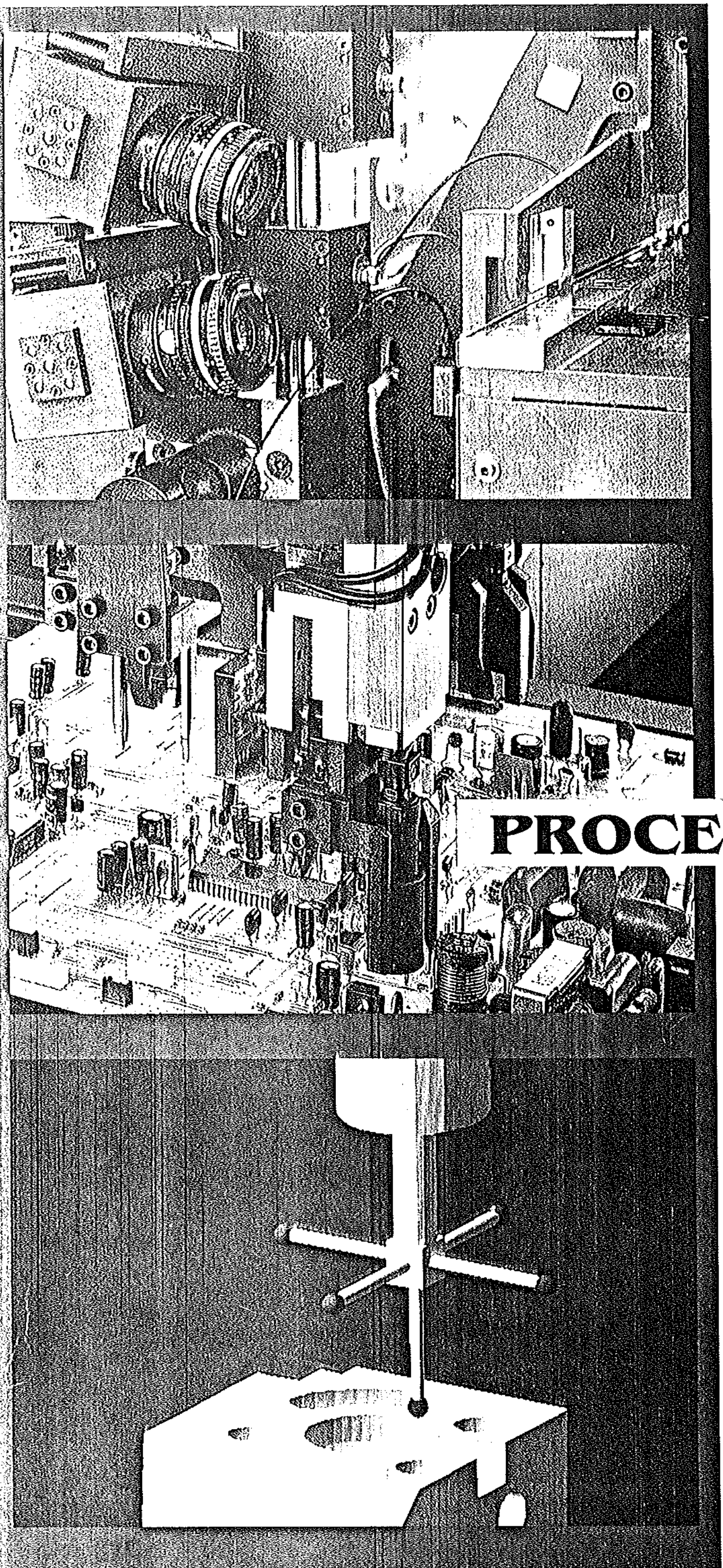

ICARCV '92

SECOND INTERNATIONAL CONFERENCE ON AUTOMATION ■ ROBOTICS AND ■ COMPUTER VISION

PROCEEDINGS

VOLUME 1 OF 3

ORGANISED BY

School of Electrical and
Electronic Engineering
Nanyang Technological University
Singapore

The Institution of Engineers,
Singapore

16-18 September 1992

A MEDICAL ROBOTIC APPLICATION OF MONOCULAR STEREOVISION - FILTERING AND MATCHING TECHNIQUES.

