A Novel Segmentation Strategy Based on Colour Channels Coupling[*]

Alberto Ortiz and Gabriel Oliver

Department of Mathematics and Computer Science,
University of the Balearic Islands, Spain
{alberto.ortiz, goliver}@uib.es

Abstract. A segmentation method based on a physics-based model of image formation is presented in this paper. This model predicts that, in image areas of uniform reflectance, colour channels keep coupled in the sense that they are not free to take any intensity value, but they depend on the values taken by other colour channels. This paper first enumerates and analyzes a set of properties in which this coupling materializes. Next, a segmentation strategy named C^3S and based on looking for violations of the coupling properties is proposed. Segmentation results for synthetic and real images are presented at the end of the paper.

1 Introduction

When shading is hardly noticeable in the image, as well as specularities and other optical phenomena such as inter-reflections, areas of the scene of uniform reflectance appear as regions of more or less constant colour if shadows are avoided. Nevertheless, when curved objects are imaged, the scene curvature and the objects glossiness, among others, give rise to noticeable changes in image intensity not necessarily related to object boundaries. In order to cope with these intensity variations in a more suitable way, physics-based segmentation methods embed into the logic of the algorithm a physics-based model of image formation. Among the several physics-based segmentation strategies which have been proposed so far, some of them are based on estimating directly the reflectance of the surfaces present in the scene (see for instance [1]), others look for certain configurations of clusters in colour space, as it is predicted by the Dichromatic Reflection Model proposed by Shafer [2] (by way of example, see [3,4,5]), and, finally, others use photometric invariants in their different forms ([6,7], among others).

C^3S, the segmentation method proposed in this paper, is also based on the Dichromatic Reflection Model, but it does not use any of the approaches mentioned above. Its main point, which is the most important contribution of the paper, is the use of the fact that, in uniform reflectance areas, colour channels are coupled by the reflectance of the surface material, while, in reflectance transition zones, such coupling can be broken in a number of ways. Consequently, material changes can be found by looking for violations of this coupling, which allows computing an edge map from which a segmentation of the image can be obtained. The work presented in this paper is the continuation of a research line which started in [8].

The rest of the paper is organized as follows: section 2 presents the image formation model considered in this work, and comments on the properties of uniform reflectance areas according to that model; sections 3 and 4 describe C^3S; section 5 presents some segmentation results for synthetic and real images; and, finally, conclusions appear in section 6.

2 Image Formation Model

2.1 General Description

It is generally accepted that objects reflection is an additive composition of body or diffuse reflection and interface or specular reflection [2]. Besides, this model is enhanced by a further term accounting for non-directional or ambient lighting which interacts with the scene increasing objects radiance irrespectively of local surface geometry. All in all, radiance at a scene point p and for a wavelength λ can be summarized as indicated in equation 1:

$$L(p;\lambda) = \overbrace{L_a(\lambda)\rho_a(p;\lambda)}^{L_a(p;\lambda)} + \overbrace{m_b(p)\left[L_d(\lambda)\rho_b(p;\lambda)\right]}^{L_b(p;\lambda)} + \overbrace{m_i(p)\left[L_d(\lambda)\rho_i(p;\lambda)\right]}^{L_i(p;\lambda)}, \quad (1)$$

where: (i) $L_a(\lambda)$ represents light coming from all directions in equal amounts while $L_d(\lambda)$ represents directional lighting; (ii) ρ_a, ρ_b and ρ_i are surface material reflectances expressing the fraction of the incoming light which is conveyed by the corresponding reflection component (ambient, body and interface, respectively), being ρ_a assumed a linear combination of the body and interface reflectances, ρ_b and ρ_i; (iii) m_b and m_i are terms dependent on local surface geometry such that $m_b, m_i \in [0, 1]$. After the process of photoelectrons integration performed in the sensor, the intensity at colour channel k can be approximated by $I^k(i,j) = C_a^k(i,j) + m_b(i,j)C_b^k(i,j) + m_i(i,j)C_i^k(i,j)$ where the composite reflectances C_a^k, C_b^k and C_i^k represent the joint contributions of lighting and the different material reflectances to the corresponding colour component.

