

Image Segmentation in the Presence of Shadows and Highlights

Eduard Vazquez, Joost van de Weijer, and Ramon Baldrich

Computer Vision Center \ Dpt. Ciències de la Computació. Edifici O Universitat Autònoma de Barcelona. 08193 Cerdanyola del Valles, Barcelona, Spain

Abstract. The segmentation method proposed in this paper is based on the observation that a single physical reflectance can have many different image values. We call the set of all these values a dominant colour. These variations are caused by shadows, shading and highlights and due to varying object geometry. The main idea is that dominant colours trace connected ridges in the chromatic histogram. To capture them, we propose a new Ridge based Distribution Analysis (RAD) to find the set of ridges representative of the dominant colour. First, a multilocal crease-ness technique followed by a ridge extraction algorithm is proposed. Afterwards, a flooding procedure is performed to find the dominant colours in the histogram. Qualitative results illustrate the ability of our method to obtain excellent results in the presence of shadow and highlight edges. Quantitative results obtained on the Berkeley data set show that our method outperforms state-of-the-art segmentation methods at low computational cost.

1 Introduction

Image segmentation is a computer vision process consisting in the partition of an image into a set of non-overlapped regions. A robust and efficient segmentation is required as a preprocessing step in several computer vision tasks such as object recognition or tracking. On real images the varying shapes of the objects provoke several effects related with the illumination such as shadows, shading and highlights. These effects, are one of the main difficulties that have to be solved to obtain a correct segmentation.

There exist several different methods covering a broad spectrum of points of view. The work presented by Skarbek and Koschan [1], draws the basis of the current classifications of segmentation methods. Some other comprehensive surveys of colour segmentation techniques are presented in [2] and [3], where a similar schema is followed. From these works segmentation methods are divided in four main categories: feature-based, image-based, physics-based and hybrid approaches. Feature-based approaches are focused on the photometric information of an image represented on its histogram [4],[5]. Image-based approaches exploit the spatial information of the colour in an image, named *spatial coherence* [6]. Physics-based methods use physics and psychophysics information to

perform the segmentation. Finally, hybrid techniques combine methods of the previous categories.

This paper introduces a method that exploits exclusively the photometric information of an image on its histogram. Therefore, it belongs to the category of feature-based segmentation methods. This category can be further split in three main categories, i.e., 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 [7]. Clustering techniques, also named hard clustering, perform a partition of the feature space under different criteria such a distance measure as k-means or ISODATA, probabilistic/statistical approaches, such as Mean Shift [8], or the spectral analysis of the data [9], based on the Karhunen-Loeve transformation. Fuzzy clustering includes methods such as fuzzy k-means, Gath-Geva clustering, or mixture models [10], [11] which are a way to look for areas of high density. The most related technique with the one introduced in this paper, is the Mean shift, which will be commented and compared with our method in section 4. Each technique has its own advantages and drawbacks. There is a difficulty shared between all these methods, i.e., their behaviour in the presence of shadows, shading, and highlights. Furthermore, the work presented by Martin *et al.* in [12], points out the existence of strong edges related with this physical effects in an image that are not considered in a human segmentation. These edges are detected for both image and feature based methods.

Our approach. 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 dominant colours (DC), where each dominant colour is described by a distribution in histogram-space. Each DC is related to a semantic object in the image. For example, in Figure 1 we distinguish between four different DC's, namely: red for the pepper, green and brown for the branch and black for the background.

A DC generates many image values due to geometrical and photometric variations. Our main aim is to find a good representation of the topologies which DC's are likely to form in histogram space. For this purpose, consider the distribution of a single DC as described by the dichromatic reflection model [13]:

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

in which $\mathbf{f} = \{R, G, B\}$, \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 DC 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.

The two parts of the dichromatic reflectance 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 DC forms a ridge-like structure in histogram space.

