the recognition task may be carried out by parts or by functions. For example, if the task is to recognise a cylindrical container then three parametric models are needed, two to verify *containment* and one to verify the cylindrical shape. Again the more features tracked the greater the confidence that the particular hypothesis is correct.

Although a function may involve a number of parametric models, in this paper the maximum number of unknown parameters in the parametric models considered is two.

2.3 The Method Used to Determine the Parametric Model

The images are initially processed to extract edges, corners or other features. The local energy edge detector [9] was used with a 3×3 mask for edge detection. As the analysis of these images is based on the knowledge of the camera motion, feature points in the plane in which the camera is moved encompassing the foveation point, are obtained. Initially all the parametric models are in contention and are considered simultaneously. All the detected feature points have to be checked to verify which of the parametric models will best relate to the object in the scene. The combination of feature points from a sequence of images is large forming a large search space requiring much computation. Thus there is a need to determine an efficient solution.

With some initial estimates, the method of recursive least squares (RLSQ) is applied to a series of images in a sequence so that a more accurate estimate of the parameters can be obtained as the vision system acquires more images. The method of RLSQ is mathematically equivalent to that of Kalman filtering except that the parameters being estimated are not a function of time [5, 10]. The only problem is that the function primitives are not linear and need to be linearised.

One or more images may be use in the updating process (currently one image is used). The images may be captured one at a time and each time a new image is obtained, the information from the image is used to update the parameter values as well as the belief in having obtained the correct parametric model.

2.3.1 Obtaining Initial Estimates of the Various Parametric Models

For simplicity, this discussion first considers solving for the unknown parameters in a parametric model that has only one unknown requiring only one image. For each feature point found, its distance from the foveation point F_{ij} [1] is calculated and substituted into the parametric model along with the CFP distance (R), the focal length (f) and the tilt of the camera (θ_i). For example, in the function primitive used for verifying the function *flatness*, there is only one unknown to be solved (i.e. D):

$$d = f \tan \left(\theta - \arctan \left(\frac{R \sin \theta}{R \cos \theta + D} \right) \right). \qquad (4)$$

As there are a number of feature points within an image, a number of values for D will be obtained from one image (each associated with one value of F_{ij}) and only one of these values will be correct. It is not possible to know which of these is correct from one image and thus a number of images from different viewpoints are required. For each of these images and for each of the values of F_{ij} in each of the images, there are associated values of D. A strategy is required to determine which of these initial values is the correct value of D.

However, if two or more unknowns are to be determined within a parametric model then two or more images would

[1] Each image is processed only in the plane P, in which the camera is moved and which encompasses the foveation point. The symbol F_{ij} in this paper denotes the distance of the j^{th} feature from the foveation point within plane P in the i^{th} image.

be required. If there are two unknown parameters, then it is necessary to obtain values of F_{ij} from a pair of images. If n feature points are found in one of the images and m in the other image then the total number of possible correspondences that need to be considered is mn. The more unknown parameters within a parametric model to be solved, the more images are required and the greater the numbers of ways to combine the values of F_{ij} from these images.

2.3.2 The Proposed Strategy Using the Recursive Least Squares Method (RLSQ)

The recursive least squares method (RLSQ) is applied to update the values of the parameters for each new observation. The advantages of this method are: (1) there is no need to store past observations, and (2) the estimate obtained by this method takes into account the effect of past observations.

The initial estimates of the parameter values within the specific parametric models are obtained in the manner as described in Section 2.3.1. As a small number of images are used, the belief that the obtained value(s) related to the parametric models is low. Each unknown parameter in the parametric model has a variance associated with it. If the belief of having obtained the correct value is low, then the associated variance of the unknown parameter is high.

