Describing Reflectances for Colour Segmentation Robust to Shadows, Highlights and Textures

Eduard Vazquez, Ramon Baldrich, Joost van de Weijer, and Maria Vanrell

Abstract—The segmentation of a single material reflectance is a challenging problem due to the considerable variation in image measurements caused by the geometry of the object, shadows, and specularities. The combination of these effects has been modelled by the dichromatic reflection model. However, the application of the model to real-world images is limited due to unknown acquisition parameters and compression artifacts.

In this paper, we present a robust model for the shape of a single material reflectance in histogram-space. The method is based on a multilocal creaseness analysis of the histogram, which results in a set of ridges representing the material reflectances. The segmentation method derived from these ridges is robust to both shadow, shading and specularities, and texture in real-world images. We further complete the method by incorporating prior-knowledge from image statistics, and incorporate spatial coherence by using multi-scale color contrast information. Results obtained show that our method clearly outperforms state-of-the-art segmentation methods on a widely used segmentation benchmark, having as a main characteristic its excellent performance in the presence of shadows and highlights at low computational cost.

Index Terms—Segmentation, Color.

1 Introduction

Image segmentation is a computer vision process which aims to partition an image into a set of non-overlapped regions, called segments. A robust and efficient segmentation is required as a preprocessing step in several computer vision tasks such as image classification or object detection and recognition. In real images changes due to illumination, shadow, shading and highlights provoke image measurements to vary significantly. These effects, are one of the main difficulties that have to be solved to yield a correct segmentation.

Image segments caused by a single material reflectance form complex shapes in histogram-space, due to shading effects and specularities. The fact that these physical effects lead to undesired image segments is also confirmed by Martin *et al.* in [28]. He points out the existence of strong edges caused by such physical effects which are not considered in human segmentations, but which tend to be detected by current segmentation methods. Previous work on image segmentation robust to shading effects and specularities is based on a reflection model of light. These methods, called physics-based, predominantly based on the dichromatic reflection model (DCM) [3] [6] are aimed to explain the behavior of light in a scene. Thus, from a theoretical point of view, these models are able to explain the formation of shadows and specularities. These methods are based on several assumptions which severely limit their applicability. The main problem is the presence of artifacts introduced by acquisition conditions, clipped highlights or image compression. A second set of segmentation methods are feature-based [24] [27]. These methods are not based on prior assumptions of the underlying physics and are therefore more flexible to mentioned problems. However, ignorance of the physical process often leads to incongruences in the presence of shadows and highlights.

In this article, we aim to combine the strengths of physics and feature-based methods. The presented method is based on the observation that the distribution of single material reflectance can be robustly represented by a single connected ridge in histogram space. The method is named Ridge-based Analysis of Distributions (RAD). The detection of these ridges is based on a creaseness analysis of the histogram. This technique connects the shadows in the dark parts of the object, to the brighter regions, and further up to the highlights (see Fig. 1). Furthermore, the ridges are capable to correctly connect single material textures, such as grass or sand. The advantage over previous physics-based methods is that our method does not assume a parametric shape, and is therefore robust for non-linear acquisition and image compression.

We propose two further extensions to the basic method.

Firstly, to suppress those ridges in the less probable orientations and favor those ridges in the probable ones, we extend the method to exploit the image statistics of

- *E. Vazquez is with Computer Vision Center / Department of Computer Science, Universitat Autonoma de Barcelona, 08193 Cerdanyola del Valles, Barcelona (Spain),*
  E-mail: eduard@cvc.uab.cat
- *R. Baldrich, J. van de Weijer and Maria Vanrell are with Computer Vision Center, Universitat Autonoma de Barcelona.*

ridge orientations. This extension is called physics-based RAD (pRAD). Secondly, ridges on the histogram can just cope with those segments derived from single materials. Segments formed by more than a material will be represented by different ridges in the histogram. These textures tend to be present at certain scales, but display weak contrast at other scales. This fact is exploited by the multi-contrast representation of the image, in which texture contrast is suppressed. This method is called spatial RAD (sRAD).

This paper is organized as follows: in section 2 we explain the related work in image segmentation. Afterwards, in section 3 we explain the theoretical basis and motivations of our approach. Subsequently in sections 4, 5 and 6 we explain RAD, pRAD and sRAD respectively. A comparison with Mean Shift segmentation and a performance evaluation of our approach is done in section 7. Finally, conclusions of the current work are given in section 8.

2 Related work

There exist several different methods covering a broad spectrum of points of view. The work presented by Skarbek and Koschan [13], draws the basis of the current classifications of segmentation methods. Some other comprehensive surveys of colour segmentation techniques are presented in [24] and [27], where a similar schema is followed. From these works segmentation methods are divided in four main categories: image-based, feature-based, physics-based and hybrid approaches. Feature-based approaches are focused on the photometric information of an image represented on its histogram [44] [64]. Image-based approaches exploit the spatial information contained in the image, named *spatial coherence*. Physics-based methods use the knowledge about the physical formation of the scene (light, surfaces reflectance), to perform the segmentation. Finally, hybrid techniques combine methods of the previous categories. As stated before, this paper introduces a method that performs an analysis of the histogram (feature-based method, RAD) exploiting the statistics of the ridges (physics-based, pRAD) and adding as a final step the spatial coherence of the image (image based, sRAD) . Therefore, the segmentation method presented belongs to the category of hybrid methods.

Regarding image-based methods, these include region and boundary information [35] [58] graph-based approaches as nCuts [25] or the efficient graph-based image segmentation [37], region growing algorithms [7] [48] or segmentation based on watershed [52] [45] and in general topological approaches [42]. These basic techniques are either mixed [16] or complete by means of markov random fields [12] [53] or neural network approaches [34].

Feature-based methods can be further split in three main categories: histogram thresholding, clustering and fuzzy clustering. Histogram thresholding techniques assume that there exist a threshold value that isolates all pixels representative of an object in a scene. This basic concept is exploited in several ways as explained in [41]. Clustering techniques, perform a partition of the feature space under different criteria as described in [44]. Such criteria include distance measures as k-means or ISODATA [22], probabilistic/statistical approaches, such as Mean Shift [1], or the spectral analysis of the data [29], based on the Karhunen-Loeve transformation. Fuzzy clustering includes methods such as fuzzy c-means [21] [38], Gath-Geva clustering [4], or mixture models [49] [62] which are a way to look for areas of high density. From all of them, the most related work with RAD is Mean shift. Both, Mean Shift and RAD, use topological information (modes and gradients for Mean Shift and structural tensor, creaseness and ridges for RAD) to perform the classification of colors in the histogram space.

Physics techniques model the behavior of the light in the scene. For instance, with an statistical approach based on b-splines fitting in the HSV [66], or by means of a generalized Hough transform method, gradient descent method, and eigenvectors method as suggested in [39]. The main contribution to physics-based techniques was done by S.A. Shafer in 1985 with the introduction of the dichromatic reflection model (DCM) [3]. DCM, has been the basis of several segmentation techniques [5] [6]. Its limitations regarding different materials (metals and inhomogeneous dielectrics) geometry and non Lambertian surfaces has been investigated [15] [18].

DCM explains from a theoretical point of view the sort of shapes that a single surface can form in the histogram due to illumination interactions. The fact that these shapes do not correspond with the common feature-based clustering techniques such as Mean Shift [1] [33] is the other observation that forms the basis of our proposal. Mean Shift can not extract the elongated shapes described by the DCM, whereas pRAD performs an analysis of the histogram space focused on the extraction of elongated shapes that correctly follow the directions of the DCM.

In addition to these approaches we include within physics-based approaches those models of color spaces proposed to cope with shadows and highlights. The first good proposal was the Ohta space [2] proposed in 1980 which is a linear transformation of the RGB space that has been used in several approaches for images segmentation. Other interesting proposals for color spaces robust to, or that deal with, shadows and highlights, include an eigen color representation [43], an illuminant independent log-opponent representation [23] and a specific model to deal with color distortion [40].

Finally hybrid approaches combine techniques of the three previous categories. For instance, by adding image spatial constraints (spatial coherence) to a clustering technique such as k-means [9] or more recently with fuzzy c-means [50]. The JSEG segmentation method [26] is a two-step schema following a similar idea. First, a

clustering of the color space is performed. Afterwards, a criterion of *good* segmentation is applied using the spatial coherence of the image. Another schema proposes that a good segmentation region should be formed by strongly connected pixels with homogeneous colors [54].

In this work, we use the spatial coherence of the image to build a multiscale information-based chromatic contrast map, called *multicontrast image*. This map will guide a procedure to combine a set of sub-segmentations computed from a single image at different feature-space scales. Hence, we use the multicontrast image, to determine the goodness of a segmentation (or a segment) [65] [51]. This extension of the method is called sRAD.

3 Our approach: Theoretical Foundations

Our approach to colour image segmentation is based on the insight that the distributions formed by a single-colored object have a physically determined shape in colour histogram-space. We model an image as being generated by a set of segments, each of which corresponds with a material reflectance (MR) described by a distribution in histogram-space. Each MR is related to a semantic object in the image. For example, in Figure 1 we distinguish between four different MRs, namely: red for the pepper, green and brown for the branch and black for the background.

A MR generates many image values due to geometrical and photometric variations. Our main aim is to find a good representation of the topologies which MR's are likely to form in histogram space. For this purpose, consider the distribution of a single MR as described by the dichromatic reflection model (DCM) [3]:

$$\mathbf{f}(\mathbf{x}) = m^b(\mathbf{x})\,\mathbf{c}^b + m^i(\mathbf{x})\,\mathbf{c}^i \tag{1}$$

in which $\mathbf{f} = \{R, G, B\}$, $\mathbf{x}$ is the spatial image coordinate, $\mathbf{c}^b$ is the body reflectance, $\mathbf{c}^i$ the surface reflectance, m^b and m^i are geometry dependent scalars representing the magnitude of body and surface reflectance. Bold notation is used to indicate vectors. For one MR we expect both $\mathbf{c}^b$ and $\mathbf{c}^i$ to be almost constant, whereas $m^b(\mathbf{x})$ and $m^i(\mathbf{x})$ are expected to vary significantly. Hence, as for this definition, a MR, is formed by a single body reflectance $\mathbf{c}^b$ and a surface reflectance $\mathbf{c}^i$.

