FR9600993

cea

DEMANDE D'AUTORISATION N°
EN VUE D'UNE PUBLICATION OU D'UNE COMMUNICATION

NIG n° 316

Direction : ....DPSA.............
Centre : ....FAR ....
Réf émetteur : .STR/94-1575 ..

DL IJ21 N. FR 9600 993

Titre original du document : REGION SEGMENTATION ALONG IMAGE SEQUENCE

Titre traduit en anglais : REGION SEGMENTATION ALONG IMAGE SEQUENCE

Titre traduit en français : SEGMENTATION DE REGIONS DANS UNE SEQUENCE D'IMAGES

AUTEURS	AFFILIATION [1]	DEPT/SERV/SECT	VISA (d'un des auteurs)	DATE
MONCHAL L.	CEA	DPSA/STR/LTO	[signature]	
AUBRY P.	CEA	DPSA/STR/LTO		

9110125

CEA-CONF-- 12165

Nature du document [2] :

ERIODIQUE ☐ CONF/CONGRES ☐ POSTER ☐ RAPPORT ☐ THESE ☐ COURS ☐ MEMOIRE DE STAGE ☐ Chapitre d' OUVRAGE ☐

Pièces jointes : RESUME ☐ TEXTE ☐

CONGRES
CONFERENCE
Nom : 8th INT'L CONF. ON IMAGE ANALYSIS AND PROCESSING
Ville : SAN REMO Pays : ITALIE Date du : 13.09.95 au : 15.09.95
Organisateur : LEILA DE FLORIANI, DISI, UNIV. OF GENOVA, VIALE BENEDETTO XV, 3, I-16132 GENOVA, ITALY

DOMAINES :

LANGUE :

N° EPAC :

SUPPORT : Disquette ☐ Papier ☐

ERIODIQUE
Titre :
Comité de lecture : oui ☐ non ☐

OUVRAGE
Titre :
Éditeur :

THESE
MEMOIRE DE STAGE
COURS
Université / Établissement d'enseignement :

MOTS-CLES : VISION PAR ORDINATEUR, SEQUENCES D'IMAGES, REGIONS

Les visas portés ci-dessous attestent que la qualité scientifique et technique de la publication proposée a été vérifiée et que la présente publication ne divulgue pas d'information brevetable, commercialement utilisable ou classée.

	SIGLE	NOM	DATE	VISA	OBSERVATIONS	REF
CHEF DE SERVICE	STR	J.M.DETRICHE	28-12	[signature]		
CHEF DE DEPARTEMENT	DPSA	A.GOUFFON	2/1	[signature]		

Date limite d'envoi du résumé : Date limite d'envoi du texte : 02/01/1995 Date limite d'envoi du poster :.../.../...

Destinataires:
Les correspondants publication des départements se chargent de transmettre à l'INSTN/MIST/CIRST (Saclay) copies des demandes d'autorisation de publication, du résumé et du texte définitif.

1 Entité d'appartenance de l'auteur. Ex. : CEA, CNRS, INSERM ...
2 Cocher la case correspondante

Region segmentation along image sequence.

Laurent MONCHAL, Pascal AUBRY
CEA/DTA/CEREM/DPSA/STR
CEN/FAR, BP6, F-92265 Fontenay-Aux-Roses CEDEX, France

Abstract **A method to extract regions in sequence of images is proposed. Regions are not matched from one image to the following one. The result of a region segmentation is used as an initialisation to segment the following image and to track the regions along the sequence. The image sequence is exploited as a spatiotemporal event.**

I. INTRODUCTION

A 3D geometrical interactive modelling system called PYRAMIDE [1] has been developed in the Teleoperation and Robotics Service (STR) at CEA (French Nuclear Agency). This system, among other goals, is to create a 3D geometrical model of the environment, to localise the robot in the modelled environment and to monitor telerobotics tasks. To reconstruct a geometrical model, a CCD camera is used.

Computation of the structure from a monocular camera requires a sequence of images. Most of the time this estimation may be split into three main steps. The first step consists in extracting relevant features in the images, then these features or primitives are matched along the given sequence. Finally through these matches motion and structure are computed. For example in [2] and [3] Motion and Structure are computed from line correspondence with least square based techniques. Zhang [4], Ayache [5], Faugeras, Lustman and Toscani [6] compute Motion and Structure from point and line matches, using a Kalman filter approach. Matching points or lines is then an important step.

When matching is not necessary, it is attempted to use the continuity of the sequence. In this case data are rather dense: these are methods based on optical flow as it is done in [7] or on the use of the sequence as a spatio-temporal event (see figure 1) as Bolles, Baker and Marimont did in [8].

