Parallel Coordinates-based Fuzzy Color Distance and its Applications

Constantin Vertan

Image Analysis and Processing Laboratory, Politehnica University of Bucharest
Bucharest, Romania
Email: cvertan@alpha.imag.pub.ro

Abstract— *Although simple and based on physical realizable color primaries, the RGB color space cannot allow the direct definition of a topology-preserving, perceptually-compatible inter-color distance. The classical normalized inter-color distances are rather complex to implement and tune and do not account for the natural incertitude and confusion regarding both the numerical color values and the color perception. We propose the use of a planar color representation plot, derived from multivariate data visualization, which allows the introduction of a luminance-invariant, geometry-based fuzzy inter-color distance, used with good results in the implementation of distance-based color filters and edge detectors.*

Keywords— fuzzy distance, fuzzy color, fuzzy image filtering.

1 Introduction

The classical inter-color distances are based on standard colorimetric representations, being metrics (or almost metrics) in the color gamut space (a subset of $\mathbf{R}^3$). We propose to investigate a new approach, inspired by the reduced ordering principle of Barnett [1] and also used in multivariate data visualization [2]: namely, the multivariate data (colors) are mapped to some familiar, two-dimensional objects, that can be grouped, compared, and plotted with more ease and are more suited for human perception. Examples of such mappings are the Chernoff faces [3], the Andrews curves [4] and their possible extensions [5], the basic and modified parallel coordinates [6] and the star glyphs [2]. Previous work showed that we could extend the geometric distance between simplified, plane color representations (such as the star glyphs) to color distances [7].

This work proposes the embedding of the uncertainty regarding the exact values of the color components describing the color into a fuzzy color distance, inspired by the two-dimensional geometrical representation of the color by means of the parallel coordinates.

The remainder of the paper is organized as follows: section 2 describes the basic parallel coordinates representation of multivariate data and its application to the representation of colors, section 3 introduces the distance between colors based on their parallel coordinates representation, section 4 introduces the fuzzy extension of the inter-color distance and, finally, section 5 presents some applications of the proposed fuzzy distance in non-linear color image filtering and color edge extraction.

2 Parallel coordinates: the basics

The parallel coordinates representation is a visualization technique that basically allows plotting n-dimensional points and patterns into the bi-dimensional plane. Thus, the parallel coordinates representation transforms multidimensional problems into two-dimensional patterns, without loss of information.

Let us consider a n-dimensional data point $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, originally represented in a Cartesian coordinate system. If each of the n axes of the coordinate system are lined up in parallel with all others into the plane, separated by a fixed distance Δ, we obtain the parallel coordinates system (as presented in Fig. 1) [6]. Obviously, the approach is scalable with respect to the data dimension, n. Visualization is facilitated by viewing the two-dimensional representation of the n-dimensional data points as lines, crossing the n parallel axes, each of them representing one dimension of the original feature space. A n-dimensional point $\mathbf{x}$ is represented in the plane by the series of $n-1$ connected line segments defined by the end points $[(kD, x_1), ((k{+}1)D, x_2)]$, $[((k{+}1)D, x_2), ((k{+}2)D, x_3)]$, , $[((k{+}n{-}2)D, x_{n-1}), ((k{+}n-1)D, x_n)]$ (as shown in Fig. 1).

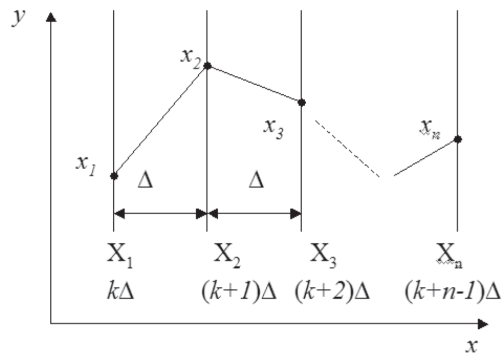


Figure 1: Basic parallel coordinates representation of a n-dimensional data point.

