Engineering, **2011, 3, 203-214**
doi:10.4236/eng.2011.33024 Published Online March 2011 (http://www.SciRP.org/journal/eng)

Unsupervised Segmentation Method of Multicomponent Images Based on Fuzzy Connectivity Analysis in the Multidimensional Histograms

Sié Ouattara[1], Georges Laussane Loum[1], Alain Clément[2]
[1]*Laboratoire d'Instrumentation*, *d'Image et de Spectroscopie* (*L2IS*), *Institut National Polytechnique Félix Houphouët-Boigny* (*INPHB*), *Yamoussoukro*, *Côte d'Ivoire*
[2]*Laboratoire d'Ingénierie des Systèmes Automatisés* (*LISA*), *Institut Universitaire de Technologie*, *Angers Cedex*, *France*
E-mail: *sie_ouat@yahoo.fr*
Received January 20, 2011; *revised February* 9, 2011; *accepted February* 15, 2011

Abstract

Image segmentation denotes a process for partitioning an image into distinct regions, it plays an important role in interpretation and decision making. A large variety of segmentation methods has been developed; among them, multidimensional histogram methods have been investigated but their implementation stays difficult due to the big size of histograms. We present an original method for segmenting *n*-D (where *n* is the number of components in image) images or multidimensional images in an unsupervised way using a fuzzy neighbourhood model. It is based on the hierarchical analysis of full *n*-D compact histograms integrating a fuzzy connected components labelling algorithm that we have realized in this work. Each peak of the histogram constitutes a class kernel, as soon as it encloses a number of pixels greater than or equal to a secondary arbitrary threshold knowing that a first threshold was set to define the degree of binary fuzzy similarity between pixels. The use of a lossless compact *n*-D histogram allows a drastic reduction of the memory space necessary for coding it. As a consequence, the segmentation can be achieved without reducing the colors population of images in the classification step. It is shown that using *n*-D compact histograms, instead of 1-D and 2-D ones, leads to better segmentation results. Various images were segmented; the evaluation of the quality of segmentation in supervised and unsupervised of segmentation method proposed compare to the classification method k-means gives better results. It thus highlights the relevance of our approach, which can be used for solving many problems of segmentation.

Keywords: Multicomponent Images, Unsupervised Segmentation, *n*-D Histogram, Fuzzy Connected Components Labelling, *n*-D Compact Histogram, Evaluation of Segmentation Quality

1. Introduction

In many fields such as medicine, food, robotics, security systems, etc, the acquisition and the analysis of images are essential processes for taking decisions [1-4], as it is often difficult to perform them manually, it is necessary to automate certain tasks by computing. Thus in this paper, we present, a new vectorial segmentation method based more particularly on the full *n*-D compact histograms by a fuzzy connexity analysis.

Segmentation is an important step in image processing and automatic pattern recognition processes based on image analysis as subsequent extracted data are highly dependent on the accuracy of this operation. In general, the automated segmentation is one of the most difficult tasks in image analysis [5], because a false segmentation will cause degradation of the measurement process and therefore the interpretation may fail. The segmentation methods based on the analysis of color histograms are facing the difficulty of treating a huge quantity of information: for a color image of resolution $N \times M$ with each component coded on 8 bits, a classical 3D histogram is an array of 2^{24} cells, the number in each cell being coded on at least $\log_2(MN)$ bits in order to store the number of pixels of each color allowed. In the case where $M = N = 256$, the standard 3-D histogram requires 128 Mo.

Copyright © 2011 SciRes.

A few years ago, we have proposed a new way of coding the *n*D histograms, leading to the so-called compact histogram [6], in which only the occupied cells of the classical histogram are memorised. This reduces the memory space required drastically: it is lowered to a value of 500 ko for a 256 × 256 image with color components coded on 8 bits, without any loss of color information, it shows its efficiency in previous articles [7,8].

Using histograms for classifying colour pixels can be achieved through four different strategies.

The first strategy proceeds in a marginal way. Each component of the histogram is examined separately [9,10]. The method is easy to implement, but it does not take into account the correlation between colorimetric components.

In the second strategy, each colorimetric component is requantized on q bits ($q < 8$) in order to reduce the histogram size [11]. The method is efficient, but it makes an *a priori* color classification.

The third strategy that many authors have developed proceeds by projection of the 3-D histogram on two of the three colorimetric axes. W have developped this strategy in a previous paper [12] and used also in other articles such as [13-15]; The histogram, with populations requantized on 256 levels, is considered as a gray level image and processed by a watershed algorithm. The correlation between components is partially taken into account, but the requantization of the populations alters the true histogram, and the projection can result in ignoring some significant classes.