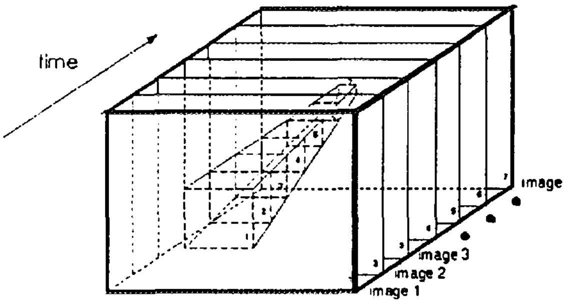

Figure 1: Image Sequence as a spatio-temporal event

Continuity is also used in matching processes: it is one of the constraint (along with uniqueness constraint) of the Marr-Poggio algorithm [9].

Moreover, as matching is necessary for stereo vision, constraint on Epipolar lines is widespread. Such a constraint is of course inadequate for a sequence of monocular images: in that case Epipolar constraint can't be applied. So matching points and edges found along a monocular sequence is more difficult than matching tokens between two stereo images.

Other constraints have been introduced to match points along a sequence. For example proximity constraint implies that trajectories of points are smooth and uniform (see [10]). To simplify the problem, regions can be chosen as primitives to be matched. For, region-based stereo analysis is easier as there are fewer regions than edges or points in an image and discrimination between regions is easier than between edges. Moreover region-based matching is seen as more accurate and reliable, especially in case of occlusion, according to [11] and [12]. The advantage of a region segmentation of an image is to obtain a dense description of the image. This description can be of good help to match less dense primitives such as points or edges. Through region segmentation along a sequence we have a more cohesive description of the sequence. That's why it is intended here to establish a denser map of the image by segmenting a sequence of images in regions.

Region segmentation requires time. that's why we intend to do it once and to track it in the image sequence. Segmentation is carried out in the first image of the sequence in a classical way. In the following image, segmentation is initialised by the labels of the previous region. The regions are transformed to fit the new image. Old regions are the "seeds" to recover regions in the present image.

II. INITIALISATION OF THE METHOD

A. *The sequence*

We are in the case where the camera observes a rigid scene with unknown motion. As we are not working under the assumption of an optical flow-based method, it is not necessary to have a very dense sequence of images. But the sequence must not be very scarce (about 10 pixels disparity between images). That is to say that motion of the camera must not be too large between images. We will see that it depends on the size of the largest regions.

B. *First segmentation*

To initialise the region segmentation of the sequence, region segmentation is done in a classical way for the first image. Pixels are gathered in a region according to a classical merging process [10]. A region is an homogeneous group of connected pixels. Each region is labelled.

The problem is that there is not necessarily always a link between a region and a face of an object. It is not very disturbing when the face of an object is divided into different regions. It can be an inconvenience when a region is the gathering of parts of different objects.

III. SEGMENTATION OF A SEQUENCE

Region segmentation is done in a classical way only for the first image. It is not attempted to extract region in the next image and to match them with the regions of the previous one. It is more attempted to "track" the regions in the following image.

The segmentation is done at different levels: a basic level with large regions and other levels with more little regions that can be repaired by the larger regions.

A. *Basic Level Regions*

When superimposing regions of the previous image on the next one there will be no "correlation" between the position of the region in the previous image and the pixels of the next one, for little region. Whereas large regions have large shared areas from one image to the following one, as it is shown in figure 2.

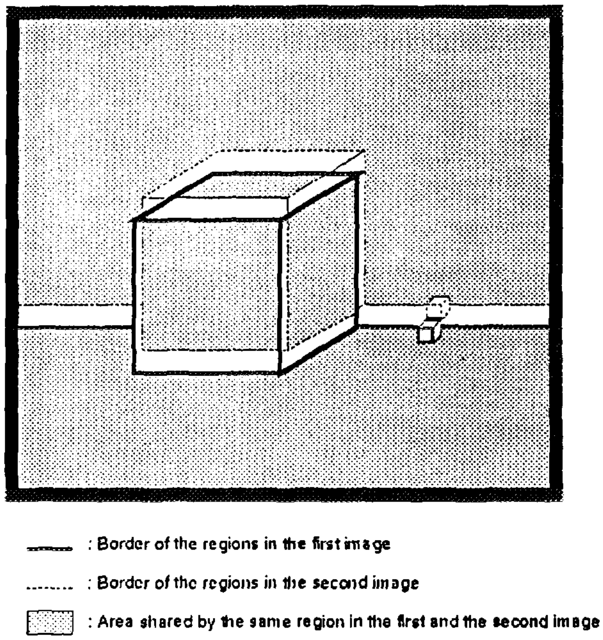

Figure 2: Regions in two successive images.