A dynamic window is used for searching for correspondence among feature points, where the size and the position of the window are determined by the information obtained from each step in the updating process. Based on the initial estimates of the parameters, the specific parametric model, the camera parameters, and the covariance matrix associated with the parameters, the size and the position of the window for the next image in the sequence are determined. Feature points found within that window will be considered as the most likely candidates and thus most of the feature points in the plane in which the camera is moved are eliminated. The candidates are ranked using the residual values returned by the RLSQ method. The best candidate is the one with the lowest absolute residual value. Initially, the best candidate will be used to update the values of the parameters. The updated value will again be used to calculate the position and size of the dynamic window for the next image to be captured. The whole process is repeated as new images are captured.

However, if no feature points are found in the new image, the next most likely candidate from the previous image is used to recalculate the values of the parameters. Using the recalculated parameter values, the position and size of the dynamic window for the image are also recalculated. Again the strategy will look for feature points within the dynamic window.

The strategy is able to handle a situation where wrong feature points have been selected. When this happens, the values of the estimated parameters will diverge. The next point on the ranked list will be used to recalculate the values of the parameters. These values will again converge to the correct values in subsequent images if the correct feature points are selected. The ranking of the feature points as mentioned earlier uses the residual values obtained in the RLSQ method. The residuals are the differences between the predictions and observations and are used to alter the estimates of the parameters. This implies that the observations that are closely related to the current parametric model will have small residuals and thus will be ranked first.

The size and position of the dynamic window is determined in the following manner. Using the covariance matrix of the parameters, the standard deviations associated with each of the parameters are obtained. Based on a 95% confidence, the range of values of the particular parameter is determined. Then this range of values is used to determine the size of the dynamic window. The minimum and maximum values of the parameter is used to predict the location of the tracked feature points in the subsequent image. The absolute difference between the two predicted locations of the tracked feature points is the size of the window. The centre of the window corresponds to the most likely predicted location of the tracked feature point.

As more images are processed, the belief that the parameter values are correct increases and the variance associated with the parameter values decreases. The dynamic window which is calculated based on the variances associated with the parameters will also become progressively smaller. The minimum window size will be one pixel in height although this implies little inaccuracy in feature detection. Hence in practice, a minimum size window of a few pixels (e.g. five pixels) is chosen. When the parameter values differ by a small number ε in consecutive steps of the updating process, it is then considered that we have obtained the values of the parameters of the object and these values are then used in the tracking process to confirm whether the object fulfils the required function. In the tracking stage, as new images are obtained, the parameter values will still be checked for consistency.

3 Experimental Results

The action and movement of the camera are modelled using physical modelling (i.e. raytracing) so as to avoid the issues of active control of the camera [11]. Raytracing is acceptable as an image formation technique as we are mainly interested in the geometry and not the camera control issues or generating perfect representations of the real world. Using raytracing, we can accurately control the parameters of the objects and parametric models and repeat experiments under different known conditions. This is much more difficult when using real cameras and robots. To demonstrate the effectiveness of our strategy, some real images are used in the determination of the parametric models and their associated parameters.

If prior knowledge of the task or context (top-down or task-driven) is available, the task is made easier as only a specific parametric model needs to be considered. As the parametric model is known, knowledge concerning the number of parameter values to be determined is also known. This paper deals with the case where there is no available knowledge of the context (i.e. bottom-up or data-driven) and thus, all the parametric equations currently of interest are to be considered. The number of unknown parameters that need to be considered here would be equal to the maximum number of unknown parameters among the parametric models under consideration.

The parametric models and their associated parameters have been determined successfully under the following two conditions: (1) feature points have been consistently found at the location predicted by the models in a number of images, and (2) the values of the associated parameters have been consistent. The checking of whether consecutive values of each parameter is less than some tolerance ε is only carried out after the updating process has been repeated two to three times. This is to allow for the fact that the initial value of the parameter value(s) may be very inaccurate. The RLSQ method can converge quite quickly to a value close to the expected value.