Here we propose a fourth strategy. It is fully vectorial, which is allowed by the use of the compact *n*-D histogram. Its principle consists in finding the peaks of classic multidimensional histogram using the compact multidimensional histogram by the consideration of a model of fuzzy neighbourhood. The kernels of the built classes correspond to the peaks retained using the algorithm for labelling connected components with binary fuzzy neighbourhood that we realized in this work, it accepts two input parameters which are the degree of fuzzy connectedness expressing the similarity between the vector points of the compact *n*-D histogram and a population threshold limiting the class size. The choice of the fuzzy connectedness here is justified by the difficulty of defining a net similarity between the vector points attribute or "colors" of the pixels of the image [16,17]. This leads to different results of segmentation of the same number of classes for different values of pairs of thresholds. Assessing the quality of segmentation can choose the most relevant segmentation. The results of our multidimensional segmentation method are compared to those of the method of classification *K*-means [13,18, 19] in order to evidence the performance of our approach using supervised and unsupervised [20-23] evaluation methods of the segmentation quality.

2. A Fast and Compact Multidimensional Histogram

We consider multicomponent images (for example multispectral images) with n components $I_i(x_1, x_2)$, for i = 1 to n, each I_i has a tonal resolution of Q possible values and each pixel of spatial coordinate (x_1, x_2) takes a value among the Q possibilities. A n-dimensional histogram of such an image would comprise Q^n cells. For an image with N pixels of resolution, only at most N of these Q^n cells can be occupied, meaning, as n grows, most of the cells of the n-dimensional histogram are in fact empty. For example, for a commun 512 × 512 RGB color image with $n = 3$ and $Q = 256$, there are $Q^n = 256^3 \approx 16 \times 10^6$ colorimetric cells of n-tuples with at most only $3N = 512^2 = 262144$ of them which can be occupied. The idea of a compact representation of the n-dimensional histogram [6], where only those cells that are occupied are coded. The n-dimensional histogram is coded as a linear array where the entries are the n-tuples (the colors) present in the image and arranged in lexicographic order of their components ($I1$, $I2$, $\cdots$, In). To each entry (in number $\leq N$) is associated the number of pixels in the image having this n-value (this color). An example of this compact representation of the n-dimensional histogram is shown in **Table 1**.

The practical calculation of such a compact histogram starts with the lexicographic ordering of the N n-tuples corresponding to the N pixels of the image. The result is a linear array of the N ordered n-tuples with their population. With a dichotomic quick sort algorithm to realize the lexicographic ordering, the whole process of calculating the compact multidimensional histogram can be achieved with an average complexity of $O(N \log N)$, independent of the dimension n. Therefore, both compact

Table 1. An example of compact coding of the 3-dimensional histogram of an RGB color image with Q = 256. The entries of the linear array are the components ($I1$, $I2$, $I3$) = (R, G, B) arranged in lexicographic order, for each color present in the image, and associated to the population of pixels having this color.

R	*G*	*B*	population
0	0	3	11
0	0	201	5
...	...	...	...
255	40	0	32
255	251	254	78

Copyright © 2011 SciRes.

representation and its fast calculation are afforded by the process for the multidimensional histogram. For example, for a 9-component 838 × 762 satellite image with Q = 256, the compact histogram was calculated in about 5 s on a standard 1 GHz clock desktop computer, with a coding volume of 1.89 Moctets, while the classic histogram would take 3.60 × 10^{16} Moctets completely unmanageable by today's computers.

The advantages of the compact multidimensional applied to multicomponent images (see **Figure 1**) are shown in **Table 2**.

3. Fuzzy Connected Components Labeling of Compact Histogram

3.1. Formalism of Fuzzy Logic and Concept of Fuzzy Neighbourhood

We present here some aspects of the theory of fuzzy logic, fundamental to understanding the algorithm of Fuzzy Connected Component Labeling (FCCL) compact multidimensional histograms, developed in following sections. The reader will find a complete overview in [24].

X denote the universe of reference, consisting of elements x, and place us first in the theory of classical ensembles, that is to say net. Then any subset A of X net is completely defined by its characteristic function μ_A, defined on the set {0,1} score, by:

$$\mu_A(x) = \begin{cases} 1 & ssi \quad x \in A \\ 0 & \text{sinon} \end{cases} \tag{1}$$

If the set score is now the continuum [0,1], A becomes a fuzzy subset of X, and μ_A is its membership function. The subset A is then defined by:

