FIVE '96

Framework for Immersive Virtual Environments

ESPRIT Working Group 9122

Proceedings of the 2nd FIVE International Conference
Palazzo dei Congressi 19-20 December 1996

Advances, Applications and Impact of Immersive Virtual Environments

edited by

Massimo Bergamasco

PERCRO

Scuola Superiore S. Anna Pisa-Italy

REAL-TIME TRACKING MOVEMENTS USING A VIDEO CAMERA AND A POSITION SENSOR

F Gaudillat, V Rodin, J Tisseau

Ecole Nationale d'Ingénieurs de Brest
Technopôle Brest-Iroise - CP 15
29608 Brest Cedex, France

Abstract

This paper describes a real-time tracking movement system running on a Silicon Graphics Indy station It deals more particularly with the data fusion from a position sensor and a standard color camera so as to speed up processing Variations of the shape of the hand are taken into account using an active contour model, and the motion of the hand is detected by the position sensor
We describe a method to find local deformations by segmentation, filtering and moving of an active contour We will also account into details for the use of global deformations allowing to follow the user's fast moves. That last point implies a calibration of the camera-sensor couple to know the location of the sensor in the image at all times
Keywords *Active Contour, Data Fusion, Hand Tracking, Image Processing, Position Sensor*

1.Introduction

This paper describes a real-time tracking movement system running on a Silicon Graphics Indy station equipped with a standard color camera. Our system also uses a position sensor [1] (Flock of Birds) allowing acceleration of processing, locating the hand in the image at all times. Please note that to detect the hand more easily in the image, the user wears a green glove.
We will use an active contour model [2] moving in a region image, so as to find the outline of the hand in an image. The active contour model is also used in combination with the data from the position sensor to perform an initial image search in order to locate the hand and avoid failure at the time of fast movements.
In our system, the hand is taken as a deformable object which is characterized by local deformations (variations of the shape of the hand) and global deformations (rotation, translation, scale factor) The local deformations are taken into account by the active contour model whereas the position sensor provides information about global deformation.
This describes a method to find local deformations by segmentation, filtering and moving of an active contour. We will also account into details for the use of global deformations allowing to follow the user's fast moves. That last point implies a calibration of the camera-sensor couple to know the location of the sensor in the image at all times.

2.Local Deformations

One has defined a discrete, closed, active contour model moving around in a region image, so as to track the outline of the hand
On the first hand, we will perform a simple color filtering of the image allowing extraction of the green colored regions (same color as the user's glove). To perform the filtering, we take a pixel as belonging to the green color if this equation proves correct :

$$\text{Green} - \text{Blue} - \text{Red} \geq 0$$

A median filter [3] is then applied so as to remove the residual noise in the region image.
Eventually, we will use an active contour which moves towards the boundaries of the region corresponding to the hand.

3.Global Deformations

Suppose that location of the hand does not vary much from an image to another. Then the contour which has been detected in an image can be used to initialize the search in order to locate the contour in the following image.
A first contour detection must be carried out to initialize the model. The contour is then projected from an image to another and must adapt to the new outline of the hand each time. That is an effective method, also allowing to follow slow movements in real time.
Unfortunately, some problems turn up when fast movements of the hand are performed. The gap between the projected outline and the new outline of the hand becomes more important indeed, and the convergence of the active outline is longer. The system then cannot keep up anymore and fails.

A new element will be introduced, so that the question of movement speed limit can be solved : it is an electromagnetic position sensor which can give its orientation and location in space.
The sensor lies on the user's hand for knowing, at all times, the position of the hand in the image. The main advantage in using that sensor is that the active contour which has be found in an image can be projected to a position which is very near the new contour to be taken out of the following image The initializiation of the outline detection in the following image is as follows:

$$\vec{X}_i = \vec{X}_{i-1} + \vec{T}_i$$

Where : $\vec{X}_{i-1}$, is the model which has been taken out of the previous image,

$\vec{X}_i$, is the model which has been projected into the current image,

$\vec{T}_i$, is the translation calculated using the position sensor.

It should be noted that, the calculations being fast and of an elementary nature, only the translation of the hand is taken into account.

4. Calibration and Detection of Sensor in Image

The calculation of the sensor position in the image needs a calibration of the camera-sensor system. The selected calibration method first goes through modeling of the camera using the pin-hole model [4] (figure1) Please note that there are other more complex models (thin lens model, thick lens model), and that each can be replaced by an equivalent pin-hole model.