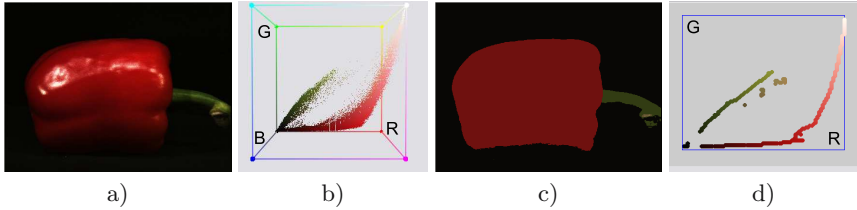


Fig. 1. (a) An image from [14] and (b) its histogram. The effects of shading and highlights are clearly visible in the red colours of the histogram. (c) Segmented images using RAD. (d) Ridges found with RAD. Note that the three branches of the red pepper are correctly connected in a single ridge.

To illustrate the difficulty of extracting the distributions of DCs 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 DC: 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.

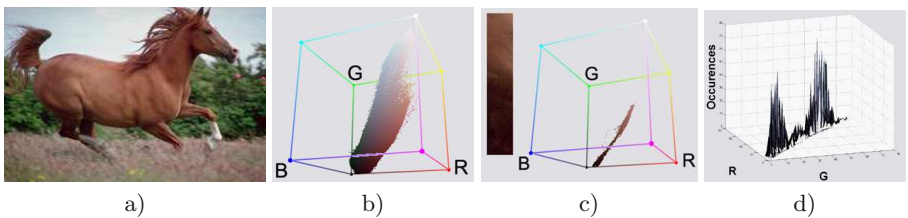


Fig. 2. (a) An image and (b) its 3D RGB histogram. (c) A patch of a) and its RGB histogram. (d) 2D histogram of c) to illustrate the discontinuities appearing on a DC.

In literature several methods have explicitly used the dichromatic model to obtain image segmentation, e.g. [15]. 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. 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.

This paper is organized as follows: in section 2 RAD is presented as a feature space analysis method. Afterwards, in section 3 RAD is introduced as a segmentation technique. The results obtained and a comparison with Mean Shift and various other state-of-the-art methods on the Berkeley dataset is presented in section 4. Finally, conclusions of the current work are given in section 5.

2 A Ridge Based Distribution Analysis Method (RAD)

In this section we present a fast algorithm to extract DCs from histogram space. Here we propose a method to find dominant structures (DS) for a d -dimensional feature space. In the context of this paper the dominant colours are the dominant structures of the 3D chromatic histogram. The proposed method is divided in two main steps. First, we propose a method to extract ridges as a representative of a DS. Afterwards a flooding process is performed to find the DSs from its ridges.

2.1 First Step: Ridge Extraction

To extract a DS descriptor we need to find those points containing the most meaningful information of a DS, *i.e.*, the ridge of a DS. We propose to apply a multilocal creaseness algorithm to find the best ridge point candidates. This operator avoids to split 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.

Multilocal Creaseness: Finding Candidates and Enhancing Connectivity. In order to deal with this commonly heavily jagged DS (see Figure 2d), we propose to apply the MLSEC-ST operator introduced by Lopez *et al.* in [16] to enhance ridge points. This method is used due to its good performance compared with other ridge detection methods [16] 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 ST field S is defined as:

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

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 $\bar{w}(\mathbf{x}, \sigma)$ in a neighbourhood of size proportional to σ_i centered at \mathbf{x} is:

$$\bar{w}(\mathbf{x}, \sigma) = \mathbf{sign}(\mathbf{w}^t(\mathbf{x}, \sigma) \cdot \nabla^t\Omega(\mathbf{x}, \sigma_d))\mathbf{w}(\mathbf{x}, \sigma) \quad (3)$$

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, namely n_k , on the r -connected neighbourhood of size proportional to σ_i . That is:

$$k(\mathbf{x}, \sigma) = -\text{Div}(\bar{\mathbf{w}}(\mathbf{x}, \sigma)) = -\frac{d}{r} \sum_{k=1}^r \bar{\mathbf{w}}^t(\mathbf{k}, \sigma) \cdot \mathbf{n}_k \quad (4)$$

where d is the dimension of $\Omega(\mathbf{x})$. The creaseness representation of $\Omega(\mathbf{x})$ will be referred hereafter as Ω^σ .