The tolerance value of 1.0 is used to eliminate the other parametric models quickly. A value of 1.0 is chosen because the values of the associated parameters change by quite large values in the incorrect parametric models. This is indicated in the experimental results. To illustrate this process, various test objects were used. The aim is to obtain the correct parametric models that pertain to the object. Since there is no prior knowledge, a number of parametric models will be considered initially. These are represented by the following symbols: B_1, F_1 and T_1. Listed below are the interpretations of these symbols:

- B_1 represents the case of the tracked feature point being on an occluding boundary. One unknown parameter is associated with this model.

- F_1 represents the case of the tracked feature point being on the same horizontal level as the foveation point. One unknown parameter is associated with this model.

- T_1 represents the case of the tracked feature point being some horizontal distance as well as some vertical distance below the foveation point. Two unknown parameters are associated with this parametric model.

Using the first two images of the sequence, the initial estimates of the parameters will be determined using the method of simultaneous equations. Two images were used as the maximum number of parameters in the parametric models is two. Values of F_{ij} are obtained from the images and substituted into each of the parametric models resulting in a number of sets of simultaneous equations. The result of this initial process is a list of possible initial estimates for the parameters. As these values are determined from only one (or one set of) image(s), the variances associated with the unknown parameters are high. Every value in this list will be used as a starting value in the process of obtaining the parametric model and its associated parameter values.

The next image is then obtained. Using the values of F_{ij} and all the associated camera parameters, the parameter values associated with each of the parametric models under consideration is updated. The process of capturing an image from the next viewpoint and then using information from the image to update the parameter values of the parametric models is repeated for a few images (normally two to three).

The process described will be repeated for each set of experiments described below. Various test objects were used. In addition, the investigation also uses real images to evaluate the performance of the strategy.

3.1 Parametric Model and the Associated Parameter Values for an Hemispherical Object

From the initial processing, two possible values of 49.6 and 1562.2 are obtained for the two parametric models with only one unknown (B_1 and F_1). Values of 49.6 and 50.0 are the values for the model with two unknown parameters (T_1). All of these will be considered as the initial estimates for the parametric models in contention.

Table 1 illustrates the changing parameter values of the two initial values that are updated as new images are acquired. In the case where the initial value is 1562.2 units, both the function primitives B_1 and T_1 drop out of contention after one image and thus the only function primitive that remains is F_1 (see the column labelled as Case 2 in Table 1). Subsequent processing will show that the function primitive F_1 is out of contention after the second image because no feature points can be found in the dynamic window in the subsequent images.

When the initial value is 49.6 units, there is no updated parameter values for the case of T_1 as new images are acquired as no feature points are found within the dynamic window and the associated parametric model drops out of contention. As the initial belief is low, the size of the window will be large. Thus if no feature point can be found within the window it can be inferred that the parametric model in question is not the correct model.

However, it is possible that the feature point may be lost in that particular image owing to errors in processing (e.g. in edge detection and corner detection). Thus the subsequent images will also be checked to see if there are feature points that will fit the failed models. This form of checking will only be carried out with the first few images which enables the confirmation that the particular model definitely is not related to the objects in the scene.

Notice that the value of the parameter for the function primitive B_1 is consistent compared to that of the function primitive F_1. Thus the function primitive B_1 will be selected as the most likely model after three images are used. As new images are acquired, the selected function primitive and its associated parameter values are used to predict the location of features and track them. If features from a large number of subsequent images can be tracked using this function primitive and its associated parameter values then the belief that the model is the correct one can be strengthened. Note that the confirmation of whether the parametric model in question is correct occurs in the tracking stage [7].

3.2 Parametric Models and the Associated Parameter Values for a Cup

Using the images illustrated in Figure 3, the list of possible initial values of the parameter values are obtained. Table 2 shows the result of using one of the initial values as the starting point and using the images shown in Figure 3. Notice that in this set of images, the internal bottom edge and the top opening of the cup are present. Note that the cup is 80 units high, the internal width is 77 units and the external width is 80 units. As depicted in Table 2, at the end of processing the fourth image, there are still two parametric models in contention. One is for the top edge and the second is for the bottom internal edge. The parameter values of both the parametric models for consecutive images are varying less than the tolerance and thus is considered to be consistent and correct.

