Non Supervised Goodness Measure for Image Segmentation

Eduard Vazquez and Ramon Baldrich

Computer Vision Centre, Universitat Autònoma de Barcelona, Spain

ABSTRACT: In this article we propose a new measure of goodness to evaluate a segmentation in a fully-non-supervised way. If the proper function of goodness can be found, then is possible to automatically tune the parameters of any segmentation method, solving one of the most challenging problems in the segmentation field. In this article, the goodness is determined by using two different methods which compute features such as color distinctiveness and chromatic contrast based on Shannon's Information Theory. The main idea, present in both methods, is that a good segment (that is a region of a segmented image) have a high color distinctiveness and contrast. This article analyses this two methods and compare its performance using a benchmark. Results obtained shows that both methods are useful to compute the goodness of a segmentation.

1 INTRODUCTION: In Computer Vision, there are a great amount of algorithms which requires a good segmentation as a preprocessing step. Nonetheless, existing segmentation methods require to adapt its parameters to each problem, and even each image, to yield good results. Therefore, general purpose segmentation, understood as a process without human supervision, is still a challenging problem with no solution. This paper addresses this problem presenting a combination of features which are aimed to give a measure of correctness of a segmentation when no benchmark exist. Consequently, yielding a segmentation without human supervision. The aim of this article is not to propose a solution for this problem, but to analyze the behavior of some existing methods.

One of the earliest works to decide the goodness of a segmentation is the JSEG segmentation method introduced in [2]. JSEG is a two-step segmentation schema. First, a clustering of the color space is performed. Afterwards, a criterion of *good* segmentation is applied using the spatial coherence of the image, *i.e.*, the information of the spatial relation existing between the pixels in the image space. Another schema proposes that a good segmentation region should be formed by strongly connected pixels with homogeneous colors [8]. This second approach follows a similar idea of the of the work introduced by Heidemann in [12], which uses the color distinctiveness as a measure of goodness. Other measures to describe the correctness of a segment are the homogram proposed in [4], a calculus based in the Bhattacharyya distance [11], a probabilistic approach as explained in [9] or a graph-based method as explained in [13]. The methods studied in this article belong to those family methods that use contrast as a criteria of the goodness of a segmentation (e.g. [7][12]).

In this article, we analyze the performance of two methods. First, the one introduced in [12]. Such method proposes a goodness function for color segmentation, which allows to predict whether the segmented regions will be stable against noise, variation of lighting, and change of viewpoint. As such a measure, color saliency defined from average border contrast is proposed. The second method to be analyzed, is Color Boosting, introduced by van de Weiejer *et al.* in [10]. This method, is based on the self information of the chromatic transition (first order derivatives of the image). It is shown in [10] that Color Boosting improve the color distinctiveness in a framework of interest points detection. In the present work, we want to analyze if this characteristic can be also useful to decide the correctness of a segmentation.

To decide if these methods are able to take a correct decision, we compare their results using a benchmark. Thus, a good method of correctness should choose among different segmentations the one that have the best score in the benchmark. Such a coincidence between the non-supervised methods and the benchmark would imply that the measure of correctness applied would be a useful tool to perform a non-supervised segmentation, by automatically selecting the best set of parameters for each image.

This article is organized as follows: in section 2 we explain the method color saliency proposed in [12]. Afterwards, in section 3 we explain Color Boosting. Finally, sections 4 and 5 presents results obtained and a discussion of this article respectively.

2 HEIDEMAN'S COLOR SALIENCY:

The approach of Heidemann introduced in [12] proposes a goodness function for color segmentation, which allows to predict whether the segmented regions will be stable against noise, variation of lighting, and change of viewpoint. Color saliency defined from average border contrast of the segmented image. This work points out that the idea to maximize contrast is segmentation has been previously considered. Nevertheless, Heidemann analyses the problem and shows how it leads to improve region stability. Experiments for three different algorithms show that the effect is independent of the particular functional principle of segmentation. Thus, the measure can be applied for the automatic and context-free optimization of segmentation parameters.

The measure proposed is based of color distinctiveness of the regions of the segmented image. Thus, as far is the (euclidean) distance between neighboring regions, the better is the segmentation. Given an image I having three chromatic channels for each pixel (x, y), we compute a segmentation from which I is divided in N_R non-overlapped regions. The *region color* is defined as the mean color of this region in the original image.

