Estimation of the Color Image Gradient with Perceptual Attributes

[1]Philippe Pujas , [2]Marie-José Aldon

[1]Institut Universitaire de Technologie de Montpellier, Université Montpellier II,
17 quai Port Neuf, 34500 Béziers, FRANCE
[2]LIRMM - UMR C55060 - CNRS / Université Montpellier II
161 rue ADA, 34392 Montpellier Cedex 05, FRANCE

Abstract

Classical gradient operators are generally defined for grey level images and are very useful for image processing such as edge detection, image segmentation, data compression and object extraction. Some attempts have been made to extend these techniques to multi-component images. However, most of these solutions do not provide an optimal edge enhancement.

In this paper we propose a general formulation of the gradient of a multi-image. We first give the definition of the gradient operator, and then we extend it to multi-spectral images by using a metric and a tensorial formula. This definition is applied to the case of RGB images. Then we propose a perceptual color representation and we show that the gradient estimation may be improved by using this color representation space. Different examples are provided to illustrate the efficiency of the method and its robustness for color image analysis.

1   Introduction

This paper addresses the problem of detecting significant edges in color images. More specifically, given a scene including objects which are characterized by homogeneous colors, we want to detect and to extract their contours in the image. Our objective is to propose a solution which overcomes the problems of shades and reflections due to lighting conditions and objects surface state, in order to achieve an adequate image segmentation.

The paper is divided into four parts. In the first part, we propose a general formula to define the gradient of a multi-spectral image. In the second part, we apply this definition to color images whose components are described in the RGB space. We present a brief overview of the classical approaches consisting of an evaluation of the gradient as a simple function of its three components. Thirdly, we show that it is possible to improve this gradient estimation by using an Euclidean metric in the RGB space.

We then introduce the perceptual color space representation which separates the intensity and chromatic components. We improve the classical HSV model by introducing a new parameter, the "chromaticity degree" γ, which allows the separation of chromatic and achromatic areas within the image. In the fourth part of the paper, we use the $HV\gamma$ representation to propose a gradient estimator which avoids enhancement of non significant edges created by object shades or reflections. Results obtained with true images illustrate the advantages of this approach.

2   Multi-spectral image gradient

2.1   Gradient definition

Let f(M) be a scalar potential field in R^2 (for instance a monochrome image), with M = M(x,y). The gradient of f is referred to as ∇f :

$$\nabla f = \begin{pmatrix} \dfrac{\partial f}{\partial x} \\ \dfrac{\partial f}{\partial y} \end{pmatrix} \qquad (1)$$

In the image plane the gradient ∇f may be represented by a 2D vector; its orientation corresponds to the direction along which f has the maximum rate of change and its magnitude is the absolute value of this maximum rate of change. Let . denote the dot product in R^2 and : $dM = (dx, dy)^T$. Then we can write:

$$\nabla f. \ dM = df \qquad (2)$$

Currently, different kinds of methods are available to compute the gradients' components in a monochrome image. The earliest of them are based on simple discrete approximations of continuous derivatives [1,2]. More recent operators have been designed by taking into account a model of the edges to be detected and a quantitative definition of the performance of the edge detector [3,4,5].

2.2   Application to multi-images

The previous gradient definition (1) cannot be applied to a multi-image such as a color image. Such an image being described by a set of components cannot be modelled by a scalar potential field but by a 2D vector field [6]. Let C be the multi-spectral image and ∇C its gradient. Let $K=(k_1,k_2,...k_n)^T$ be the image representation in a given space. The problem discussed in this section is the way to combine the elementary gradient components $\nabla k_i \ i \in [1,n]$ in order to obtain the best gradient estimation ∇C :

$$\nabla C = F(\nabla k_1, \nabla k_2, ... \ \nabla k_n) \qquad (3)$$

Several solutions have been proposed to deal with color images (n=3). The straightforward approaches make use of linear combinations of the ∇k_i, the simplest method consisting in a vectorial sum of the ∇k_i. Another approach is to estimate the resultant gradient magnitude at point M(x,y) as an Euclidean distance between the color vectors K(M) and K(M'). More sophisticated methods make use of distance between averaged vector values. As explained in [7], in all of these approaches the image components k_i do not cooperate with one another. To avoid this drawback, Di Zenzo proposes a solution based on the use of a tensor gradient of the multi-image which is defined as a vector field. This formula has been adopted by Chapron [8].