V.RODIN, A.AYACHE.

ENSEEIHT - IRIT - UA 1399 CNRS
Laboratoire "Vision par Calculateur A. Bruel"
2 Rue C. Camichel
31071 Toulouse Cedex - France

Phone : (33) 61 58 82 72
Fax : (33) 61 58 82 09
e-mail : ayache@enseeiht.fr

Abstract : In this paper, using the concept of axial stereovision, we propose a matching method to recover 3D informations about non-polyhedric objects. Experiments on real images are presented. This work is related to medical robotics, applied to surgical endoscopy.

I.- INTRODUCTION

In real images processing, several matching methods have been proposed to reconstruct polyhedric objects by the use of axial stereovision systems. But to our knowledge, the homologous computation is not yet solved in order to reconstruct the curved surfaces of non-polyhedric objects.

In our laboratory, works made upon axial stereovision [4][5] have allowed both to reconstruct objects from the polyhedric world and to modelise an axial stereovision camera.

In this paper, we detail an original matching algorithm to get 3D information from non-polyhedric objects. This matching algorithm is based on the study of the curves formed by the grey levels along the epipolar lines in the two images "Front" and "Back". The two main points of our method are as follows:

- a filtering process applied on the noisy curves,
- a correlation process to match characteristic points of the curves of the epipolar lines in the two images.

The 3D points computed, irregularly distributed on the scene, can be used to reconstruct the shapes of non-polyhedric objects by a surface interpolation process.

This work is related to medical robotics, applied to surgical endoscopy. Until now, when operating, the surgeon uses an endoscope and a video camera, but doesn't get any 3D information about the analysed cavity. The fact that the surgeon can't possibly use more than one camera, has led us to use the concept of axial stereovision in order to compute the shapes of organs.

II.- PREVIOUS WORKS ON AXIAL STEREOVISION

These last years, a great number of research groups thought that a camera with a zoom system [4],[7],[8],[11],... or a camera moving along its optical axis [2],[3],[5],... would make it possible to recover the 3D informations. The relative plainness of the geometrical characteristics of these systems make them very interesting for industrial applications and simplifies the calibration and matching process.

In order to solve the calibration problem, several approaches have been proposed [1],[6],[9],[10],... Some of them have given highly accurate results. [4][5] have studied the influence of calibration errors over the 3D points computation.

The matching process has also been studied. Let us quote the works of [2] who, for synthetic scene, have proposed an algorithm based on the study of the radiance in the scene and who have also solved the problems caused by the occlusions. The problem of this algorithm is that the illumination characteristics of the scene as well as the reflectance properties of the objects contained in the scene, must be known. We propose an algorithm which is released from these constraints, working under a natural light and with objects whose characteristics of reflectance are not known beforehand.

