SEMI-AUTOMATED EDGE SEGMENT SPECIFICATION FOR AN INTERACTIVE MODELLING SYSTEM OF ROBOT ENVIRONMENTS

Philippe EVEN, Anne MALA VAUD
Atomic Energy Commission, Teleoperation and Robotics Department, (CEA/STR), PO Box 6, F-92265 Fontenay-Aux-Roses, France, Philippe.Even@cea.fr, Anne.Malavaud@cea.fr

KEY WORDS: Modelling, Semi-automation, User interfaces, On-line, Robots.

ABSTRACT

Interactive methods are well suited to telerobotics purposes. Based on the super-imposition of the model on video images, they provide a friendly way to acquire or update the environment model from an on-board CCD camera. It is a flexible way to cope with a priori model inaccuracies and any possible mission incidents. Modelling time is a key requirement which can be fulfilled through the integration of semi-automatic assistances. This paper deals with the semi-automation of edge segment specification tasks. Edge segments combinations define 3-D lines or planes, which are used to orient the modelled objects. Their accurate specification is a tiring and time consuming task, which can not be fully automated. In the semi-automatic mode we set up, the operator quickly draws a line over the image. This line is automatically attracted towards the nearest extracted edge contour. The automatic attraction function is based on the Hough transform. This semi-automatic assistance has been integrated into the Pyramide interactive 3-D modelling system, which has been developed at CEA/STR. Evaluations on realistic sites showed its high flexibility and efficiency.

1 INTRODUCTION

Within the last decade, CEA/STR has developed an interactive system, called Pyramide, to acquire a 3-D model of a teleoperated robot environment using video images provided by an embedded CCD camera (Even and Marcé, 1988). Its principle consists in superimposing solid primitives on the images, and interactively adjusting their position and parameters until they match the objects contours. It has been designed for on-line applications to provide computer graphics assistances during robotics missions. Fast acquisition and user-friendly operating mode are important requirements. In order to fulfil these requirements, several assistances based on image processing techniques or structural and functional knowledge have been integrated. A major one is a semi-automated edge segment specification assistance, which brought noticeable improvements.

Pyramide provides a generic modelling module aiming at modelling any structured environment as an assembly of basic volumes (block, cylinder, cone, ...). 3-D features (plane, line, point, orientation) are determined from image features (edge segments or points) or already defined objects (planes, edges, vertices), and displayed as manipulation frames. They are then used to position the selected primitives in the 3-D space. For instance, the selection of the same contour on two oriented images defines a 3-D line used to constrain the position of one of the object edges. In order to speed up that process, the structural or functional knowledge on some particular environment is exploited inside specialised modules featuring dedicated primitives, constraints and modelling methods. In the case of a piping module, a straight pipe is modelled by specifying its contours in two images or more. So edge segment specification is often required as well in the generic module as in specialised ones.

Careful matching of 2-D segments on image edges in manual mode is long and visually tiring. On the over hand, its full automation is often doomed to failure, because of poor visibility, bad scene illumination, numerous reflects on stainless steel surfaces, frequent occlusions, ... Therefore a semi-automated mode has been implemented. Each time the operator draws a segment over an image, a local edge detection function is automatically executed, and attracts the segment on the nearest edge found in the image. Hence only a coarse specification is required from the operator. Of course the segment may be attracted towards some disturbing edge close to the desired one. The operator then selects one of the segment ends and drags it towards the relevant edge. New edge points are thus extracted on a different image area. This is the only way to act on the automatic function. In case of repeated failures, the operator can come back to the manually specified segment with a simple undoing action. When it happens too often, the automatic function can be disabled.

The edge detection function consists in extracting edge points inside a rectangular image area around the selected segment. For each point, possible edges are accumulated inside a Hough table. A vote is then performed to provide the best candidate. The sides of the detection area are parallel to the image borders. The distance between the initial segment and
the area perimeter is a system parameter which can be fixed through the interface. Actually we consider that the operators should not need to know how this function works. A default value of 5 pixels was assigned. It does not require an accurate initial matching and affords some discrimination possibilities well appreciated when disturbing segments are close to the desired one.