2.2 Properties of Uniform Reflectance Image Areas

Within an image area of pixels corresponding to the same scene material, the ambient, body and interface composite reflectances are constant if the light distribution is approximately uniform throughout the area. In a noiseless environment, colour changes between image locations are, thus, only due to changes in the geometrical factors m_b and m_i, common to all the colour channels. The study of the image formation model has allowed us to express this coupling in the form of the following properties:

Property P1. *Colour channels do not cross one another.*

Property P2. *Colour channels vary in a coordinated way: when one changes, so do the others, and in the same sense, all increase or all decrease.*

Property P3. *As the intensity in one colour channel decreases, so does the difference between colour channel intensities; the opposite happens when the intensity in one channel increases.*

Table 1. Dependence of the fulfillment of properties P1-3 on the particular instantiation of the image formation model. **TP** and **AL** stand for *type of pixel* and *ambient lighting*, respectively. The NIR model assumes that $\rho_i(\lambda) = \rho_i, \forall \lambda$.

TP	AL	properties fulfilled
matte	no	P1-3 always
matte	yes	P2 always; P1 and P3 if $L_a(\lambda) = kL_d(\lambda), k > 0$
glossy	no	P1 under the NIR model and white-balanced images
glossy	yes	P1 under the NIR model, white-balanced images and $L_a(\lambda) = kL_d(\lambda), k > 0$

As it is indicated in table 1, these properties hold always in case the illumination is purely directional and the pixels considered do not show specular reflection. In the rest of cases, their general fulfillment depends on the satisfaction of the indicated conditions (see [9] for the formal proofs). However, most times all three properties hold within areas of uniform reflectance, since their unfulfillment takes place only for particular combinations of material reflectances and precise values of m_b and m_i.

Essentially, properties P1-3 lead to a set of necessary compatibility conditions between pixels, in the sense that if two pixels fail to satisfy any of the three properties for any of the two-by-two combinations between colour channels, then both pixels cannot correspond to the same scene material. Given a certain combination of colour channels, say C_1 and C_2, these three conditions can be stated geometrically, considering the space of intensity values taken by both channels, as it is shown in figures 1(a)-(c), for, respectively, properties P1-3. In all the figures and for case $I_a^1 \geq I_a^2$, the shaded areas represent the set of points (I_b^1, I_b^2) which are compatible with a certain point (I_a^1, I_a^2), in the sense of the satisfaction of the coupling properties; in case $I_a^1 \leq I_a^2$, the compatibility zone would be the complement of the shaded areas of figures 1(a) and (c) for properties P1 and P3. Therefore, if for a given (I_a^1, I_a^2), (I_b^1, I_b^2) is not inside the compatibility zone, then the corresponding pixels cannot correspond to the same scene material. Finally, figure 1(d) shows the compatibility area for all three properties simultaneously (triangles correspond to the case $I_a^1 \leq I_a^2$, while circles are for case $I_a^1 \geq I_a^2$). In view of these graphs, a compatibility relation $\mathcal{C}$ can be summarized as: (I_b^1, I_b^2) is $\mathcal{C}$-compatible with (I_a^1, I_a^2) if: (i) $I_b^1 \geq I_b^2$ and the orientation of the vector joining both lies within $[0°, 45°]$ when $I_a^1 \geq I_a^2$; or (ii) $I_b^1 \leq I_b^2$ and the orientation of the vector joining both lies within $[45°, 90°]$ when $I_a^1 \leq I_a^2$.

3   Edge Map Computation

General Discussion. Checking, at every pixel, all three properties allows computing an edge map corresponding to reflectance transitions which can be found in the image. To this end, a pixel (i_a, j_a) is considered an edge if for any of its 8 neighbours (i_b, j_b) and at least one combination of colour channels, both pixels are not compatible with one another in the sense of the relation $\mathcal{C}$ formulated in section 2.2.