Choice of the regions

A region of the basic level is a region that have a common part which is large enough with itself between two images. More simply they are large regions. The regions of the basic level are selected in the initial image. They are chosen after having eroded twice, three or more times the regions. The remaining regions after erosion are the basic level regions. The chosen erosion for the selection depends on the motion of the camera and the size of the regions.

Finding the regions in the next image

It is assumed that for a region of the basic level the intersection between this region in image number n and image number n+1 is not empty. We must find this intersection before knowing what this region becomes. By extension of this intersection we will know what this region becomes.

Common Area

The aim, in a first step, is to find the common area a region will have to share with itself in the next image. Examples of common area is given in figure 2. To find it, elimination of pixels and erosions of the initial region are carried out.

First, regions of the previous image are imposed on the present image. Then pixels that can't belong to this region are eliminated. These are pixels whose grey scale value is not included in the range of grey scale value of the region. Pixels whose luminance is higher or lower than respectively, the maximum and the minimum of the values encountered in the previous image, are eliminated from the region. They can't belong to a common area.

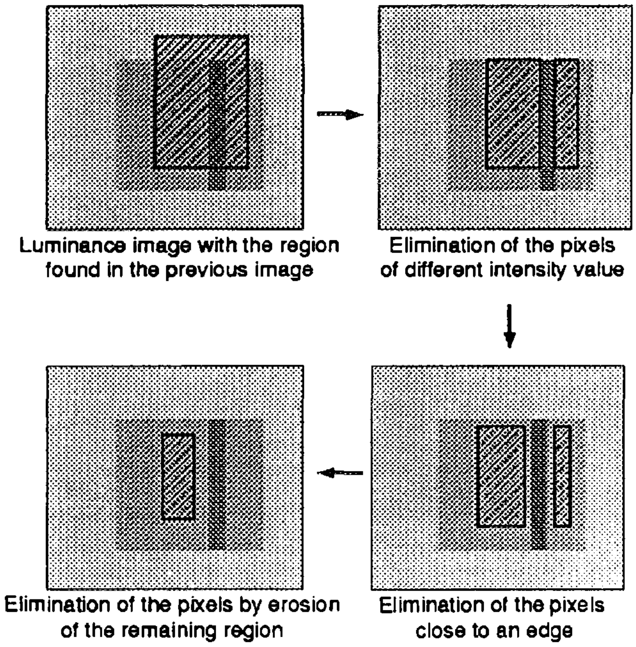

Figure 3: Finding the common area

Then the edge points of the image are extracted. The image of the edge points is dilated to obtain an image with reinforced contours. These regions created by these thick contours are also eliminated. It is done in order to eliminate noise or adjacent regions with the same intensity, regions are eroded.

The different steps are shown in figure 2. By this means it is possible to track regions which are large enough.

Recovery of the region

The goal is now to recover the region in the image from the described Common Area.

Common areas are dilated given certain rules or constraints. We try to expand the common area in order to fill the new region as it is shown in figure 3.

Dilatation of the regions is allowed on the free pixels (not belonging to a region):

- if the pixel which is chosen to be added to the region has an intensity cohesive with the previous region,
- if there is no edge point in a given neighbouring,
- if the number of pixels in the new region is not too large compared with the number of pixels contained at the previous step,
- optionally, if the shape of the new region is not too far from the shape of the preceding region (controlled by second order moments).

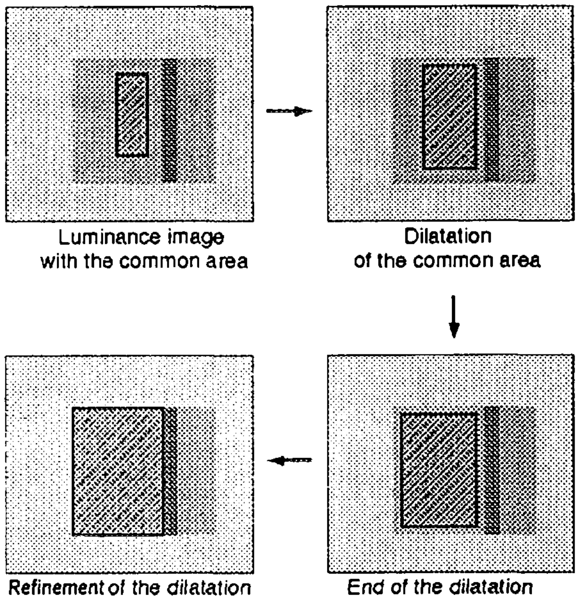

Figure 4: Dilatation of the common area

When the border of two regions meets, dilatation is done, taking into account the previous constraints and competition constraint. It is permitted for a region to dilate on an other region if the pixel at stakes incorporates very well the region. A cost function is computed to know the region a pixel most probably belongs to.