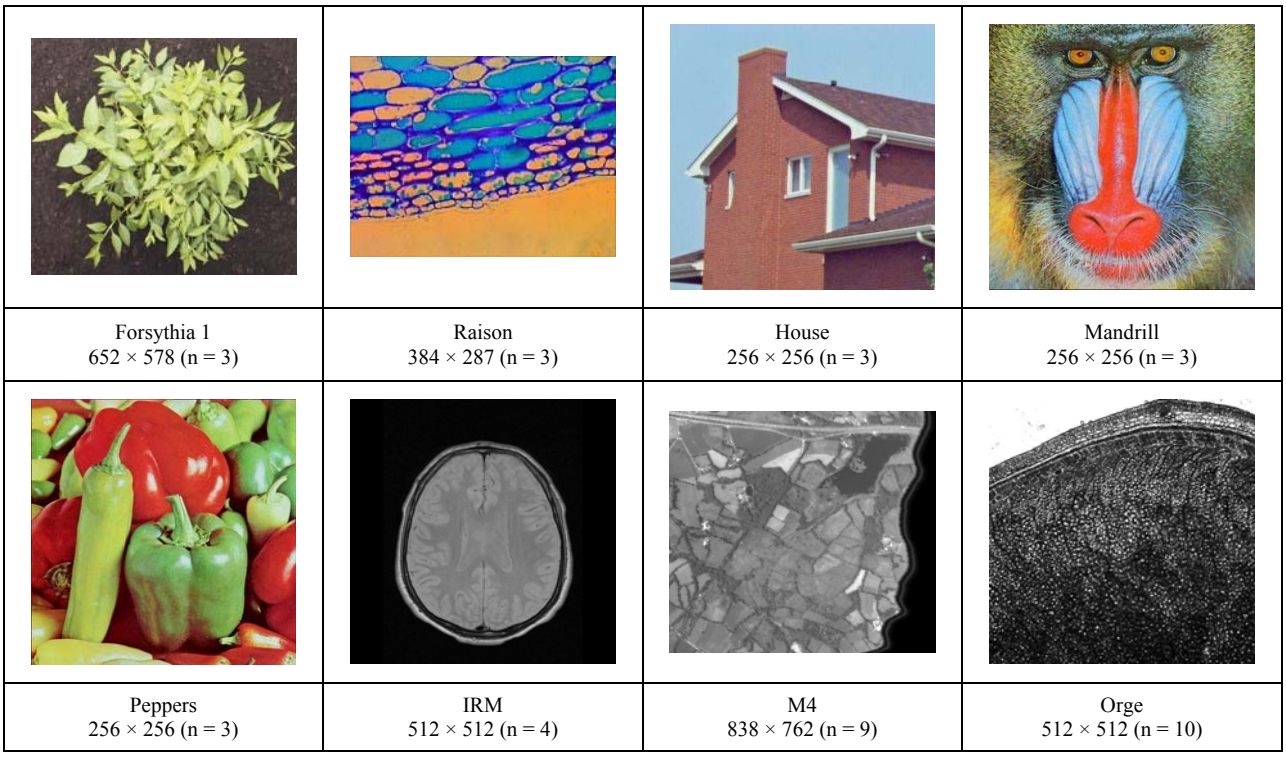
$$A = \left\{ \left[x, \mu_A(x) \right], x \in X \right\} \tag{2}$$

An α-cut of A is the net subset of items with a membership degree to A greater than or equal to α. It is noted $C_\alpha(A)$:

$$C_\alpha(A) = \left\{ x \in X \middle| \mu_A(x) \geq \alpha \right\} \tag{3}$$

The concept of fuzzy relation is a generalization with the fuzzy domain of the concept of equivalence relation defined in the net case. A fuzzy relation can be measured by a scalar in the interval [0,1], the degree to which a logical proposal is verified. With a fuzzy relation R is associated membership function, denoted by μ_R. Let X and Y be two universes of reference, the respective elements x and y. A fuzzy relation between the elements of these two worlds is formally defined as:

$$R: \begin{cases} X \times Y \to [0,1] \\ (x,y) \mapsto \mu_R(x,y) \end{cases} \tag{4}$$



Forsythia 1 652 × 578 (n = 3)	Raison 384 × 287 (n = 3)	House 256 × 256 (n = 3)	Mandrill 256 × 256 (n = 3)
Peppers 256 × 256 (n = 3)	IRM 512 × 512 (n = 4)	M4 838 × 762 (n = 9)	Orge 512 × 512 (n = 10)

Figure 1. Example of multicomponent images.

Copyright © 2011 SciRes.

Table 2. Comparative table of volumes of the classical and compact histograms of the images of Figure 1.

Images	Classical histogram data size (Mo)	Compact histogram data size (Mo)	Number of spels of compact histogram (N)
Forsythia 1	128	0,28	73278
Raisin	128	0,35	90802
House	128	0,36	33925
Mandrill	128	0,65	61662
Peppers	128	0,56	53488
IRM	32768	1,61	140881
M4	$3{,}60.10^{16}$	1,89	116425
Orge	$9{,}22.10^{18}$	4,22	246104

When X=Y, the fuzzy relation is known as binary.

In one of the following sections will be built a fuzzy connected component labeling of compact multidimensional histogram (it is him which plays the role of universe of reference X). The labelling will use a binary fuzzy relation, which we call fuzzy similarity, assessing the degree of similarity between two spells x and y of compact histogram multidimensional. The membership function of this relationship will be expressed as:

$$\mu_R(x,y) = \begin{cases} \dfrac{1}{1+d(x,y)} & si \quad d(x,y) \le M \\ 0 & \text{sinon} \end{cases} \tag{5}$$

This model of fuzzy neighborhood is used in many studies [25,26]. In our case, the threshold M is set at 7 which is the maximum distance allowed so that two spells resembles itself, and the distance d considered will be that of Chebychev or again call Queen-wise distance, given by the following equation:

$$d(x,y) = \max_{i=1}^{n} |x_i - y_i| \tag{6}$$

3.2. Classical Connected Components Labelling

We have recently achieved [8] the connected components labelling (CCL) of n-D compact histograms. It consists in sweeping all the n-tuples present in the compact histogram, in order to gather, under the same label, the n-uplets which are neighboring in the n-D colorimetric space.