Several interesting mathematical results have been proven for the parallel coordinates representation [6], [8] and several applications were proposed, such as clustering [9], fuzzy rule representation [10], air traffic control [8], and color image filtering [11] based on the independent processing of the points along the individual coordinates axes.

We may notice the strong resemblance of the parallel coordinate representation with the Andrews curves [4], [2] representation method. The Andrews curve representation maps each n-dimensional point into a continuous, analytical curve, obtained as a n-term series expansion having the point coordinates as coefficients, according to a fixed functional basis. The original method uses the Fourier expansion (sine and cosines functions), but modifications were also proposed (e.g. the use of Haar wavelets for color image filtering [5]). For a discrete representation of the Andrews curves, the continuous basis functions are sampled at regular intervals, similar to the

ISBN: 978-989-95079-6-8

use of some parallel axes.

In the following of the paper, we will focus on the representation of colors - three-dimensional data points. Thus, each data point is a triple, $\mathbf{x} = (x_1, x_2, x_3)$. This representation is straightforward for the representation of RGB color data, due to the similar nature of the RGB components; thus, the data points will be $\mathbf{x} = (R, G, B)$, as proposed in [11]. Still, we propose to use a slight modification of the classical model, by using a four-dimensional data vector for the representation of the color, namely $\mathbf{x} = (R, G, B, R)$ as shown in Fig. 2. A color will be thus represented by the polygonal line or by the subgraph (lower limited by the horizontal axis) of that polygonal line, which is itself a polygon and will be subsequently named color polygon.

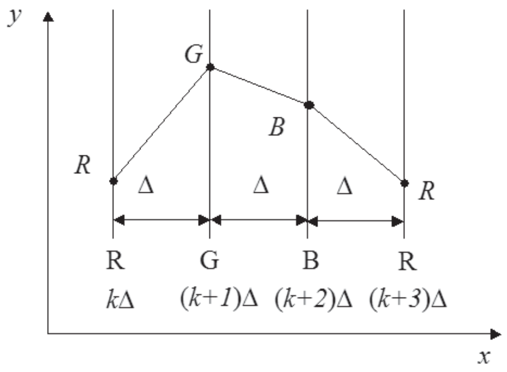


Figure 2: Proposed parallel coordinates representation of a RGB color, viewed as a 4-dimensional data vector.

3 Computing color distances

Color is represented by a three-component vector; the basic representation is the RGB model, for which the vector components are the relative amounts of normalized red, green and blue that additively mix in order to produce that color. Although simple and based on physical realizable color primaries, the RGB color space cannot allow the direct definition of a topology-preserving, perceptually-compatible inter-color distance. The same observation holds for the spectral color representations as well. As normalized by CIE (Commission Internationale de l'Eclairage), perceptual-compatible inter-color distances are obtainable from the Lab color representation. For any two colors $C_1 = (L_1, a_1, b_1)$ and $C_2 = (L_2, a_2, b_2)$, the simplest inter-color distance is the Euclidean distance in the Lab color gamut [12].

Advanced color difference formulas (such as the CMC, BFD or CIE94 [12]) have been introduced, since the basic Lab Euclidean metric do not accurately quantify small-to-medium-sized color differences. Such correctly-measured (as correlated to the subjective estimation) color difference are very complex and untractable for general-use in color image processing. Furthermore, they are not invariant to the change or luminance.

Obviously, the parallel coordinates polygon's shape is related to the nature of the color (properties like hue, saturation, colorfulness, etc.) and the overall area of the polygon relates to the luminance or brightness of the color. It can be accepted that a similarity measure between colors is given by the area of intersection of the two associated parallel coordinates polygons. In order to construct a symmetrical and normalized measure, we will follow an approach similar to the definition of the Canberra distance: the normalization is performed with respect to the average area of the two color polygons. Thus, the distance between colors C_1 and C_2, represented by their corresponding parallel coordinates polygons P_1 and P_2 can be introduced as:

$$d(C_1, C_2) = 2 \frac{Area_{P_1 \nabla P_2}}{Area_{P_2} + Area_{P_2}} \qquad (1)$$