The two parts of the dichromatic reflection model are clearly visible in the histogram of Figure 1b. Firstly, due to the shading variations the distribution of the red pepper traces an elongated shape in histogram-space. Secondly, the surface reflectance forms a branch which points in the direction of the reflected illuminant. In conclusion, the distribution of a single MR forms a ridge-like structure in histogram space.

To illustrate the difficulty of extracting the distributions of MRs consider Figure 2c, which contains a patch of the horse image. The 2D Red-Green histogram of the patch is depicted in Figure 2d to see the number of occurrences of each chromatic combination. This is done for explanation purposes. In this 2D histogram it can be clearly seen that the density of the geometric term $m_b(\mathbf{x})$ varies significantly, and the distribution is broken in two parts. However, we have an important clue that the two distributions belong to the same MR: the orientation of the two distribution is similar, which means they have a similar $\mathbf{c}^b$. We exploit this feature in the ridge extraction algorithm by connecting neighboring distributions with similar orientation.

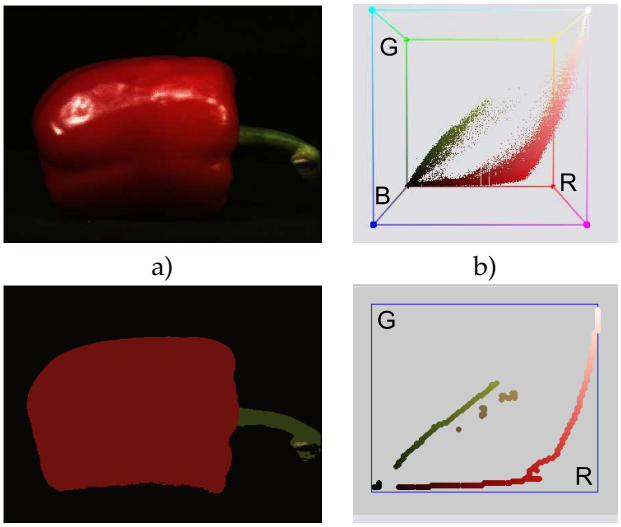


Fig. 1. (a) An image from [46] and (b) its histogram. The effects of shading and highlights are clearly visible in the red colours of the histogram. (c) Segmented images using RAD. (d) Ridges found with RAD. Note that the three branches of the red pepper are correctly connected in a single ridge.

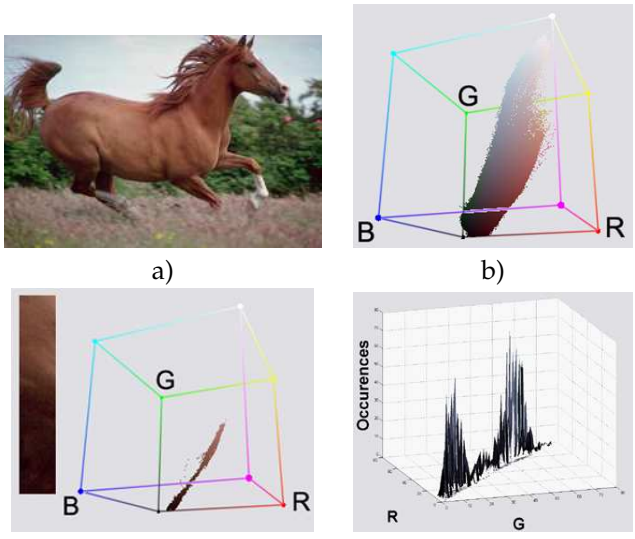


Fig. 2. (a) An image and (b) its 3D RGB histogram. (c) A patch of a) and its RGB histogram. (d) 2D histogram of c) to illustrate the discontinuities appearing on a MR.

In literature several methods have explicitly used the dichromatic reflection model to obtain image segmentation, e.g. [6]. A drawback of such methods is however that for many images Eq. 1 does only approximately model the data. This can be caused by many reasons, such as non-linear acquisition systems, clipped highlights, and image compression.

An additional problem appears when an MR is affected by more than one illuminant. An example of it is shown in Fig.3. The cast shadow on the ground causes a sudden change in the direction of the MR caused by a change in illumination. This sudden and fairly appreciable change is not modeled by Eq.1. A DCM-based segmentation would probably split the MR in two segments, as showed in Fig.3c. Figs.3e,f show a detection of this MR obtained with the method proposed in this paper. Note that the ridge corresponding with the MR of the ground copes with this sudden change of chromaticity, joining the cast shadow and the ground in a non-expected direction by the DCM.

In this article we use Eq. 1 only to conclude that objects described by this equation will trace connected ridges in histogram space. This makes the method more robust to deviations from the dichromatic model.

4 A Ridge based Distribution Analysis method (RAD)

In this section we present a fast algorithm to extract MRs from histogram space. The proposed method is divided in two main steps. First, we propose a method to extract ridges as a representative of a MR. Afterwards a flooding process is performed to find the MRs from its ridges.

4.1 First step: Ridge Extraction

To extract a MR descriptor we need to find those points containing the most meaningful information of a MR, *i.e.*, its ridge. We propose to apply a multilocal creaseness algorithm to find the best ridge point candidates. This operator avoids splitting up ridges due to irregularities on the distribution, mainly caused by the discrete nature of the data. Afterwards, we apply a ridge extraction algorithm to find the descriptor.

4.1.1 Multilocal Creaseness: finding candidates and enhancing connectivity

In order to deal with this commonly heavily jagged MR (see Fig. 2d) , we propose to apply the MLSEC-ST operator introduced by Lopez *et al.* in [20] to enhance ridge points. This method is used due to its good performance compared with other ridge detection methods [20] on irregular and noisy landscapes.

The Structure Tensor (ST) computes the dominant gradient orientation in a neighbourhood of size proportional to σ_d. Basically, this calculus enhances those situations where either a big attraction or repulsion exists in the gradient direction vectors. Thus, it assigns the higher values when a ridge or valley occurs. Given a distribution $\Omega(\mathbf{x})$, (the histogram in the current context), and a symmetric neighbourhood of size σ_i centered at point $\mathbf{x}$, namely, $N(\mathbf{x}, \sigma_i)$ the structure tensor field S is defined as:

$$S(\mathbf{x}, \sigma) = N(\mathbf{x}, \sigma_i) * (\nabla\Omega(\mathbf{x}, \sigma_d) \cdot \nabla\Omega^t(\mathbf{x}, \sigma_d)) \qquad (2)$$

where $\sigma = \{\sigma_i, \sigma_d\}$, and the calculus of the gradient vector field $\nabla\Omega(\mathbf{x}, \sigma_d)$ has been done with a Gaussian Kernel with standard deviation σ_d.

If $w(\mathbf{x}, \sigma)$ is the eigenvector corresponding to the largest eigenvalue of $S(\mathbf{x}, \sigma)$, then, the dominant gradient orientation $\overline{w}(\mathbf{x}, \sigma)$ in a neighbourhood of size proportional to σ_i centered at $\mathbf{x}$ is:

$$\overline{w}(\mathbf{x}, \sigma) = \mathbf{sign}(\mathbf{w^t}(\mathbf{x}, \sigma) \cdot \nabla\mathbf{\Omega}(\mathbf{x}, \sigma_\mathbf{d}))\mathbf{w}(\mathbf{x}, \sigma) \qquad (3)$$

The creaseness measure of $\Omega(\mathbf{x})$ for a given point $\mathbf{x}$, named $k(\mathbf{x}, \sigma)$, is computed with the divergence between the dominant gradient orientation and the normal vectors, on the *r*-connected neighbourhood. The normal vectors are computed with $\mathbf{y} - \mathbf{x}$ where $\mathbf{y}$ are the neighbors of $\mathbf{x}$ ($\mathbf{y} \in neigh(\mathbf{x})$) in an *r*-connected neighborhood. That is:

$$k(\mathbf{x}, \sigma) = -\frac{\mathbf{d}}{\mathbf{r}} \sum_{\mathbf{y} \in \mathbf{neigh(x)}} \overline{w}^{\mathbf{t}}_{\mathbf{k}}(\mathbf{x}, \sigma) \cdot (\mathbf{y} - \mathbf{x}) \qquad (4)$$

where d is the dimension of $\Omega(\mathbf{x})$. The creaseness representation of $\Omega(\mathbf{x})$ will be referred hereafter as Ω^σ. In our implementation we have used a 6-connected neighborhood in the three dimensional chromatic space.

As an example, Figure 4a shows the opponent colour 2D histogram of 4g. Its creaseness values are showed in 4b. There are three enhanced areas which corresponds with the three MRs of the original image. They appear as three mountains in 4b, clearly separated by two valleys. Note that higher creaseness values have a larger probability to become a ridge point.

4.1.2 Ridge Detection

In the previous section we have enhanced the ridge structure of the distribution with a creaseness operator. Here, we describe this distribution by a set of ridges. As a result only those points necessary to maintain the connectivity of a MR remain. These points form the ridges of Ω^σ.