Since the n-uplets are ordered in lexicographical order inside the compact histogram, labelling a n-tuple is achieved by sweeping only the $(3^n\text{-}1)/2$ n-tuples connected neighbours preceding it. For example, **Table 3** shows the four doublets of (i,j) to be swept for labelling the doublet (i,j) in the case of 2-D compact histograms.

The CCL is applied to the full n-D compact histograms of five multicomponent images (see **Figure 1**) taken on the web site of gdr-isis. **Table 4** shows the number of connected components (CC) of their histograms and, the minimum and maximum size of CC are expressed as number of pixels compared to the size of the images.

Table 3. Connected neighbours to be swept for CCL of doublet (*i, j*) in the case of a 2-D compact histogram (colorimetric axes *I* and *J*).

Plan I	Plan J
0	0
…	…
i-1	j-1
i-1	j
i-1	J+1
…	…
i	j-1
i	**j**
…	…
2^Q-1	2^Q-1

3.3. Principle of Fuzzy Connected Components Labelling

The connected component labeling with fuzzy neighborhood (CCLF) is based on the notion of fuzzy connectedness and requires the definition of a fuzzy similarity relation between the spells. Contrary to the notion of net neighborhood, a fuzzy neighborhood is defined on a fuzzy subset. In the literature, several studies take into account the spatial information in the definition of fuzzy similarity relation [27,28]. Our neighborhood fuzzy model disregards the fact that we work solely on the histogram. The model of neighborhood fuzzy (fuzzy similarity) selected is defined by Equation (5). The cost of a path between two related spells c and d in the histogram

Copyright © 2011 SciRes.

Table 4. Statistics in related connected components of the images of Figure 1.

Images	Number of CC	Minimum size of CC	Maximum size of CC
Forsythia 1	7649	1	361047
Raisin	39878	1	35158
House	6812	1	20599
Mandrill	21859	1	17686
Peppers	12443	1	21963
IRM	66538	1	152573
M4	94443	1	51147
Orge	87558	1	170465

can be defined by the fuzzy relation ξ as:

$$\mu_{\zeta}(P) = \min\left[\mu_k\left(C^1, C^2\right), \cdots, \mu_k\left(C^{m-1}, C^m\right)\right], m \geq 2 \quad (7)$$

où $P = \left\langle C^1, ..., C^m \right\rangle$ is a path between the spells $C^1 = c$ and $C^m = d$. As several paths may exist between spell c and d, the overall cost of the paths is defined as the maximum value of all path costs calculated on the set of paths. The overall cost of paths between two related spell c and d in the histogram can then be defined by the fuzzy relation ψ as:

$$\mu_{\psi}(P) = \max_{P \in Pcd}\left[\mu_{\zeta}(P)\right] \quad (8)$$

The fuzzy relation ψ is an equivalence relation. The definition of a ψ-neighborhood between two spels requires to fix a minimum threshold of similarity which we denote α. This threshold being fixed, find the predecessors of a spel c returns searching its neighbors of a minimum cost α *i.e* the spels of the net set α-cut of the fuzzy relation ψ.

Let d be an element of the set of predecessors of c in the compact histogram Hc.

$$Predecessors(c) = \left\{d \in Hc \middle| \mu_{\psi}(c,d) \geq \alpha\right\} \quad (9)$$

The principle of the CCLF is similar to that of the CCL, with for only difference the research of predecessors of a spel in compact histogram as shown by Equation (9). In practice we will vary the value of α to study its influence on the number of connected components. Two values bring back to us to particular cases:

- when $\alpha = 1$, the spels of the compact histogram are considered each one as a connected component, so the number of connected components equals the number of spels;
- when $\alpha = 0.5$ the CCLF corresponds to the CCL.

3.4. Our Algorithm of Fuzzy Connected Components Labelling

The proposed algorithm integrates the search of fuzzy predecessors, their number being limited by the fixation of overall cost α. The algorithm returns as output the vector E containing the labels of spels as they appear in the compact histogram, and the number of connected components (*NberCc*) of the histogram.

function $[E, NberCc] = FCCL(Hc, \alpha)$
//*Hc*: compact *n-D* histogram
//*E*: table of spels label
//*NberCc*: number of connected components
//*Tequiv*: management table of equivalences Label
//*taille*: function that returns the number of rows in a table
//*elimineRedondance*: removes redundant elements of an array
//*max*: function that returns the maximum of an array
//*min*: function that returns the minimum of a table
N = *taille*(*Hc*) //number of spels of n-D compact histogram
E(1) = 1 //label of the first spel of *Hc* to 1
Tequiv(1) = 1 //equivalence label 1 to 1
For *i* = 2*: N*
P = *Hc*(*i*,:) //current spel labeling
j = *i* -1
IndexPred = Ø
$d_1 = (1-\alpha)/\alpha$
For *j* =1*: i*-1
While ($Hc(j,1) - Hc(i,1) \leq d_1$)
c = *Hc*(*j*,:)
If ($\mu_k(c,d) \geq \alpha$)
IndexPred = [*IndexPred*; *j*]
EndIf
EndWhile
EndFor
If (*taille*(*IndexPred*) ≠ 0)
Etiq = *E*(*IndexPred*)//returns the labels of the predecessors
Etiq = *elimineRedondance*(*Etiq*)
EndIf
If (*taille*(*Etiq*) == 0)
E(*i*) =1 + max(*E*)
Tequiv(*E*(*i*)) = *E*(*i*)
ElseIf (*taille*(*Etiq*) ==1)
E(*i*) = *Etiq*
Else

