MVA '92 IAPR Workshop on Machine Vision Applications Dec. 7-9,1992, Tokyo

REGION-BASED TOP-DOWN SEGMENTATION CONTROLLED BY STEREO MATCHING

Sabine Randriamasy - André Gagalowicz

INRIA-Rocquencourt - Domaine de Voluceau B.P. 105
78153 Le Chesnay Cedex - FRANCE - e-mail: randria@bora.inria.fr

ABSTRACT

This paper describes experimental algorithms developed to investigate the problems of region based *segmentation* and *matching* processes on black and white and color stereo pairs of images.
Our goal is twofolds: make these operations as automatic as possible and have them cooperate with each other.
A cooperative automatic top-down segmentation algorithm is proposed. The two segmentation trees are constructed *progressively*: regions are validated if a satisfying match is found, and then their division is stopped in the trees.
The algorithm is data driven and produces usable output for the matching and 3-D reconstruction processes even from the first division steps.

INTRODUCTION

The presented algorithms participate in a tasks sequence for stereo vision where a fundamental task to solve is the visual discrimination of objects and their location in space. To achieve this goal, the main steps are successively: segmentation of a stereo pair of images, stereo matching of the regions, extracted in the 2 images, and the 3-D reconstruction of the facets corresponding to the matched regions.
It is currently admitted that no method for deriving the "perfect" segmentation exists. The basic tenet of our approach is thus twofolds:

- The success of each job is closely connected to the results of the previous one, [RG91a].
- The evaluation and improvement of segmentation relies on the goal specification of the algorithm.

Therefore a way for us to evaluate and validate *automatically* the segmentation algorithm is to use its primary goal, which, in our case, is stereo matching.

TOP-DOWN SEGMENTATION

The segmentation algorithm uses an existing thresholding technique which relies on the notion of contrast rather than homogeneity. We have developed a region recursive multi-channel version.
On each intensity channel of each dividable region, an intensity threshold s^* is computed and must maximise a function

$$AVCT(s) = RC(s)/N(s)$$

which represents the average relative contrast observed on the edges detected by a threshold s [Koh81].

$$RC(s) = \sum_{for P_a, P_b adjacent} \min(I_b - s, s - I_a)$$

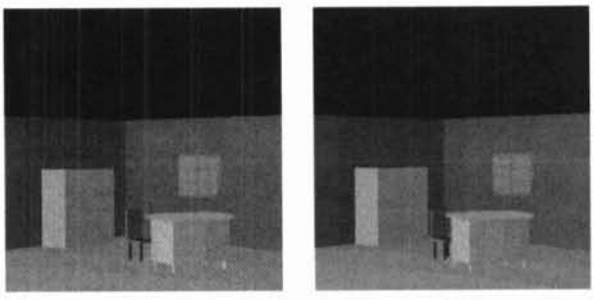
is the cumulated relative contrast on the detected edges and $N(s)$ is the number of edges detected by s.
The contrast-based approach produces accurate boundaries. The algorithm constructs contour-like regions, which are constructed around "solid" regions. Results were compared favorably with the contours produced by a Canny-like edge detector.
A noise cleaning operation is integrated into the segmentation process. It merges "tiny" regions into their spatial context and may lead to the merge of bigger ones. Both segmentation speed and similarity are improved.
Only two image scannings are necessary for each image division. The algorithm is data driven and produces usable output for the matching and 3-D reconstruction processes even at the first division steps.

SEGMENTATION RESULTS

Figure 2 shows segmentation results on an artificial colour office scene named Office1, see figure 1. Figure 4 shows results and on a natural grey level office scene named Office2, see figure 3.
The first division step thresholds one single region which is the whole image. Regions whose size is less than 4 pixels have been merged in their spatial context.

