WATERSHED APPROACHES FOR COLOR IMAGE SEGMENTATION

Jocelyn Chanussot and Patrick Lambert

Laboratoire d'Automatique et de Micro-Informatique Industrielle
LAMII/CESALP - Université de Savoie - BP 806 - 74016 Annecy Cedex - FRANCE
(CNRS G720 ISIS : Information-Signal-Image-viSion)
tel : (33) 450 666 053 - Fax : (33) 450 666 020
e-mail : {Jocelyn.Chanussot ; Patrick.Lambert}@univ-savoie.fr

ABSTRACT

This paper presents and discusses different approaches to extend the watershed transform, which is the classical segmentation tool in gray-scale mathematical morphology, to the case of color or, more generally speaking, multi-component images. Different strategies are presented and a special attention is paid to the "bit mixing approach". This method bijectively mapps multi-dimensional data into a mono-dimensional space. It thus allows to use directly the usual scalar watershed algorithm on the encoded data, nevertheless taking explicitely into account their vectorial structure.

1. INTRODUCTION

The watershed transform is the traditional segmentation technique used in gray-scale mathematical morphology [1][2][3], and an abundant literature proposes several practical implementations of the algorithm. Intrinsically, the watershed is a gray-level dedicated (i.e. : *scalar*) transform. Using the watershed approach to segment multi-component (i.e. : *vectorial*) images is thus not straightforward. In section 2, the basic principle of the scalar transform is recalled and the chosen algorithm is presented. Section 3 makes a review of the different proposed approaches for multi-component watersheds. In section 4, the "bit mixing approach" proposed in [4] for multivalued morphology is briefly exposed, and its application to color watershed is presented in section 5. Section 6 concludes.

2. THE GRAY-SCALE WATERSHED TRANSFORM

The watershed transform is intuitively best understood using the classical topographic analogy[1] for the representation of gray-scale images. In this representation, the image is considered as a topographic relief, the numerical value of each pixel determining the corresponding point elevation. Falling on a given pixel, a drop of water will naturally run on that relief following the steepest slope, until it reaches a local *minimum*. The set of all pixels leading to the same local *minimum* is called a *catchment basin*. The set of all the different catchment basins constitutes a nonintersecting partition of the image, i.e. a *segmentation* of the image. In order to obtain perceptually meaningfull regions, this algorithm is usually applied on the gradient modulus image rather than directly on the original image. Catchment basins are thus growing up from local minimal gradient seeds (original image homogeneous regions), and are delimited by gradient local *maxima* (original image edges).

To implement this transform, the classical algorithm "by swamping" proposed by Vincent [3] has been used. The relief corresponding to the gradient modulus is progressively flooded with water springing from all local *minima*. During this flooding step, each time two flooded regions coming from different springs tend to merge, a virtual infinitely tall dam is raised up to keep them separate. At the end of the flooding, when each pixel has been reached by the water, the dams network constitutes the watershed lines of the image segmentation. This step constitutes the "automatic part" of the Beaucher automatic segmentation paradigm as defined in [2]. But, when applied to natural images, this algorithm leads to a severe oversegmentation that will be illustrated in the following. To avoid this phenomenon, two different approaches can be used :

- A post-processing step can be performed after the watershed transform. Usually based on the region adjacency graph of the segmentation, the goal is to progressively merge neighbouring regions using different criteria (radiometric similarity, relative areas of the regions...).

- The other solution consists in using a pre-processing[2]

---

[1] This representation implicitely assumes that the pixels are (or can be) totally ordered : an altitude (i.e. a scalar value) is associated to each pixel. Comparing the corresponding altitudes allows to order any set of pixels.

[2] Pre- and post-processings can be applied together : the post-processing can eventually be applied to further improve a segmentation that has already been strongly simplified by the choice of an appropriate set of markers.

before the watershed transform. Basically, the idea is to reduce the number of seeds in the flooding algorithm in order to reduce the number of regions in the final segmentation. This is the *marker approach*, the set of remaining seeds being called the *set of markers*. Still using terms from [2], the selection of the most appropriate set of markers constitutes the "clever part" of the segmentation algorithm. Most of the times, this step is highly goal oriented : the sought objects have to be marked exhaustively in order to be correctly segmented, and the discriminating criteria can widely vary depending on the addressed application. We chose to implement the automatic markers extraction method presented by Crespo in [5]. This extraction is composed of three main steps : firstly, the original image is simplified using an alternating sequential filter (ASF) by reconstruction [6]. Then, the gradient of this filtered image is computed and the null-gradient zones are considered as the region markers. Finally, an area opening is performed to suppress the markers that have a too small area, corresponding to small transition regions. This set of markers is used to modify the homotopy of the gradient computed on the original unfiltered image : the markers are imposed as its only *minima* (this is performed using a geodesic reconstruction of the set of markers in the initial gradient).