Copyright © 2011 SciRes.

```
            E(i) = min(Etiq)
            Tequiv(E(IndexPred)) = E(i)
        EndIf
    EndFor
    For   i = 2:N
      E(i) = Tequiv(E(i))      //global update of labels
EndFor
```

3.5. Example of Results

The FCCL algorithm is applied on the compact n-D histograms of multicomponent images in **Figure 1**. **Table 5** illustrates the results for different values of α. It shows that when the degree of similarity α decrease, the number of connected components (CC) decrease.

4. Fuzzy Vectorial Classification of Tuples

4.1. Principle

The classification of colours is carried out in two steps: the learning step and the decision step.

The learning step is a hierarchical decomposition of populations in the compact n-D histogram. For each level of population p_n, peaks P_i are identified by the FCCL algorithm for a given value of α, which retains the connected components whose populations are greater than or equal to p_n. Each peak is then iteratively decomposed into narrower peaks, beginning from population 0. A peak is labelled as significant if it represents a population greater than or equal to a threshold S (expressed in percent of the total population in the histogram). The procedure is illustrated in part a of **Figure 2** (drawn in one dimension for clarity). We shall name kernels K_i the peaks corresponding to circled leaves in part b of **Figure 2**. In other words, kernels are significant peaks (part a of **Figure 2**) without descendants in the hierarchical decomposition tree (part b of **Figure 2**) (e.g., **Figure 2** shows five significant peaks P_i (i = 0 to 4) and three kernels K_i (i = 2, 3, 4)). The number of classes N_c is taken equal to the number of kernels (the class corresponding to kernel K_i is noted C_i). Therefore N_c depends

Table 5. Statistics in related fuzzy connected components of the images of Figure 1.

Images	Number of CC						
	$\alpha = 0.50$	$\alpha = 0.33$	$\alpha = 0.25$	$\alpha = 0.20$	$\alpha = 0.17$	$\alpha = 0.14$	$\alpha = 0.125$
House	6812	2000	684	252	108	59	28
Mandrill	21859	5939	2016	736	279	106	57
Peppers	12443	3277	1130	514	261	153	110
IRM	66538	33594	16572	7687	3718	1704	837
M4	94443	30373	8583	3020	1307	678	379
Orge	87558	55312	43431	35493	29236	24808	21423

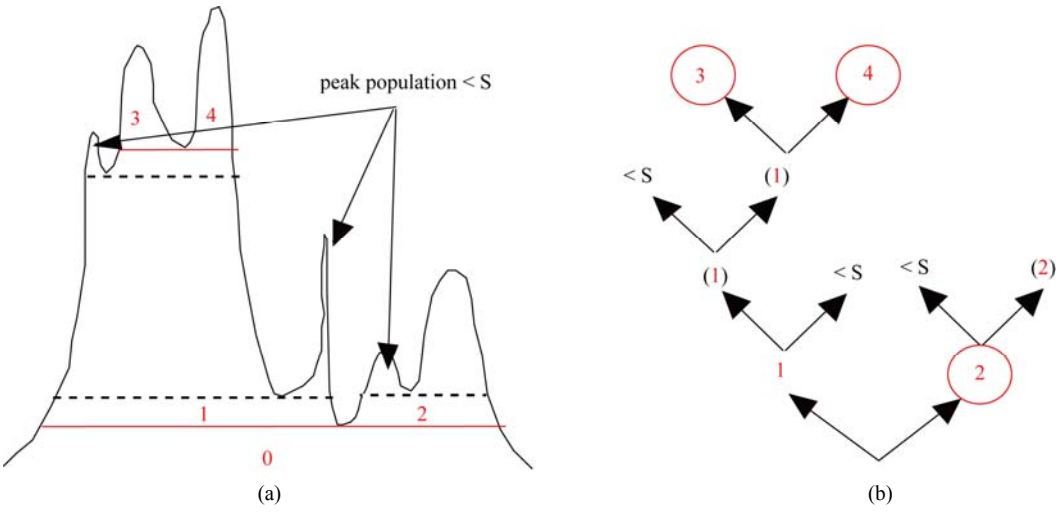


Figure 2. An example of hierarchical decomposition with α = 0.5. The circled leaves (part (b)) correspond to significant peaks as obtained at the end of the iterative decomposition (solid lines in part (a)), whereas leaves marked < *S* (part (b)) correspond to insignificant peaks (dotted lines in part (a)).