B. Higher Level

It is necessary to track other regions that are smaller than the basic level Regions, to have a denser description of the image. They are thin or little regions.

Higher Level Regions (HLR) can't be found in the same way than Basic Level Regions (BLR). For, in general, no Common Area can't be found. But we know the relative position of the BLR toward the HLR in the previous studied image. As we know the position of the BLR in the previous and the present image, we can interpolate the position of the HLR in the present image as it is shown in figure 5.

For the interpolation of the motion of a little region we choose four big regions surrounding it. To do that, we divide the image around the region taken into account into four parts. In each part we take the closest region. The motion of the little region is a linear combination of the motion of the larger regions weighted by the inverse of the distance between the little region and the big one taken into account.

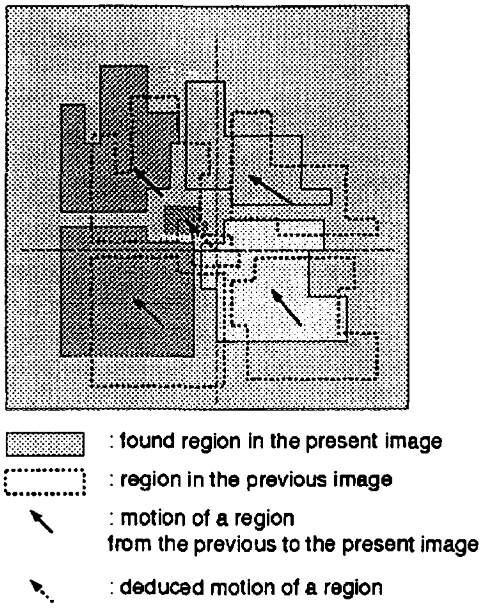

Figure 5 : Interpolation of a HLR

This approximates rather well the motion of the small region. So we can locate roughly the new small regions and apply for them procedures of erosion and dilatation in the same way as for the BLR.

IV. RESULTS

We tested our method on different kind of sequence of images. As expected it is well adapted to images with large regions and translations: large regions are well tracked whereas little regions are well deduced from larger ones.

A result is presented in figure 6, 7 and 8. There are many little regions that have no shared regions from one image to the following one. As their motion is locally a translation that can be deduced from the surrounding large regions they can be tracked.

Few little regions disappear. Most of them are not distorted; their shapes are not very different from the beginning.

