

A Basin Morphology Approach to Colour Image Segmentation by Region Merging

Erchan Aptoula and Sébastien Lefèvre

UMR-7005 CNRS-Louis Pasteur University
LSIIT, Pôle API, Bvd Brant, PO Box 10413, 67412 Illkirch Cedex, France
{aptoula,lefevre}@lsiit.u-strasbg.fr

Abstract. The problem of colour image segmentation is investigated in the context of mathematical morphology. Morphological operators are extended to colour images by means of a lexicographical ordering in a polar colour space, which are then employed in the preprocessing stage. The actual segmentation is based on the use of the watershed transformation, followed by region merging, with the procedure being formalized as a *basin morphology*, where regions are “eroded” in order to form greater catchment basins. The result is a fully automated processing chain, with multiple levels of parametrisation and flexibility, the application of which is illustrated by means of the Berkeley segmentation dataset.

1 Introduction

Automatic, robust and efficient colour image segmentation is nowadays more indispensable than ever, since numerous image depositories have been formed and continue to grow with an increasing speed. As far as the human vision system is concerned, edge information is primarily contained within the luminance component. Hence colour is regarded as an invaluable, yet auxiliary component when it comes to image segmentation and generally object recognition. The problem of its efficient exploitation in this context remains to be resolved, because not only the principles of human colour vision have not yet been fully understood, but because it also introduces additional parameters in the already elusive problem of general purpose image segmentation.

Specifically, one of the major questions is the representation of colour vectors and the choice of the associated colour space. Since the desired segmentation outcome is almost always based on the human interpretation of objects, it is deemed natural to attempt to emulate the sensitivities of human colour vision. That is besides the reason why polar colour spaces have been gaining popularity in this regard. However as it will be elaborated in section 2, these spaces also suffer from considerable drawbacks.

Among the approaches developed to resolve the problem of colour segmentation, mathematical morphology offers a different perspective from the mostly statistical and clustering based methods, since it is an algebraic image processing framework capable of exploiting not only the spectral, but spatial relationships

of pixels as well. In this paper, we present a fully automated colour segmentation procedure, designed for polar colour spaces, based on morphological operators. In particular, the proposed approach consists primarily of manipulating the catchment basins resulting from a watershed transformation, through their interpretation as the new processing units for morphological operators. Hence leading to a “*morphology of basins*”. A hierarchy of attributes is thus organised, making it possible to merge these regions based on arbitrary characteristics such as mean colour and texture. The resulting method is tested using the Berkeley segmentation dataset [1].

The rest of the paper is organised as follows. Section 2 discusses initially the crucial choice of polar colour space. Then in section 3 the proposed segmentation approach is elaborated and its individual stages are detailed. Finally section 4 is devoted to concluding remarks.

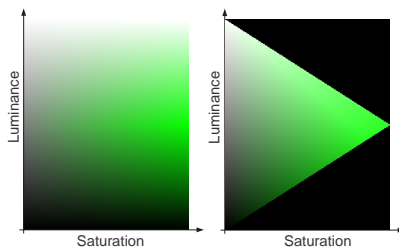


Fig. 1. Vertical semi-slice of the cylindrical HLS (left) and bi-conic IHLS (right) colour spaces.

2 Choice of Colour Space

For the reasons mentioned in the previous section, here we concentrate on 3d-polar colour spaces, that have appeared as the result of attempts to describe the RGB cube in a more intuitive manner, from the point of view of human interpretation of colour, in terms of *luminance*, *saturation* and *hue*. While luminance $L \in [0, 1]$ accounts for the amount of light, saturation $S \in [0, 1]$ represents the purity of a colour. The values of the periodical hue interval $H \in [0, 2\pi[$ on the other hand, denote the dominant wavelength, with 0 corresponding to red.