Copyright © 2011 SciRes.

on the threshold S, i.e. on the precision the image colors are analyzed with and the value of α the degree of similarity between the spels.

In the decision step, the mass center $\mu(K_i)$ of each kernel K_i is calculated in the feature multidimensional space. Let us denote by β the color corresponding to the point of coordinates $(g_1, g_2, \cdots, g_n)$ in the feature space. Two cases appear: if $(g_1, g_2, \cdots, g_n)$ belongs to K_i, color ß is attributed to class C_i; if not, let us denote by P_k the peak which belong to $(g_1, g_2, \cdots, g_n)$; color β is attributed to class C_i corresponding to kernel K_i, son of P_k, such that $d[\mu(K_i), (g_1, g_2, \cdots, g_n)]$ is minimum, where $d[y, z]$ is the Euclidean distance between y and z.

We give a name of this method that we call *HierarchieFuzzy_nD*.

4.2. Results of Segmentation

The following figure denote **Figure 3** shows an example of segmented images by HierarchieFuzzy_*n*D and K-means methods.

5. Methods for Assessing the Quality of Segmentation

Image segmentation is a fundamental process in image and video analysis. Several approaches have been put forward in the literature [22,29]. We have the region, contour and texture approaches, but this work interests the region approach. It is often used to partition an image into separate regions, which ideally correspond to different

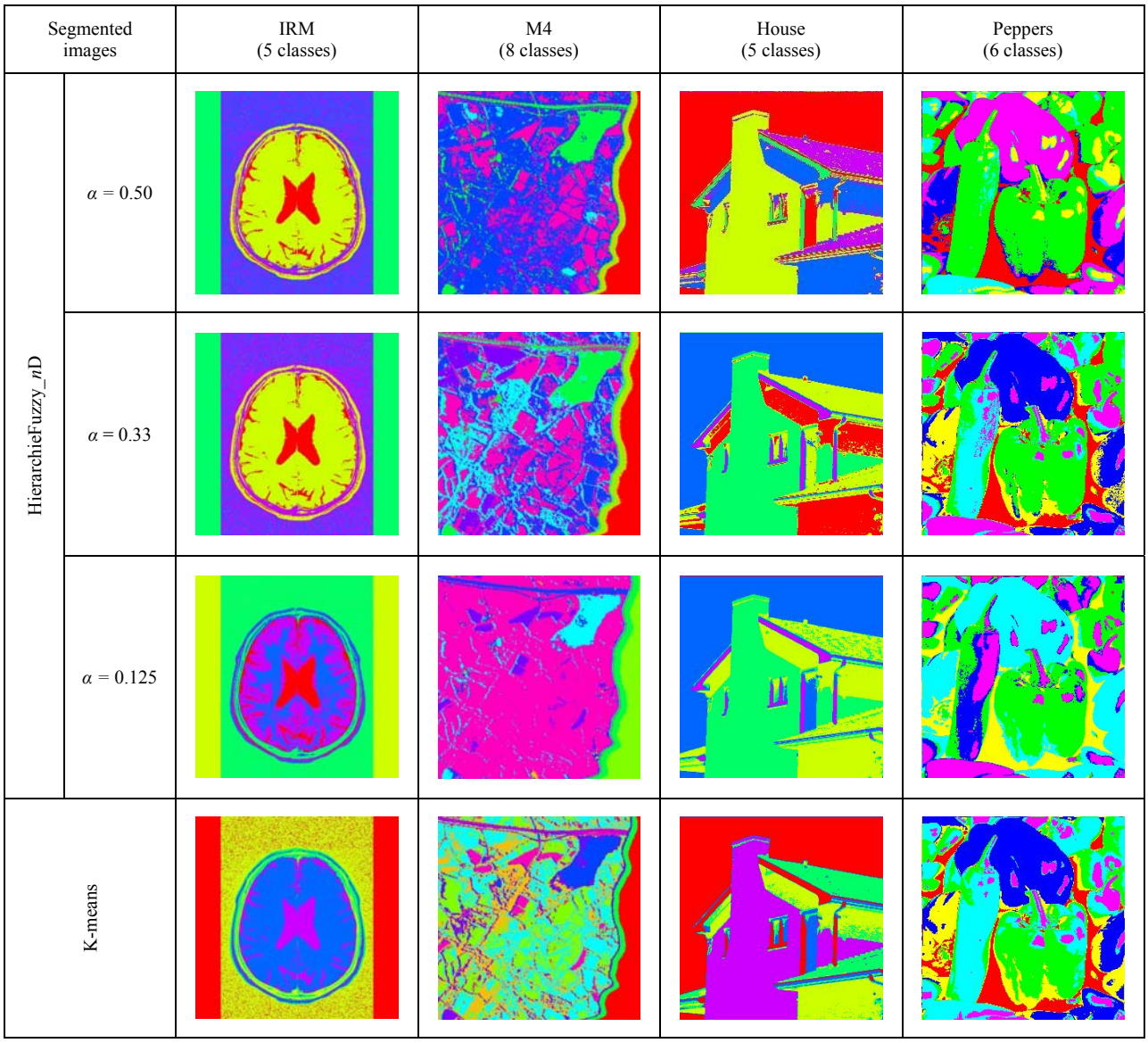


Figure 3. Example of segmented images by HierarchieFuzzy_*n*D and K-means methods.