This semi-automatic edge segment specification has been evaluated on images with different kinds of difficulties. Image size was 768x576 pixels and experiments were performed on a Silicon Graphics Indigo workstation. The computation time depends on the selected segment length and angle. In the worst cases, a latent period of 8 seconds has been observed. But they are rather rare and in most of the encountered cases, the computation time was around one second. Edge segment specification tasks are largely speeded up. Their accuracy is comparable with a careful segment specification in manual mode. We have also evaluated the repercussion of this operation improvement on a whole modelling task. The tests performed on several types of industrial environments showed a large reduction of the global modelling time.

2 THE TELEROBOTICS CONTEXT

A 3-D geometric model of the remote environment is required at the different stages of a telerobotics mission. In particular during the execution stage, possible incidents and unforeseen events add specific requirements to the programming and supervision tasks. Based on the availability of a consistent model, Virtual Reality techniques (Sayers and Paul, 1994, Natonek et al., 1994, Thibout et al., 1994, Bejczy, 1995, Burdea, 1999) and Augmented Reality techniques (Oyama et al., 1993, Cannon and Thomas, 1997, Kim, 1999) can provide flexible solutions to enhance the remote control safety and to reduce the mission execution time.

Of course the model acquisition or update may be performed during some previous inspection step. But it is sometimes difficult to provide a precise specification of the mission requirements. Moreover the environment may evolve during the mission. Some robotics tasks may cause unpredictable objects displacements. The availability of on-line modelling techniques using information feedbacks from on-board sensors reduce the model inconsistency problems. The time devoted to the modelling task must be short compared to the time spared from the model exploitation.

The kind of sensor used for the modelling is selected according to the mission constraints. It depends on the robot payload capacity and energy autonomy, but also on the environment features. Especially in the case of missions inside radioactive cells, tolerance to radiation effects, easy remediation and low cost are important criteria. The CCD camera already used to pilot the robot and to supervise the tasks execution seems quite convenient.

On the other hand automatic modelling methods fail to provide the appropriate model, because of the scene complexity, disturbing reflections due to highly specular surfaces, and low contrasts from bad lighting conditions. On the other hand, manual methods are time consuming. An interactive 3-D modelling principle (fig. 1) has been proposed to meet the telerobotics field requirements (Even and Marcé, 1988). It relies on a man-machine cooperation through the super-imposition of 3-D solid primitives over the camera images.

![Interactive 3D Modelling Principle](image)

Figure 1: The interactive 3-D modelling principle.

The operator brings his structural and functional knowledge on the remote environment. He controls the data collection, selects the appropriate primitives and matches them on the image details. This mixing of the virtual world with the real data is a reliable way to ensure the model consistency. The operator settles the final level of detail according to the tasks to be performed and the available modelling time.

CEA/STR has developed the Pyramide interactive modelling system as an integration platform for the new modelling methods. Images from an on-board camera are used. The CCD camera parameters are identified during a previous calibration step. Wide angle lenses are required for piloting or supervision activities. Thus distortion parameters are also identified and used to remove the distortion effect from the acquired images. This image reconstruction step takes
time and restricts the camera use to static views. Moreover it inserts some blur in the final image. In order to ensure a correct match between the model and the images, the camera viewpoint must be computed from the mobile robot relative localization module (Even et al., 1989) or from the manipulator arm joint values (Even et al., 1999). Pyramide's interest has been verified through many experimental campaigns in realistic conditions.

Time is the most important limitation factor for this technique. Therefore most of the works performed these last years deal with solutions to speed up the interactive modelling process. Structural knowledge on particular environments has been integrated into specialized modules, based on dedicated primitives handled through an optimized modelling operating mode and automatic constraints management. A piping module has been developed and evaluated (Even et al., 2000). Considerable time is spared for the modelling of piping elements. The remaining objects are modeled using the generic module. A large improvement is also obtained by setting the description level to the mission needs. The functional classification proposed in (Even et al., 1999) shares the objects into one of the three following types: out of reach objects are coarsely modelled, just letting some distinct features for navigation purposes; potential obstacles are composed of surrounding volumes, their accuracy being directly connected to the security distance of a collision avoidance module; manipulated objects are more detailed with regards to the required geometric features. The basic idea is to provide the operator with some task-oriented requirements in order to avoid useless model refinements.