We propose here a general gradient definition which can be applied to a color image, or to any kind of multi-image. Given a point $K=(k_1,k_2,...k_n)$ in the image space, we define a metric dC^2 in this space such as:

$$dC^2 = dK^T \ g \ dK \qquad (4)$$

where g represents the metric tensor and dK an elementary displacement in the considered space. Notice that g is a non negative defined tensor, in order to define a positive metric. K being a function of the pixel position M, it may be considered as a vector field. Then:

$$dK = \frac{\partial K}{\partial M} dM \qquad (5)$$

$$dC^2 = dM^T (\frac{\partial K}{\partial M})^T g (\frac{\partial K}{\partial M}) dM \qquad (6)$$

We note $\tilde{g}$ the non negative defined tensor associated to this metric:

$$\tilde{g} = (\frac{\partial K}{\partial M})^T g (\frac{\partial K}{\partial M}) \qquad (7)$$

Let θ be the direction of the maximum image change, i.e. the gradient direction. An elementary displacement along this direction can be noted:

$$dM = dl \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \qquad (8)$$

Equation (6) can be reformulated as:

$$dC^2 = dl^2 (\tilde{g}_{11} \cos^2(\theta) + (\tilde{g}_{12} + \tilde{g}_{21}) \cos(\theta) \sin(\theta) + \tilde{g}_{22} \sin^2(\theta)) \qquad (9)$$

The gradient direction corresponds to the value of θ which maximizes the function $F(\theta)$, such as:

$$F(\theta) = \tilde{g}_{11} \cos^2(\theta) + (\tilde{g}_{12} + \tilde{g}_{21}) \cos(\theta) \sin(\theta) + \tilde{g}_{22} \sin^2(\theta) \qquad (10)$$

The gradient norm is obtained from equation (2) and can be expressed as a function of $F(\theta)$ [9].

3 Application to RGB Images

3.1 Gradients combination
In the (R,G,B) color space, the color image is noted:
$$f(x,y) = (R(x,y), G(x,y), B(x,y)) \qquad (11)$$
As seen in section 2.2, the simplest method to obtain the gradient of a color image consists in adding the elementary gradient components:
$$\nabla C = \nabla R + \nabla G + \nabla B \qquad (12)$$
Using equation (2), we obtain the color differential :
$$dC = dR + dG + dB \qquad (13)$$
This equation expresses the fact that points of the (R,G,B) space which belong to a plane parallel to the R+G+B=0 plane have the same color. In other words, the iso-chroma (r,g,b) is a plane whose equation is R+G+B=r+g+b. For instance, grey (1,1,1) and red (3,0,0) are identical colors! This is always true for all linear transforms.

We have applied this gradient definition to the image of figure 1a. Figure 1b presents the resultant gradient magnitude whose components have been obtained with a Sobel operator [1]. This result show that (18) can provide an over-segmentation. For instance internal edges appear in the front face of the upper orange cube. Due to shadows, the lower left purple cube has been separated in two regions. One of these regions has been merged with another one belonging to the green neighbor cube.

3.2 Euclidean metric

In [7] and [8], the gradient estimator is implicitly based on the use of an Euclidean metric in the RGB space. Here, the metric tensor is a unit tensor : $g = I_d$. Then, $\tilde{g}$ is a symmetric tensor such as:

$$\tilde{g} = \begin{pmatrix} (\frac{\partial R}{\partial x})^2 + (\frac{\partial G}{\partial x})^2 + (\frac{\partial B}{\partial x})^2 & \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y} \\ \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y} & (\frac{\partial R}{\partial y})^2 + (\frac{\partial G}{\partial y})^2 + (\frac{\partial B}{\partial y})^2 \end{pmatrix} \quad (14)$$

Computing the maxima of the function $F(\theta)$ requires the estimation of the partial derivatives of each component of $f(x,y)$. These derivatives are obtained using a classical operator [1,2,3,4,5].