The *region saliency* $S_R(R_i)$ is defined as the average color difference of R_i to the neighboring regions. Concretely, let the boundary of R_i be given as a set $B(R_i)$ consisting of $N_B(R_i)$ different pixels. Then $S_R(R_i)$ is calculated along the boundary as

$$S_R(R_i) = \frac{1}{N_B(R_i)} \sum_{(x,y)\in B(R_i)} \frac{1}{N_{diff}(x,y)} \times \sum_{R_j(x',y')|(x',y')\in Neigh4(x,y)} \| \overline{C}(R_i) - \overline{C}(R_j) \| . \quad (1)$$

Here, $\|\|$. denotes the color distance measure for the particular segmentation algorithm used. For color spaces such as RGB, $L*u*v*$ or $L*a*b*$ the Euclidean distance is used.

The first sum in Eq. 1 is over all boundary pixels (x, y). The second sum goes over the pixels (x', y') within a 4-neighborhood of (x, y), the 4-neighborhood being denoted by $Neigh4(x, y)$. To each neighboring pixel (x', y') the corresponding region $R_j(x', y')$ has to be found, so that the Euclidean distance between the region colors $\overline{C}(R_i)$ and $\overline{C}(R_j)$ can be calculated. $N_{diff}(x, y)$ denotes the number of pixels of $Neigh4(x, y)$ that belong to a different region, not to R_i. This factor is introduced to avoid âĂŸâĂŸdilutionâĂŹâĂŹ of the average distance in case there is, e.g. only one neighboring pixel which belongs to a different region. $N_{diff}(x, y)$ is at least 1 since (x, y) is part of the boundary, the maximum value is $N_{diff}(x, y) = 4$ in the case that (x, y) is a region consisting of an isolated pixel.

The Saliency measure of an image I denoted by $S(I)$ is given by the average over all its regions

$$S(I) = \frac{1}{N_R} \sum_{R_i\in I} S_R(R_i) \quad (2)$$

Summarizing, $S(I)$ is a measure of the color distinctiveness of the regions of a segmented image.

In the next section we explain a another way to include color distinctiveness to and contrast to decide the goodness of a segmentation.

3 COLOR BOOSTING:

The second method that will be analyzed is Color Boosting, as introduced in [10]. This method, is based on the self information of the chromatic transition (first order derivatives of the image). It is shown in [10] that Color Boosting improve the color distinctiveness in a framework of interest points detection. In the present work, we want to analyze if this characteristic can be also useful to decide the correctness of a segmentation.

The color saliency method by Van de Weijer *et al.* [10] is inspired by the notion that a feature's saliency reflects its information content. Consider an image $\mathbf{f} = (R, G, B)^{\mathbf{t}}$. The information content, I, of an image derivative $\mathbf{f_x}$, according to information theory, is given by the logarithm of its probability p:

$$I = -log(p(\mathbf{f}_x)). \tag{3}$$

Hence, color image derivatives which are equally frequent have equal information content. To map image derivatives to a saliency map, a function g is required for which the following holds:

$$p(\mathbf{f}_x) = p(\mathbf{f}'_x) \leftrightarrow |g(\mathbf{f}_x)| = |g(\mathbf{f}'_x)|. \tag{4}$$

The saliency function g transfers color image derivatives to a space where their norm is proportional to their information content.

In Fig. **??**, the distribution of color derivatives for the COREL dataset is given. The derivatives form an ellipsoid-like distribution, of which the longest axis is along the luminance direction. This indicates that equal displacements are more informative along the color directions (perpendicular to the luminance) than in the luminance direction. The saliency transformation in [10] is restricted to a transformation based on known color spaces. Now we propose a more general transformation to compute g in that it is not fixed to a pre-defined color space.

Let the distribution of the ellipsoid to be described by the covariance matrix $\mathbf{M}$:

$$\mathbf{M} = \overline{\mathbf{f}_\mathbf{x} \left(\mathbf{f}_\mathbf{x}\right)^\mathbf{t}} = \begin{pmatrix} \overline{R_\mathbf{x} R_\mathbf{x}} & \overline{R_\mathbf{x} G_\mathbf{x}} & \overline{R_\mathbf{x} B_\mathbf{x}} \\ \overline{R_\mathbf{x} G_\mathbf{x}} & \overline{G_\mathbf{x} G_\mathbf{x}} & \overline{G_\mathbf{x} B_\mathbf{x}} \\ \overline{R_\mathbf{x} B_\mathbf{x}} & \overline{G_\mathbf{x} B_\mathbf{x}} & \overline{B_\mathbf{x} B_\mathbf{x}} \end{pmatrix} \tag{5}$$