3.3 Parametric Models and the Associated Parameter Values for a Box Using Real Images

Some real images were used to investigate the performance of the technique. Using the same processing as before, the changes to the parameter values in each of the function primitives under consideration are shown in Table 3 as new images are processed. One of the initial values of the parameters to be considered was 98.3 units. Again the function primitive B_1 was out of contention after the first image in the sequence was processed. By the end of four images, two function primitives are still under consideration. This is reasonable as can be seen from the images. The two function primitives are for the rear top and the internal bottom edge of the box. The images in Figure 4 show the location of the tracked features using both of the function primitives under consideration. Their associated parameter values are shown in Table 3.

4 Improving the Hypothesis by Using Multiple Slices

The main features used so far have been individual edge points. As only one or two points are involved, the hypothesis involving a specific model would be a weak one without strong *a priori* assumptions about the objects. One of the ways to increase the confidence in the hypothesis is to use a number of sets of edge points simultaneously. The more edge points that are tracked and found to fit the general form of the parametric model, the more reliable the determination of the decision concerning the current hypothesis.

The process of tracking multiple feature points has been discussed in Lam *et al* [6]. An example of this is shown in Figure 5 which illustrates the result of tracking multiple feature points using seven 1D slices across the images. The foveation point is the single thick white cross in the centre of the image. The feature points being tracked are marked with thin white crosses.

5 Conclusion

This paper has described a technique for recognising a 3D shape given parametric models describing function. As a consequence the dimensions of the object are determined.

The strategy was able to select the correct parametric model as well as obtain the associated parameter values in a bottom up approach when images of different objects were used.

The RLSQ method has been used as it allows, via the error measures, the use of dynamic search windows to reduce the search space when looking for correspondence between features in adjacent images in time. Results on different objects, both simulated and real, with and without Gaussian noise, demonstrate the technique works and chooses the correct model quickly in all cases. Extensions to the methods described in this paper include using different features, and testing on different objects.

References

[1] D. H. Ballard. Eye Movements and Spatial Cognition. Technical Report TR 218, University of Rochester: Computer Science, November 1987.

[2] C. H. Chen and P. G. Mulgaonkar. CAD-Based Feature - Utility Measures For Automatic Vision Programming. *IEEE Proceedings Workshop on Directions in Automated CAD-Based Vision*, page 106:114, 1991.

[3] S. J. Dickinson, A. P. Pentland, and A. Rosenfeld. From Volumes to Views: An Approach to 3D Object Recognition. *IEEE Proceedings Workshop on Directions in CAD-based Vision*, pages 85–97, 1991.

[4] P. Freeman and A. Newell. A Model for Function Reasoning in Design. *Proceedings of International Joint Conference Artificial Intelligence*, pages 621–634, 1971.

[5] E. J. Krakiwsky. A Synthesis of Recent Advances In the Method of Least Squares. Surveying Engineering, University of New Brunswick, May 1975.

[6] C. P. Lam. *Investigation Into Using Function-Based Primitives and Active Vision Strategies for Hypothesis Refinement*. PhD thesis, School of Computing, Curtin University of Technology, Perth, W. Australia, 1995.

[7] C. P. Lam, S. Venkatesh, and G. A. W. West. Hypothesis Verification Using Parametric Models and Active Vision Strategies. *CVGIP: Image Understanding*, 68(2):209–236, 1997.

[8] M. C. Robey. *An Investigation into the Applicability of Physical Modelling in Aiding 3D Object Recognition*. PhD thesis, School of Computing, Curtin University of Technology, W. Australia, 1994.