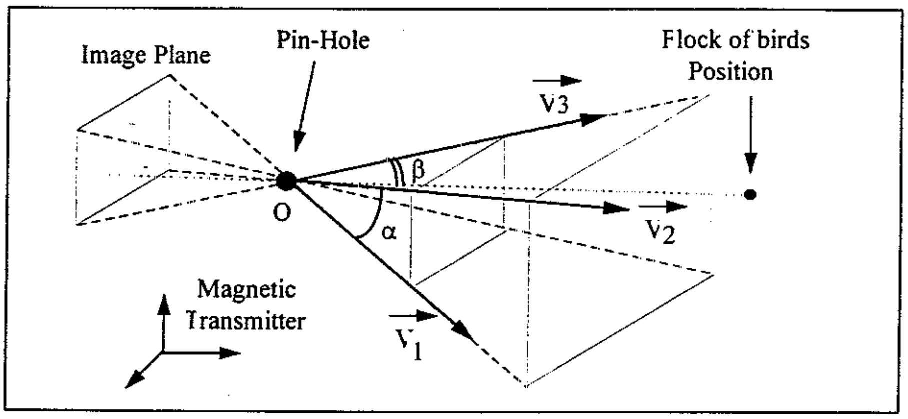

Figure 1. Modeling of the camera-sensor couple.

The aim of calibration is to determine angles α and β corresponding to the angles of the camera field of view (figure 1). Calculation of the sensor position in the image will involve calculation of angles α_M and β_M representing the sensor coordinates in the cone of the camera (figure 2). It is quite simple to locate the sensor in the image from those angles

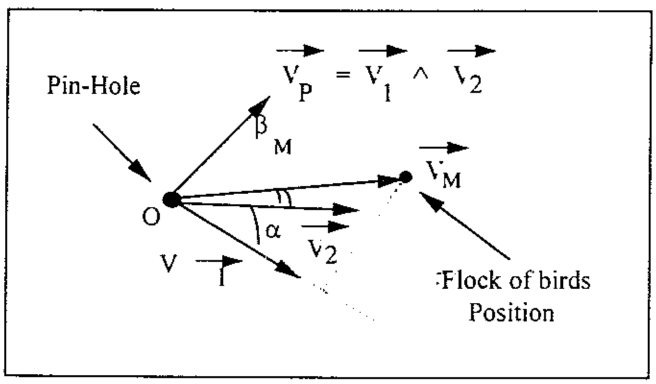

Figure 2 Determination of angles α_M and β_M.

4.1 Determination of Angles α and β

To calculate angles α and β, we determine vectors $\vec{V}_1$, $\vec{V}_2$ and $\vec{V}_3$ (figure 1) from four measures of the sensor position (sensor fixed on camera lens and in three different corners of the image). Angles α and β are then defined by the scalar products :

$$\vec{V}_1 \cdot \vec{V}_2 = \left\|\vec{V}_1\right\| \cdot \left\|\vec{V}_2\right\| \cos\alpha$$
$$\vec{V}_2 \cdot \vec{V}_3 = \left\|\vec{V}_2\right\| \cdot \left\|\vec{V}_3\right\| \cos\beta$$

4. 2. Determination of Angles α_M and β_M

Given vector $\vec{V}_M$ defined by the sensor position in space and fixed position of lens at the time of calibration (figure 2)
Angle β_M corresponds to the angle between the straight line defined by $(O, \vec{V}_M)$ and plane P defined by $(\vec{V}_1, \vec{V}_2)$
Given $\vec{V}_P = \vec{V}_2 \wedge \vec{V}_1$ a normal vector at plane P. According to figure 2, angle β_M may be deduced according to the following relation :

$$\vec{V}_M \cdot \vec{V}_P = \left\|\vec{V}_M\right\| \cdot \left\|\vec{V}_P\right\| \cos\left(\frac{\pi}{2} - \beta_M\right)$$
$$= \left\|\vec{V}_M\right\| \cdot \left\|\vec{V}_P\right\| \sin\beta_M$$

α_M can be valued using the same method.
Angles α_M and β_M now known to us, the sensor position in the image can be determined by :

$$i = \frac{\alpha_M}{\alpha} \cdot nbLines\ ,\ j = \frac{\beta_M}{\beta} \cdot nbColumns$$

5.Conclusion

This study has described a tracking movement system. It deals more particularly with the data fusion from a position sensor and a color camera so as to speed up processing. Variations of the shape of the hand are taken into account using an active contour model, and the motion of the hand is detected by the position sensor.
The system can process from 15 to 28 images a second (according to the size of the hand in the image and the wanted tracing precision). It runs smoothly with a complex light-colored background (dark color could be interpreted as another shade of green). However, segmentation mistakes can originate from reflections and backlit effects on the user's hand. The system is fairly immune to noise though, but does not allow clearing up of partial stoppages, for we do not use a model of the hand.

References

[1] Ascension Technology Corporation BOP 527, Burlington, Vermont 05402 « The Flock of Birds™, Position and orientation measurement system », 1995.

[2] M. Kass, A. Witkin, D. Terzopoulos. « Active contour models » International Journal of Computer Vision Vol 1, pp 312-331, 1988

[3] J.W. Tukey « Nonlinear (nonsuperposable) methods for smoothing data » In Conf Rec. EASCON, p. 673, 1974.
[4] B.K. Horn. « Robot Vision » The MIT Press, 1986.