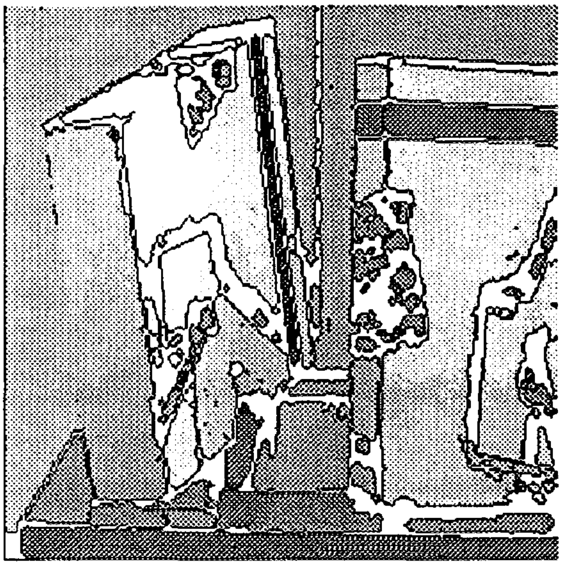

Figure 6 : Regions at the first step

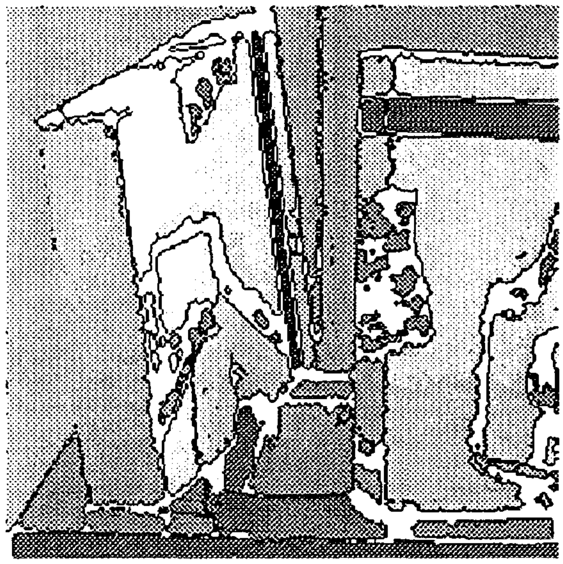

Figure 7 :Regions at the second step

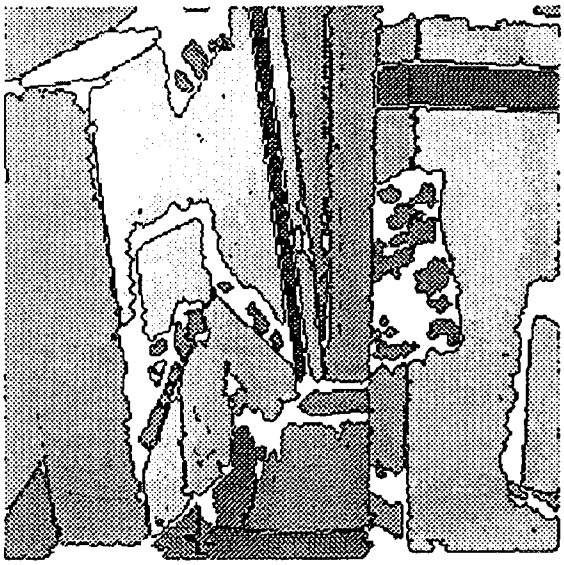

Figure 8 :Regions at the seventh step

V. CONCLUSION

We presented here a method to obtain a dense description of a spatio temporal event, which was our goal. We have a description of the sequence of images as time evolving regions. To avoid time consuming in extracting regions in each image and then matching them between images we "track" regions. It gave good results that will be used in the next step of our work: tracking and matching of contours.

REFERENCES

[1] P. Even and L. Marcé, *"PYRAMIDE: An interactive tool for modelling of teleoperation environments"* in IEEE Int. Workshop on Intelligent Robots and Systems, (Tokyo, Japan), pp. 725-730, '88 IROS, October 31 - November 2, 1988

[2] Y. LIU, T. S. Huang *"A linear Algorithm for Motion Estimation Using Straight Line Correspondences"* CVGIP 44, p. 35-37, 1988

[3] J. WENG, T. S. HUANG, N. AHUJA *"Motion and Structure from Line correspondences: Closed-Form solution, Uniqueness and optimisation"* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 14, no 3, March 1992

[4] Z. ZHANG *"Analyse du mouvement d'une séquence de scènes stéréoscopiques et applications à la robotique mobile"* Thèse de l'Université de Paris Sud, centre d'Orsay, 5 Octobre 1990

[5] N. AYACHE *"Construction et fusion de représentations visuelles 3D, applications à la robotique mobile"* Paris Sud, centre d'Orsay, Thesis 1988

[6] O. D. FAUGERAS, F. LUSTMAN, G. TOSCANI *"Motion and Structure from Point and Line Matches"* IEEE, Proc. 1st Int. Conf. Computer Vision, (London, England), pp. 25-34, June 1987

[7] B.K.P. HORN, E.J. WELDON JR *"Direct Methods for Recovering Motion"* International Journal of Computer Vision, 2, 51-76, 1988

[8] R. C. BOLLES, H. H. BAKER, D. H. MARIMONT *"Epipolar-Plane Image Analysis: An Approach to Determining Structure from Motion"* International Journal of Computer Vision, 1, 7-55, 1987

[9] D. MARR, T. POGGIO *"A computational theory of human stereo vision "* Proc. Royal Society London B.204, 301-328, 1979

[10] Y. SHIRAI *"Three-Dimensional Computer Vision"* Springer-Verlag 1987

[11] K. RANGARAJAN, M. SHAH *"Establishing Motion Correspondence"* CVGIP: IMAGE UNDERSTANDING, Vol. 54, No. 1, July, pp. 56-73,1991

[12] L. COHEN, L. VINET, A. GAGALOWICZ, P. T. SANDER *"Hierarchical Region Based Stereo Matching"* Proceedings of Computer Vision and Pattern Recognition, San Diego, 1989