From simple plane geometry considerations one can easily show that the distance in (1) can be approximately put in the form of:

$$d(C_1, C_2) = (D_{RG} + D_{GB} + D_{BR})/3 \qquad (2)$$

where D_{RG} is given in (3) and D_{GB} and D_{BR} have similar forms.

$$D_{RG} = 2 \begin{cases} \frac{|R_1 - R_2| + |G_1 - G_2|}{R_1 + R_2 + G_1 + G_2} & if \Delta R \Delta G \geq 0 \\ \frac{(R_1 - R_2)^2 + (G_1 - G_2)^2}{(R_1 + R_2)^2 - (G_1 + G_2)^2} & if \Delta R \Delta G < 0 \end{cases} \qquad (3)$$

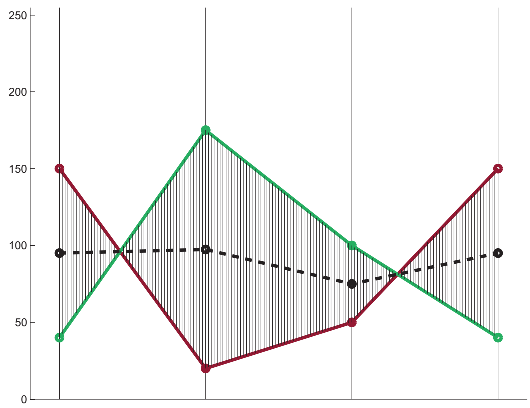


Figure 3: Proposed parallel coordinates distance for two colors (red with $RGB = (150, 25, 50)$ and green with $RGB = (40, 170, 110)$): the hashed area is the uncommon area of the parallel coordinates polygons associated to the colors; the dashed central line is the parallel coordinates polygon having the average area.

It can be easily shown that for unsaturated colors (grays), the proposed inter-color distance is the Canberra distance. The distance in (2) is normalized in the $[0; 1]$ range. It can be easily shown that the proposed distance is invariant to luminance changes, that is $d(C_1, C_2) = d(\alpha C_1, \alpha C_2), \forall \alpha \in \mathbf{R}$.

4 The fuzzy color distance

All image processing algorithms must deal with the imprecision and vagueness that naturally arise in the digital representation of visual information. Noise, quantization and sampling errors, the tolerance of the human visual system are the cause of this imprecision. This strongly suggests that fuzzy models may be used for taking them into account.

In the case of color images, color attributes and color differences play a particularly important role in the perception of

ISBN: 978-989-95079-6-8

object boundaries. The process of measuring color differences must be designed to maintain a balance between the computed and the perceived difference. Still, the simple use of any given color representation does not account for the similarity perception and the visual confusion of colors. We propose to deal with this factors in the framework of a fuzzy representation of the basic RGB color components and obtain a fuzzy distance between colors based on the parallel coordinate representation of the fuzzy colors.

The assumed model means that we will consider each of the R, G, and B color components as fuzzy numbers, centered at their ideal (correct, crisp) value. Thus, the polygonal line representing the given color in the parallel coordinates system will evolve into some fuzzy polygonal line, composed by fuzzy line segments.

We will define the fuzzy distance $\widetilde{D}$ between two fuzzy colors by means of its α-cuts. The corresponding α-cuts of any fuzzy color component are closed intervals along the corresponding parallel coordinate axes. Under these circumstances we can easily determine (from basic geometrical considerations) the minimal and maximal common areas of the corresponding color parallel coordinates polygon and thus to determine the lower and upper bounds of the corresponding α-cut of the fuzzy distance $\widetilde{D}$.

In the particular case of modelling the fuzzy RGB components by fuzzy triangular numbers, the limits of the α-cut intervals of the fuzzy distance are particulary simple to obtain. We will assume that the fuzzy color components $\widetilde{V_i}$ (where V_i is any of the RGB components) are defined by triangular fuzzy membership functions around their crisp value V_{i0}:

$$\mu_{\widetilde{V_i}}(V) = max(0, 1 - \frac{|V - V_{i0}|}{2 * \delta}) \quad (4)$$

In equation (4) above, δ is the support of the membership function, measuring its width around the crisp (correct) value and being thus the direct measure of the incertitude regarding the correct component value V_{i0}.