Although the checking of relation $\mathcal{C}$ allows covering a considerable amount of reflectance transitions, there exist others which make manifest in such a way that the
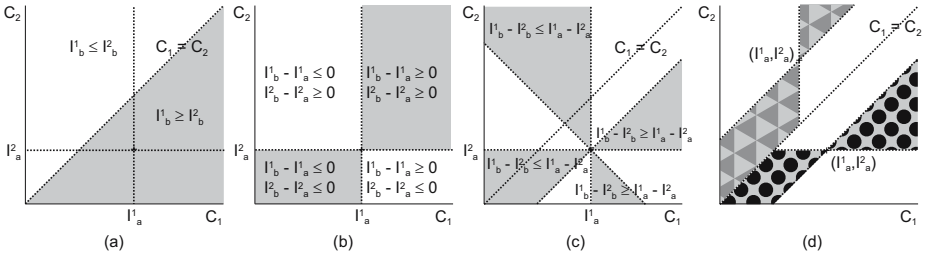
Fig. 1. Geometrical interpretation of properties P1-3

variation in the intensity does not violate any of properties P1-3 despite a reflectance transition is arising. On those cases, there is no way to decide whether the corresponding intensity changes in the colour channels are due to a reflectance transition or merely a change in m_b and/or m_i, and, thus, an ambiguity results. In other words, the interrelation between colour channels is not powerful enough on those cases so as to disambiguate the source of the intensity variation. Accordingly, knowledge from the application domain level or data from another sensor, if available, should be used to make the decision. For instance, if objects with smooth surfaces are expected in the scene, gradient information can result practical to locate this sort of reflectance transitions. In particular, LOG zero-crossings have become particularly interesting during the experiments performed. This is because LOG zero-crossings correspond to concavity-convexity changes in intensity due to the second derivative order nature of LOG, and those changes correspond quite well to the aforementioned reflectance transitions.

Implementation Issues. In order to counteract image noise in an adaptive way when computing the edge map, uncertainties related to sensor noise are used for every intensity level of every colour channel. Therefore, every digital noisy level D^k outputting the camera is associated to an interval $[D^k - \delta(D^k), D^k + \delta(D^k)]$. In this expression, the uncertainty $\delta(D^k)$ is defined as it is indicated in [9], which takes as a basis the noise model proposed in [10] and estimates the sensor performance parameters by means of a radiometric calibration of the camera [11]. With the introduction of these uncertainties, either $\mathcal{C}$ and the strategy for locating LOG zero-crossings are redefined:

- With regard to $\mathcal{C}$, given noisy intensities (D_a^1, D_a^2), a rectangular uncertainty area around (D_a^1, D_a^2) covering t uncertainties is defined, as it is depicted in figure 2(a). $\mathcal{C}$ is redefined then as: (D_b^1, D_b^2) *is $\mathcal{C}$-compatible with* (D_a^1, D_a^2) *if any of the points belonging to the uncertainty area of* (D_b^1, D_b^2) *falls within the union of the compatibility zones of the points belonging to the uncertainty area of* (D_a^1, D_a^2) *(shaded area of figure 2(a)).* The pixel corresponding to (D_a^1, D_a^2) is thus considered an edge only if there is no possibility of (D_b^1, D_b^2) being $\mathcal{C}$-compatible with (D_a^1, D_a^2).
- As for LOG zero-crossings, whenever a zero-crossing is detected, it is classified as relevant if the positive and negative peak LOG values along the direction of detection are larger than t times the respective uncertainties. Those uncertainties are calculated using standard uncertainty propagation rules, by which, if the output of the operator is calculated as $f = \sum_x D^k(x)m(x)$, where $m(x)$ would be the LOG
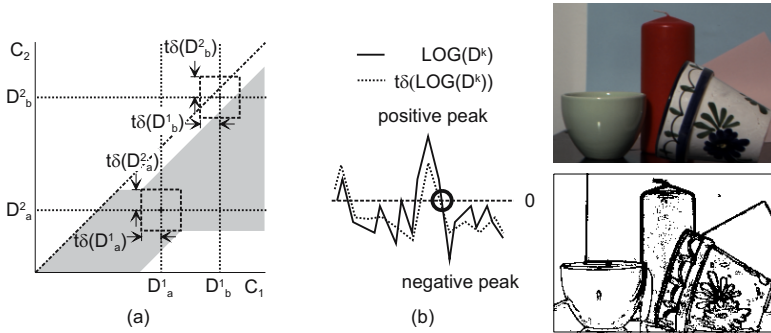
Fig. 2. Left: (a) Redefinition of $\mathcal{C}$ (the uncertainty areas have been magnified for illustration purposes); (b) Only the LOG zero-crossing behind the circle is considered relevant. Right: Edge map for image on top.