As an example, Figure 3a shows the opponent colour 2D histogram of 3g. Its creaseness values are showed in 3b. There are three enhanced areas which corresponds with the three dominant colours of the original image. They appear as three mountains in 3b, clearly separated by two valleys. Note that higher creaseness values have a larger probability to become a ridge point.

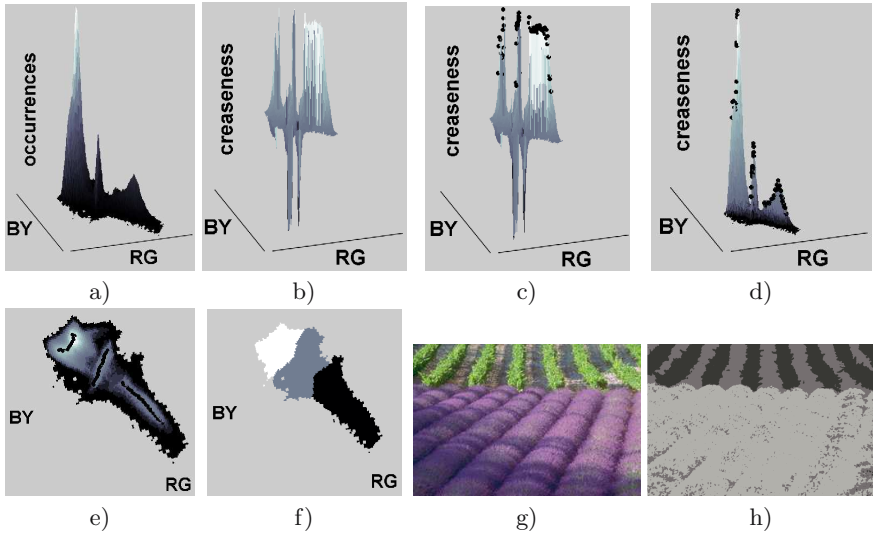


Fig. 3. A graphical example of the whole process. (a) Opponent Red-Green and Blue-Yellow histogram $\Omega(\mathbf{x})$ of g . (b) Creaseness representation of a). (c) Ridges found in b). (d) Ridges fitted on original distribution. (e) Top-view of d). (f) Dominant structures of a). (g) Original image. (h) Segmented image.

Ridge Detection. In the previous section we have detected a set of candidate ridge points. In this section we discard superfluous points. As a result only those points necessary to maintain the connectivity of a DS 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 $\Omega(x, y)$ be a continuous 2D surface and $\nabla\Omega(x, y)$ be the gradient vector of the function $\Omega(x, y)$. 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 $|\lambda_1| \leq |\lambda_2|$. Then, for the 2D case:

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

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

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

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, $RP(\Omega(x, y)) = LMP \cup TRP \cup SP$. A further classification of ridges and its singularities can be found in [17] and [18].

A common way to detect RP 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 [18]. 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. A further explanation can be found in [19].

Figure 3c depicts the RP found on Ω^σ with black dots. Figures 3d,e show a 3D view and a 2D projection view respectively of how these RPs fit in the original distribution as a representative of the three DS. Finally, from the set of RPs of a distribution we can perform the calculus of each DS. 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.

2.2 Second Step: DS Calculus from Its RPs

In this final step we find the DS 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 [20] 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 RPs found and then, DSs are found using the algorithm proposed in [20] applied on the inverse Ω^σ distribution. The proposed algorithm, is not based on the gradient vectors of a landscape [21] 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 [16], we force the flooding process to begin at the same time in all DS descriptors, on the smoothed $\Omega(\mathbf{x})$ distribution with a Gaussian kernel of standard deviation σ_d (already computed on the ST calculus). Then, we define RAD as the operator returning the set of DS of Ω^σ using RPs as marks:

$$RAD(\Omega(x)) = W(\Omega^\sigma, RP(\Omega^\sigma)) \quad (8)$$

Following this procedure, Figure 3f depicts the 2D projection of the DSs found on 3a.