We classify ridge points in three categories. First, Transitional Ridge Points (TRP): when there is a local maximum in a single direction. Second, Saddle Points (SP): when there is a local maximum in one direction and a local minimum in another one. Third, Local Maximum Points (LMP). Formally, let $\Omega(\mathbf{x})$ be a continuous 2D surface and $\nabla\Omega(\mathbf{x})$ be the gradient vector of the function $\Omega(\mathbf{x})$. We define ω_1 and ω_2 as the unit eigenvectors of the Hessian matrix and λ_1 and λ_2 its corresponding eigenvalues with $\mid \lambda_1 \mid \leq \mid \lambda_2 \mid$. Then, for the 2D case:

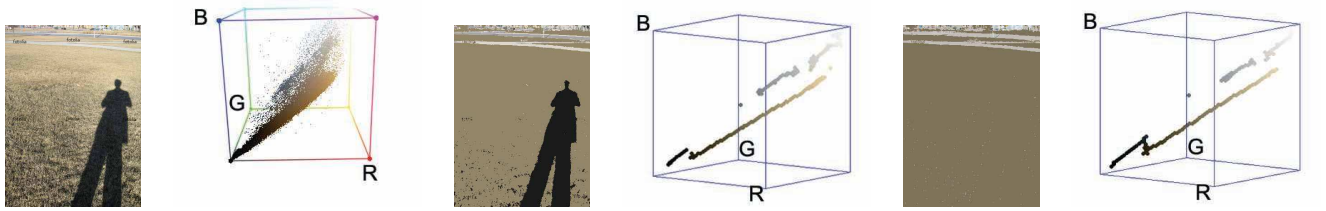


Fig. 3. Effects on the MR shape when illuminant color changes. (a,b) Original image and its histogram. (c,d) Possible segmentation based on DCM. (e,f) Segmentation proposed by our method where a single MR is found for the ground.

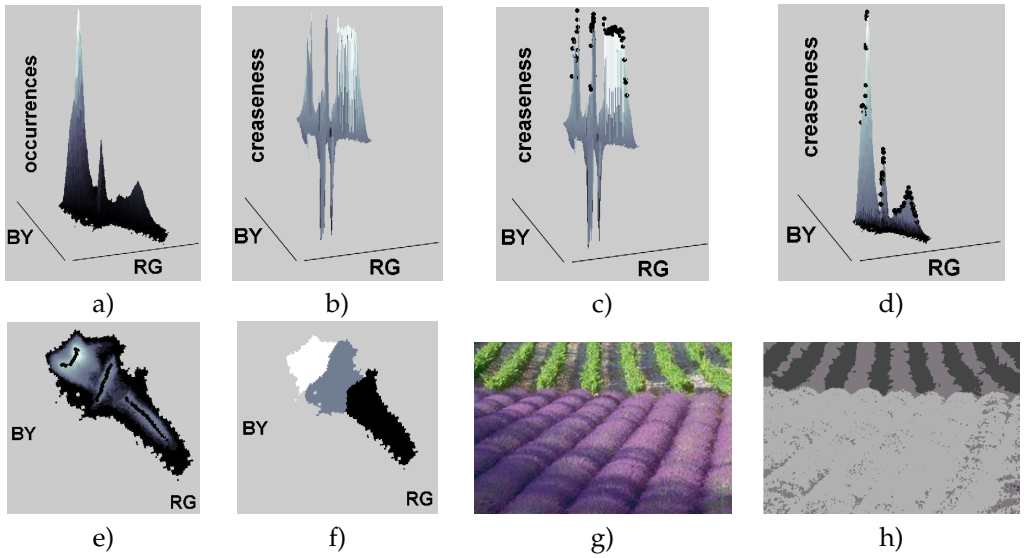


Fig. 4. A graphical example of the whole process. (a) Opponent Red-Green and Blue-Yellow histogram $\Omega(\mathbf{x})$ of g). (b) Creaseness representation of a). (c) Ridges found in b). (d)Ridges fitted on original distribution. (e) Top-view of d). (f)MRs of a). (g) Original image. (h)Segmented image.

$$LMP(\Omega(\mathbf{x})) = \{(\mathbf{x})|(\|\nabla\Omega(\mathbf{x})\| = 0), \lambda_1 < 0, \lambda_2 < 0\} \quad (5)$$

$$\begin{aligned} TRP(\Omega(\mathbf{x})) = \{(\mathbf{x})|&\|\nabla\Omega(\mathbf{x})\| \neq 0, \lambda_1 < 0, \nabla\Omega(\mathbf{x}) \cdot \omega_1 = 0, \\ &\|\nabla\Omega(\mathbf{x})\| \neq 0, \lambda_2 < 0, \nabla\Omega(\mathbf{x}) \cdot \omega_2 = 0, \\ &\|\nabla\Omega(\mathbf{x})\| = 0, \lambda_1 < 0, \lambda_2 = 0\} \end{aligned} \quad (6)$$

$$SP(\Omega(\mathbf{x})) = \{(\mathbf{x})|\|\nabla\Omega(\mathbf{x})\| = 0, \lambda_1 \cdot \lambda_2 < 0\} \quad (7)$$

This definition can be extended for an arbitrary dimension using the combinatorial of the eigenvalues. Hereafter we will refer to these three categories as *ridge points* (RP). Thus, $RP(\Omega(\mathbf{x})) = LMP \bigcup TRP \bigcup SP$. A further classification of ridges and its singularities can be found in [11] and [30].

A common way to detect ridge points is to find zero-crossing in the gradient of a landscape for a given gradient direction. Thus, we need to compute all gradient directions and detect changes following the schema proposed in [30]. In our case, we propose a way to extract a ridge without the need to calculate the gradient values for all points in the landscape. We begin on a local maxima of the landscape and follow the ridge by adding the higher neighbours of the current point, if there is a zero-crossing on it, until it reaches a flat region. This method can be easily applied to an arbitrary dimension. Formally, let $neigh(\mathbf{x}, \Omega^\sigma)$ be the set of neighbours of a point $\mathbf{x} \in \Omega^\sigma$, and $Cneigh(\mathbf{x}, \mathbf{y}, \Omega^\sigma)$ be the set of common neighbours between point $\mathbf{x} \in \Omega^\sigma$ and $\mathbf{y} \in \Omega^\sigma$. We also define a function $\mu(\mathbf{x}, \Omega^\sigma)$ as follows:

$$\mu(\mathbf{x}, \Omega^\sigma) = \sharp\{\mathbf{y} \in neigh(\mathbf{x}) \mid \Omega^\sigma(\mathbf{y}) \geq \Omega^\sigma(\mathbf{x})\} \quad (8)$$

where $\sharp$ is the cardinality of the set. Therefore, $\mu(\mathbf{x}, \Omega^\sigma) = 0$ means that $\mathbf{x}$ is a local maximum. Finally, we define μ' as:

$$\mu'(\mathbf{x}, \mathbf{y}, \Omega^\sigma) = \sharp\{\mathbf{z} \in Cneigh(\mathbf{x}, \mathbf{y}) \mid \Omega^\sigma(\mathbf{z}) \geq \Omega^\sigma(\mathbf{y})\} \quad (9)$$

$\mu(\mathbf{x}, \mathbf{y}, \Omega^\sigma) = 0$ means that $\mathbf{y}$ is a local maxima in the common neighbours between $\mathbf{x}$ and $\mathbf{y}$. To extract

ridges we propose an iterative process beginning on local maxima, that is

$$RP_0(\Omega^\sigma) = \mathbf{x} \in \Omega^\sigma \| \mu(\mathbf{x}, \Omega^\sigma) = 0 \quad (10)$$

Then, we just have to follow ridges starting on $RP_0(\Omega^\sigma)$ until its ending.

$$RP_{\mathbf{z}}(\Omega^\sigma) = RP_{\mathbf{z}-1}(\Omega^\sigma) \cup \{\mathbf{n} \in neigh(\mathbf{l}) \mid \mathbf{l} \in RP_{\mathbf{z}-1}(\Omega^\sigma), \mu'(\mathbf{l}, \mathbf{n}, \Omega^\sigma) = 0\} \quad (11)$$

Fig. 4c depicts the ridge points found on Ω^σ with black dots. Figs. 4d,e show a 3D view and a 2D projection view respectively of how these ridge points fit in the original distribution as a representative of the three MRs. Finally, from the set of ridge points of a distribution we can compute each MR. A second example is shown in Figure 1. The complicated colour distribution of the pepper, caused by shading and highlight effects, is correctly connected in a single ridge.

4.2 Second step: MR Calculus from its ridge points

In this final step we find the MR belonging to each ridge found. From topological point of view, it implies finding the portion of landscape represented by each ridge. These portions of landscape are named *catchments basins*. Vincent and Soille [8] define a catchment basin associated with a local minimum M as the set of pixels p of Ω^σ such that a water drop falling at p flows down along the relief, following a certain descending path called the downstream of p, and eventually reaches M. In our case, M are the set of ridge points found and then, MRs are found using the algorithm proposed in [8] applied on the inverse Ω^σ distribution. The proposed algorithm, is not based on the gradient vectors of a landscape [10] but on the idea of *immersion* which is more stable and reduces over-segmentation. Basically, the flooding process begins on the local minima and, iteratively, the landscape sinks on the water. Those points where the water coming from different local minima join, compose the watershed lines. To avoid potential problems with irregularities [20], we force the flooding process to begin at the same time in all MR descriptors, on the smoothed $\Omega(\mathbf{x})$ distribution with a Gaussian kernel of standard deviation σ_d (already computed on the structure tensor calculus). Then, we define RAD as the operator returning the set of MRs of Ω^σ using ridge points as marks:

$$RAD(\Omega, \sigma) = W(\Omega^\sigma, RP(\Omega^\sigma)) \quad (12)$$

Following this procedure, Fig. 4f depicts the 2D projection of the MRs found on 4a.

5 Adding Physical Preference (pRAD)

The dichromatic model predicts pixels of a single colored object to form a line passing through the origin as long as no specular reflection is present. In case of specular reflection, it models pixels of both body and specular reflection to form a plane. However, applying these geometrical models to the pixel values often leads to unsatisfying results because of the many deviations causing the body reflectance pixels neither to lie on a line, nor the combined body and specular pixels to lie in a plane. In the previous section, we therefore proposed a method to extract ridges from histogram space, based on the observation that ridges capture the essential structure predicted by the dichromatic model while being more robust to slight deviations from the ideal case. These ridges are allowed to have any orientation. However, the dichromatic results suggests the orientation of body reflection and of specular reflection, to be more likely than others. In this section, we will incorporate this additional information into the RAD method, and propose the *physic-based RAD* called *pRAD*.