III.- GEOMETRICAL MODEL OF AXIAL STEREOSCOPIC SYSTEMS

III.1.- GEOMETRICAL MODEL

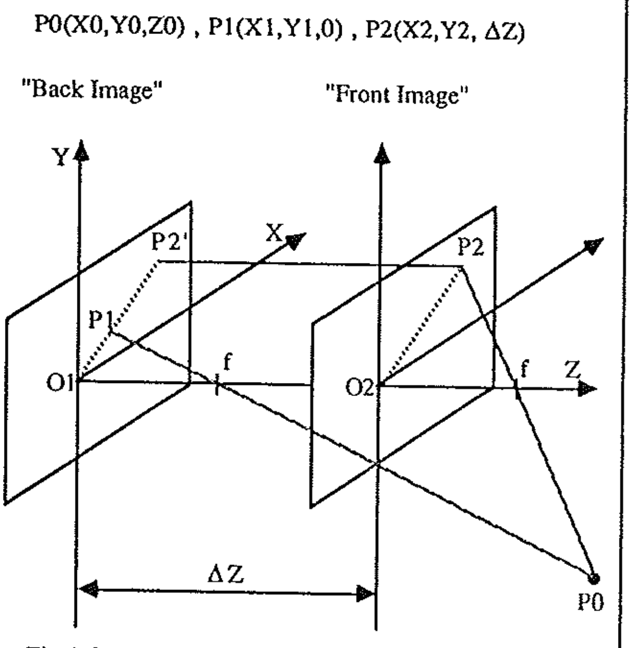

Fig.1 Geometrical Model of axial stereovision system

A system of axial stereovision can be seen as a set of two identical cameras that have common optical axes and their image planes parallel and separated by a distance ΔZ. Such a system can be simulated in two ways:

- A camera moving along its optical axis. In this case, there are mechanical problems to keep thoroughly under control the moving of the camera along its axis.
- A zoom. The crucial point of this system is the accurate calibration of the zoom which is a compulsory process in order to reconstruct a 3D scene.

We have decided to use a zoom to simulate an axial system in order to avoid the mechanical problems.

If we assume that the distance "objects-camera" is large with respect to the focal length, the "pin hole" model can be adopted as the model of the image formation process (see fig.1).

Considering a cartesian system located on the Back Image plane which has the Z axis coinciding with the optical axis, the projection of the 3D point (X0,Y0,Z0) is (X1,Y1,0) in the Back Image plane and (X2,Y2,ΔZ) in the Front Image plane.

Therefore, we have got the following geometrical equations:

$$\frac{X0}{X1}=\frac{Y0}{Y1}=\frac{-(Z0-f)}{f} \qquad (1)$$

and

$$\frac{X0}{X2}=\frac{Y0}{Y2}=\frac{-(Z0-\Delta Z-f)}{f} \qquad (2)$$

From (1) and (2) we have got the 3D coordinates:

$$X0=\frac{X1.X2.\Delta Z}{f(X1-X2)}.\,Px \qquad (3)$$

$$Y0=\frac{Y1.Y2.\Delta Z}{f(Y1-Y2)}.\,Py \qquad (4)$$

and

$$Z0=f+\frac{X2.\Delta Z}{(X2-X1)}=f+\frac{Y2.\Delta Z}{(Y2-Y1)} \qquad (5)$$

where X1,Y1,X2,Y2 are expressed in pixels unities, Px Py (pixel dimensions), ΔZ, f (focal length) are expressed in metric unities.

III.2.- GEOMETRICAL PROPERTIES

Some geometrical properties due to the system have been demonstrated in [2][5] in order to restrict the research area during the matching process. Let us quote more particularly:

- The orthogonal projection of O2 (Front Image center) on the Back Image plane is O1 (Back Image center).
- The Front and Back Image planes are parallel, the P0 P1 P2 plane is perpendicular to the image planes and contains the [P1 O1] and [P2 O2] segments. Therefore [P1 01] and [P2 O2] are parallel and are contained in the same quadrant in Back and Front Image planes [2].
- The field of view of the back camera completely covers the field of view of the front camera.
- [2][5] have demonstrated that the [O2 P2] segment of the Front Image has got a length at least equal to the one of the [O1 P1] segment of the Back Image.

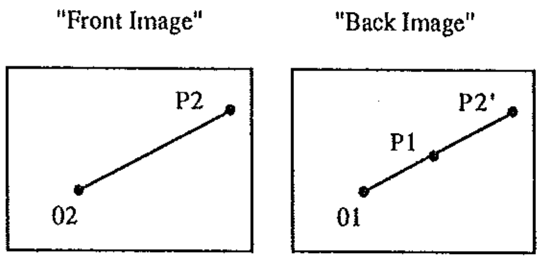

Fig.2 Homologous search area of a point P2 in the Back Image

Therefore the homologous point of P2 on the Back Image will be on the [O1 P2'] segment, P2' being the orthogonal projection of P2 on the Back Image plane (see fig.2).

The (O2 P2) , (O1 P1) epipolar lines are radial, starting from the center of the images.