The α-cuts of any of the terms that sum to the color distance defined in (2) are bounded intervals, like $[D^-_{XY}(\alpha); D^-_{XY}(\alpha)]$. These upper and lower bounds of the α-cuts intervals of color distance terms are given by the expressions in (5) and (6) (for the D_{RG} term given in (3)), according to the cases when the line segments within two successive parallel coordinate axes intersect (corresponding to $\Delta R \Delta G < 0$) or not ($\Delta R \Delta G \geq 0$).

$$D^{\pm}_{RG}(\alpha) = 2\frac{|R_1 - R_2| + |G_1 - G_2| \pm 4(1-\alpha)\delta}{R_1 + R_2 + G_1 + G_2} \quad (5)$$

$$\begin{aligned} P_1 &= (R_1 + R_2 + G_1 + G_2 \mp 4(1-\alpha)\delta) \\ P_2 &= (|R_1 - R_2| + |G_1 - G_2| \mp 4(1-\alpha)\delta) \\ U_1 &= (R_1 - R_2 \pm 2(1-\alpha)\delta)^2 \\ U_2 &= (G_1 - G_2 \pm 2(1-\alpha)\delta)^2 \\ D^{\pm}_{RG}(\alpha) &= 2\frac{U_1 + U_2}{P_1 P_2} \end{aligned} \quad (6)$$

The fuzzy color distance is obtained by its α-cuts, by summing the α-cuts of the corresponding D_{XY} terms in (2). The choice of the width δ of the triangular membership function that models the RGB color components and the discrete nature of the component values imply that the number of distinct α-cuts of the color distance is limited to the rounded value of $\delta + 1$. For instance, the fuzzy distance between the two colors represented in parallel coordinates in figure 3 (red with $RGB = (150, 25, 50)$ and green with $RGB = (40, 170, 110)$) computed according to $\delta = 2.2$ is defined by three α-cuts: 2.53 for $\alpha = 1$, $[2.49; 2.58]$ for $\alpha \in [0.55; 1)$ and $[2.44; 2.62]$ for $\alpha \in [0.09; 0.55)$.

5 Applications

A significant part of nonlinear filters for color images are distance-based, i.e. they rely on the computation of inter-color distances between the colors selected by the filtering window. We shall prove the use of the proposed fuzzy inter-color distance in both smoothing filters (such as the vector median filter) and edge extraction filters, such as standard derivative-based filters.

5.1 *Vector median filtering using the fuzzy inter-color distance*

The classical VMF (Vector Median Filter) [13] defines the vector (or color) median as the vector characterized by a minimal aggregated distance with respect to all other vectors in the filtering neighborhood; this ordering of the color vectors is an instance of the reduced ordering principle [1]. Usually, the distance between color is computed as a L^k norm of the RGB color vectors. We will show here that the use of the proposed fuzzy inter-color distance provides good filtering results.

The fuzzy aggregate distance associate to each color vector within the filtering window is computed based on the α-cuts of the individual inter-color distances. Finally, the aggregated distances (which are also fuzzy numbers defined by their α-cuts) are ranked, and the minimal fuzzy aggregated distance corresponds to the median color.

The ranking of the fuzzy numbers (the fuzzy aggregated distances) is performed according to the classical total integral value. The total integral value of a fuzzy number $\widetilde{D}$ was introduced in [14] as:

$$I_T(\widetilde{D}) = \int_0^1 \mu^{-1}_{\widetilde{D}}(y)dy \quad (7)$$

According to the ranking proposed in [14], the smallest fuzzy number has the smallest total integral value (we may also notice that this ranking is an instance of Barnett's reduced ordering principle [1]).