The general structure which a single colored object forms in histogram space, is a ridge in the radial direction caused by shadow and shading variations with in the higher intensity regions of the RGB cube some branches in the illuminant direction caused by specularities. Changes in the chromatic direction, perpendicular to these two directions are seldom. Due to blurring effects, caused by for example out of focus, relative motion between camera and object, and aberrations in the optical system, ridges in the chromatic direction are formed between different surface reflectances. These ridges which might result in undesired segmentation results. To suppress ridges in the less probable orientations and favor ridges in probable ones, we propose to exploit the image statistics of ridge orientations. This statistic is captured by computing the normalized tensor representation $\hat{S}$ of the color histograms generated by a set of images in a train data set with,

$$\begin{aligned} \overline{\hat{S}(\mathbf{x}, \sigma)} &= \sum_{i \in T} \hat{S}_i(\mathbf{x}, \sigma) \\ \hat{S}_i(\mathbf{x}, \sigma) &= \frac{S(\mathbf{x}, \sigma)}{\|S(\mathbf{x}, \sigma)\|} \end{aligned} \quad (13)$$

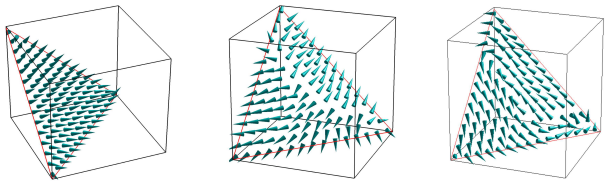
where T is the set of indexes of the train data. We normalize the tensors with

$$\|S(\mathbf{x}, \sigma)\| = N(x, \sigma_i) * \left(\nabla \Omega^{\mathbf{t}}(\mathbf{x}, \sigma_d) \cdot \nabla \Omega(\mathbf{x}, \sigma_d)\right) \quad (14)$$

since we are only interested in the orientation of the ridges, not their strength (note that the transpose operates on the first gradient here, whereas in in Eq. 2 it operates on the second). The outcome $\overline{\hat{S}}$ is a tensor field, which for each RGB value in the histogram indicates the relative likelihood of the orientations of ridges passing through this point.

The tensor field $\overline{\hat{S}}$, which does not require human segmentation, is learned on the complete COREL dataset of over 40.000 images [19]. In Fig.5 we have depicted the three eigenvectors of the tensor field for a slice of the RGB-cube, namely the chromatic plane ($R+G+B=1$). The dominant orientation (Fig.5a) coincides with the intensity direction. The orientation of the second and

third eigenvector is less clear. However, in general the second eigenvector points outwards from the center of the chromatic plane (Fig.5b). This is called the saturation direction, since changes in this direction cause colors to become more or less saturated. The least variations is found in the angular direction in the chromatic plane (Fig.5c), coinciding with hue changes. This is what we expect since most physical changes such as shadows, shading, and specularities (when white) do not cause any hue changes of the MR.



a) b) c)

Fig. 5. From left to right: first, second and third dominant orientations of the tensor field computed using 40.000 images of the COREL dataset.

This prior knowledge can be incorporated in the ridge extraction framework proposed in the previous section by using instead of Eq. 2 the following equation,

$$S^{\lambda}(\mathbf{x},\sigma) = (1-\lambda)\, S(\mathbf{x},\sigma) + \lambda \left\| S(\mathbf{x},\sigma) \right\| \overline{\hat{S}(\mathbf{x},\sigma)} \quad (15)$$

where $\lambda = [0,1]$ regulates the influence of the prior knowledge represented by $\overline{\hat{S}}$. For example, $\lambda = 0.25$, indicates that 75% of the strength of tensors in S^{λ} is based on the image to be segmented, and 25% of strength is defined by prior knowledge. The regulation parameter λ is learned from a training data set. In our experiments on the Berkeley training set. We found $\lambda = 0.33$ to yield the optimal results (optimization is based on the GCE score). Further results obtained with pRAD are detailed in section 7.

In conclusion, we proposed a method to favor ridges in orientation commonly seen in real-world images, and suppress ridges in the less probable orientations. It is important to note that the extra computational cost of pRAD is negligible with respect to RAD, since $\overline{\hat{S}}$ is precomputed.

6 Multi-scale segmentation adding image spatial coherence (sRAD)

With the addition of the physical information (pRAD) we add robustness to the method, which has a better behavior in those cases where the geometry of the objects and the light causes shadows, shading and highlights for a single colored object. In this section, we propose two further improvements. Firstly, the optimal parameter setting was found to vary for each image. To automatically obtain a good segmentation, we propose to combine the segmentations at various parameter settings. Secondly, RAD tends to oversegmentation in textures formed by multiple chromaticities. To overcome this problem, we propose to use mutliscale contrast (see section 6.2). We call this method *sRAD* (spatial RAD) or in combination with pRAD, it is called *spRAD*.

The idea to combine different sub-segmentations to build a combined segmentation has been investigated before [59] [55] [56]. The objective is to take the strengths of each segmentation while avoiding its weakness. Roughly, it implies to determine a measure of the goodness of a segment. JSEG algorithm [26], for instance, propose the *J-measure* that is based on the variance of the pixels belonging to a class-map (color quantization). This measure of goodness is computed at different scales forming the set of images that have to be combined. This is achieved by a region growing algorithm. The resulting image (based on goodness, not on chromaticity) tends to be oversegmented. Hence, a merge algorithm based on Euclidean distance of the histogram of each neighboring region is applied. Actually, this way to merge regions is commonly applied, e.g. also for Mean Shift segmentation, another algorithm which tends to oversegment. A graph-based approach to merge oversegmented images is presented in [56]. Other measures to describe the correctness of a segment are the homogram proposed in [31], the spatial-color compactness degree [54], a calculus based in the Bhattacharyya distance [59] or a probabilistic approach as explained in [55]. The method introduced in this article to combine sub-segmentations belongs to those methods that use contrast as a criteria of the goodness of a segmentation (e.g. [51][65]).

6.1 Combining sub-segmentations

With RAD and pRAD, we can segment at different feature-space scales by changing σ_d (feature-space smoothness). The optimal σ_d value varies depending the image. Therefore, whereas a single value can be found as the optimal when considering a whole dataset, results can be improved by selecting a different value depending on the image. Moreover, good segments can be found at different scales. Hence, we propose a method to consider segments at different feature-space scales for any single image. For the selection of segments we use a multi-scale contrast representation of the image which suppresses texture edges. Its computation is further explained in section 6.2.

The procedure follows two steps. First, we perform a set of segmentations at different feature-space scales (named sub-segmentations) of the same image. An example of these sub-segmentations is showed in Fig. 6, second column. The number refers to the value of σ_d used. Afterwards, we select the best segments of these sub-segmentations using the multiscale contrast image (Fig. 6) to build the final segmentation. The selection is based on a ranking which is computed by summing the contrast underlying the edges of the segments normalized for the perimeter of the segment. This operation, a combination of each sub-segmentation with the multiscale image is represented with $\otimes$. In Fig. 6, we show

the ranking for every segment selected by spRAD with a gray-value codification: as lighter the color, the higher the rank position. Once the most contrasted segment (first in the ranking) has been added to the combined segmentation, the contrast already contained in this segment is removed from the multiscale image and the ranking is done again. The numbers appearing in the images of Fig. 6 (third row and spRAD segmentation), illustrate at what scales the segments were selected that form the combined segmentation (spRAD).

In this article, sRAD is evaluated on the Berkeley dataset using the GCE error measure (a more detailed explanation of this error measure follows below). Comparison based on GCE requires methods to have a similar number of segments [28]. In the Berkeley dataset, human segmentations have an average of eight segments. Additionally, the results of all methods presented in section 7 have a similar number of segments (between 7 and 10). Therefore, we will select the nine first ranked segments to build the final segmentation. It is interesting to note that the combined segmentation was found to outperform all of the sub-segmentations from which it was formed, showing the validity of the approach.

6.2 Multiscale Color Contrast

In the previous section, we introduced a multiscale color contrast image as a selection criteria to combine various segmentations. The relevance of the segments is computed by summing the contrast underlying the edges of the segment (normalized for the perimeter of the segment). To obtain a good segmentation, we need the contrast-image to suppress shadow and specular edges, as well as spurious texture edges.

Textures tend to be present at certain scales, but exhibit weak contrast at other scales. For this reason, we propose to use a multi-scale contrast image. This multiscale chromatic contrast is computed as a linear combination of the Gaussian pyramid image, commonly used in saliency (*e.g.* [17]), according to:

$$s(\mathbf{x}) = \sum_{\sigma \in \Sigma} \sum_{\mathbf{x}' \in N(\mathbf{x})} \left\| \mathbf{M}^{\sigma} \left(\mathbf{f}^{\sigma}(\mathbf{x}) - \mathbf{f}^{\sigma}(\mathbf{x}') \right) \right\|^2 \qquad (16)$$

where $\mathbf{f}^{\sigma}$ denotes the Gaussian smoothed image at scale σ chosen from $\Sigma = \{1, 2, 4, 6, 8, 10, 12, 14\}$. $N(\mathbf{x})$ is a 9x9 neighborhood window. The approach is similar to [61].