Basically, polar colour spaces achieve this transformation by representing colours with respect to the achromatic axis of RGB. Nevertheless, several implementational variants are available for this single transformation, e. g. HSV, HSB, HLS, HSI, etc [2]. According to Hanbury and Serra [3], all of the aforementioned colour spaces were developed primarily for easy numerical colour specification, while they are ill-suited for image analysis. Specifically, although they were initially designed as conic or bi-conic shaped spaces, later on their cylindrical versions were employed in practice, in order to avoid the computationally expensive

(for that period) checking for valid colour coordinates. The passage from conic to cylindrical shape however resulted in many inconsistencies within these spaces, for instance by allowing fully saturated colours to be defined in zero luminance. Extensive details on this topic can be found in [3].

Here we adopt the suggestion made in [3], and make our colour space choice in favour of the *improved HLS* space (IHLS), which employs the original biconic version of HLS, hence limiting the maximal allowed value for saturation in relation to luminance (figure 1). Further advantages of IHLS with respect to its counterparts include the independence of saturation from luminance, thus permitting the use of any luminance expression (e. g. RGB average, perceptual luminance, etc) and the comparability of saturation values.

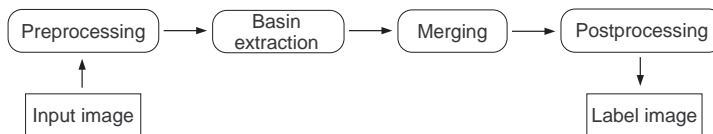


Fig. 2. Summary of the proposed processing chain.

3 Proposed Approach

In an ideal world all images would have the same resolution, colour number and overall complexity. Unfortunately it is not the case. The problem of segmentation in its most general form is highly difficult to resolve, as it aims to detect the semantic regions of extremely heterogeneous input. Moreover semantic level segmentation requires naturally some a priori information of semantic nature, hence rendering it feasible only for domain specific applications, since no ontology incorporating all types of objects exists. Consequently a more “practical and realistic” aim, also adopted here, is to attempt to detect the *principal regions* of images with respect to homogeneity, which is a task most important for content based image retrieval.

Considering the vast heterogeneity of image data, an equally high degree of adaptability is crucial, which will take into account the different types of border information contained within an image, e. g. spectral, textural, etc. To this end, we propose the processing chain summarised in figure 2. Briefly, the input image is first simplified using border preserving morphological operators and then through the combination of a colour gradient and the watershed transformation, the catchment basins are obtained. Next, an iteratively applied hierarchical fusion is carried out, providing a rough approximation of the sought borders, that are finally refined in the last stage by means of a marker based watershed transformation. Details on each step follow.

3.1 Preprocessing

This first step aims to simplify the input image, and eliminate any “excessive” details. The morphological toolbox offers a rich variety of operators for this purpose, however, several issues arise. The first concerns the extension of grayscale morphological operators to colour images, a theoretical problem stemming from the need to impose a complete lattice structure on the pixel intensity range, which in the case of multivalued images, is equivalent to the need to order vectorial data [4, 5]. Several approaches have been developed to this end, a survey of which may be found in [6]. Here, it has been chosen to order the colour vectors of the IHLS space by means of a lexicographical ordering:

$$(h_1, s_1, l_1) < (h_2, s_2, l_2) \Leftrightarrow \begin{cases} l_1 < l_2, \text{ or} \\ l_1 = l_2, s_1 < s_2 \text{ or} \\ l_1 = l_2, s_1 = s_2 \text{ and } h_1 < h_2 \end{cases} \quad (1)$$

where $l_1, l_2, s_1, s_2 \in [0, 1]$ and $h_1, h_2 \in [0, 2\pi]$. As the hue component is a circular value, an angular distance from a reference hue h_0 [7] is employed for their comparison:

$$h \div h_0 = \begin{cases} |h - h_0| & \text{if } |h - h_0| < \pi \\ 1 - |h - h_0| & \text{if } |h - h_0| \geq \pi \end{cases} \quad (2)$$

which for the sake of simplicity is set as $h_0 = 0.0$. The hue values are then ordered according to their distances from h_0 :

$$\forall h, h' \in [0, 2\pi], \quad h < h' \Leftrightarrow h' \div h_0 < h \div h_0 \quad (3)$$

where hues closer to h_0 are considered greater. Hence, with the luminance components compared first, this ordering leads to operators that process particularly this channel, containing the majority of the total variational information.