IV.- THE MATCHING METHOD

In the described approach, the image center is determined by the computation of the Focus Of Expansion (FOE) of [6]. We suppose that the scene is made up of smooth and sufficiently textured surfaces.

In axial stereovision, if ΔZ is small with respect to the objects-camera distance, the grey levels of two homologous points are very close one to the other. This observation allows us, when using the geometrical properties (cf III.2.) and a correlation algorithm working on the intensity profiles along the epipolar lines, to match points of these lines. These points are local minima and maxima of the intensity curves as well as outline points.

The direct correlation of the profiles is only made possible after efficient filtering by minimizing the loss of information. Indeed the noise would disturb the correlation causing wrong matchings and therefore important errors in the computation of the 3D points (see [5]).

Our filtering process is a key point of our method.

IV.1.- THE FILTERING PROCESS OF NOISY CURVES

We propose a filtering of the profiles which minimizes the loss of information while smoothing them in a significant way. This method is mainly based on the use of the median filter a variant of which we propose that we call the "Threshold Median Filter".

Fig. 3 shows the general organisation of our filtering process.

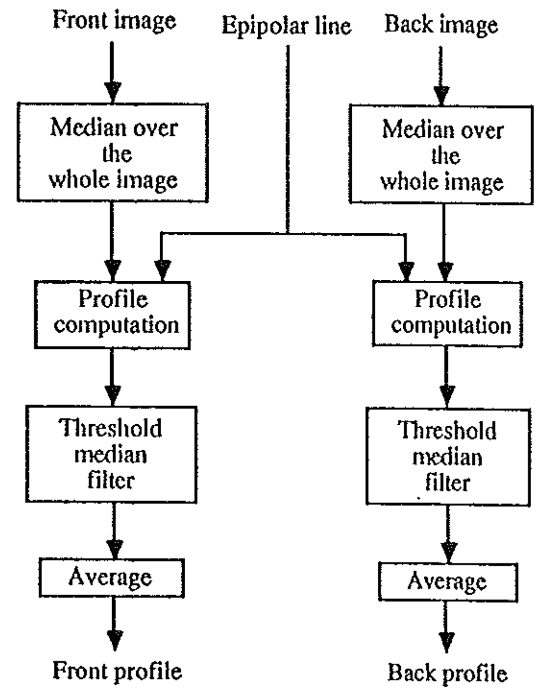

Fig.3 General organisation of our filtering process

Let us recall the classical definition of the median value: X_{med}.

Let V_O be the original value to filter.

$X_{med} \in \{X_0,..., X_i, ..., X_n\}$ such as $X_0 \le .. \le X_i \le .. \le X_n$,
as $P(X \ge X_{med}) \ge 1/2$
and $P(X \le X_{med}) \ge 1/2$

The V_O value is then replaced by X_{med}.

We define the "Threshold Median Value": X_{medT}.
Let T be the maximum difference between a noisy point and the same point made unnoisy.

Let V_O be the original value to filter.

$X_{medT} \in \{X_0,...,X_i, ..., X_n\}$ such as $X_0 \le .. \le X_i \le .. \le X_n$,
as $P(X \ge X_{medT}) \ge 1/2$
and $P(X \le X_{medT}) \ge 1/2$
if $|X_{medT} - V_O| > T$ then $X_{medT} = V_O$

The V_O value is then replaced by X_{medT}.

This filtering allows to keep the intensity peaks narrow and significant.
Fig.4 shows the results of the "Threshold Median Filter" with respect to the classical median filter on a noisy profile.