[9] S. Venkatesh. *A Study of Energy Based Models for the Detection and Classification of Image Features*. PhD thesis, Department of Computer Science, The University of Western Australia, 1990.

[10] D. E. Wells and E. J. Krakiwsky. The Method of Least Squares. Department of Surveying Engineering, University of New Brunswick, May 1971.

[11] G. A. W. West. Advances in Active Vision Research by Simulation Using Raytracing. *Second International Conference on Modelling and Simulation*, 1993.

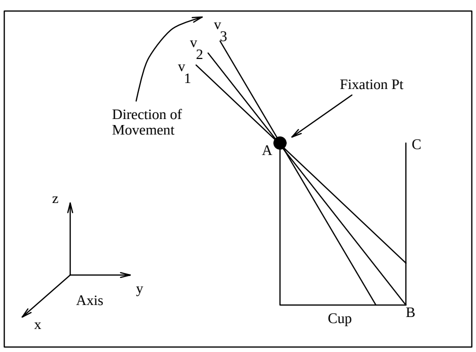


Figure 1: Camera moves in a circulatory trajectory in a vertical plane about the foveation point (from [7]).

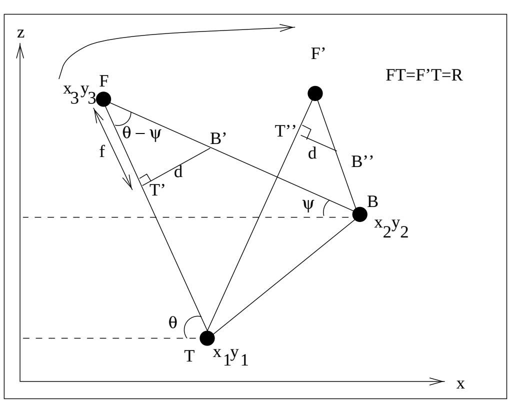


Figure 2: Foveating on a stationary feature point and tracking another stationary feature point (from [7]).

Image No.	Hypothesized Parametric Model	Inferred Values of Associated Parameter	
		Case 1	Case 2
0	ALL	49.6 49.6 49.6, 50.0	1562.2
1	B_1 F_1 T_1	52.6 32.0	1610.1
2	B_1 F_1 T_1	53.2 14.3	
3	B_1 F_1 T_1	52.7 5.7	

Table 1: A comparison of the values of the parameter associated with each of the parametric models, calculated as each image is obtained. The parametric model with a consistent value for its parameters is the parametric model that represents the object in the image.

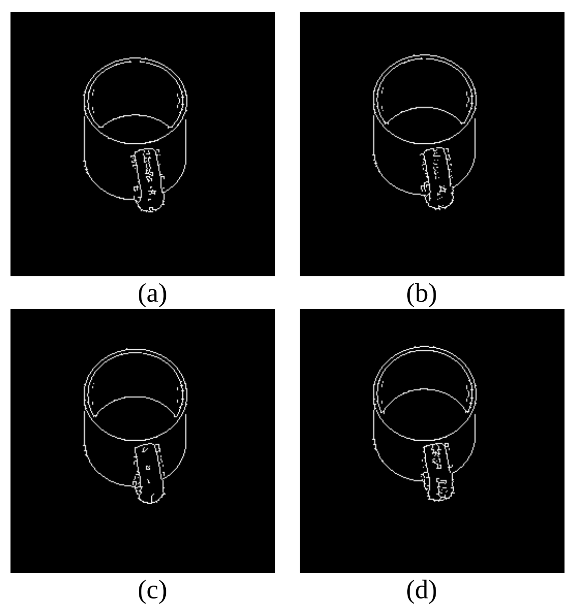
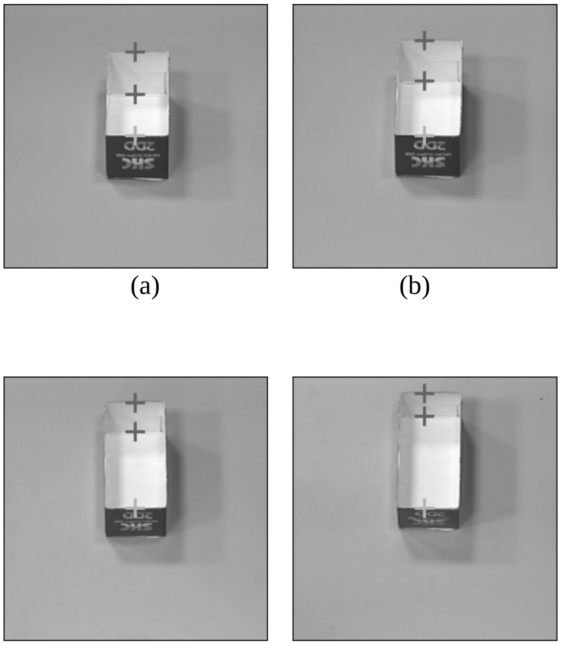