Results of the algorithm using this Euclidean metric are illustrated in [9]. Compared with the previous results, it provides a best edge enhancement; however, there are always over-segmentations. The drawbacks of this solution are partly due to the used color space representation. It cannot represent color information like it is perceived by a human. Moreover, this space is not adapted to direct color comparison, because equal geometric distances in the RGB frame does not correspond to equal perceptual changes in color [10].

4 Perceptual Color Space Representation

In [11], Nevatia noticed that most of the edge information was in the intensity component of the image. The RGB model being enable to separate the chromatic and luminance information in a color image, Nevatia proposed an edge detector in a new space defined by the intensity. Later, other perceptual spaces have been defined [10,12]. The classical HSV representation [12] makes use of three attributes to describe a color: the hue H, the saturation S and the value (or intensity) V.

4.1 The HSV frame

All the perceptual frames are defined by a non-linear diffeomorphism. Our HSV frame is obtained by the following equations:

$$V = \sup(R, G, B)$$

$$S = 256 \frac{V - \inf(R, G, B)}{V}$$

$$H = \begin{bmatrix} \dfrac{G - B}{V - \inf(R, G, B)} & \text{if } V = R \\ \left(2 + \dfrac{B - R}{V - \inf(R, G, B)}\right) & \text{if } V = G \\ \left(4 + \dfrac{R - G}{V - \inf(R, G, B)}\right) & \text{if } V = B \end{bmatrix} \quad (15)$$

In these equations we suppose the dynamic range of the image signal to be 256. This coordinate transform requires few operations and is simpler than most of the other perceptual representations. For instance, the Luv transform [13] is five time more consuming.

4.2 Some difficulties with the perceptual frames

In a previous paper [15], we specified two kinds of problems connected with the definition of the hue:

- H is 6-periodic
- H is undefined if S=0 (15).

The first problem leads to computational difficulties for estimating hue averages or differences. It is easily solved using a suitable algebra [15]. The second problem makes the interpretation more difficult when using the hue to describe the image. We call this case the "low saturation effect" because in experiments, it begins to appear when the value of S is low. In order to solve this problem, we consider the existence of achromatic and chromatic areas in the perceptual color space [15]:

- In the achromatic area, H is undefined (or badly defined), so that it should not be used; the color is only described by S and V.
- In the chromatic area, previous research works [14] confirm that hue is generally the most discriminant attribute. It can be used for an efficient image analysis.

In order to optimize the analysis algorithms, it is necessary to define a way for classifying pixels into chromatic and achromatic zones. We have proposed the concept of "chromaticity degree" [15]. The chromaticity degree γ of a pixel is a scalar function depending on S and V in the range $[0,1]$. The closer γ is to one, the more chromatic the pixel is; the closer γ is to zero, the more achromatic the pixel is. The identification of γ is done by manual classification of a representative set of pixels. We have proposed in [15] different identification methods.

4.3 The HVγ representation

γ being a function of S and V and $\dfrac{\partial \gamma}{\partial S}$ being different of zero, the HVγ transform is a diffeomorphism. Consequently, it defines a new perceptual color space. In this frame, H is still periodic and undefined if $\gamma = 0$. The interest of the HVγ representation is that it does not require any external variable to perform the analysis of the color image, unlike HSV which also needs the chromaticity degree.

5 Image Gradient in a Perceptual Color Frame

One of the most important advantages of a perceptual space is that it allows direct color comparison based on geometric distance estimation.

5.1 Linear gradients combination

Using a linear combination of the three gradient components in the HSV or in the HVγ space cannot provide a correct gradient estimation. In fact, hue being not defined inside achromatic areas, the component ∇H has no significance here. Moreover, across transitions between chromatic and achromatic zones, a large hue variation induces a wrong edge.