3 Colour Image Segmentation Using RAD

Once RAD has been applied we need to assign a representative colour to each DS found. Thus, let $DS_n = \{\mathbf{x}_1, \dots, \mathbf{x}_r\}$ be the n th DS of $\Omega(\mathbf{x})$, and $\Omega(\mathbf{x}_i)$ the function returning the number of occurrences of \mathbf{x}_i in Ω . Then, the dominant colour of DS_n , namely, $DC(DS_n)$ will be the mass center of $\Omega(DS_n)$:

$$DC(DS_n) = \frac{\sum_{i=1}^r \mathbf{x}_i \cdot \Omega(\mathbf{x}_i)}{\sum_{i=1}^r \Omega(\mathbf{x}_i)} \quad (9)$$

The segmented image will have as many colours as the number of DSs found. Figure, 3h shows the segmentation obtained with RAD from 3g. This segmentation has been performed in the opponent colour histogram. Although RAD can be applied to any chromatic representation of an image such as CIE, RGB, Ohta spaces or 2-dimensional ones such as Opponent or normalized RGB.

4 Results and Performance Evaluation

In the experiments we qualitatively and quantitatively evaluate the proposed segmentation method. Firstly, RAD is compared with Mean Shift (MS) [8], [22]. MS has been chosen because it is widely used, has a public available version, the EDISON one [23] and it has demonstrated its good performance [24]. Additionally, Mean Shift is a feature space analysis technique, as well as RAD, and yields a segmentation in a rather reasonable time, in opposition to other set of methods such as the Graph-Based approaches [25], (with the exception of the efficient graph-based segmentation method introduced in [26]). Secondly, our method is compared on the Berkeley data set against a set of state-of-the-art segmentation methods.

The MS method [22], 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 near of a given threshold are joined

in one unique mode. Finally, the basis of attraction of these modes will compose a dominant colour of the image. Mean Shift has two basic parameters to adapt the segmentation to an specific problem, 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 [24] and [5]. 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 $(\sigma_d, \sigma_i) = \{(0.8, 0.05), (0.8, 0.5), (0.8, 1), (0.8, 1.5), (1.5, 0.05), (1.5, 0.5), (1.5, 1.5), (2.5, 0.05), (2.5, 0.5), (2.5, 1.5)\}$. 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 has been performed on the publicly available Berkeley image segmentation dataset and benchmark [12]. We use the Global Constancy Error (GCE) as an error measure. This measure was also proposed in [12] 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

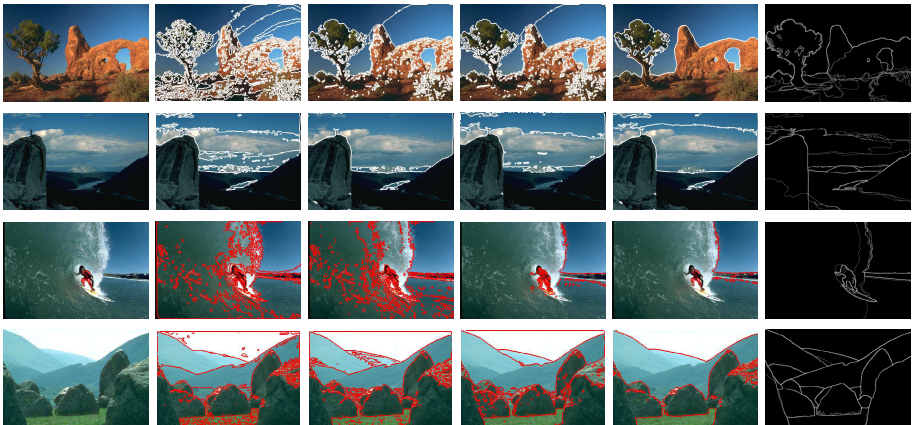


Fig. 4. Examples of segmentation. Original image. Columns from 2 to 5: segmentation for RAD on RGB with $(\sigma_d, \sigma_i) = \{(0.8, 0.05), (1.5, 0.05), (2.5, 0.05), (2.5, 1.5)\}$. Last column: human segmentation.