To prevent the re-introduction of shadow and specular edges, we apply a color boosting matrix $\mathbf{M}$ in Eq. 16 [57]. This approach was originally proposed to amplify salient chromatic edges in the image, and thereby indirectly suppressing shadows and specularities. The boosting matrix is computed with

$$\mathbf{M}^{\sigma} = \left(\operatorname{diag}\left(\bar{\mathbf{o}}^{\sigma}_{\mathbf{x}} \right) \right)^{-1} \mathbf{U}, \qquad (17)$$

where

$$\begin{aligned} \mathbf{o}^{\sigma}_{\mathbf{x}}(\mathbf{x}) &= \mathbf{U}\mathbf{f}^{\sigma}_{\mathbf{x}}(\mathbf{x}) \\ \bar{\mathbf{o}}^{\sigma}_{\mathbf{x}} &= \sqrt{\sum_{\mathbf{x} \in \mathrm{X}} \left(\mathbf{o}^{\sigma}_{\mathbf{x}}(\mathbf{x}) \right)^2} \end{aligned} \qquad (18)$$

where the summation is over all pixels in the data set X, and $\mathbf{U}$ the transformation from RGB to opponent color space is given by

$$\mathbf{U} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}. \qquad (19)$$

The boosting matrix $\mathbf{M}$ normalizes the derivatives in each of the opponent color channels with the average derivative energy in that opponent channel in the data set. As was shown in [57] most derivative energy is along the intensity axis (the third opponent axis $O3$) and only little variations in the chromatic directions (the first and second opponent axes $O1$ and $O2$). Therefore, multiplication with the boosting matrix emphasizes salient chromatic edges and suppresses achromatic edges. We found color boosting to improve segmentation results [1].

The multiscale approach helps to minimize oversegmentation in textured parts of the image. Fig.7 shows two examples of the improved behavior when adding spatial coherence to the method. We can see how the flower-texture is assigned to a single segment by sRAD, whereas RAD and pRAD assign multiple labels to this texture. The same occurs in the second row with the plants.

RAD pRAD spRAD

Fig. 7. Examples of the performance of spRAD in textured images. spRAD assigns a single segment to all the flowers. A similar effect occurs for the plants of the images shown in the second row.

7 Results and Performance Evaluation

In this section we analyze and compare RAD, pRAD, sRAD and spRAD. Firstly, RAD is compared with Mean Shift (MS) [1], [33] qualitatively and quantitatively. Secondly, the performance of RAD is compared with pRAD, sRAD and spRAD. Finally, our method is compared on the Berkeley data set against a set of state-of-the-art segmentation methods.

1. Applying boosting improved the CGE score on the Berkeley data set from 0.1738 to 0.1678. See section 7 for a further explanation about the GCE score.

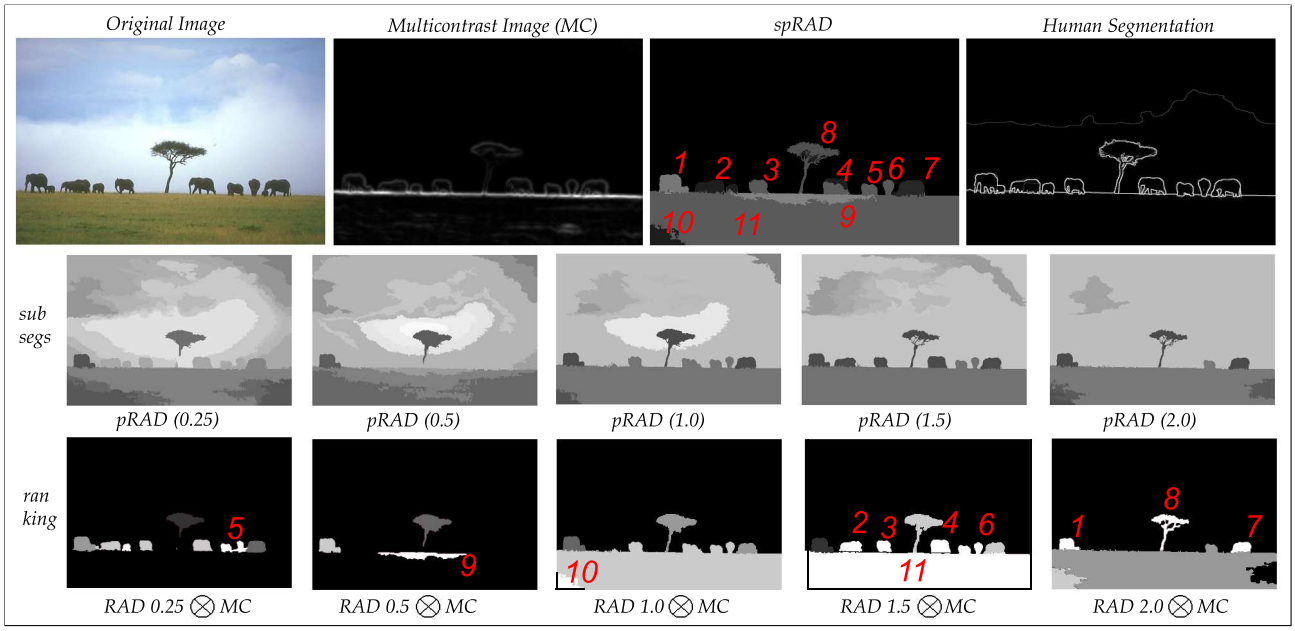


Fig. 6. spRAD segmentation: from an original image, we generate a multiscale image (MC) and a number of sub-segmentations with different parameters of pRAD (second row). The goodness of a segment is computed by summing the contrast underlying the edges of the segments normalized for the perimeter of the segment ($RAD \otimes MC$). The best segments form the combined segmentation (spRAD).

7.1 Performance evaluation of RAD

In this section we compare RAD with MS which has been chosen because it is widely used, has a public available version, the EDISON one [32] and it has demonstrated its good performance [47]. Additionally, Mean Shift is a feature space analysis technique, which uses topological contained in the histogram to perform a partitioning of the space, as well as RAD does. Furthermore, MS yields a segmentation in a rather reasonable time, in opposition to other methods such as the Graph-Based approaches [60] , (with the exception of the efficient graph-based segmentation method introduced in [37]).

The MS method [33], consists of finding the modes of the underlying probability function of a distribution. The method finds the Mean Shift vectors in the histogram of an image that point to the direction of higher density. All values of the histogram attracted by one mode compound the basis of attraction of it. In a second step, the modes which are closer than a threshold are joined into one unique mode. Finally, the basis of attraction of this mode will form a segment. Mean Shift has two basic parameters to adapt the segmentation to a specific image, namely, h_s, which controls a smoothing process, and h_r related with the size of the kernel used to determine the modes and its basis of attraction. To test the method, we have selected the set parameters $(h_s, h_r) = \{(7,3),(7,15),(7,19),(7,23),(13,7)(13,19),(17,23)\}$ given in [47] and [64]. The average times for this set of parameters, expressed in seconds, are 3.17, 4.15, 3.99, 4.07, 9.72, 9.69, 13.96 respectively. Nevertheless, these parameters do not cover the complete spectrum of possibilities of the MS. Here we want to compare RAD and MS from a soft oversegmentation to a soft undersegmentation. Hence, in order to reach an undersegmentation with MS, we add the following parameter settings $(h_s, h_r) = \{(20,25),(25,30),(30,35)\}$. For these settings, the average times are 18.05, 24.95 and 33.09 respectively.

The parameters used for RAD based segmentation are (σ_d,σ_i)={ (0.8,0.05) , (0.8,0.5) , (0.8,1) , (0.8,1.5) , (1.5,0.05) , (1.5,0.5) , (1.5,1.5) , (2.5,0.05) , (2.5,0.5) , (2.5,1.5) }. These parameters vary from a soft oversegmentation to an undersegmentation, and have been selected experimentally. The average times for RAD are 6.04, 5.99, 6.11, 6.36, 6.11, 5.75, 6.44, 5.86, 5.74 and 6.35. These average times, point out the fact that RAD is not dependent of the parameters used. In conclusion, whereas the execution time of Mean Shift increases significantly with increasing spatial scale, the execution time of RAD remains constant from an oversegmentation to an undersegmentation.

The experiments have been performed on the publicly available Berkeley image segmentation dataset and benchmark [28]. We use the Global Constancy Error (GCE) as an error measure. This measure was also proposed in [28] and takes care of the refinement between different segmentations. For a given pixel p_i, consider the segments (sets of connected pixels), S_1 from the benchmark and S_2 from the segmented image that contain this pixel. If one segment is a proper subset of the other, then p_i lies in an area of refinement and the error measure should be zero. If there is no subset relationship, then S_1 and S_2 overlap in an inconsistent manner and the error is higher than zero, (up to one in the worst possible

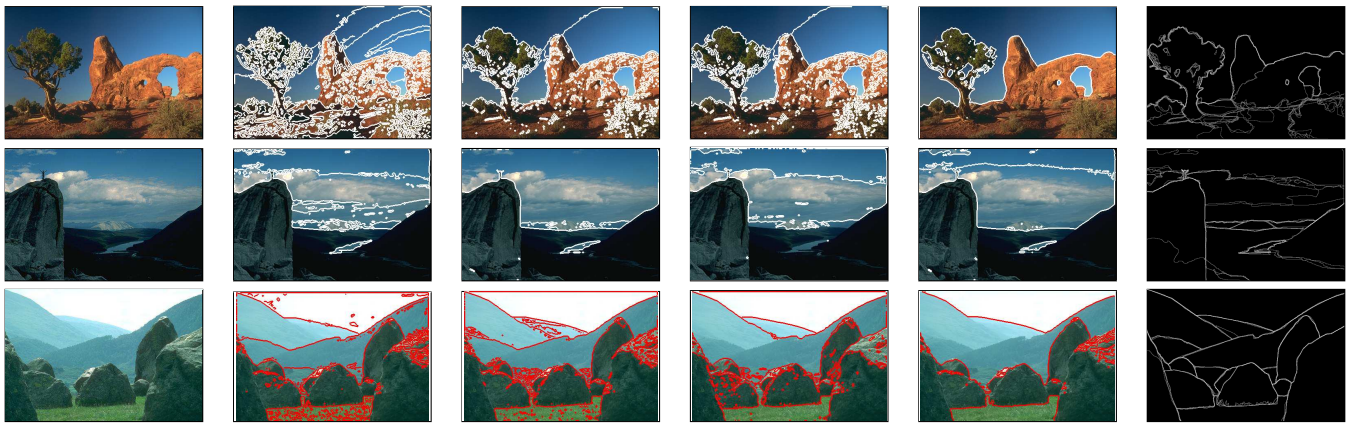


Fig. 8. Examples of segmentation. Original image. Columns from 2 to 5: segmentation for RAD on RGB with (σ_d,σ_i)={(0.8,0.05),(1.5,0.05),(2.5,0.05),(2.5,1.5)}. Last column: human segmentation.

case).