Figure 1c illustrates this problem. We present the magnitude of the ∇H gradient component which has been obtained with the Sobel operator. Erroneous edges appear in the achromatic areas and along the reflections boundaries. In the next section we propose a solution to this problem by taking into account the chromaticity degree.

5.2 Non Linear gradients combination

In the HSV space, we propose the following definition of the resultant gradient:

$$\nabla C = \gamma(\nabla H + \nabla S) + (1-\gamma)\,\nabla V \qquad (16)$$

This definition may be improved to solve the reflections problem using:

$$\nabla C = \gamma \nabla H + (1-\gamma)\,\nabla V \qquad (17)$$

In the HVγ space the non linear gradient combination takes the form:

$$\nabla C = \gamma \nabla H + (1-\gamma)\,\nabla V + \nabla \gamma \qquad (18)$$

We have shown, in [9], that this solution improves the previous results. However some edges are not correctly enhanced.

5.3 A metric for the perceptual space

We have defined a new metric in the perceptual space which is based on a tensorial formula:

$$dC^2 = \gamma\,dH^2 + (1-\gamma)\,dV^2 + d\gamma^2 \qquad (19)$$

This definition means that in chromatic zones the distance between two pixels is a function of the hue H, while it is a function of the intensity V in achromatic zones. The last term $d\gamma^2$ allows the separation of chromatic and achromatic zones. This yields the metric tensor:

$$\tilde{g} = \begin{pmatrix} \gamma(\dfrac{\partial H}{\partial x})^2 + (1-\gamma)(\dfrac{\partial V}{\partial x})^2 + (\dfrac{\partial \gamma}{\partial x})^2 & \gamma\dfrac{\partial H}{\partial x}\dfrac{\partial H}{\partial y} + (1-\gamma)\dfrac{\partial V}{\partial x}\dfrac{\partial V}{\partial y} + \dfrac{\partial \gamma}{\delta x}\dfrac{\partial \gamma}{\delta y} \\[3mm] \gamma\dfrac{\partial H}{\partial x}\dfrac{\partial H}{\partial y} + (1-\gamma)\dfrac{\partial V}{\partial x}\dfrac{\partial V}{\partial y} + \dfrac{\partial \gamma}{\partial x}\dfrac{\partial \gamma}{\partial y} & \gamma(\dfrac{\partial H}{\partial y})^2 + (1-\gamma)(\dfrac{\partial V}{\partial y})^2 + (\dfrac{\partial \gamma}{\partial y})^2 \end{pmatrix} \qquad (20)$$

In figure 1d we give the multilevel gradient image obtained with this new metric. Now, homogeneous color regions are well separated, with few over-segmentations. Their edges are correctly enhanced. This gradient estimator is less sensitive to shadow and reflection effects.

In figure 2a we present the original color image of a natural landscape. The following normalized gradient images have been obtained respectively with a sum of the elementary gradient components in the RGB space (18) (Fig. 2b), and by using the perceptual metric in the HVγ space (20) (Fig. 2c). We can see that with this second metric, the road and mountains edges are more correctly enhanced. We have also compared these gradient estimators with the color image of two birds (fig. 3a). The same performance can be observed for edge detection (fig. 3b, fig 3c).

6 Conclusion

In this paper we have proposed a new definition based on a spatial metric, for the gradient of a multi-spectral image. Our main goal being to obtain a perceptual gradient, we have defined a new color representation based on the concept of chromaticity degree. We have proposed a metric adapted to this color space and deduced a perceptual gradient formulation. Experimental results show the performance of this gradient estimation. They illustrate the interest of the chromaticity degree concept and its robustness. Moreover, this concept can be extended with other

perceptual color representations in order to define a perceptual gradient. It is obvious that such a color gradient definition may be applied in classical gradient-based edges detectors or in segmentation methods like those using region growing [15].