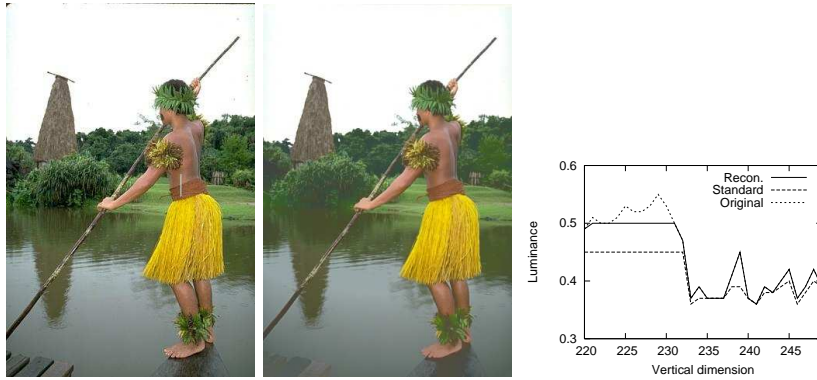


Fig. 3. From left to right, the original image (#101087), its preprocessed form and the intensity transition of the white line in the original image, for a reconstructive and standard processing based leveling.

Equipped with this ordering, erosion (ε), dilation (δ) and all the deriving grayscale morphological operators may be extended to colour data. Nevertheless, a second issue in this regard is the need for border preserving operators. That is why, it was chosen to employ a morphological leveling $\Lambda(f, m)$ [8], which provides a simplified version of the input image f , by applying iterative geodesic erosions and dilations to the marker m , a rough approximation of the input, until idempotence, i. e., $\Lambda(f, m)^i = \sup \{ \inf [f, \delta^i(m)], \varepsilon^i(m) \}$, until $\Lambda(f, m)^{i+1} = \Lambda(f, m)^i$. The marker image is obtained by means of a reconstruction based opening followed by a reconstruction based closing. The result is a “leveled” image, of which the details smaller than the structuring element’s (SE) size have been removed while also preserving perfectly all the region borders (figure 3). The size of the SE, typically a square of 7×7 pixels is determined with respect to the dimensions of the input image.

3.2 Basin extraction

Having simplified the input, this step consists in computing a first segmentation map of the image using the watershed transformation. As this powerful operator can be applied only to a scalar input representing the topographic relief of the image, it has been chosen to combine the colour channels by means of a channel wise maximum of marginal gradients:

$$\rho_{HLS}(h, s, l) = \max \{ \rho(l), \rho(s), \rho_H(h) \} \quad (4)$$

where $\rho = f - \varepsilon(f)$ is the standard internal morphological gradient. Although the components of the polar colour spaces are highly intuitive, their combination is relatively problematic. In particular, hue is of no importance if saturation is “low”, while the bi-conic shape of the colour space assures that no high saturation levels exist, if luminance is not “high enough”. Hence the hue gradient needs to be weighted with a coefficient that has a strong output only when both compared saturation values are “sufficiently high”:

$$\rho_H(h) = \max_{i \in B} \{ j(s, s_i) \times h \div h_i \} - \min_{i \in B} \{ j(s, s_i) \times h \div h_i \} \quad (5)$$

where B is the local 8-neighborhood and $j(\cdot, \cdot)$ a double sigmoid controlling the transition from “low” to “high” saturation levels:

$$j(s_1, s_2) = \frac{1}{(1 + \exp(\alpha \times (s_1 - \beta))) \times (1 + \exp(\alpha \times (s_2 - \beta)))} \quad (6)$$

where $\alpha = -10$ and the offset $\beta = \mu_S$ is set as the mean saturation of the image, hence making it possible to adapt the gradient’s sensitivity to colour, according to the image’s overall colourfulness level. The application of the watershed transformation on the newly computed gradient leads to the result depicted in figure 4.