Copyright © 2011 SciRes.

real-world objects.

Many segmentation methods exist in literature, but it is difficult to evaluate the efficiency and to make an objective comparison of different segmentation methods. This general problem has been addressed for the evaluation of a segmentation result [20]. There are two main approaches:

- There are supervised evaluation criteria based on the computation of a dissimilarity measure between a segmentation result and ground truth. Baddeley's distance [30], Vinet's measure [31], or Hausdorff's measure [32] are examples of supervised evaluation criteria. In practice the ground truth of natural images is a segmentation results manually made by experts.
- There are unsupervised evaluation criteria that enable the quantification of the quality of a segmentation result without any a priori knowledge. These criteria generally compute statistical measures such as standard deviation or the disparity of each region or class in the segmentation result [33-36].

Currently in practice, no evaluation criterion appears to be satisfactory in all cases [22]. However, we choose here the best criterion for each type of evaluation, *i.e.* Vinet measure for supervised evaluation and Zeboudj measure in unsupervised case [22].

One potential benefit of supervised methods over unsupervised methods is that the direct comparison between a segmented image and a reference image is believed to provide a finer resolution of evaluation, and as such, discrepancy methods are commonly used for objective evaluation. However, manually generating a reference image is a difficult, subjective, and time-consuming task [20].

5.1. Vinet Measure

The Vinet's measure [31] that is a supervised criterion which corresponds to the correct classification rate is used as reference for the analysis of the synthetic images. In this case, the ground truth is available. This criterion is often used to compare a segmentation result I_R with a ground truth $I_{R^{\mathrm{Re}f}}$ in the literature. We compute the following superposition table:

$$T\left(I_R, I_{R^{\mathrm{Re}f}}\right) = \sum_{(i,j)} card\left\{R_i, R_j^{Ref}\right\} \qquad (10)$$

Where $card\left\{R_i \cap R_j^{Ref}\right\}$ is the number of pixels resulting from the intersection of regions R_i and R_j in the ground truth. The best match between R_i and R_j^{Ref} is one that maximizes $T\left(I_R, I_{R^{Ref}}\right)$. Vinet's measure gives a dissimilarity measure, it is computed as follows:

$$Vinet\left(I_R, I_{Ref}\right) = \frac{Card\left(I_R\right) - Card\left(R_i \cap R_j^{Ref}\right)}{Card\left(I_R\right)} \qquad (11)$$

This criterion is often used to compute correct classification rate of the segmentation result of a synthetic image.

5.2. Zeboudj Criterion

Zeboudj [35] proposed a measure based on the combined principles of maximum interregions disparity and minimal intraregion disparity measured on a pixel neighborhood. Zeboudj's criterion is defined by:

$$Zeboudj\left(I_R\right) = \frac{\sum_i Card\left(R_i\right) \times C\left(R_i\right)}{Card\left(I_R\right)} \qquad (12)$$

Where $C\left(R_i\right) \in [0,1]$ is the disparity of the region R_i.

This criterion is suitable for evaluating segmentation of homogeneous and little texture images.

6. Results and Discussion

The tests were performed on many multicomponent images. A synthetic illustration of the evaluation of our method of segmentation is applied to images of **Figure 1** with a variation in the number of classes per image to be segmented.

Indexed **Table 6(a)** to **(f)** contains the results of the unsupervised evaluation applied to different images of which we doesn't ground-truth. However **Table 7** illustrates the results of the supervised evaluation obtained from 24 images of forsythia and 1 grape image of which we have their ground-truth.

The relevance of our method of segmentation is made with respect to k-means known in the world of image processing because of its simplicity and its performance.

The **Table 7** shows that in supervised evaluation, the fuzzy multidimensional hierarchical method of segmentation HierarchieFuzzy_*n*D outperforms k-means for each image. This is explained by its ability to detect colorimetrically or spectrally similar classes.

The indexed **Table 6** by **(a)** to **(e)** evaluates in a synthetic way HierarchieFuzzy_*n*D comparing it systematically to k-means and by putting forward the evaluation of HierarchieFuzzy_*n*D in the case of the classical connexity between spels *i.e* for α = 0,5. Thus can be drawn the following conclusions:

- The segmentation method HierarchieFuzzy_*n*D proposed gives better results overall compared to k-means as well for color images as multicomponent images in unsupervised evaluation. This is justified by taking into account the uncertainty in

Copyright © 2011 SciRes.

the spectral data by solving the problem of colorimetric or spectral similarity between spels of the multidimensional histogram.

- While limiting oneself to alpha = 0,5 corresponding to the traditional connexity, we note that for the images M4 and Orge of which the number of spectral plans is high, that HierarchieFuzzy_*n*D is systematically considered to be less powerful than k-means. This lower performance is a consequence of the phenomenon of over-segmentation justified by the **Table 4** which shows the widespread problem of the spels of the *n*D histograms in a multidimensional space when *n* is large, hence the interest of the introduction of fuzzy connectedness between spels characterized by the parameter α, highlighting the relevance of HierarchieFuzzy_*n*D. The major drawback of this last is to be costly in computing time.