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 case). MS segmentation has been done on the CIE Luv space since this is the space used in [24] and [5]. RAD based segmentation has been done on the RGB colour space for two reasons. First, the Berkeley image dataset does not have calibrated images and, consequently, we can not assure a good transformation from sRGB to CIE Luv. Second, because the size of L, u and v, is not the same and the method will require six parameters, instead of two, that is, $\vec{\sigma}_L$, $\vec{\sigma}_u$ and $\vec{\sigma}_v$. Nonetheless, for the sake of clarity, we also present some results of RAD on CIE Luv to directly compare results with MS.

Figure 4 depicts a set of examples for RAD on RGB. From left to right: original image, RAD for $(\sigma_d, \sigma_i) = \{ (0.8, 0.05), (1.5, 0.05), (2.5, 0.05), (2.5, 1.5) \}$ and human segmentation. Figure 5 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 DS of a histogram, whereas MS is ignorant to the physical processes underlying the structure of the DSs as Abd-Almageed and S. Davis explain in [10]. Graphically, the set of images depicted in the first row of Figure 5, shows this behavior in 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 a rock or for the mountain, whereas RAD, as shown in Figure 4, is able to do.

A danger of RAD is that for some parameter settings it is prone to undersegmenting. Consequently it finds only one dominant colour for the whole image. This happens in some cases for $(\sigma_d, \sigma_i) = \{ (2.5, 1), (2.5, 1.5) \}$, as Figure 6 illustrates. In the first example, the aircraft has a bluish colour similar to the sky, as well as the fish and its environment in the second example.

Additional examples related to the presence of physical effects, such as shadows, shading and highlights are shown in Figure 7. The good performance of RAD in these conditions can be clearly observed for the skin of the people, the elephants and buffalos, as well as for the clothes of the people.

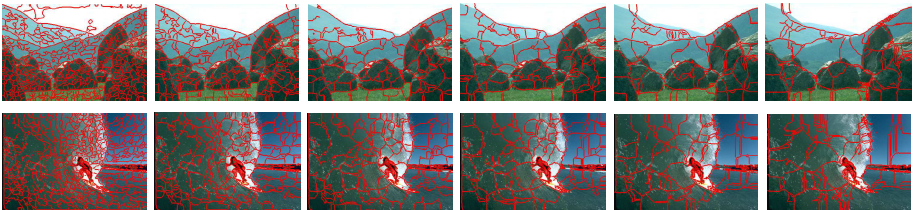


Fig. 5. MS segmentation examples for different parameters. Columns from 1 to 5: $(h_s, h_r) = \{ (7, 15), (13, 19), (17, 23), (20, 25), (25, 30) \}$.

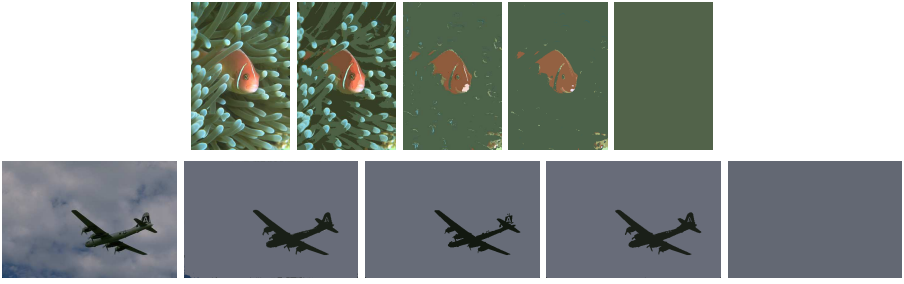


Fig. 6. Examples of undersegmentation. Original image. Columns from 2 to 5: segmentation for RAD with $(\sigma_d, \sigma_i) = \{(0.8, 0.05), (1.5, 0.05), (2.5, 0.05), (2.5, 1.5)\}$.