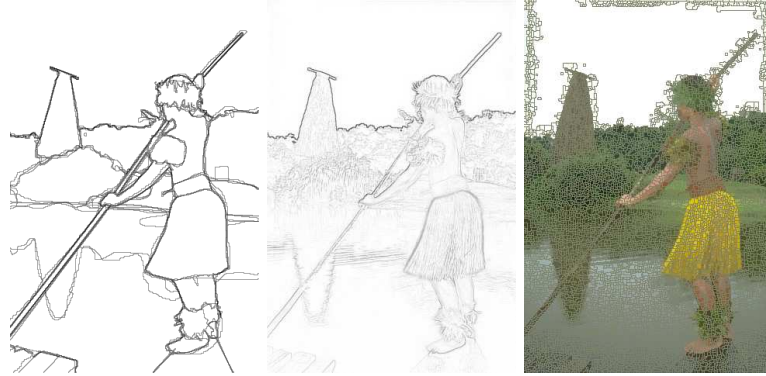


Fig. 4. From left to right, the hand reference segmentation, the proposed colour gradient and its oversegmented watershed transformation result, superposed on the original image.

3.3 Merging

Given the sensitivity of the internal gradient, the oversegmented result has been expected in the previous step. Considering that the sought borders are contained within this complex of adjacency relations, from this point on all efforts are concentrated on eliminating the unwanted borders, and thus increasing the sizes of the catchment basins. Merging the mosaic of basins obtained by watershed transformation is a well known technique in automated image segmentation [9, 10]. Here we follow a graph based formalisation for this procedure.

As each basin represents a locally homogeneous region, despite the level of oversegmentation, by providing spectrally atomic regions, the entire watershed procedure greatly reduces the volume of clustering to be carried out in the later stages. At this point, based on the atomicity of each basin, one can proceed by manipulating the image content with the catchment basins being the new processing “image units”, instead of pixels. Hence the image can be viewed as an undirected graph of basins, where each node is characterised by a set of spectral and other properties (e. g. mean colour, variance, etc) as well as its set of adjacent basins, or *neighbours*.

With this point of view, the merging procedure, can be defined as an operator on this graph, which propagates labels, and modifies adjacency relations. Furthermore, by imposing a complete lattice structure on the “value interval” of basins, one can define morphological operators, hence leading to a *basin morphology*. In particular, by formulating the merging of basins, as the replacement of each node, by its closest with respect to a certain metric, the operator becomes intuitively similar to an erosion. While a dilation would dually replace each node with its most distant neighbour. However this option is of no interest on its own. Consequently, the basin erosion (ε_b) and dilation (δ_b) of a graph $G = (V, E)$ can

be defined respectively as:

$$\varepsilon_b(G) = \left\{ G \mid \forall V_i \in V, \text{label}(V_i) = \text{label}\left(\underset{V_j \in N(V_i)}{\text{argmin}} d(V_j, V_i)\right) \right\} \quad (7)$$

$$\delta_b(G) = \left\{ G \mid \forall V_i \in V, \text{label}(V_i) = \text{label}\left(\underset{V_j \in N(V_i)}{\text{argmax}} d(V_j, V_i)\right) \right\} \quad (8)$$

where $N(V_i)$ is the set containing the neighbours of V_i . Several operators may be derived from the combinations of these two, their efficiency however in this context is strongly related to the metric of similarity in use ($d(\cdot, \cdot)$). Consequently, one can implement a rich variety of merging strategies, where each is based on a different basin similarity metric, exploiting some of their properties.