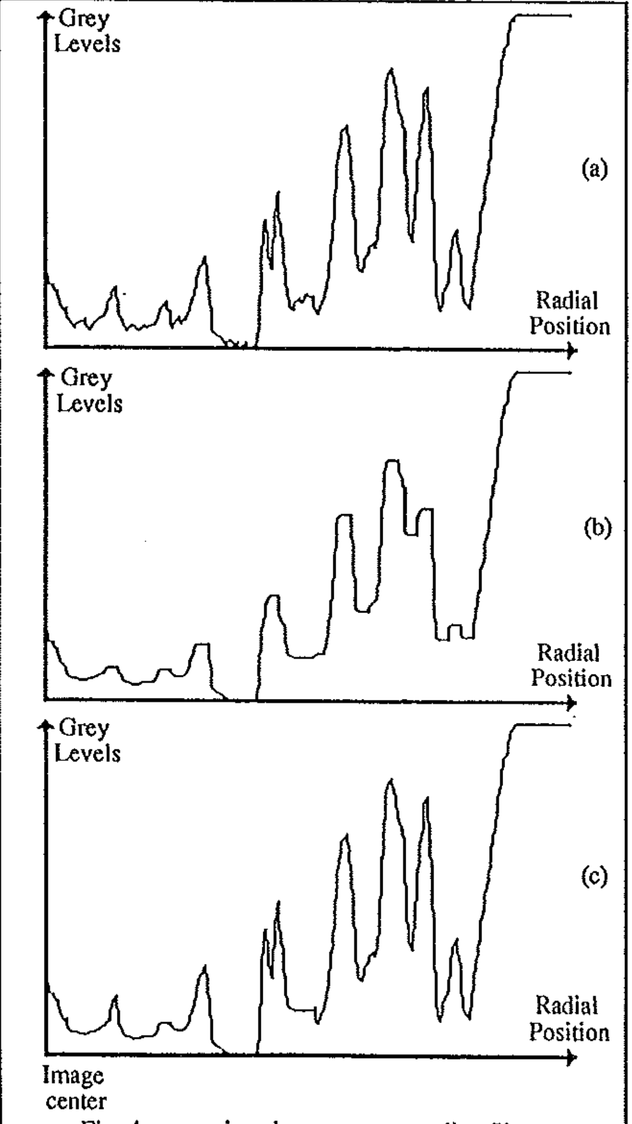

Fig. 4 comparison between two median filters.
(a) Original profile. (b) Result of the classical median filter.
(c) Result of the"Threshold Median Filter"

Fig.5 shows the results,obtained on an epipolar line, after application of our filtering process.

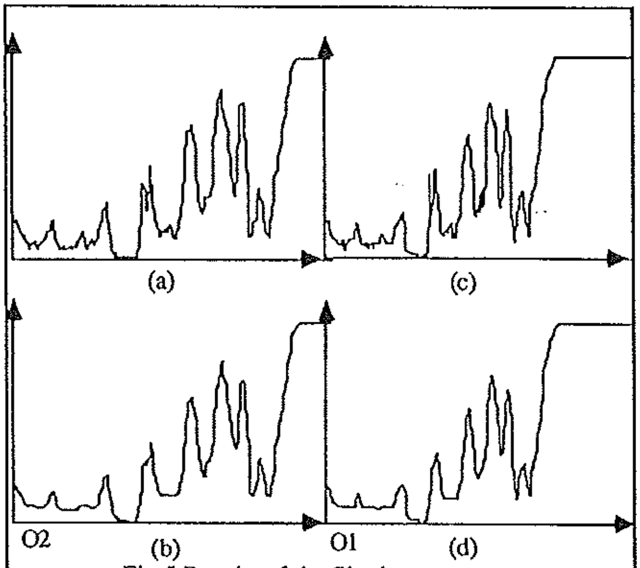

Fig.5 Results of the filtering process.
(a) Original Front profile, (b) Filtered Front profile,
(c) Original Back profile, (d) Filtered Back profile.

IV.2.- THE MATCHING ALGORITHM

This algorithm works with the front and back profiles got out of an epipolar line. By rotating the line, it is made possible to cover all the areas of the image.
After filtering of the front and back profiles, a correlation algorithm can be applied. Our algorithm has got two phases:

- a coarse segmentation,
- a matching algorithm.