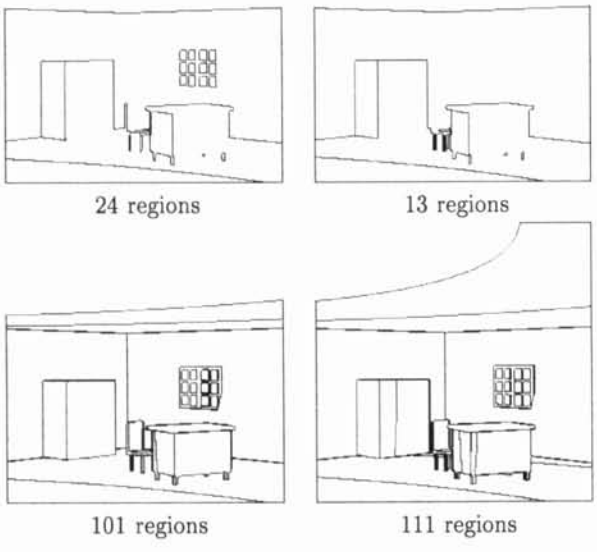
SEGMENTATION DRIVEN BY STEREO MATCHING

Another fundamental advantage of the top-down approach is that image divisions may be alternated with region matching, allowing a *direct* cooperation between these two processes.
For segmentation, we use a goal oriented local validation criterion which is: a left region and a right region are well segmented if they match and if this match is valid [Ran92].
Since there is generally no optimal segmentation level, we have chosed a hierarchical approach, where all the already produced segmentation levels are stored, and the



left right

Figure 1: *The original stereo pair of artificial colour images* Office1



24 regions 13 regions

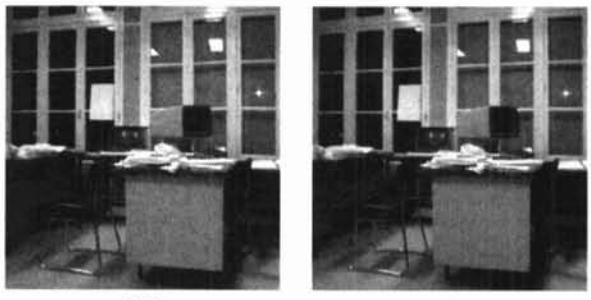
101 regions 111 regions

Figure 2: *2nd and 4th division steps on the pair* Office1

final regions of each image, are selected at different levels.
The matching process *drives* directly the segmentation in the following way:
matches are searched after each image division, and *across* the different already produced levels of segmentation, see figure 5.
Thus, segmentation is driven by its immediate goal, which becomes its main convergence criterion. The final segmentation is derived from the validated matches rather than the contrary.

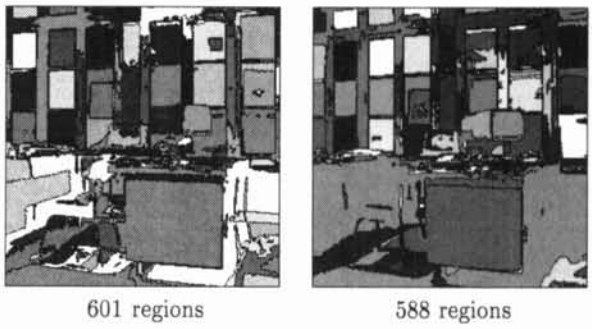
STEREO MATCHING WITH BACKTRACKING

The region matching technique uses an epipolar constraint and the attribute similarity of the regions to be matched. Color and boundary regularity are very important. The cooperative algorithm is top-down and automatic. The search for matches begins at the coarsest segmentation level. So, there is a risk for undersegmented regions to be similar enough to be matched, see



left right

Figure 3: *The original stereo pair of natural images* Office2



601 regions 588 regions

Figure 4: *4th division step on the pair* Office2

figure 6.
In order to avoid the selection of non coherent regions, correspondences found at a given segmentation level may be reconsidered during a fixed number Δ_m of later segmentation levels. The substitution of matches is performed according to the following rule:

- Let $p_N = (R_l^{kg}, R_r^{kr})$ be a region pair found at the iteration $N = \max(kl, kr)$ of the cooperative segmentation algorithm.
- If R_l^{kl} or R_r^{kr} has at least one descendent performing a better match

  Then p_N will be replaced by all its matched descendents in the match image.