Figure 1 presents the corresponding generic processing synopsis. The "clever" part of the segmentation consists in the marker extraction and the homotopy modification of the original gradient, and the "automatic" part consists in applying the watershed algorithm to that modified gradient.

3. MULTI-COMPONENT WATERSHEDS

Multi-component image processing has become an important and very active field of research (color, or, more generally speaking, multi-spectral images, multi-date data and s.o.). To apply the watershed technique to these multivalued images, different approaches are possible[3].

It is possible to perform a single gray-level segmentation of the most significative component, such as the intensity in the Hue-Saturation-Intensity representation of color images. The choice of the processed component can be made using prior knowledge about the different available components. It also can be made adaptively with respect to the addressed application or to the processed image (different criteria, such as the dynamics or the standard mean deviation can be used to quantify the relative importance of the different components and to choose which is the most significative). Theoretically, this method may lead to inaccurate segmentation since two different regions may not be discriminated using one single component. Nevertheless, in the case of color images, processing the luminance alone leads most of the times to satisfactory results [8].

It is also possible to perform the watershed transform of each component separately and to aggregate the results afterwards to get the segmented color image. This can be done for instance by taking the union of all the watershed lines obtained on the different components (this assumes that common edges are located at the very same position, leading otherwise to an oversegmentation). More elaborate fusions immediately become difficult to implement and lead to time consuming algorithms.

In both cases, the watershed is applied to scalar images and it thus does not require any modification of the algorithm in itself.

Another approach consists in first computing the norm (for instance the sum or the *max* of the different gradient components) of a vector gradient and in performing the watershed on the so obtained gradient image [5]. Shafarenko [9] proposes to use the LUV color space and to calculate the *max* of the chrominance gradient components in order to exclude potential changes of illumination from consideration.

In [10], Crespo proposes an extension of his "flat zone approach" to the case of color images by forcing the inclusion between the flat zones (i.e. piecewise-constant regions) of each component and the regions of the final segmentation.

Meyer [11] proposed a seeded region growing method that directly segments the input color image and that can also be considered as a watershed transform. Some authors have also proposed different clustering methods of the color histogram based on the watershed algorithm to perform multispectral image partitioning [12][13]. Finally, for examples of post-processings based on the region adjacency graph applied after a color watershed transform to reduce the oversegmentation, see for instance [8][14]. In the first case, the watershed is computed on the luminance gradient image, and in the second case, a fusion is performed either on the gradient components before the watershed or on the segmented components after three scalar watersheds.

4. THE BIT MIXING APPROACH

In this paper, we propose to extend the watershed segmentation tool to the multi-dimensional case by using the "bit mixing" approach, that is described and justified in [4]. In this frame, the color image is processed directly as a whole (i.e. in a *vectorial* way), but only simple mono-dimensional processings are involved. The method is thus very fast and easy to implement.

We now briefly recall the principle of this approach. Firstly, the image is encoded in a scalar form, without any loss of information. Although the proposed method can

---

[3] A quite complete presentation and classification of the different approaches to color watershed is proposed by Saarinen in [7].

be extended to any multi-dimensional data set in a very straightforward way, we restrict this paper to color images represented in the Red-Green-Blue base, where each component is coded with 8 bits. To encode one pixel, the binary representation of each component is used : the 3*8 available bits are mixed up together, taking alternatively one bit from the red, the green and the blue component. For each pixel we build the corresponding 24 bits long integer and, thanks to this mapping, we get a mono-dimensional "gray level-like" representation of the color image, with a range going from 0 to (2^{24} - 1). More details and corresponding equations are given in [4]. In particular, it is shown that the definition of such a bijective mapping is both sufficient and necessary to generate a total ordering relation on the input vector data (to compare two vectors, one just has to compare their scalar transformed values). It is also shown that any vectorial total ordering relation can be represented by a space filling curve (i.e. a monodimensional curve that goes one single time through every point of the input data space). Comparing vectors comes to compare the corresponding curvilinear abscissas along the curve. Furthermore, the study of the geometric properties of the curve (anisotropy, fractal structure...) helps to characterize the chosen ordering relation[4].

After this coding step, it is possible to process the equivalent scalar representation of the color image as a simple mono-dimensional picture. Figure 2 presents the corresponding processing synopsis.

This method turned out to be well suited to extend processings involving compositions of rank order filters, such as the morphological filters [4][17][5]. But, let us recall Netto's theorem, quoted by Sagan [15]. This theorem shows that any bijective mapping between two spaces with different dimensions is necessarily discontinuous and, as a consequence, necessarily non linear. This results in a space topology distortion : the neighbourhood notion is not preserved. Equivalently, that implies that such a bijective transform will not commute with linear data combinations. This theoretically affects the computation of the morphological gradient which is calculated as the difference (i.e. a linear combination) of a dilation and an erosion.