IV.2.1- Coarse segmentation. First of all, there is a coarse segmentation of the two profiles in order to extract the features to be matched:

- the increasing and decreasing zones of the profiles,
- the discontinuity points of these zones corresponding to a change of the gradient sign along the profiles.

This first phase allows to initialize the matching process by supplying a segmentation which is coarse but nevertheless significant of the two profiles.
Fig.6 shows an example of coarse segmentations of the two profiles.

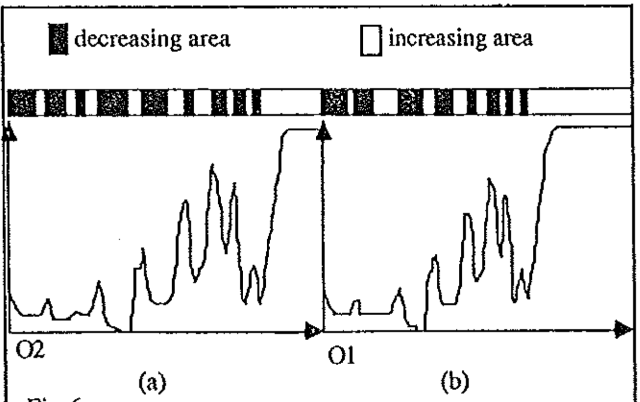

Fig.6
(a) Coarse segmentation of the Front profile,
(b) Coarse segmentation of the Back profile.

IV.2.2- The matching process. The features of this algorithm are :

- the increasing and decreasing areas of the profiles,
- the intensity profiles to compute the features of these areas (average, variance, max-min value, ...),
- the discontinuity points to be matched.

The main idea is a parallel processing of the two coarse segmentations starting from the image center.

Let SG_F be the coarse segmentation of the front profile.

Let SG_B be the coarse segmentation of the back profile.

Next-area(SG_i) is a function returning the area (increasing or decreasing) following SG_i.

Let Z_F and Z_B be the current areas of the two profiles.

Let compatibility(Z_F,Z_B) be a function which checks up the compatibility of the current front and back areas. The compatibility is defined as:

- the geometrical compatibility: the width of a Front area is at least equal to the width of its corresponding area in the Back Image.
- the radiometric compatibility: Two corresponding areas must have radiometric characteristics close enough.

This function is defined as follows:

```
compatibility($Z_F$,$Z_B$)
    { if width($Z_F$) > width($Z_B$)
          and
        Average($Z_F$) and Average($Z_B$) are close enough
          and
        Max-Min($Z_F$) and Max-Min($Z_B$) are close enough
      then return true
      else return false
    }
```

Our algorithm can be explained as folllows:

```
1) $Z_F$ <- Next-area($SG_F$); $Z_B$ <- Next-area($SG_B$)
2) if $Z_F$ = NIL then stop.
3) if compatibility($Z_F$,$Z_B$)
    then - match the discontinuity points following $Z_F$ and $Z_B$
         - goto 1)
    else - modify $Z_F$ and $Z_B$ to obtain a compatibility by:
              a) analyzing the neighborhood of the current
                 front and back areas,
              b) minimizing the radiometric chararcteristics to
                 choose the right segmentation modification.
         - if it is possible goto 3)
           else stop
```

This algorithm supplies:

- a matched points list,
- a finer segmentation of the two profiles.

Fig.7 shows an example of matched points on the profiles:

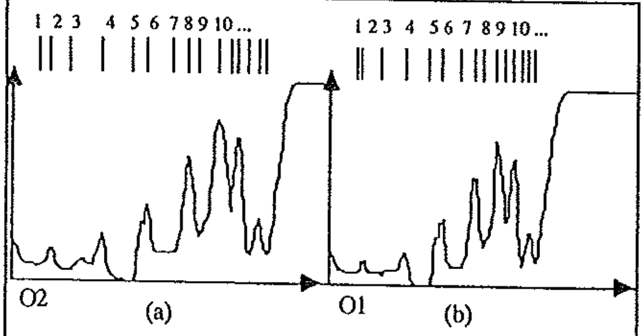

Fig.7 Matched points on the two profiles
Front (a) and Back (b)

By applying the same algorithm on the gradient profile, the outline points of the objects are matched.