MS segmentation has been done on the CIE Luv space since this is the space used in [47] and [64]. RAD based segmentation has been done on the RGB colour space because the Berkeley image dataset does not have calibrated images and, consequently, we can not assure a good transformation from sRGB to CIE Luv. Nonetheless, for the sake of clarity, we also present some results of RAD on CIE Luv to directly compare results with MS. Furthermore, we apply gamma-correction to the images of the Berkeley dataset, assuming the images to be in sRGB.

Figure 8 depicts a set of examples for RAD on RGB. From left to right: original image, RAD for (σ_d,σ_i)={ (0.8,0.05) , (1.5,0.05) , (2.5,0.05) , (2.5,1.5) } and human segmentation. Figure 9 shows some results for the mean shift segmentation, corresponding to $(h_s, h_r) = \{(7,15),(13,19),(17,23),(20,25),(25,30),(30,35)\}$. These results point out the main advantage of RAD in favor of MS, namely, the capability of RAD to capture the MR of a histogram, whereas MS is ignorant to the physical processes underlying the structure of the MRs as Abd-Almageed and S. Davis explain in [49]. The set of images depicted in the first row of Figure 9, shows this behavior for a practical case. In the last column, MS joins rocks with the mountain, and the mountain with the sky, but is not able to find one unique structure for rock or mountain, whereas RAD, as shown in Figure 8, is able.

Additional examples related to the presence of physical effects, such as shadows, shading and highlights are shown in Figure 10. The good performance of RAD in these conditions can be clearly observed for the skin of the people, the elephants and buffaloes, as well as for the clothes of the people.

The histogram of the mean GCE values versus the percentage of images for each GCE value are shown in Figs. 11a,b for RAD on RGB and MS respectively. As more bars are accumulated on the left, the better is the method. Figures 11c,d show the standard deviation (blue line) along with the maximum and the minimum GCE values (red lines) for each of the 10 sets of parameters for RAD on RGB and MS. Note that the behaviour of both methods in this sense is almost the same. A low and similar standard deviation along all parameters means that the method has a stable behaviour. Figure, 11e depicts the mean GCE index for each image ordered by increasing index for MS (green), RAD on RGB (black) and RAD on Luv (red). This plot shows, not only the good performance of RAD, but also that RAD has a similar behavior on RGB and CIE Luv spaces, even with the aforementioned potential problems on Luv. Figure 11f plots the GCE index differences for each image between RAD on RGB and MS. Values lower than zero indicate the number of images where RAD performs better than MS. The same but for RAD on Luv versus MS, and RAD on RGB versus RAD on Luv is depicted on Figure 11g,h.

Fig. 10. Examples of segmentation in presence of shadows and highlights.

7.2 Results obtained with pRAD, sRAD and spRAD

As explained before, the addition of physical-based prior knowledge requires to select a proper value for λ. We obtained that the best value for λ is 0.33. This prior knowledge is added following equations 14 and 15. The addition of the prior knowledge aims to favor ridges following the statistically expected directions of a surface reflectance (mainly achromatic changes), at the same time that we suppress those ridges formed by different surface reflectances. These effects can be observed in Fig. 7, first row. Whereas RAD joins the purple flowers with

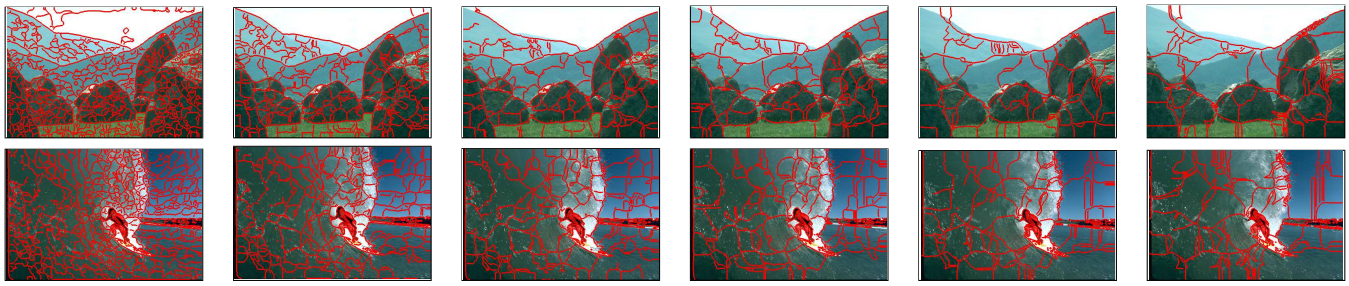


Fig. 9. MS segmentation examples for different parameters. Columns from 1 to 5: $(h_s, h_r) = \{(7,15),(13,19),(17,23),(20,25),(25,30)\}$.

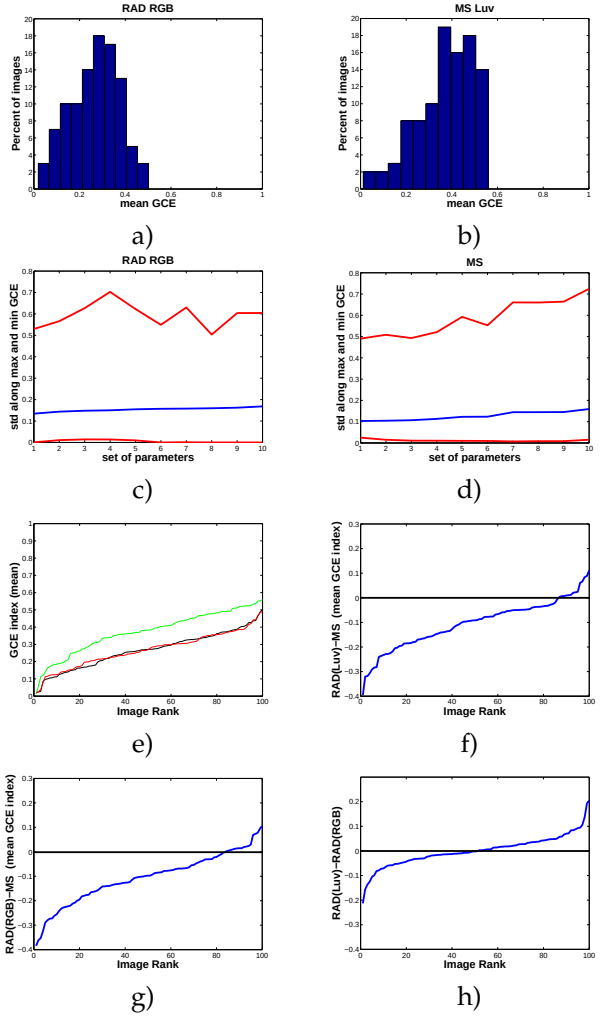


Fig. 11. (a,b) Mean GCE values for each set of parameters. (c,d) Standard deviation of GCE along maximum and minimum values for each set of parameters. (e) Mean GCE values for each image sorted form lowest to highest. (f) Comparison of RAD Luv with MS. Values higher than zero: images where MS performs better. (g,h) The same as f) but for MS and RAD Luv and for RAD RGB versus RAD Luv.

the grass, pRAD correctly finds a ridge for purple flowers and another for the grass. More qualitative examples are showed in Fig. 12. First row: pRAD is able to find a segment for the gray little stones in the top-left part of the images. Second and fourth rows: in both cases, pRAD finds a single surface reflectance for all the rocks, whereas RAD clearly oversegment these rocks.

As a second adaptation to RAD we propose sRAD which uses the information contained in the image to yield a segmentation enhancing the multicontrast of the image. It can be performed using subsegmentations generated either by RAD (then the method is called sRAD), or by pRAD (then the method is called spRAD). This segmentation is less affected by textures, since they have a weak effect in a multiscale analysis. Fig. 7 shows two clear examples of it. We can see how, RAD segments incorrectly the red and yellow flowers and oversegment the grass on the second row. pRAD find better segments in both cases, but still with an oversegmentation. spRAD, instead, find a single segment for the red flowers, the purple ones, and the floor of the second row, that is, produces a non-oversegmented images. In the examples showed in Fig. 12 the same effect can be observed. For instance, in the second row, we can see how spRAD is able to find a single segment for the trees and segment for the rocks. In the third row it is showed an example of the improvement achieved by combining multiple segmentations. RAD and pRAD, can not find a segment for every mountain due to a clear blurring effect, whereas spRAD produces a better segmentation. We point out that results obtained with sRAD are worse than the ones obtained with spRAD, as can be observed in these examples. This is the expected result, since the subsegmentations used by spRAD are better than the ones used by sRAD.

Quantitative results using the GCE score are presented in section 7.3.