Figure 3: Edge-detected images of a cup with the internal bottom edge that are used to determine the function-based parametric models and their associated parameter values.

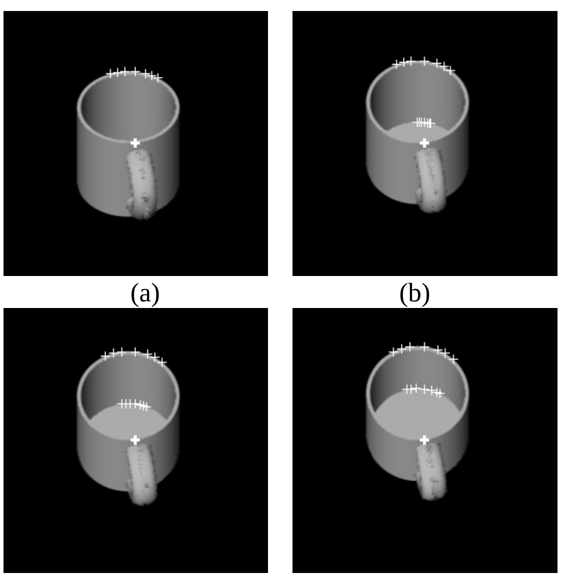


(a) (b) (c) (d)

Figure 4: Images illustrating the feature points selected using the parametric model and its associated parameter values. The cross in the centre of the image indicates the foveation point on the object. The selected feature points are marked by the other crosses.

Image No.	Hypothesized Parametric Model	Inferred Values of Associated Parameter
0	ALL	80.3
		80.3
		80.3, 83.4
1	B_1	
	F_1	79.1
	T_1	80.4, 83.5
2	B_1	
	F_1	79.3
	T_1	79.1, 82.3
3	B_1	
	F_1	79.0
	T_1	78.1, 81.3
4	B_1	
	F_1	79.0
	T_1	77.7, 80.9

Table 2: A comparison of the values of the parameter associated with each of the parametric models, calculated as each image is obtained. The parametric model with a consistent value for its parameters is the parametric model that represents the object in the image.



(a) (b) (c) (d)

Figure 5: A sequence of images showing the tracking of feature points on a cup. The technique of multiple slices was employed. The thick cross indicates the foveation point. The thin crosses indicate the feature points tracked.

Image No.	Hypothesized Parametric Model	Inferred Values of Associated Parameter
0	ALL	98.3
		98.3
		98.3, 50.4
1	B_1	
	F_1	106.5
	T_1	96.3, 40.3
2	B_1	
	F_1	104.6
	T_1	95.6, 41.8
3	B_1	
	F_1	102.7
	T_1	99.4, 40.7
4	B_1	
	F_1	102.5
	T_1	101.8, 40.9

Table 3: A comparison of the values of the parameter(s) associated with each of the parametric models, calculated as each image is obtained.