Some matching errors can occur if the presegmentation of the profile is too regular (this is made possible in such textures as the one of the orange fruit) and the displacement too important. Some matching errors can also occur if the profiles are too different (in this case the filtering has not been sufficient).

Fig. 8 shows the general organisation of the matching process.

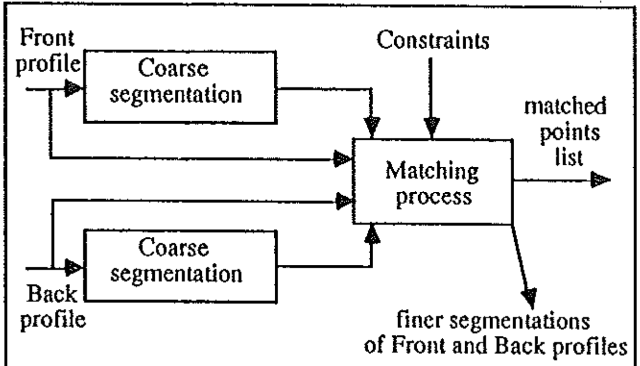

Fig.8 General organisation of the matching process

V.- EXPERIMENTAL RESULTS

The development of the specific endoscopic captor being in progress, we have been working on natural scenes of non-polyhedric objects. By applying our algorithm on some epipolar lines, we obtain the results shown fig .9.

These results show the matched points inside the surfaces.

We can note that where classical techniques would not have allowed to match points, our analysis allows to match small details present in the images.

Nevertheless there remain some problems:

- if the image is very little textured, very few points are matched. This will set a problem during the surface reconstruction.
- a too large displacement of the camera lets appear too many details in the Front Image with respect to the number of details that appear in the Back Image. In such a case wrong homologous are found.

VI.- CONCLUSION

The results obtained on real images allow to make possible a 3D reconstruction of non-polyhedric objects, by using a zoom.

There remain nevertheless some unsolved problems:

- The necessity of developping highly accurate calibration techniques.
- The necessity of realising a surface interpolation algorithm taking into account that the points that are on a surface are distributed irregularly.
- When the displacement is too important, our algorithm does not work well enough. But having an important displacement allows to be more accurate during the 3D points computation. We think that this problem can be solved by controlling the zoom mechanically and by taking images while varying the focal length a little at each step.

This work is part of medical robotics application to the surgical endoscopy. A 3D representation of the analysed cavity must allow to make the medical hand motion more steady and accurate.