7.3 Comparison to State of the Art

In this section we show more quantitative results obtained with our segmentation method. Table 1 shows GCE values for several state-of-the-art methods. These values are taken from [55] and [67]. These experiments have been performed using the train set of 200 images. For both RAD and MS we present the results obtained with the best parameter settings. For RAD, best results were obtained with (σ_d,σ_i)={(2.5,0.05)}. The

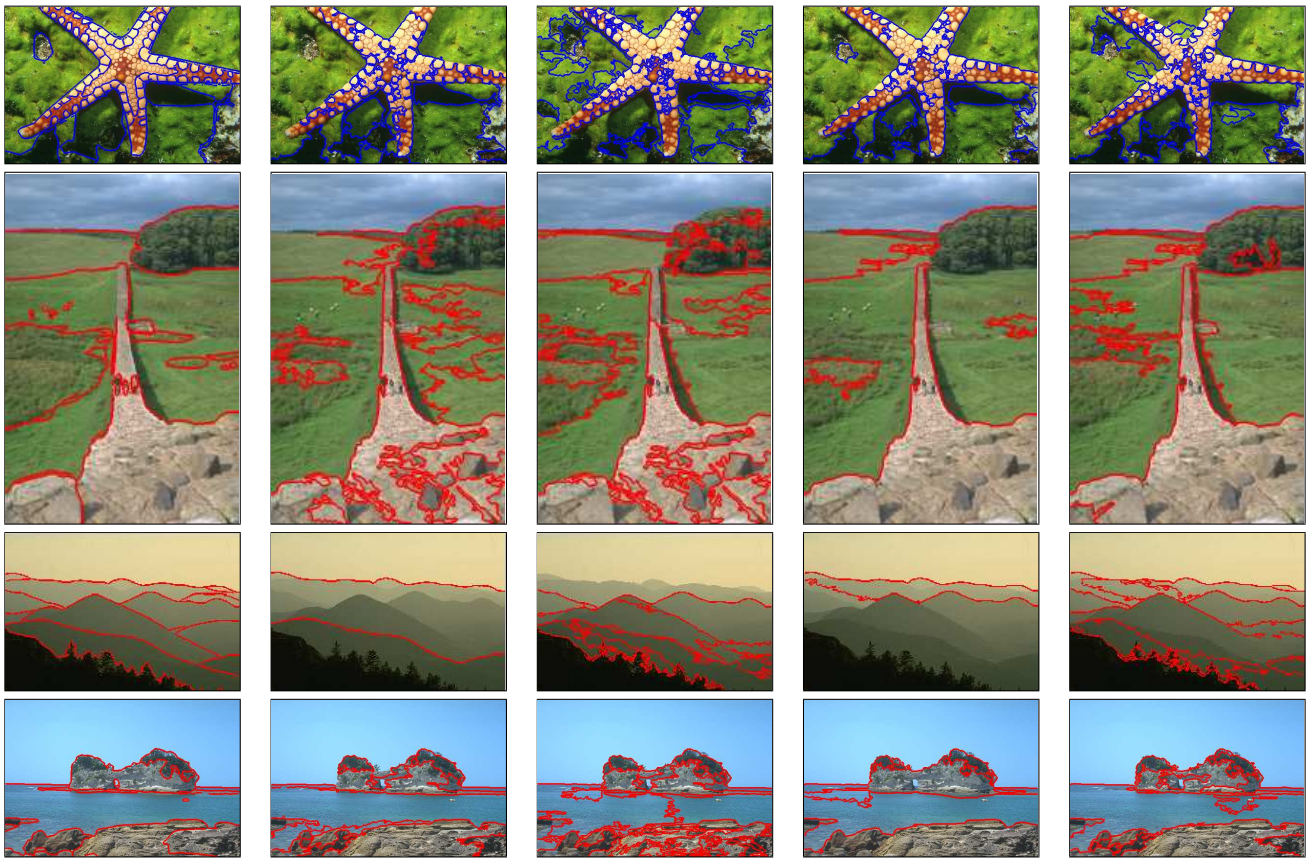


Fig. 12. Examples of segmentation. First column: From first to last column, respectively: Human segmentation, RAD, sRAD, pRAD and spRAD. It can be observed that when adding spatial coherence, the segmentations have a closer similarity with human segmentation.

TABLE 1
Global Constancy Error for several state-of the-art methods: seed [55], fow [36], MS, and nCuts [25]. Values taken from [55] and [64].

	human	spRAD	pRAD	sRAD	RAD	seed	fow	MS	nCuts
GCE index	0.080	0.1678	0.1780	0.1860	0.2048	0.209	0.214	0.2598	0.336

mean number of SR found using RAD had been 5, but it is not directly translated in 5 segments on segmented images. Often, some segments of few pixels appear due to chromaticity of surfaces as can bee seen in figure 4h. CGE evaluation favors oversegmentation [28]. Hence, to make feasible a comparison with other methods using GCE, we have performed the segmentation without considering segments of an area lower than 2% of the image area. In this case, the mean number of segments for the 200 test images is 6.98 (7 segments). The number of segments for the other methods varies from 5 to 12. The same occurs with pRAD. Finally, for RAD and spRAD we show results obtained by generating a combined segmentation having 9 segments. Furthermore, we stand out that results obtained with spRAD, outperform all results obtained with its sub-segmentations. These sub-segmentations, have GCE values going from 0.1780 to 0.2205.

As can be seen our method obtains the best results. Furthermore, it should be noted that the method is substantially faster than the $seed$ and the $nCuts$ [25] method. In addition, the results obtained with the MS need an additional step. Namely, a final combination step, which requires a new threshold value, is used to fuse adjacent segments in the segmented image if their chromatic difference is lower than the threshold (without pre- an postprocessing MS obtains a score of 0.2972).

Finally, when comparing the different versions of RAD, we can see how, each of them improve in a coherent way the results obtained with the basic version of RAD. Thus, we can see how pRAD clearly outperforms results obtained with RAD, at the same computational cost. It makes pRAD, the best version when looking for a fast method of segmentation. Further, spRAD outperforms all the other methods. Nonetheless, its computational cost is much higher, since it computes five

subsegmentation, a multicontrast image and a ranking of all the segments obtained.

8 Conclusions and Future Work

This paper introduces a new segmentation method, called pRAD, that extracts the ridges formed by a surface reflectance. This method is robust against discontinuities appearing in image histograms due to compression and acquisition conditions. Furthermore, those strong discontinuities, related with the physical illumination effects are correctly treated due to the topological treatment of the histogram and the addition of prior knowledge. As a consequence, the presented method yields better results than Mean Shift on a widely used image dataset and error measure. Additionally, even with neither preprocessing nor postprocessing steps, pRAD has a better performance than the state-of-the-art methods. Furthermore, we have proposed an improvement of pRAD, called spRAD, consisting in the addition of the spatial coherence to be less affected by texture edges and avoiding oversegmentation. spRAD outperforms results obtained with pRAD but at higher computational cost. Results obtained with pRAD point out that the chromatic information is an important cue on human segmentation. Additionally, the elapsed time for pRAD is not affected by its parameters. Due to that it becomes a faster method than Mean Shift and the other state-of-the-art methods.

For future work, using the fact that RAD is a fast segmentation method, we are specially interested to apply it for the computation of superpixels. Superpixels are used as a prior step in many high-level computer vision tasks, such as object recognition. Controlling oversegmentation using parameter setting could give us a starting point to superpixel processing. Moreover the extension of RAD to video sequences could yield interesting results since the RAD representation can handle with temporal variations due to camera movements, slight illuminant changes (as expected between subsequent frames) and moving shadows. This extension can be driven by a fifth dimension or by spatio-temporal coherence of extracted ridges.

Acknowledgments

This work has been partially supported by projects TIN2007-64577, TIN2009-14173 and Consolider-Ingenio 2010 CSD2007-00018 of Spanish MEC (Ministery of Science), the Marie Curie Reintegration grant (ERG 224737) of the European Union, Generalitat de Catalunya (2009 SGR669) and the Ramon y Cajal Program.

References

[1] Fukunaga, K., Hostetler, L.D.: The estimation of the gradient of a density function, with applications in pattern recognition. Information Theory, IEEE Transactions on **121**(1) (1975) 32–40

[2] Y. Ohta and Takeo Kanade and T. Sakai: Color Information for Region Segmentation. Computer Graphics and Image Processing **13**(3) (1980) 222–241

[3] Shafer, S.: Using color to seperate reflection components. Color research and application **10**(4) (1985) 210–218

[4] Gath I. and Geva A. B.: Unsupervised Optimal Fuzzy Clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence **11**(7) (1989) 773–780

[5] Bajcsy, R. and Lee, SW and Leonardis, A.: Color image segmentation with detection of highlights and localillumination induced by inter-reflections. 10th International Conference on Pattern Recognition **1** (1990) 785–790

[6] Klinker, G., Shafer, S.: A physical approach to color image understanding. Int. Journal of Computer Vision **4** (1990) 7–38

[7] Pavlidis, T. and Liow, Y.T.: Integrating Region Growing and Edge Detection. **12**(3) (1990) 225–233

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

[9] Pappas, T. N.: An adaptive clustering algorithm for image segmentation. IEEE Transactions on Signal Processing **40**(4) (1992) 901–914

[10] Gauch, J.M., Pizer, S.M.: Multiresolution analysis of ridges and valleys in grey-scale images. IEEE Transactions on Pattern Analysis and Machine Intelligence **15**(6) (1993) 635–646

[11] Wang, L., Pavlidis, T.: Direct gray-scale extraction of features for character recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence **15**(10) (1993) 1053–1067

[12] Bouman, CA and Shapiro, M.: DA multiscale random field model for Bayesian image segmentation. IEEE Transactions on Image Processing **3**(2) (1994) 162–177

[13] Skarbek, W., Koschan, A.: Colour image segmentation — a survey. Technical report, Institute for Technical Informatics, Technical University of Berlin (October 1994)

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

[15] Maxwell, B.A. and Shafer, S.A.: Physics-based segmentation of complex objects using multiple hypotheses of image formation. Computer Vision and Image Understanding **65**(2) (1997) 269–295

[16] Haris, K. and Efstratiadis, SN and Maglaveras, N. and Katsaggelos, AK: Hybrid image segmentation using watersheds and fast region merging. IEEE Transactions on Image Processing **7**(12) (1998) 1684–1699