mask constants, then $\delta(f) = \sqrt{\sum_x \delta(D^k(x))^2 m^2(x)}$ [12]. By way of example, see figure 2(b), where only the zero-crossing behind the circle would be considered relevant.

To finish about the computation of the edge map: (1) although a reflectance transition should theoretically involve just two pixels, in real images they tend to span along several image cells because of real cameras' aliasing, so that the edge map must in general be expected to consist of thick edges; and (2) pixels around specularities can be labelled as edges, because, although not always edges related to the $\mathcal{C}$ relationship are found near specularities, relevant LOG zero-crossings can clearly be found, particularly for "steep" specularities. The two above-mentioned facts can be easily observed in figure 2(right).

4   Colour Channel Coupling-Based Segmentation (C^3S)

In general, segmenting an image consists in grouping pixels in homogeneous regions, generally according to a certain perceptually-based homogeneity criterion. Most times, one is interested in the image regions corresponding to the same perceptual colour. In physical terms, this means grouping pixels in uniform reflectance areas.

The experiments performed have shown that the closed regions contained in the edge map described in section 3 tend to show uniform reflectance. Therefore, a first partition of the image can be obtained if connected components not including edge pixels are found and edges are then added to the nearest most similar connected component. To perform this last task, *Principal Component Analysis* is used to characterize the cluster in colour space corresponding to every connected component. As the edge map is presumed to separate pixels corresponding to "steep" specularities from matte pixels, and according to the Dichromatic Reflection Model, the corresponding regions are expected to be describable by linear or point clusters in colour space.

In order not to mix pixels corresponding to different scene materials, connected components without a linear or point-wise shape in colour space are discarded. Those
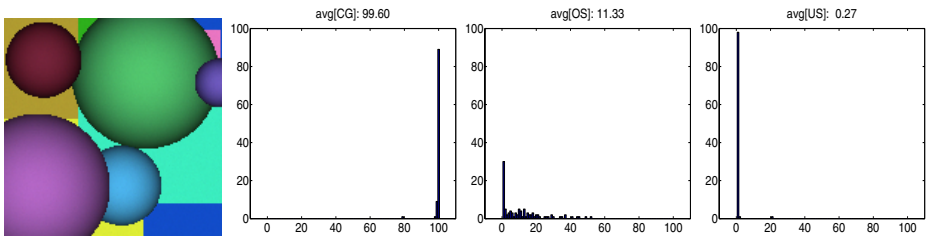
Fig. 3. (1st) example of noisy synthetic image; (2nd,3rd,4th) CG, OS and US histograms

connected components result when a contour is not finished due to noise and a path of non-edge pixels between two regions of different reflectance is created. On those cases, the corresponding pixels are grouped again controlling the growth of the corresponding cluster in colour space as the pixels are added. When the cluster dimensionality exceeds that of a line, the growing of the connected component stops and a new connected component is created and grown. The process continues until no more pixels of the original connected component are unlabelled.