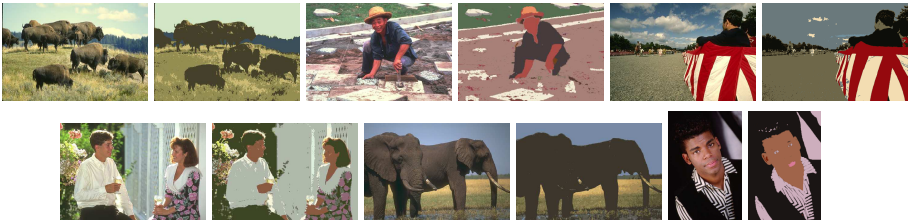


Fig. 7. Examples of segmentation in presence of shadows and highlights

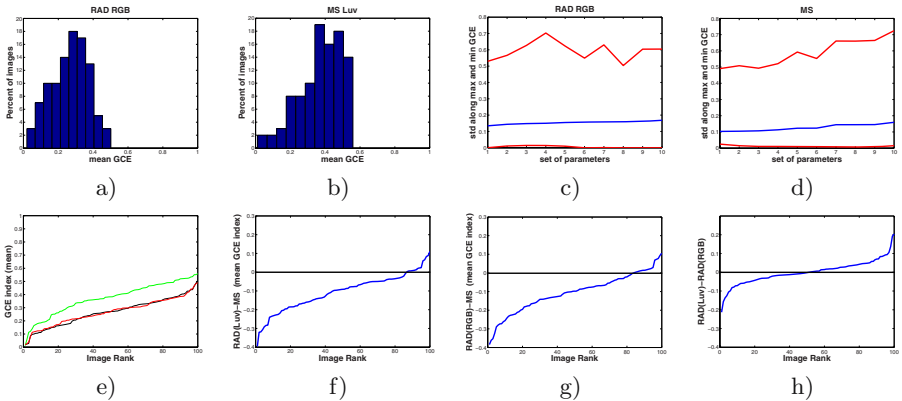


Fig. 8. (a,b) Mean GCE values for each set of parameters. (c,d) Standard deviation of GCE along maximum and minimum values for each set of parameters. (e) Mean GCE values for each image sorted from lower to higher. (f) Values higher than zero: images where MS performs better RAD. (g,h) The same as f) but for MS and RAD Luv and for RAD RGB versus RAD Luv.

Table 1. Global Constancy Error for several state-of-the-art methods: seed [27], fow [28], MS, and nCuts [29]. Values taken from [27] and [5].

	human	RAD	seed	fow	MS	nCuts
GCE	0.080	0.2048	0.209	0.214	0.2598	0.336

The histogram of the mean GCE values versus the percentage of images for each GCE value are shown in Figures 8a,b for RAD on RGB and MS respectively. As more bars are accumulated on the left, the better is the method. Figures 8c,d show the standard deviation along 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, 8e 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 that RAD has a similar behavior on RGB and CIE Luv spaces, even with the aforementioned potential problems on Luv. Figure 8f 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 8g,h.

Additionally, table 1 shows GCE values for several state-of-the-art methods. These values are taken from [27] and [5]. 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 our method the best results were obtained with $(\sigma_d, \sigma_i) = \{(2.5, 0.05)\}$. The mean number of dominant colours 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 be seen in figure 3h. CGE evaluation favors oversegmentation [12]. 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.

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* [29] 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). For our RAD method we do not apply any pre- or postprocessing steps.

5 Conclusions

This paper introduces a new feature space segmentation method that extracts the Ridges formed by a dominant colour in an image histogram. 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. 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, RAD has a better performance than the state-of-the-art methods. It points out that the chromatic information is an important cue on human segmentation. Additionally, the elapsed time for RAD 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.

The next step is to add spatial coherence to help the method in those areas which are not well-represented by a dominant colour. Furthermore, improvement is expected by looking for dominant colours only in interesting regions in the image instead of in the whole image at once.

Acknowledgements

This work has been partially supported by projects TIN2004-02970, TIN2007-64577 and Consolider-Ingenio 2010 CSD2007-00018 of Spanish MEC (Ministry of Science) and the Ramon y Cajal Program.