[17] Itti, L. and Koch, C. and Niebur, E.: A Model of Saliency-Based Visual Attention for Rapid Scene Analysis. IEEE Transactions On Pattern Analysis And Machine Intelligence **20**(11) (1998) 1254–1259

[18] Ong, C.K. and Matsuyama, T.: Robust color segmentation using the dichromatic reflection model. 14th International Conference on Pattern Recognition **1** (1998) 780–784

[19] O. Chapelle, P. Haffner, V. Vapnik: Support Vector Machines for Histogram-Based Image Classification. IEEE Trans. on Neural Networks **10**(5) (1999) 1055–1064

[20] López, A.M., Lumbreras, F., Serrat, J., Villanueva, J.J.: Evaluation of methods for ridge and valley detection. IEEE Transactions on Pattern Analysis and Machine Intelligence **21**(4) (1999) 327–335

[21] Schmid P.: CSegmentation of digitized Dermatoscopic Images by Two-Dimensional Color Clustering. IEEE Trans. on Medical Imaging **18**(2) (1999) 164–171

[22] Takahashi K. and Abe K.: Color Image Segmentation Using ISO-DATA Clustering Algorithm. rans. of the Institute of Electronics, Information and Communication Engineers **J82D-II**(4) (1999) 751–762

[23] Berens J. and Finlayson G.D.: Log-Opponent Chromaticity Coding of Color Space. Pattern Recognition, 2000. Proceedings. 15th International Conference on **01** (2000) 206–211

[24] Cheng, H., Jiang, X., Sun, Y., Wang, J.: Color image segmentation:advances and prospects. Pattern Recognition **34**(6) (2001) 2259–2281

[25] Shi, J., Malik, J.: Normalized cuts and image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **22**(8) (2000) 888–905

[26] Deng Y. and Manjunath B.S.: Unsupervised segmentation of Color-Texture Regions in Images and Video. IEEE Transactions on Pattern Analysis and Machine Intelligence **23**(8) (2001) 800–810

[27] Lucchese, L., Mitra, S.: Color image segmentation: A state-of-the-art survey. In: INSA-A: Proceedings of the Indian National Science Academy. (2001) 207–221

[28] 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

[29] Verma, D., Meila, M.: A comparison of spectral clustering algorithms. technical report uw-cse-03-05-01, university of washington.
[30] Bishnu, A., Bhowmick, P., Dey, S., Bhattacharya, B.B., Kundu, M.K., Murthy, C.A., Acharya, T.: Combinatorial classification of pixels for ridge extraction in a gray-scale fingerprint image. In: ICVGIP. (2002)
[31] 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
[32] Christoudias, C., Georgescu, B., Meer, P.: Synergism in low level vision. International Conference on Pattern Recognition **4** (2002) 150–155
[33] 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
[34] Egmont-Petersen, M. and de Ridder, D. and Handels, H.: Mean smage processing with neural networksa review. Pattern Recognition **35**(10) (2002) 2279–2301
[35] Freixenet, J., Munoz, X., Raba, D., Mart, J., Cuf, X.: Yet another survey on image segmentation: Region and boundary information integration. In: ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part III, London, UK, Springer-Verlag (2002) 408–422
[36] Fowlkes, C., Martin, D., Malik, J.: Learning affinity functions for image segmentation: combining patch-based and gradient-based approaches. Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on **2** (2003)
[37] Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. Intl. Journal of Computer Vision **59**(2) (2004)
[38] Kim, D.W. and Lee, K.H. and Lee, D.: A novel initialization scheme for the fuzzy c-means algorithm for color clustering. Pattern Recognition Letters **25**(2) (2004) 227–237
[39] Nikolaev, D.P. and Nikolayev, P.P.: Linear color segmentation and its implementation. Computer Vision and Image Understanding **94**(1-3) (2004) 115–139
[40] Omer, I. and Werman, M.: Color Lines: Image Specific Color Representation. IEEE Computer Society Conference on Computer Vision and Pattern Recognition **2** (2004) 946–953
[41] Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative J. Electron. Imaging **13**(1) (2004) 146 – 165
[42] FJ. Staal and M. D. Abramoff and M. Niemeijer and M. A. Viergever and B. van Ginneken: EffRidge-based vessel segmentation in color images of the retina IEEE Transactions on Medical Imaging **23**(4) (2004) 501–509
[43] Alaa E. Abdel-Hakim and Aly A. Farag: Color segmentation using an Eigen color representation Information Fusion, 2005 8th International Conference on **2**(2-3) (2005) 1576–1583
[44] Agarwal, S., Madasu, S., Hanmandlu, M., Vasikarla, S.: A comparison of some clustering techniques via color segmentation. In: ITCC '05: Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II, Washington, DC, USA, IEEE Computer Society (2005) 147–153
[45] Cates, J.E. and Whitaker, R.T. and Jones, G.M.: Case study: an evaluation of user-assisted hierarchical watershed segmentation. Medical Image Analysis **9**(6) (2005) 566–578
[46] Geusebroek, J.M., Burghouts, G.J., Smeulders, A.W.M.: The Amsterdam library of object images. Int. J. Comput. Vision **61**(1) (2005) 103–112
[47] Pantofaru, C. and Hebert, M.: A comparison of image segmentation algorithms. Technical Report CMU-RI-TR-05-40, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (September 2005)
[48] Shih, F.Y. and Cheng, S.: Automatic seeded region growing for color image segmentation. Image and Vision Computing **23**(10) (2005) 877–886
[49] Abd-Almageed, W. and Davis, L.: Density Estimation Using Mixtures of Mixtures of Gaussians. 9thEuropean Conference on Computer Vision (2006)
[50] Chuang, K.S. and Tzeng, H.L. and Chen, S. and Wu, J. and Chen, T.J.: Fuzzy c-means clustering with spatial information for image segmentation. Computerized Medical Imaging and Graphics **30**(1) (2006) 9–15
[51] 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
[52] Hanbury, A. and Marcotegui, B.: Waterfall Segmentation of Complex Scenes. Lecture Notes in Computer Science **3851** (2006) 888
[53] Kato, Z. and Pong, T.C.: A Markov random field image segmentation model for color textured images. Image and Vision Computing **24**(10) (2006) 1103–1114
[54] 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
[55] Micusık, B., Hanbury, A.: Automatic image segmentation by positioning a seed. Proc. European Conference on Computer Vision (ECCV) (2006)
[56] Pichel, J.C. and Singh, D.E. and Rivera, F.F.: mage segmentation based on merging of sub-optimal segmentations. Pattern recognition letters **27**(10) (2006) 1105–1116
[57] 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
[58] Cremers, D. and Rousson, M. and Deriche, R.: A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape. International Journal of Computer Vision **72**(2) (2007) 195–215
[59] 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
[60] Ge, F., Wang, S., Liu, T.: New benchmark for image segmentation evaluation. Journal of Electronic Imaging **16** (2007) 033011
[61] Liu, T. and Sun, J. and Zheng, N.N. and Tang, X. and Shum, H.Y.: Learning to Detect A Salient Object. Proc. of IEEE Conference on Computer and Vision Pattern Recognition (2007)1–8
[62] Ma, Y. and Derksen, H. and Hong, W. and Wright, J.: Segmentation of Multivariate Mixed Data via Lossy Data Coding and Compression. IEEE Transactions on Pattern Analysis and Machine Intelligence **29**(9) (2007) 1546–1562
[63] Vazquez, E., Baldrich, R., Vazquez, J., Vanrell, M.: Topological histogram reduction towards colour segmentation. In: 4477-0055 - LNCS - Pattern Recognition and Image Analysis. (2007) 55–62
[64] Yang, Y., Wright, J., Sastry, S., Ma, Y.: Unsupervised segmentation of natural images via lossy data compression. (2007)
[65] Heidemann, G.: Color segmentation robust to brightness variations by using B-spline curve modeling. Image and Vision Computing **26**(2) (2008) 211–227
[66] Kim, C. and You, B.J. and Jeong, M.H. and Kim, H.: Color segmentation robust to brightness variations by using B-spline curve modeling. Pattern Recognition **41**(1) (2008) 22–37
[67] 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
[68] Wattuya, P. and Rothaus, K. and Prassni, J.S. and Jiang, X.: A random walker based approach to combining multiple segmentations. Proc. International Conference on Pattern Recognition (2008) 1–4

Eduard Vazquez is a Teaching Assistant in the Universitat Autonoma de Barcelona (UAB). He received his MSc in the UAB in 2007. He is currently finishing his PhD in Computer Science in the Computer Vision Center (UAB). His research interests are colour, segmentation, saliency, perception and medical imaging with an special emphasis on multidisciplinary approaches.

Ramon Baldrich is an Associate Professor in the Computer Science Department of the Universitat Autonoma de Barcelona and is attached to the Computer Vision Center as a researcher. He received his Phd in Computer Science from the Unviversitat Autonoma de Barcelona in 2001. His research interest is mainly focused in colour treatment in computer vision problems, including colour segmentation, colour induction and image shadows, as well as technological transfer.

Joost van de Weijer received his MSc (Delft University of Technology) in 1998 and his PhD (University of Amsterdam) in 2005. He was Marie Curie Intra-European fellow in INRIA Rhone-Alpes. He was awarded the Ramon y Cajal Research Fellowships in Computer Science by the Spanish Ministry of Science and Technology. He is based in the Computer Vision Center in Barcelona (Spain) since 2008.

Maria Vanrell is an Associate Professor of Artificial Intelligence in the Computer Science Department of the Universitat Autnoma de Barcelona (UAB). In 1996, she received her PhD degree at the UAB for her work on texture analysis and representation. She is an affiliate researcher at the computer Vision Center, where she leads several industrial projects and coordinates the Color Research group. She is also an active member of the Image Analysis and Pattern Recognition Spanish Association (IAPR). Her current research interests are computational models of color perception from a multidisciplinary approach.