In order to control the matching process, a $L1$ planar analytical approximation [Abd75] of the intensity function of the matched regions is computed. The approximation error is used to detect under-segmentations and helps to control the validation of the performed matches. The bactracking on the matches will only start if the following constraint $C_{APP}(R_l, R_r)$ is verified.

$$\begin{cases} N_{R_l}(p_e) & < & S(R_l) * (0.20) \\ N_{R_r}(p_e) & < & S(R_l) * (0.20) \end{cases}$$

where $N_{R_l}(p_e)$ and $N_{R_r}(p_e)$ are the number of non-isolated error pixels in each region R_l and R_r. Due to the results produced by the segmentation method, one bactracking level $\Delta_m = 1$ is mostly sufficient.

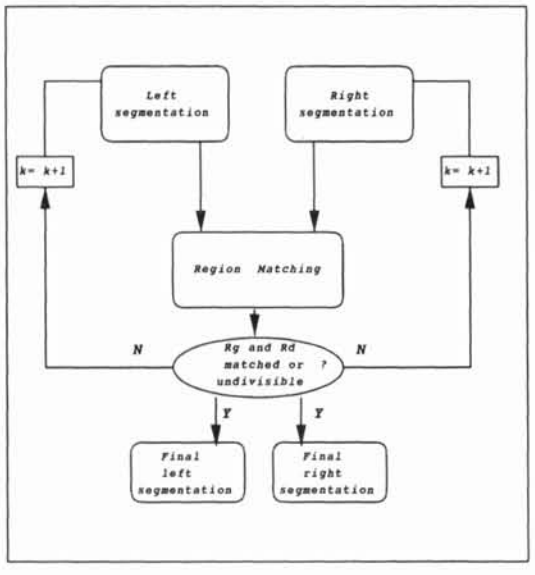


Figure 5: *Region division driven by stereo matching*

COOPERATIVE SEGMENTATION RESULTS

The above algorithms have been implemented in C language on a SPARC/station and tested on natural and artificial images of office scenes.

We first present, results on the artificial pair Office1, which illustrate the principle of the presented algorithms, on figure 6. On the matching images, the non-matched areas are in white. Whereas 6 divisions were necessary to make the single segmentation process converge, producing respectively 130 regions for the left image and 141 for the right one, the cooperative segmentation converged at the $4th$ division, and 23 regions where finally selected in each image.

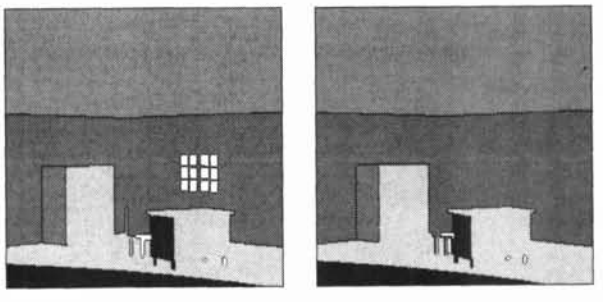
Note that without backtracking, the hierarchical matching algorithm would stop at the first division step, where most of the regions are obviously under-segmented.

Figure 7 shows the matches selected at step 5, *without* backtracking. Figure 8 shows results on the natural office scene Office2 with backtracking. Note the difference for the desk.

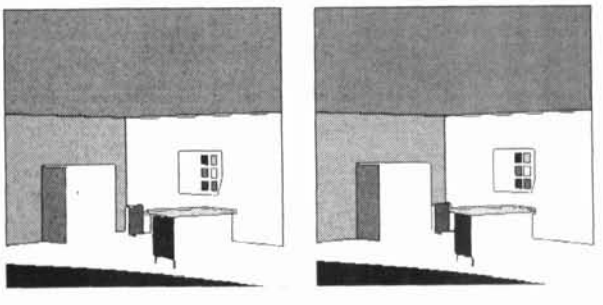
CONCLUSION

The cooperative top-down segmentation algorithm, produces directly and incrementally the $2-D$ information necessary for the application. This approach avoids both the computational complexity for the matching process and the need to construct and prune the whole segmentation trees.