For the marker selection, this problem is of no consequence : the use of the ASF to simplify the image is valid (since the bijective scalar coding induces a total ordering, our approach even prevent the potential appearence of new vectors during that step, and thus prevents the appearence of new spurious discontinuities in the image). Furthermore, the null gradient regions are not modified by the coding (they actually correspond to the flat zones defined by Crespo [10]).

But, the use of the gradient itself during the flooding step may be more problematic. Since we process the color image as a mono-dimensional signal, the output of the morphological gradient is mono-dimensional too, and, as a consequence, the topographic analogy is directly respected without any more data aggregation. But, due to the unavoidable distortion of the neighboorhouds, this gradient may not fit with the psychovisual intuition (quite similar encoded data can correspond to very different input data, and the dual problem can also appear). Neverthless, our approach tends to statistically reduce these unavoidable problems [4].

5. APPLICATION TO COLOR WATERSHED

Figure 3 presents the segmentation results obtained on a color image. Figure 3-a- is the luminance of the original image (*still life on pedestal table*, P. Picasso, oil on canvas, Picasso Museum, Paris - size : 248*188).

Figure 3-b- and -c- present the segmentation results obtained with the proposed approach (scalar coding of the input data followed by the scalar processing of the encoded data) respectively without the marker approach (the dramatic oversegmentation is clearly visible) and with the marker approach (using an ASF by reconstruction of size 9 and an area filtering of size 20). The segmentation is greatly simplified and the main objects are correctly segmented. Some small or thin regions are missing.

Figure 3-d- corresponds to the segmentation result obtained by processing the luminance component alone (with the same parameters values).

Figure 3-e- corresponds to the segmentation result obtained with a hybrid method. The bit mixing approach is only used when it is purely licit, i.e. for the image simplification with the ASF. But, for the gradient, the sum of the marginal gradients computed on each component separately has been used to avoid the topological distortion.

The results obtained with these three methods are qualitatively very similar.

6. CONCLUSION

In this paper, we intended to validate the approach of color image segmentation *via* bijective scalar coding of the vector input data. We did not want to find the most effective watershed algorithm, which is, most of the times, application dependant. Compared with other methods, our approach has not significantly improved the segmentation of color images. Nevertheless, since only one single scalar processing is involved, its main advantage lies in its simplicity and rapidity of implementation. Furthermore, due to

---

[4] For another application of the equivalence existing between any total vectorial ordering relation and its corresponding representation as a space filling curve, see also [16] where a vector-median like filter is proposed using the same principle.

[5] For further discussion on definition and application of morphological filters to multivalued data in the general case, see [4][17][18][19][20][21][22].

the bijectivity of the used transform, the method avoids to unwillingly merge dissimilar regions. For the same reason, and because of the induced topological distortion, detected edges position may be degraded. But, none of these phenomema appeared clearly during the experimental study on real images.

7. REFERENCES

[1] F. Meyer and S. Beucher. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1):21–46, 1990.

[2] S. Beucher and F. Meyer. *Mathematical morphology in image processing*, chapter The morphological approach to segmentation : the watershed transformation, pages 433–481. E. Dougherty, marcel dekker, new york edition, 1998.

[3] L. Vincent and P. Soille. Watersehds in digital spaces : an efficient algorithm based on immersion simulations. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(6):583–598, June 1991.

[4] J. Chanussot and P. Lambert. Total ordering based on space filling curves for multivalued morphology. In *4th International Symposium on Mathematical Morphology and its Applications (ISMM'98)*, pages 51–58, June 1998. Amsterdam, The Netherlands.

[5] J. Crespo. *Morphological connected filters and intraregion smoothing for image segmentation*. PhD thesis, Georgia Institute of Technology, 1993.

[6] J. Crespo, J. Serra, and R. Schafer. Theoretical aspects of morphological filters by reconstruction. *Signal Processing*, 47:201–225, 1995.

[7] K. Saarinen. Watershed in color image segmentation. In *IEEE International Workshop on Non Linear Signal and Image Processing*, pages 14–17, June 1995. Neos Marmaras Haldikiki , Greece.

[8] S. Ji and H.W. Park. Image segmentation of color image based on region coherency. In *IEEE International Conference on Image Processing*, 1998. Chicago, Illinois.

[9] L. Shafarenko, M. Petrou, and J. Kittler. Automatic watershed segmentation of randomly textured color images. *IEEE Trans. on Image Processing*, 6(11):1530–1544, November 1997.