A large enhancement can also be expected from the integration of automatic assistances based on computer vision techniques. For instance the automatic fitting of a primitive on image features has been integrated into several modelling systems (Hsieh, 1995, Debevec et al., 1996, Kim, 1999, Gülch et al., 1999). The operator very quickly provides an coarse initial solution. Manual fine matching is tiring and takes a lot of time. This task is autonomously performed using an optimisation algorithm based on the minimisation of a distance between the drawn segment and the extracted edge segments in the image. Pyramide has also been improved by integrating such primitive fitting in semi-automatic mode (Bonneau and Even, 1993).

3 EDGE SEGMENT SPECIFICATION PURPOSE

The basic modelling principle in Pyramide consists in matching solid primitives on the relevant image features. However the fine control of an object orientation in space using a 2-D input device such as a mouse and only visual feedbacks is a difficult task. Computing the object orientation from image features can help a lot. Once the object is well oriented, its matching is easily obtained with the mouse. Pyramide provides several methods to compute an orientation from the available edge segments in the images (fig. 2).

- **A)** Extraction of a vector from the projection of parallel lines in one view. The intersection of the edge segments is the vanishing point associated to the 3-D lines common direction. It is used to compute the lines unit vector. For a better accuracy, wide angle lens and distant edge segments in the image are recommended (Shufelt, 1999b).

- **B)** Extraction of a line from its projection in two views. For each view, the edge segment defines an interpretation plane passing through the projection center. The 3-D line lies at the intersection of both edge segments interpretation planes. This method is used to align an object on a given edge. The object is then moved along and rotated around its edge. Accuracy requirements deal with the selected edge segment length, with its sharpness and with the angle between the interpretation planes.

![Figure 2: Available methods for the extraction of 3-D features from edge segments specification.](image-url)
C) Extraction of a line from the contours of a shape of revolution in two views or more.
This method is a generalization of the previous one. On each image, a median plane to the contours interpretation plane is computed. The intersection of each median plane provides the 3-D shape revolution axis. Cylinders and cones are classical shapes of revolution, but more complex shapes can also be modelled with surrounding volumes by setting the specified 2-D lines upon the shape outline (see for example fig. 3).

D) Extraction of a plane from the projections of coplanar features in two views.
The plane is computed from the specification of the projection of three points, one point and one line, or two lines, according to the visible features in the image. The edge segments are interpreted the same way than in method B. An optical ray is computed for each image point and the image projection center. Because of the inaccuracies on the viewpoint localization and on the points specification, the optical rays of homologous points do not intersect. The 3-D point is approximated using a weighted mean.

E) Extraction of an orientation from the projection of three orthogonal lines in one or multiple views.
This method is a generalization of an analytic solution to the interpretation of three edge segments corresponding to perpendicular lines in space (Horaud et al., 1989). Two possible orientations are provided. The operator visually detects the wrong one. An important requirement for a better accuracy is to avoid edge segments which interpretation plane normal vector is nearly orthogonal to the view optical axis (grazing view).

All these methods could be enhanced if more image features were used. The information redundancy could be exploited to achieve a better accuracy. But time is as essential requirement as precision for on-line modelling. Therefore we rather try to reduce the amount of operator’s actions.

4 SEMI-AUTOMATIC EDGE SEGMENT SPECIFICATION REQUIREMENTS

Edge segments specifications are often required when modelling with Pyramide. Manual specification is a tedious task. Accurate drawing of a line segment over the image feature is time consuming and visually tiring. Therefore a semi-automatic assistance has been integrated. One solution consists in first automatically extracting the edge segments in the image, then manually selecting appropriate ones (Lang and Förstner, 1996). But the edge segment density is sometimes so high that the selection task requires a fine accuracy and becomes tedious. In the solution that we investigated, a line segment is quickly drawn close to the expected edge segment. It is then automatically attracted towards the best edge segment extracted in the vicinity of the initial solution.