References

1. I. Sobel, "Neighbourhood coding of binary images for fast contour following and general array binary processing", Computer Graphics and Image Processing, vol. 8,1978, pp. 127-135.
2. J.M.S. Prewitt, "Object enhancement and extraction", Picture Processing and Psychopictorics, B.S Lipking and A. Rosenfeld, editors", Academic Press, New York, 1970, pp. 75-149.
3. J.F. Canny, "A Computational approach to edge detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 8, November 1986, pp. 769-798.
4. R. Deriche, "Using Canny criteria to derive an optimal edge detector recursively implemented, The International Journal of Computer Vision, vol. 2, April 1987, pp. 167-187.
5. J. Shen and S. Castan, "An Optimal linear operator for edge detection", Conference on Computer Vision and Pattern Recognition, Miami Beach, Florida, USA, 1986, pp. 109-114.
6. R. Machuca and K. Philips, "Application of Vector Fields to Image Processing", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 5, n°3, may 1983, pp. 316-329.
7. S. Di Zenzo, "Note on the gradient of a multi-image", Computer Vision, Graphics, And Image Processing, vol. 33, 1986, pp. 116-125.
8. M. Chapron, "A New Chromatic Edge Detector Used for Color Image Segmentation", 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, 1992, pp. 311-314.
9. P. Pujas: Analyse de scènes exploitant des images couleur et 3D, PhD Thesis, University Montpellier II, France, February 1996.
10. R. Taylor and P. Lewis, "Color Image Segmentation Using Boundary Relaxation", 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, 1992, pp. 721-724.
11. R. Nevatia, "A Color Edge Detector and Its Use in Scene Segmentation", IEEE Trans. on Systems, Man, and Cybernetics, vol. 7, 1977, pp. 820-826.
12. A.R. Smith", "Color gamut transform pairs", SIGGRAPH'78, Atlanta, USA, august 1992, pp. 721-724.
13. Y. Ohta and T. Kanade and T. Sakai, "Color Information for Region Segmentation", Computer Graphics And Image Processing", vol. 13, 1980, pp. 222-241.
14. D. Tseng and C. Chang, "Color Segmentation Using Perceptual Attributes", 11th IAPR International Conference on Pattern Recognition, The Hague, vol. 3, August 1992, pp. 228-231.
15. P. Pujas and M.J. Aldon, "Robust Color Image Segmentation", 7th ICAR, Sant Feliu de Guixols, Catalonia, Spain, September 1995, pp. 145-155.
16. S. Tominaga, "Color Image Segmentation Using Three Perceptual Attributes", CVPR, 1996, pp. 628-630.

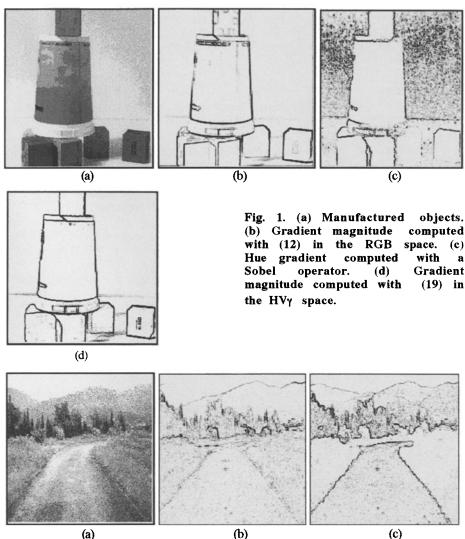(a)

(b)

(c)

(d)

Fig. 1. (a) Manufactured objects. (b) Gradient magnitude computed with (12) in the RGB space. (c) Hue gradient computed with a Sobel operator. (d) Gradient magnitude computed with (19) in the HVγ space.

(a)

(b)

(c)

Fig. 2. (a) Landscape. (b) Gradient magnitude computed with (12) in the RGB space. (c) Gradient magnitude computed with (19) in the HVγ space.
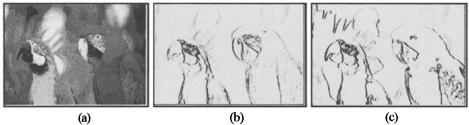

(a)

(b)

(c)

Fig. 3. (a) Birds. (b) Gradient magnitude computed with (12) in the RGB space. (c) Gradient magnitude computed with (19) in the HVγ space.