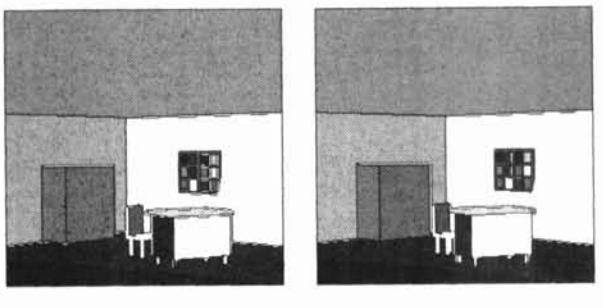
Segmentation is evaluated and driven by its immediate goal which is stereo matching. Other resulting images are available and directions for further study are outlined.



level 2: 6 created matches

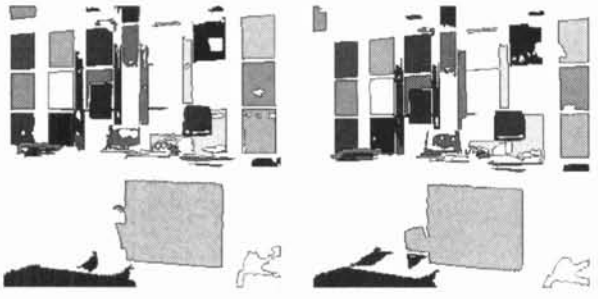


level 3: 14 matches: 2 validations, 2 substitutions, 10 created and substituted pairs, 1 cancelled match.



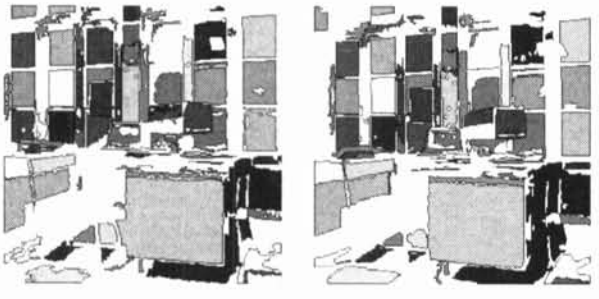
level 4: 23 matches: all the pairs have been validated, among which 9 are new or substituted.

Figure 6: *Stereo matching with one backtracking level on the pair* Office1



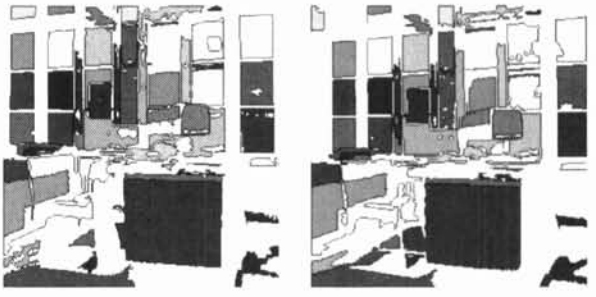
level 4: 38 matches

level 5: 58 matches



level 6: 73 matches

Figure 8: *Stereo matching with one backtracking level on the pair* Office2



level 5: 50 matches

Figure 7: *matches selected without backtracking at level 5 on the pair* Office2

References

[Abd75] Nabih N. Abdelmalek. An efficient method for the discrete linear l1 approximation problem. *Mathematics of Computation*, 29(131):844–850, july 1975.

[Koh81] Ralph.R. Kohler. A segmentation based on thresholding. *Computer Graphics and Image Processing*, 15:319–338, 1981.

[Ran92] Sabine Randriamasy. *Segmentation descendante coopérative en régions de paires d'images stéréoscopiques.* PhD thesis, Université Paris 9 - Dauphine, Septembre 1992.

[RG91a] S. Randriamasy and A. Gagalowicz. Region based stereo matching oriented image processing. In *IEEE conference on Computer Vision*