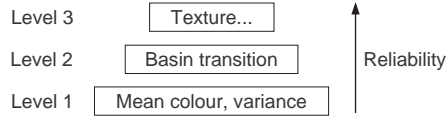


Fig. 5. Hierarchy of merging criteria.

We propose a hierarchical approach in this regard, as illustrated in figures 5 and 6, which consists in employing various properties of the basins in different scales, in order to compute their distances and hence realise their erosions (i. e. mergings) by means of equation (7). Specifically, it begins with a series of thresholded erosions based on their mean colour, where only the basins that are closer than a predefined limit are taken into account. This first step aims to merge only spectrally similar basins. The colour distance in use is:

$$\begin{aligned} \forall c_1 = (h_1, s_1, l_1), c_2 = (h_2, s_2, l_2), \\ d(c_1, c_2) = j(s_1, s_2) \times h_1 \div h_2 + (1 - j(s_1, s_2)) \times |l_1 - l_2| \end{aligned} \quad (9)$$

By means of factor $j(\cdot, \cdot)$, a saturation based continuous transition of priority is realised between hue and luminance. This low level step is carried out iteratively with thresholds starting from $t_0 = 0.01$ and increasing until t_{max} , which doubles the initial intra-basin variances. This process results in a preliminary segmentation map, where relatively homogeneous regions appear. Next, we modify the distance metric so as to eliminate intensity gradients, and apply it using the same threshold. For this purpose, the erosions are computed by taking into account only the bordering pixels of basins. Whereas at the third step, higher level merging criteria are employed. Specifically, in order to calculate the textural similarity of basins, their mean covariance vector is used:

$$K(f) = \text{Vol}(\varepsilon_{P_{2,v}}(f)) / \text{Vol}(f) \quad (10)$$

where $P_{2,v}$ is a pair of points separated by a vector v and Vol the volume, i. e. sum of pixel values. Of course one is by no means limited with these criteria; for instance border geometry may be further exploited. The threshold in this case is fixed as the mean covariance vector of the entire image.



Fig. 6. From left to right, the three levels of merging using the principle illustrated in figure 5.

3.4 Post-processing

Once this stage is reached, the principal regions of the input are expected to have been formed. As a last touch one can eliminate all regions inferior to a certain surface, by merging them with their closest neighbor. A more serious problem however, concerns the possibility of local deviations from the sought borders, since according to the definition of the proposed erosion operator, basin processing has been carried out so far using only the immediate neighborhood of each basin.

To counter this phenomenon, one can employ for instance multiple scales by modifying the size of the processed neighborhood, or in other words the shape and extent of the SE. Another possibility is to profit from the topological properties of the marker based watershed transform, which by limiting the flooding sources, provides an absolute control over the number of regions that are formed. As to the markers, one can very simply erode the binary region map, while preserving their connectivity. Thus, flexibility areas are formed among them which make it possible to realise topological border corrections (figure 7).

4 Discussion and conclusion

In this paper, an unsupervised and input specific colour image segmentation method has been presented. It has been developed for the improved HLS space,



Fig. 7. From left to right, the segmentation result using the jump connection algorithm [11], the marker image and the final marker based watershed result.

and constitutes an attempt to integrate the spatial sensitivities of morphological operators with spectral image properties. Furthermore, a graph based morphology approach has been formulated in order to manipulate the catchment basins produced by the watershed transform. This formulation aims mainly to provide a more efficient and flexible exploitation framework for the wealth of topological information provided by the aforementioned transform.

Pertinent results have been obtained with the Berkeley dataset (figures 8), even by using the basic erosion definition in combination with a hierarchy of multiple merging criteria, ranging from mean colour to covariance based texture features. While more sophisticated operators (e.g. geodesic reconstruction of basins, etc), as well as the exploration of further basin metrics remain to be investigated. Its execution speed and adaptivity, along with its capacity to provide the “main” borders of its input, render this approach suitable for applications where precision is of secondary importance, and a fast and robust segmentation is prioritised (e.g. content based image retrieval).