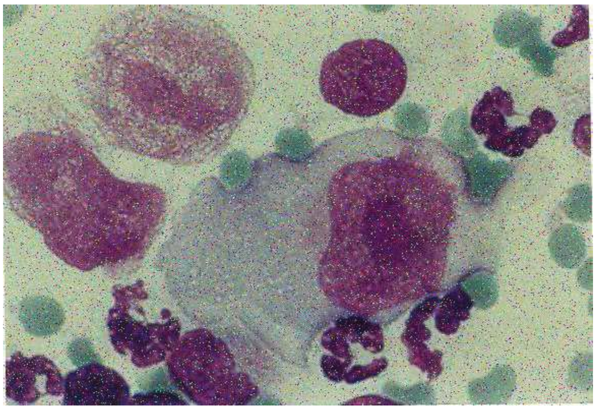
Figure 4 presents the result of a fuzzy-VMF filter applied for impulsive noise reduction in a color image.

5.2 *Fuzzy color edge detection*

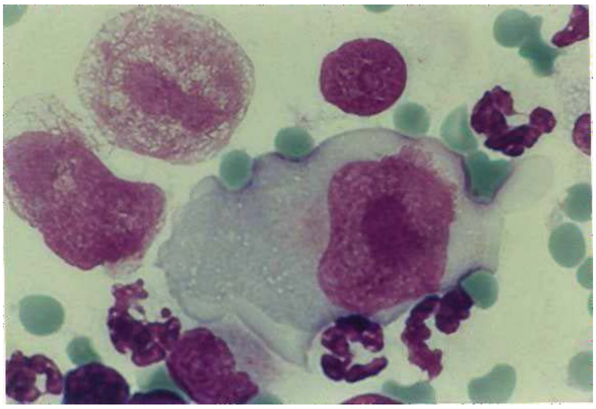
Basically, all edge detection operators rely on the computation of an edge intensity map, which is further thresholded in order to obtain a binary edge map. The edge intensity map exhibits important values for the pixels that are on the boundaries of uniform regions, characterized by a discontinuity (or variation) of their colors.

One of the simplest contour extraction methods is the image Laplacian. We shall use the luminance-invariant fuzzy inter-color distance introduced in (2) for the implementation

ISBN: 978-989-95079-6-8



a)



b)

Figure 4: a) 5% Impulsive noise degraded image and b) fuzzy-VMF filtered image within a 3×3 window, with a fuzzy color model characterized by $\delta = 2.2$.

of a fuzzy color Laplacian operator L. The proposed operator is a modification of the classical (derivative-type) V4-neighborhood Laplacian operator; the color Laplacian is the average inter-color distance within the color at the current processed location (i, j) and its immediate neighboring colors from the color image f. Mathematically we can express the proposed fuzzy Laplacian at location (i, j) as:

$$\widetilde{L}(i,j) = \frac{1}{4} \sum_{(k,l) \in V_4} \widetilde{D}(f(i+k, j+l), f(i,j)) \qquad (8)$$

The fuzzy Laplacian will be defined by the use of fuzzy color distances in (8). The summation of the fuzzy color distance is performed via the corresponding summation of the limits of their α-cut intervals, yielding an α-cut definition of the fuzzy color Laplacian at each image location.

The extraction of the binary edge map image implies the thresholding of the fuzzy Laplacian edge intensity map. According to the desired level of detail that we want to extract form the image, we may use as color edge intensity map either the lower or the upper limit of a specified α-cut interval, as shown in figures 5 and 6. The choice of a particular value of α acts as trimming parameter for the extraction of a more accurate and detailed or a more rough binary edge map, while keeping a same fixed general threshold. This is not possible when using a non-fuzzy color distance, either in RGB or Lab or other color representations.

6 Conclusions

In this contribution we presented a new fuzzy inter-color distance measure, derived from geometrical considerations related to a plane, two-dimensional, reversible color representation. This color representation by parallel coordinates (representing a color as a an open polygonal line) was primarily used in multivariate data representation. The proposed fuzzy inter-color distance can be used with good results in the median (non-linear) filtering of color images and fuzzy edge detection.

Acknowledgment

This work was supported in part by the Romanian National Agency for Scientific Research under the PNCDI2 51-021.2/2007 and 12-136/P2/2008 research grants.