where the matrix elements are computed by

$$\overline{R_\mathbf{x} R_\mathbf{x}} = \sum_{i \in S} \sum_{\mathbf{x} \in X^i} R_\mathbf{x}(\mathbf{x})\, R_\mathbf{x}(\mathbf{x}) \tag{6}$$

where S is a set of images, and X^i is the set of pixels coordinates $\mathbf{x}$ in image i. Matrix $\mathbf{M}$ describes the derivatives energy in any direction $\hat{v}$. This energy is computed by $E(\hat{v}) = \hat{v}\mathbf{M}\hat{v}^\mathbf{t}$. Matrix $\mathbf{M}$ can be decomposed into eigenvector matrix $\mathbf{U}$ and eigenvalue matrix $\mathbf{\Lambda}$ according to $\mathbf{M} = \mathbf{U\Lambda U^t}$. This provides us with the saliency function g:

$$\mathbf{g}\left(\mathbf{f}_\mathbf{x}\right) = \mathbf{\Lambda}^{-1}\mathbf{U}^\mathbf{t}\mathbf{f}_\mathbf{x}. \tag{7}$$

Substitution of Eq. 7 into Eq. 5 yields

$$\mathbf{g}\left(\mathbf{f}_\mathbf{x}\right)\left(\mathbf{g}\left(\mathbf{f}_\mathbf{x}\right)\right)^\mathbf{t} = \mathbf{\Lambda}^{-1}\mathbf{U}^\mathbf{t}\mathbf{U\Lambda U}^\mathbf{t}\mathbf{U\Lambda}^{-1} = I, \tag{8}$$

meaning that the covariance matrix of the transformed image is equal to the identity matrix. This implies that the derivative energy in the transformed space is equal in all directions.

Using this procedure, we can generate an image where the energy of the information of the first order derivatives is considered. Fig. 2c shows an example.

4 RESULTS OBTAINED: In this section we explain the procedure to follow in order to determine the correctness of the methods detailed in sections 2 and 3. Afterwards we show prior results obtained.

To figure out the validity of both methods, we use a widely used dataset and benchmark, introduced in [3]. First, we take an image that will be segmented using two different methods. One of them, is the Mean Shift (MS) [5]. It have a public available version, called EDISON [5] and has been widely used and studied. As a second segmentation method, we use the one introduced in [14], as a new method which outperforms Mean Shift. The next step, is to compute an error measure. The one that we propose, is the Boundary Displacement Error (BDE) [1] since fits to a segmentation problem where there is no control about the number of segments. Having all these segmentations we compute their saliency index $S(I)$ following equations 1 and 2. Afterwards, we apply color boosting to the original image (Fig. 2a) and we generate the energy image of it (energy of its first order derivatives). The energy image is showed in Fig. 2c. Note its great similarity with human segmentation (Fig. 2b). Then, we compute the intersection between the energy image and the borders of the regions of each segmentation. That is, the overlapping of each region normalized by the number of points of the region's border $N_B(R_i)$.