Issues that require further attention include improving the estimation of arguments as well as the use of additional high level merging criteria, such as border geometry.

References

1. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th International Conference on Computer Vision. Volume 2., Vancouver, Canada (2001) 416–423
2. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Second edn. Addison-Wesley, New York (1992)
3. Hanbury, A., Serra, J.: Colour image analysis in 3d-polar coordinates. In: International Conference on Image Processing and its Applications, Magdeburg, Germany (2003)

4. Serra, J.: Image Analysis and Mathematical Morphology Vol I. Academic Press, London (1982)
5. Ronse, C.: Why mathematical morphology needs complete lattices. *Signal Processing* **21**(2) (1990) 129–154
6. Aptoula, E., Lefèvre, S.: A comparative study on multivariate mathematical morphology. *Pattern Recognition* (2007) (doi:10.1016/j.patcog.2007.02.004).
7. Hanbury, A., Serra, J.: Morphological operators on the unit circle. *IEEE Transactions on Image Processing* **10**(12) (2001) 1842–1850
8. Gomila, C., Meyer, F.: Levelings in vector spaces. In: *Proceedings of the IEEE Conference on Image Processing, Kobe, Japan* (1999)
9. Chen, Q., Zhou, C., Luo, J., Ming, D.: Fast segmentation of high-resolution satellite images using watershed transform combined with an efficient region merging approach. In: *Proceedings of the IWCIA, Auckland, New Zealand* (2004) 621–630
10. Garrido, L., Salembier, P., Garcia, D.: Extensive operators in partition lattices for image sequence analysis. *Signal Processing* **66**(2) (1998) 157–180
11. Angulo, J., Serra, J.: Modelling and segmentation of colour images in polar representations. *Image and Vision Computing* (2006) (doi:10.1016/j.imavis.2006.07.018).

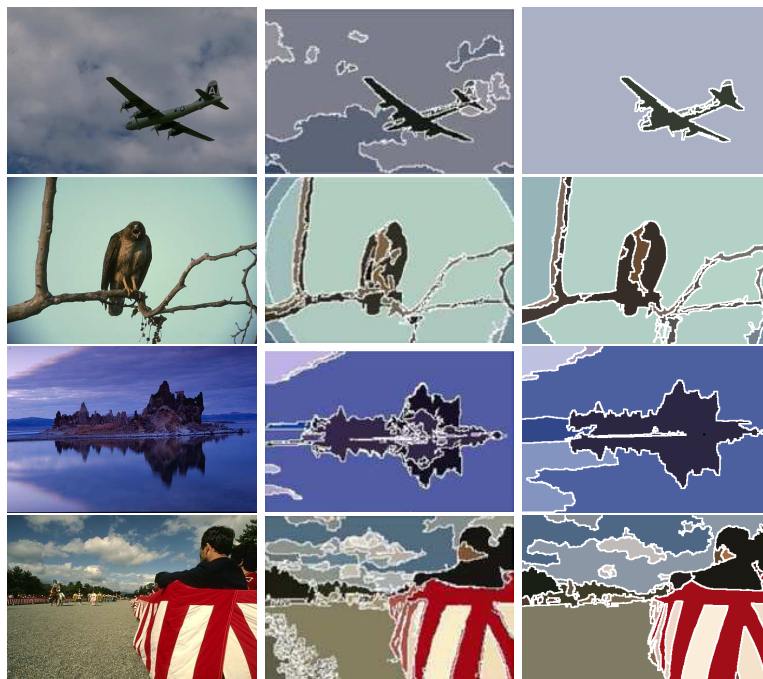


Fig. 8. From left to right, the original images, their segmentations based on jump connection [11] and based on the proposed approach (top to bottom: #3096, #42049, #143090 and #145086).