Several requirements should be fulfilled to ensure the efficiency of the integrated assistance (Hsieh, 1995). The automatic attraction should be robust, and discard disturbing edge contours that could keep the segment away from the solution. It should also be fast enough in order to provide a good man-machine cooperation. This notion of acceptable response time depends mostly on the operator’s current activity. In the present case, the operator waits on the result to proceed on the modelling work. Therefore the average acceptable limit is about one second.

The time devoted to the possible settings of the automatic assistance should not exceed the expected benefits. The possible failures and their recovery should not compromise the assistance over-all efficiency on the whole modelling. The assistance should not demand a strong expertise to the operator. If the implementation requires some particular operating mode, its learning should be counterbalanced by large benefits for the modelling task.

Figure 3: Examples of virtual edges (shape of revolution outlines) or nearly virtual edges (partly occluded segments).
The automatic process limitations should not affect the application domain of the modelling system. Specially virtual or nearly virtual edges are sometimes required (fig. 3). Their specification should be left possible. In the case of multiple failures, the assistance should be possibly deactivated.

The assistance performance should be equivalent for similar images. This is important to make it predictable so that the operator can decide as soon as possible if he uses it or not.

5 AUTOMATIC EDGE SEGMENT ATTRACTION

The simultaneous extraction of multiple edges which are linked together through geometric constraints - global energy minimisation (Fuia and Leclerc, 1988), part of a solid primitive (Läbe and Gülich, 1998), collinear line grouping (Shufelt, 1999a) - is a way to provide robustness. But in Pyramide, lines are processed separately. A gradient based approach is used in (Debevec et al., 1996) to attract a manually defined line segment to the nearest edge contour. Because of the uncertain quality of the provided images by the on-board camera, we rather selected a more robust technique.

First of all the operator quickly draws a line segment on the image. A classical edge contour extraction is then performed in the line segment vicinity. The extraction area is a rectangle which sides are parallel to the image borders. If \( M_1(x_1, y_1) \) and \( M_2(x_2, y_2) \) are the manual segment ends, the extraction area is bounded by the line \( x_{\text{min}} = \max(0, \min(x_1, x_2) - \epsilon), \)

\( y_{\text{min}} = \max(0, \min(y_1, y_2) - \epsilon), \)

\( x_{\text{max}} = \min(\max(x_1, x_2) + \epsilon, X_{\text{im}} - 1), \)

\( y_{\text{max}} = \min(\max(y_1, y_2) + \epsilon, Y_{\text{im}} - 1), \)

where \( X_{\text{im}} \) and \( Y_{\text{im}} \) are the image size, and \( \epsilon \) is a system parameter fixed to 5 pixels. This value was found to be a good compromise between the method efficiency and its flexibility. It bounds the accepted approximation of the manual specification.

The provided edge points are then analyzed using a Hough transform. The Hough transform translates the edge segment detection problem from the Euclidean space to some parameters space. It is based on a vote technique where each extracted edge point votes for a set of possible lines. Polar coordinates are a suitable parameterization to solve this specific problem (Duda and Hart, 1972). They are given by the orientation \( \theta \) of the normal vector and the distance \( \rho \) to the image origin (see fig. 4a). The line \( \{\rho, \theta\} \) is the locus of points \( (x, y) \) defined by equation (1):

\[
\rho = x \cos(\theta) + y \sin(\theta)
\]  

(1)

Let \( V_0 \) be the neighbourhood of a given line segment \( \{\rho_0, \theta_0\} \) in the Hough space. Defined by a sampling of the local area around the line, an accumulator array is associated to \( V_0 \) (equation 2).

\[
V_0 = [\theta_0 - n_\theta \cdot \delta \theta, \theta_0 + n_\theta \cdot \delta \theta] \times [\rho_0 - n_\rho \cdot \delta \rho, \rho_0 + n_\rho \cdot \delta \rho]
\]  

(2)

The accumulator array dimension is \( N = (2n_\theta + 1) \times (2n_\rho + 1) \). Each extracted edge point \( P \) votes for the lines of \( V_0 \) it belongs to. The highest number found in the accumulator array provides the polar coordinates of the best edge segment.