[10] J. Crespo and R. Schafer. The flat zone approach and color images. In *2nd International Symposium on Mathematical Morphology and its Applications (ISMM'94)*, pages 85–92, September 1994. Fontainebleau, France.

[11] F. Meyer. Color image segmentation. In *IEE International Conference on Image Processing and its Applications*, pages 303–306, May 1995. Maastricht, The Netherlands.

[12] P. Soille. Morphological partitioning of multispectral images. *Journal of Electronic Imaging*, 5(3):252–265, July 1996.

[13] L. Shafarenko, M. Petrou, and J. Kittler. Histogram based segmentation in a perceptually uniform color space. *IEEE Trans. on Image Processing*, 7(9):1354–1358, September 1998.

[14] K. Saarinen. Color image segmentation by a watershed algorithm and region adjacency graph processing. In *IEEE International Conference on Image Processin*, volume 3, pages 1021–1025, 1994. Austin, Texas.

[15] H. Sagan. *Space filling curves*. Springer Verlag New York, 1994.

[16] K. Plataniotis, C. Regazzoni, A. Teschioni, and A. Venetsanopoulos. A new distance measure for vectorial rank order filters based on space filling curves. In *IEEE International Conference on Image Processing*, volume 1, pages 411–414, 1996. Lausanne, Switzerland.

[17] J. Chanussot. *Approches vectorielles ou marginales pour le traitement d'images multi-composantes*. PhD thesis, Universite de Savoie France, 1998.

[18] H. Talbot, C. Evans, and R. Jones. Complete ordering and multivariate mathematical morphology. In *4th International Symposium on Mathematical Morphology and its Applications (ISMM'98)*, pages 27–34, June 1998. Amsterdam, The Netherlands.

[19] J. Goutsias, H. Heijmans, and K. Sivakumar. Morphological operators for image sequences. *Computer Vision and Image Understanding*, 62(3):326–346, 1995.

[20] J. Serra. *Mathematical morphology in image processing*, chapter Anamorphoses and function lattices, pages 483–523. E. Dougherty, marcel dekker, new york edition, 1993.

[21] M.L. Comer and E.J. Delp. An empirical study of morphological operators in color image enhancement. In *SPIE Conference on Image Processing Algorithms and Techniques III*, volume 1657, pages 314–325, February 1992. San Jose, California.

[22] C. Gu. *Multivalued morphology and segmentation based coding*. PhD thesis, Ecole Polytechnique Federale de Lausanne (EPFL), 1995.

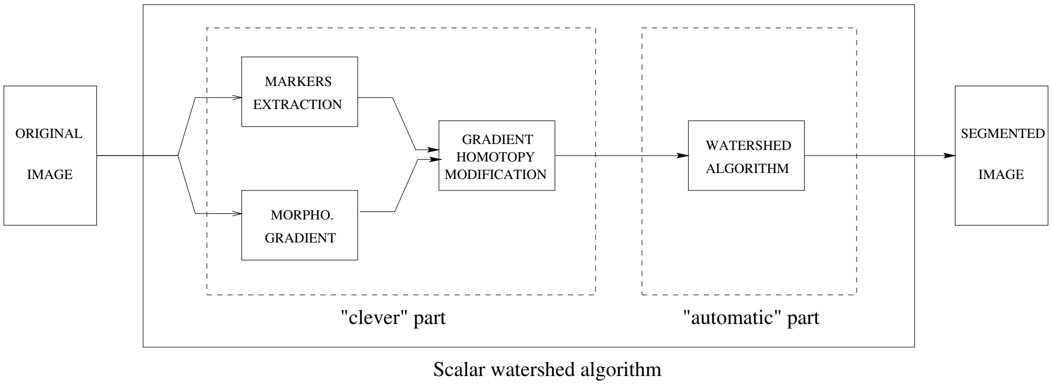

Figure 1: Scalar watershed algorithm synopsis

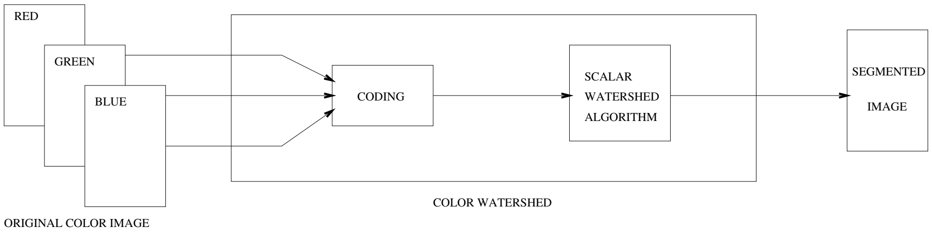

Figure 2: Color watershed algorithm synopsis

Figure 3: Results