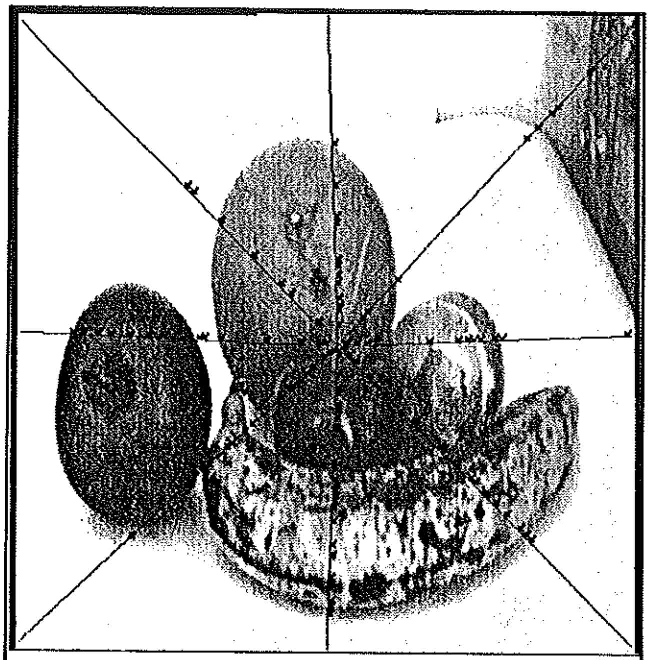

Fig. 9.a Front Image matched points on some epipolar lines

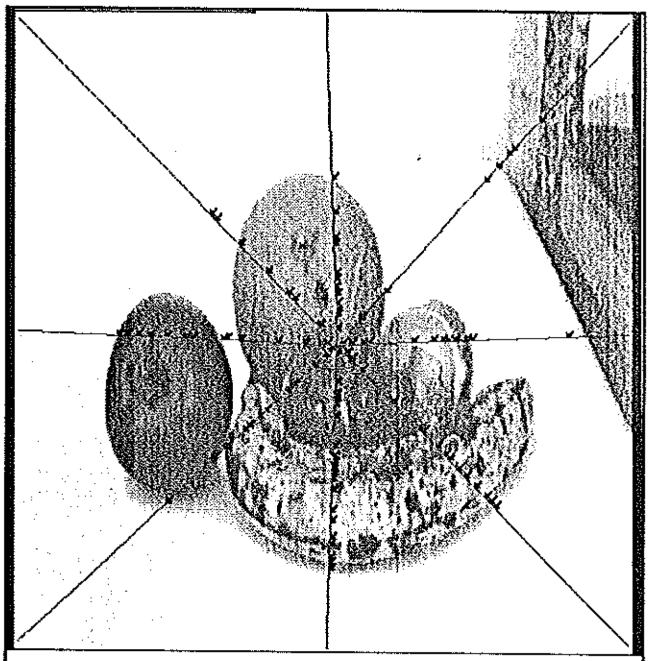

Fig. 9.b Back Image matched points on some epipolar lines

VII.- REFERENCES

[1] A.Abi-Ayad, B.Thiesse, Ph.Marthon, "Les multiples facettes du problème du calibrage", in Proc. of the Conference AFCET-INRIA, Reconnaissance de Formes et Intelligence Artificielle, tome 1. 7° Congrès, Paris 1989.

[2] N.Alvertos, D.Brzakovic,R.C.Gonzalez, "Camera Geometries for Image Matching in 3D". IEEE Trans. On Pattern Anal. Machine Intell.,vol. PAMI-11,N°9.1989.

[3] N.O'Brien, R.Jain, "Axial Motion Stereo", in Proc. IEEE Workshop Computer Vision, April 1984, pp.88-92.

[4] C.Carmona, A.Ayache, Ch.Krey, "Comparison between two axial stereovision systems", in Proc. of SPIE's 1990. Intern. Symp. on Optical, Optoelect.,Applied Science & Engineering. San Diego, July 1990.

[5] C.Carmona, A.Ayache, Ch.Krey, "A study of an axial stereovision system", in Proc.of the Intern. Conf. on Automation, Robotics and Computer Vision (ICARCV '90). 18-20 September. Singapore.

[6] R.Jain, "Direct computation of Focus of Expansion", IEEE Trans. On Pattern Anal. Machine Intell.,vol.PAMI-12, N°1. 1983.

[7] J.M.Lavest, G.Rives, M.Dhome, "Using a zoom in monocular vision for 3D reconstruction",in Proc. of the Conference AFCET-INRIA, Reconnaissance de Formes et Intelligence Artificielle, tome 1. 8° Congrès, Lyon-Villeurbanne, 25-29 November 1991.

[8] J.Ma, S.I. Olsen, "Depth from zooming. Journal of the Optical Society of America. Vol 7, N°10, October 1990.

[9] K.Tarabanis, R.Y.Tsai, D.S. Goodman, "Calibration of a computer controlled robotic vision sensor with zoom lens", in Proc. IEEE. Intern. Conference on robotic and automation. Nice, France . May 1992.

[10] R.Y.Tsai, "An efficient and accurate camera calibration technique for 3D machine vision", in Proc. Comp. Vis. Pat. Recog.Miami 1986.

[11] A.G.Wiley, K.W.Wong, "Metric aspects of zoom vision. SPIE, vol 1395. Close-Range Photogrammetry Meets Machine Vision (1990).