A region merging stage based on a possibilistic clustering algorithm [13] follows next, in order to remove the probably low, but not generally zero, degree of oversegmentation which can result.

5   Experimental Results and Discussion

Several results for synthetic and real curvature-dominated images are given below to prove experimentally the usefulness of C^3S.

On the one hand, a set of 100 synthetic colour images of different scenes consisting of spheres and planes were generated. Besides, noise was incorporated conforming with the camera model described in [11], parameterized according to the performance of the camera used in the experiments with real images. All 100 images were then segmented and the percentage of correctly grouped pixels (CG), oversegmentation (OS) and undersegmentation (US) were determined for every image (CG corresponds to pixels belonging to regions which are not spread among several regions of the reference segmentation, while OS accounts for pixels belonging to regions splitting an only true region; US, finally, covers pixels in regions spanning several true regions). Figure 3 shows an example of noisy synthetic image, together with histograms for CG, OS and US. As a global result of the experiment: $\overline{CG} = 99.60\%$, $\overline{OS} = 11.33\%$ and $\overline{US} = 0.27\%$.

On the other hand, a set of real images, standard and non-standard, were also considered. For the non-standard images, the gamma correction switch of the camera was set to 1 in order to get images whose intensity is proportional to the light striking the sensor, which is essential for physics-based vision algorithms. Although standard images do not typically satisfy this constraint, they were also considered in the experiments for comparison with other segmentation algorithms.

Figure 4 shows two non-standard and two standard images, with the region contours produced by C^3S overimposed over the original images. Besides, for comparison purposes between a physics-based and a non-physics-based approach, region contours
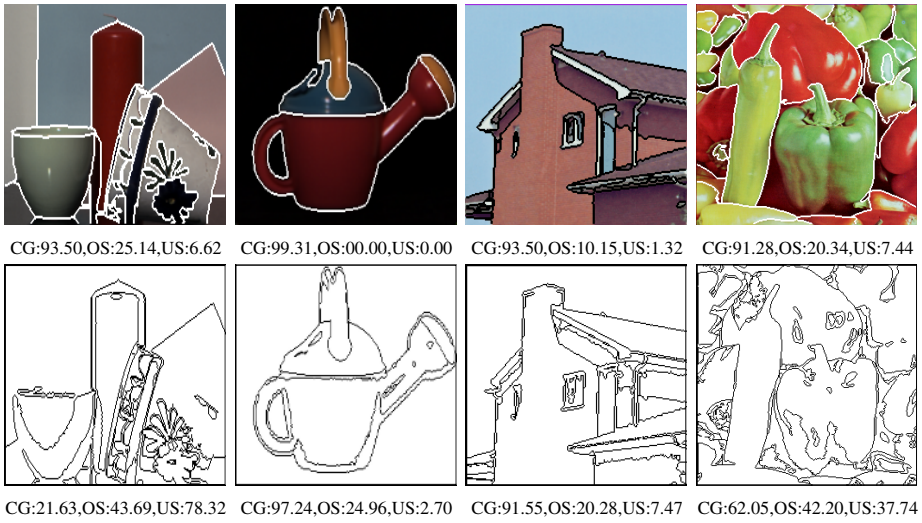
CG:93.50,OS:25.14,US:6.62   CG:99.31,OS:00.00,US:0.00   CG:93.50,OS:10.15,US:1.32   CG:91.28,OS:20.34,US:7.44

CG:21.63,OS:43.69,US:78.32   CG:97.24,OS:24.96,US:2.70   CG:91.55,OS:20.28,US:7.47   CG:62.05,OS:42.20,US:37.74

Fig. 4. (first row) C^3S; (second row) C&M

Fig. 5. More results for C^3S and real images

resulting for the recognized mean-shift based segmentation algorithm by Comaniciu and Meer [14] (C&M) are also given below each image (the undersegmentation option was used in all cases). Values for CG, OS and US are also provided for both sets of segmentation results, having used, on those cases, reference segmentations manually determined. Figure 5 shows more results for non-standard images.