References

[1] V. Barnett. The ordering of multivariate data. *J. of. Royal Stat. Soc. A*, 139(3):318–354, 1976.

[2] W. J. Krzanowski. *Principles of Multivariate Analysis: A User‘s Perspective*. Clarendon Press, Oxford, 1993.

[3] H. Chernoff. The use of faces to represent points in k- dimensional space graphically. *J. of the American Statistical Association*, 68:125–136, 1973.

[4] D. F. Andrews. Plots of high-dimensional data. *Biometrics*, 28:125–136, 1972.

[5] C. Vertan, M. Ciuc, and V. Buzuloiu. Multichannel median filtering by wavelet series expansion. In *Proc. of the 6th International Workshop on Systems, Signals and Image Processing IWSSIP 1999*, pages 77–80, Bratislava, Slovakia, 2-4 Jun. 1999.

[6] Inselberg A. and Dimsdale B. Multidimensional lines I: representation. *SIAM J. Appl. Math.*, 54(2):559–577, 1994.

[7] C. Vertan, M. Zamfir, E. Zaharescu, V. Buzuloiu, and C. Fernandez-Maloigne. Non-linear color image filtering by color to planare shape mapping. In *Proc. of ICIP 2003*, volume 1, pages 808–811, Barcelona, Spain, 10-13 Sept. 2003.

[8] Inselberg A. and Dimsdale B. Multidimensional lines II: proximity and applications. *SIAM J. Appl. Math.*, 54(2):578–596, 1994.

[9] S. Y. Chou, S. W. Lin, and C. S. Yeh. Cluster identification with parallel coordinates. *Pattern Recognition Letters*, 20:562–572, 1999.

[10] L. O. Hall and M. R. Berthold. Fuzzy parallel coordinates. In *Proc. of NAFIPS 2000*, pages 74–78, Atlanta GA, USA, 13-15 Jul. 2000.

[11] C. Vertan, M. Ciuc, and V. Buzuloiu. Parallel coordinates for color image processing. In *Proc. of International Conference on Signals, Circuits and Systems SCS 2001*, pages 469–472, Iaşi, România, 10-11 Jul. 2001.

[12] M. R. Luo. Colour science. In S. J. Sangwine and R. E. N. Horne, editors, *The Colour image processing handbook*, pages 26–66, London, UK, 1998. Chapman and Hall.

[13] J. Astola, P. Haavisto, and Y. Neuvo. Vector median filters. *Proc. IEEE*, 78(4):678–689, Apr. 1990.

[14] T. S. Lion and M. J. Wang. Ranking fuzzy numbers with integral value. *Fuzzy Sets and Systems*, 50:247–255, 1992.