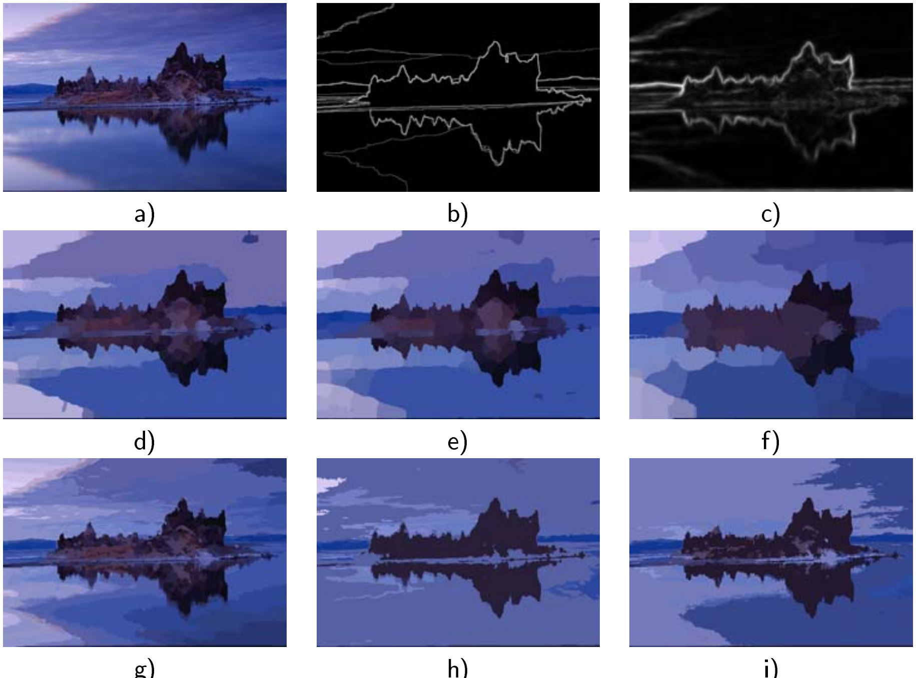

Fig. 2: (a) Original image extracted from the Berkeley dataset. (b) Human Segmentation (benchmark). (c) Energy of (a) after applying Color Boosting. (d-f) Segmentations obtained with Mean Shift.(g-i) Segmentations obtained with RAD. As can be seen h) is the segmentation that looks more similar to b) being the one selected as the best segmentation for both tested methods as well as the error measure (BDE).

Hence, if the goodness methods analyzed in this article predict that the best segmentation is the same as the one predicted using BDE and a benchmark, it would means that these methods can be effectively used to determine the goodness of a segmentation without consulting a benchmark. Table **??** summarizes results obtained. We can see how both Heidemann and Color Saliency correctly predict that the best segmentation is RAD2, as confirmed by the BDE score obtained using the benchmark.

5 DISCUSSION AND FURTHER WORK: In this article we have seen how the measure of goodness analyzed have a correct behavior. Heidemann computes the color distinctiveness of a region, whereas the method proposed using Color Boosting, introduces the chromatic information of every region obtained. The second difference is that whereas Heidemann uses the segmented image, the second approach uses the original image. Since both methods give correct results, the next steps is to combine both ideas in a single measure as well as to perform a further analysis of its performance.

REFERENCES:

[1] Huang, Q. and Dom, B.: Quantitative methods of evaluating image segmentation. IEEE International Conference on Image Processing **3** (1995) 53–56

[2] Deng Y. and Manjunath B.S.: Unsupervised segmen-

tation of Color-Texture Regions in Images and Video. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(8) (2001) 800–810
[3] 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. Proc. Eighth Int. Conf. Computer Vision **2** (2001) 416–423
[4] Cheng, H.D., Jiang, X.H., Wang, J.: Color image segmentation based on homogram thresholding and region merging. Pattern recognition **35**(2) (2002) 373–393
[5] Christoudias, C., Georgescu, B., Meer, P.: Synergism in low level vision. International Conference on Pattern Recognition **4** (2002) 150–155
[6] Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **24**(5) (2002) 603–619
[7] Ge, F. and Wang, S. and Liu, T.: Image-Segmentation Evaluation From the Perspective of Salient Object Extraction. Proc. IEEE Conference on Computer Vision and Pattern Recognition **1**(2006) 1146–1153
[8] Macaire, L. and Vandenbroucke, N. and Postaire, J.G.: Color image segmentation by analysis of subset connectedness and color homogeneity properties. Computer Vision and Image Understanding **102**(1) (2006) 105–116
[9] Micusık, B., Hanbury, A.: Automatic image segmentation by positioning a seed. Proc. European Conference on Computer Vision (ECCV) (2006)
[10] van de Weijer, J. and Gevers, T. and Bagdanov, A.D.: Boosting Color Saliency in Image Feature Detection. IEEE Transactions on Pattern Recognition and Machine Intelligence **28**(1) (2006) 150–156
[11] Donoser, M. and Bischof, H.: ROI-SEG: Unsupervised Color Segmentation by Combining Differently Focused Sub Results. Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on (2007) 1–8
[12] Heidemann, G.: Color segmentation robust to brightness variations by using B-spline curve modeling. Image and Vision Computing **26**(2) (2008) 211–227
[13] Wattuya, P. and Jiang, X. and Rothaus, K.: Combination of Multiple Segmentations by a Random Walker Approach. Lecture Notes in Computer Science **5096**(214–223) (2008) 211–227
[14] Vazquez, E. and van de Weijer, J. and Baldrich, R.: Image Segmentation in the Presence of Shadows and Highlights. 10th European Conference on Computer Vision, LNCS 5305 (2008) 1–14