As for parameters, the standard deviation for LOG was set to 1.0 in all cases. The number of uncertainties considered for detecting edges related with relation $\mathcal{C}$, t_1, and the number of uncertainties for capturing LOG zero-crossings, t_2, were set to 3 in the experiment with synthetic images, while t_2 needed a finer tuning when segmenting real images. Nevertheless, a restricted set of values (1.5 or 2) was necessary to achieve good segmentations.

As can be observed, C^3S is able to deal correctly with scene curvature, while C&M tends to produce oversegmentation on those cases. Furthermore, in cases where C&M mixes pixels from different objects, C^3S finds the correct classification (first image of figure 4). Sometimes, however, object edges give rise to a certain degree of oversegmentation in C^3S output because of the ambiguity in the reflectance transitions detected by

means of LOG zero-crossings (first image of figure 4); a similar behaviour is observed for shadows. The third image in figure 5 is also interesting because of, contrarily to expected, the right behaviour of C^3S with the textured cup. As for the standard images, both algorithms provide a similar output in the *house* image, while in the *peppers* image, where curved objects appear, C^3S clearly outperforms C&M. (See [9] for more experimental results, intermediate and final, and a more detailed discussion, not included here due to lack of space)

6   Conclusions

C^3S, a curvature-insensitive segmentation method based on a physics-based image formation model, has been proposed. The method uses the coupling between colour channels in uniform reflectance areas to locate image locations where colour channels turn out to be uncoupled because of a reflectance transition arising there. Experiments with synthetic and real images have been presented, showing the power of the approach for dealing with scenes with curved objects and different surface materials.

References

1. Lee, C.H., Rosenfeld, A.: Albedo estimation for scene segmentation. PRL **1** (1983) 155–160
2. Shafer, S.: Using color to separate reflection components. COLOR Research and Application **10** (1985) 210–218
3. Klinker, G., et al.: A physical approach to color image understanding. IJCV **4** (1990) 7–38
4. Kroupnova, N.: Method for multi-spectral images segmentation based on the shape of the colour clusters. In: SPIE - Human Vision, Vis. Proc., and Dig. Display VI. (1996) 444–453
5. Ong, C.K., Matsuyama, T.: Robust color segmentation using the dichromatic reflection model. In: Proceedings of ICPR. (1998) 780–784
6. Gevers, T.: Adaptive image segmentation by combining photometric invariant region and edge information. PAMI **24** (2002) 848–852
7. Healey, G.: Segmenting images using normalized color. SMC **22** (1992) 64–73
8. Ortiz, A., Oliver, G.: Detection of colour channels uncoupling for curvature-insensitive segmentation. In: Proc. of the Iberian Conference on Pattern Recognition and Image Analysis. (2003) 664–672
9. Ortiz, A., Oliver, G.: Segmentation of images based on the detection of reflectance transitions. Technical Report A-3-2003, Departament de Matemàtiques i Informàtica (Universitat de les Illes Balears) (2003)
10. Healey, G., Kondepudy, R.: Radiometric CCD camera calibration and noise estimation. PAMI **16** (1994) 267–276
11. Ortiz, A., Oliver, G.: Radiometric calibration of CCD sensors: Dark current and fixed pattern noise estimation. In: Proc. of ICRA. Volume 5. (2004) 4730–4735
12. Taylor, J.: An Introduction to Error Analysis. University Science Books (1997)
13. Theodoridis, S., Koutroumbas, K.: Pattern Recognition. Academic Press (1999)
14. Comaniciu, D., Meer, P.: Robust analysis of feature spaces: Color image segmentation. In: Proc. of CVPR. (1997) (The code is publicly available in http://www.caip.rutgers.edu/riul/research/papers/abstract/feature.html).