ISBN: 978-989-95079-6-8

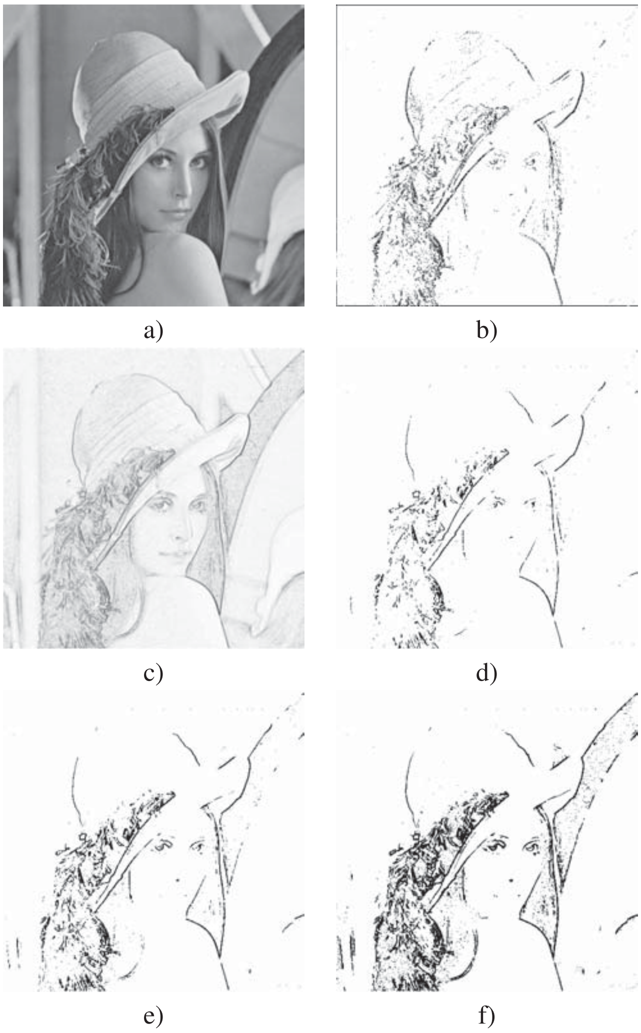


Figure 5: Fuzzy Laplacian edge extraction according to the proposed fuzzy color distance: a) Original color image; b) Classical Laplacian binary edge map; c) Fuzzy color Laplacian edge strength map computed according to (8) and (2) for $\delta = 3.2$ corresponding to an α-cut at α=1; d) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 0.66$ and thresholding of the lower bound of the α-cuts interval; e) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 1$; f) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 0.66$ and thresholding of the upper bound of the α-cut interval. The fuzzy color model is defined by $\delta = 3.2$.

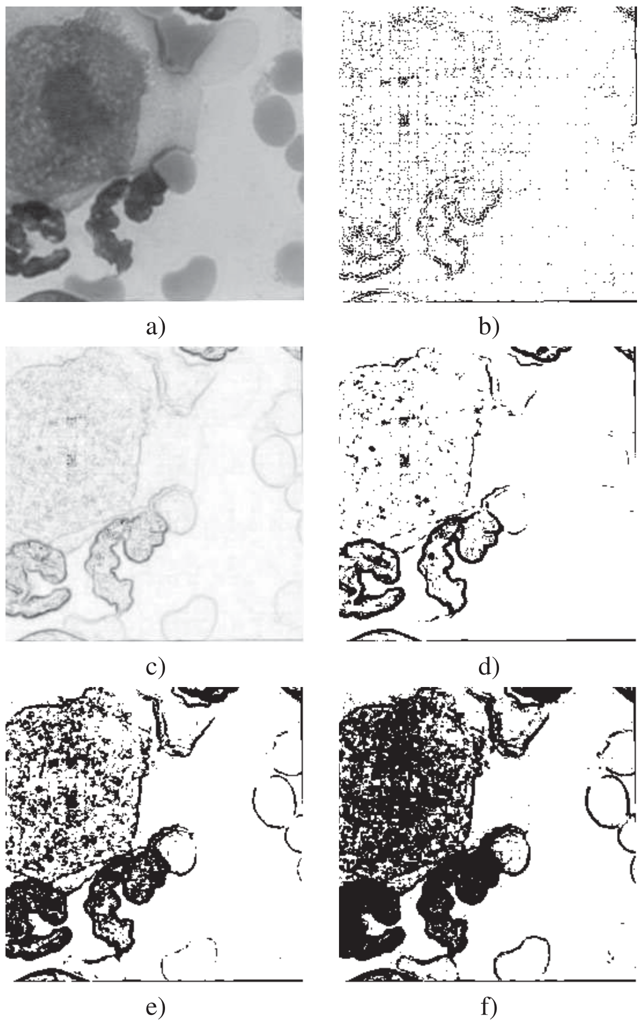


Figure 6: Fuzzy Laplacian edge extraction according to the proposed fuzzy color distance: a) Original color image; b) Classical Laplacian binary edge map; c) Fuzzy color Laplacian edge strength map computed according to (8) and (2) for $\delta = 3.2$ corresponding to an α-cut at α=1; d) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 0.66$ and thresholding of the lower bound of the α-cuts interval; e) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 1$; f) Fuzzy color Laplacian binary edge map computed from c) according to an α-cut at $\alpha = 0.66$ and thresholding of the upper bound of the α-cut interval. The fuzzy color model is defined by $\delta = 3.2$.

ISBN: 978-989-95079-6-8