The Hough transform is very robust to possible occlusions; edge points on each part of the occluded edge vote for the same line. It also lowers the influence of spurious edge points due to detection noise, disturbing reflects or close objects. For two very close edge contours, the Hough transform provides the most influential one, but not some average line between both contours. The Hough transform is applied twice, the first time with a coarse definition to lower the vote dispersion, the second time with a finer definition to provide a good accuracy.
6 SEMI-AUTOMATED EDGE SEGMENT SPECIFICATION

The automatic edge segment attraction function has been integrated within a semi-automatic mode that was called segment magnetisation. It is summarized in the flow chart of figure 5. The operator first draws an initial segment using the mouse. A press on the mouse button defines the first segment end. As long as the button is pressed, the segment is continuously displayed so that its second end follows the cursor. At the release of the button, an initial segment is provided, and the automatic attraction is immediately run. If the segment is attracted towards a spurious edge segment, the operator can click at one of the segment ends and drag it towards the desired edge. The automatic attraction is run again as soon as the segment is released. After any automatic attraction, the operator can come back to the manual solution with a simple undoing action.

![Flow chart of the semi-automated edge segment specification.](image)

The semi-automated edge segment specification was evaluated using real sites images. Both the performance of the automatic attraction function and the one of the semi-automatic mode have been evaluated. In the latter case, we recorded the time spent to model some objects with and without the assistance.

Because of the automatic function robustness, the desired segment edge is often detected at the first go. An initial solution is very quickly drawn. As most of segments length is less than 300 pixels, the average response time is less than one second on a Silicon Graphics Indigo workstation. The worst case is observed when the initial segment is one of the image diagonal. The computation time may then rise to about eight seconds. But such segments are few and the waiting situations are rather seldom so that their impact on the semi-automatic assistance remains low.

Most frequent failures are due to the presence of a better quality edge in the neighbourhood of the desired segment. When it happens, a simple modification of the segment to drag it away from the disturbing edge is most of the time sufficient. Such an example is shown in figure 6. In case of repeated failures, only a manual accurate specification can solve the problem.

![Example of a two steps edge segment specification.](image)

The operating mode of the semi-automatic edge segment specification is very intuitive and requires no learning. It is quite flexible. For instance for the virtual edges of figure (3), the assistance could be deactivated in order to avoid any disturbance of the operator’s actions. Up to now, the only images where the assistance had to be deactivated because of multiple failures are very difficult ones, featuring many piping lines made of highly specular stainless steel which produce plenty of reflects. Using all the other images from realistic sites, the time spent to the modelling tasks is greatly reduced thanks to the semi-automatic assistance. The few failures are quickly recovered, so that their influence on the over-all performance is quite low.

---

7 CONCLUSION

CEA/STR has developed Pyramide, an interactive 3-D modelling system to on-line acquire or update an environment model during telerobotics missions. Solid primitives are fit to the visible features of images provided by an on-board CCD camera. Time is a key requirement to get a benefit from these techniques. Therefore solutions have already been proposed to benefit from the operator’s structural and functional knowledge on the remote environment.

This modelling principle relies mostly on interactive tasks. The operator manipulates solid primitives in the 3-D space. In order to make that task easier, the primitives can be aligned on 3-D directions extracted from image features. Several extraction methods are implemented. They rely on the specification of some edge segments, a time consuming and visually tiring task when performed in manual mode.

A semi-automatic edge segment specification mode has been designed and integrated into Pyramide. It relies on an automatic function to attract a manually specified edge towards the nearest edge contour. Based on a Hough transform, this method is quite robust, and its response time is compatible with the semi-automatic mode requirements.

The semi-automatic assistance combines manual and automatic tasks. The operator first coarsely draws an initial solution. This line segment is immediately attracted towards the better edge in the vicinity. In case of failure due to a spurious edge, the operator drags the segment towards the correct edge. Repeated failures are overcome by deactivating the automatic function and performing an accurate manual specification.

This semi-automatic assistance performance have been evaluated on realistic images. It was found to be fast, robust, intuitive and flexible. The global modelling time is greatly reduced. Moreover the few failure cases do not entail the obtained benefits.

REFERENCES