References

1. Skarbek, W., Koschan, A.: Colour image segmentation — a survey. Technical report, Institute for Technical Informatics, Technical University of Berlin (October 1994)
2. Cheng, H., Jiang, X., Sun, Y., Wang, J.: Color image segmentation: advances and prospects. *Pattern Recognition* 34(6), 2259–2281 (2001)
3. Lucchese, L., Mitra, S.: Color image segmentation: A state-of-the-art survey. *INSA-A: Proceedings of the Indian National Science Academy*, 207–221 (2001)
4. Agarwal, S., Madasu, S., Hanmandlu, M., Vasikarla, S.: A comparison of some clustering techniques via color segmentation. In: *ITCC 2005: Proceedings of the International Conference on Information Technology: Coding and Computing (ITCC 2005)*, vol. II, pp. 147–153. IEEE Computer Society Press, Washington (2005)
5. Yang, Y., Wright, J., Sastry, S., Ma, Y.: Unsupervised segmentation of natural images via lossy data compression (2007)
6. Freixenet, J., Munoz, X., Raba, D., Mart, J., Cuf, X.: Yet another survey on image segmentation: Region and boundary information integration. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) *ECCV 2002*. LNCS, vol. 2352, pp. 408–422. Springer, Heidelberg (2002)

7. Sezgin, M., Sankur, B.: Survey over image thresholding techniques and quantitative performance evaluation. *J. Electron. Imaging* 13(1), 146–165 (2004)
8. Fukunaga, K., Hostetler, L.D.: The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory* 121(1), 32–40 (1975)
9. Verma, D., Meila, M.: A comparison of spectral clustering algorithms. technical report uw-cse-03-05-01, university of washington
10. Abd-Almageed, W., Davis, L.: Density Estimation Using Mixtures of Mixtures of Gaussians. In: 9th European Conference on Computer Vision (2006)
11. Bilmes, J.: A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. *International Computer Science Institute* 4 (1998)
12. 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: Proc. Eighth Int'l Conf. Computer Vision, vol. 2, pp. 416–423 (2001)
13. Shafer, S.A.: Using color to separate reflection components. *COLOR research and application* 10(4), 210–218 (1985)
14. Geusebroek, J.M., Burghouts, G.J., Smeulders, A.W.M.: The Amsterdam library of object images. *Int. J. Comput. Vision* 61(1), 103–112 (2005)
15. Klinker, G., Shafer, S.: A physical approach to color image understanding. *Int. Journal of Computer Vision* 4, 7–38 (1990)
16. López, A.M., Lumbreras, F., Serrat, J., Villanueva, J.J.: Evaluation of methods for ridge and valley detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 21(4), 327–335 (1999)
17. Wang, L., Pavlidis, T.: Direct gray-scale extraction of features for character recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(10), 1053–1067 (1993)
18. 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)
19. Vazquez, E., Baldrich, R., Vazquez, J., Vanrell, M.: Topological histogram reduction towards colour segmentation. In: Martí, J., Benedí, J.M., Mendonça, A.M., Serrat, J. (eds.) *IbPRIA 2007. LNCS*, vol. 4477, pp. 55–62. Springer, Heidelberg (2007)
20. 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), 583–598 (1991)
21. Gauch, J.M., Pizer, S.M.: Multiresolution analysis of ridges and valleys in grey-scale images. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(6), 635–646 (1993)
22. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* 24(5), 603–619 (2002)
23. Christoudias, C., Georgescu, B., Meer, P.: Synergism in low level vision. *International Conference on Pattern Recognition* 4, 150–155 (2002)
24. Pantofaru, C., Hebert, M.: A comparison of image segmentation algorithms. Technical Report CMU-RI-TR-05-40, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA (September 2005)
25. Ge, F., Wang, S., Liu, T.: New benchmark for image segmentation evaluation. *Journal of Electronic Imaging* 16, 033011 (2007)
26. Felzenszwalb, P., Huttenlocher, D.: Efficient graph-based image segmentation. *Intl. Journal of Computer Vision* 59(2) (2004)

27. Micusik, B., Hanbury, A.: Automatic image segmentation by positioning a seed. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952. Springer, Heidelberg (2006)
28. Fowlkes, C., Martin, D., Malik, J.: Learning affinity functions for image segmentation: combining patch-based and gradient-based approaches. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003, vol. 2 (2003)
29. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(8), 888–905 (2000)