Table 6. Unsupervised evaluation of multicomponent images segmentation; in red, segmentations considered to be the best.

(a)

M4							
3 *classes*	Zeboudj	4 *classes*	Zeboudj	8 *classes*	Zeboudj	10 *classes*	Zeboudj
α = 0.5	0.6754	α = 0.5	0.5673	α = 0.5	0.5656	α = 0,5	0.5685
α = 0.125	**0.6830**	α = 0.125	**0.6817**	α = 0.125	**0.6787**	α = 0,25	**0.6735**
k-means	0.5972	*k-means*	0.6292	*k-means*	0.6194	*k-means*	0.6172

(b)

Orge							
2 *classes*	Zeboudj	3 *classes*	Zeboudj	5 *classes*	Zeboudj	10 *classes*	Zeboudj
α = 0.5	0.6920	α = 0.5	0.6194	α = 0.5	0.6265	α = 0.5	0.6229
α = 0.125	**0.8839**	α = 0.125	**0.8839**	α = 0.125	**0.8845**	α = 0.125	**0.8894**
k-means	0.8426	*k-means*	0.8453	*k-means*	0.8334	*k-means*	0.7167

(c)

IRM							
3 *classes*	Zeboudj	5 *classes*	Zeboudj	8 *classes*	Zeboudj	10 *classes*	Zeboudj
α = 0.5	0.7325	α = 0.5	0.4869	α = 0.5	0.4285	α = 0.5	0.4356
α = 0.17	**0.7603**	α = 0.17	0.4968	α = 0.20	**0.7145**	α = 0.14	**0.7167**
k-means	0.7523	*k-means*	**0.6088**	*k-means*	0.5370	*k-means*	0,5458

(d)

Mandrill							
2 *classes*	Zeboudj	4 *classes*	Zeboudj	6 *classes*	Zeboudj	10 *classes*	Zeboudj
α = 0.5	0.4981	α = 0.5	0.3777	α = 0.5	0.3876	α = 0.5	0.3635
α = 0.33	**0.5323**	α = 0.17	**0.4408**	α = 0.17	**0.4280**	α = 0.17	**0.3845**
k-means	0.4577	*k-means*	0.4023	*k-means*	0.3783	*k-means*	0.3647

(e)

Peppers							
3 *classes*	Zeboudj	4 *classes*	Zeboudj	6 *classes*	Zeboudj	10 *classes*	Zeboudj
α = 0.5	0.6607	α = 0.5	0.6433	α = 0.5	0.5977	α = 0.5	0.5939
α = 0.25	**0.6668**	α = 0.25	**0.6515**	α = 0.125	**0.6084**	α = 0.125	**0.5847**
k-means	0.6489	*k-means*	0.5976	*k-means*	0.6052	*k-means*	0.5563

Copyright © 2011 SciRes.

Table 7. Supervised evaluation of real images segmentation; in red, segmentations considered to be the best.

		HierarchieFuzzy_*n*D	K-means
		Vinet (%)	
Forsythia (2classes)	Seg_Img01	**5.78**	8.43
	Seg_Img02	**6.11**	10.22
	Seg_Img03	**6.04**	8.98
	Seg_Img04	**5.34**	11.06
	Seg_Img05	**5.46**	9.84
	Seg_Img06	**4.34**	9.36
	Seg_Img07	**8.26**	14.06
	Seg_Img08	**5.46**	10.43
	Seg_Img09	**4.69**	10.66
	Seg_Img10	**4.85**	10.35
	Seg_Img11	**5.58**	12.21
	Seg_Img12	**3.65**	6.93
	Seg_Img13	**3.57**	6.88
	Seg_Img14	**4.00**	7.07
	Seg_Img15	**5.27**	12.08
	Seg_Img16	**5.58**	11.71
	Seg_Img17	**4.52**	8.05
	Seg_Img18	**4.21**	13.22
	Seg_Img19	**4.18**	8.57
	Seg_Img20	**4.88**	8.44
	Seg_Img21	**6.87**	18.99
	Seg_Img22	**5.04**	9.42
	Seg_Img23	**5.53**	8.52
	Seg_Img24	**4.19**	9.44
	Seg_Raisin (3 classes)	**5.76**	5.96

7. Conclusion

This work is a contribution to the classification by the analysis of multidimensional histograms. It proposes a vectorial strategy to segment color or multicomponent images and provides the tools necessary for its implementation. This work is accompanied by a more general reflexion on the principle of classification in a multidimensional space in connection with the evaluation methods of segmentation.

We have liberated the large volume of multidimensional histograms using the *n*D compact histogram for which we have proposed a fuzzy algorithm for labeling connected components. This algorithm allowed us to develop an unsupervised and nonparametric method by vectorial classification of multidimensional histograms. We made a solution to the problem of over-segmentation generated by the appearance diffuse of multidimensional histograms and evaluated our segmentation results.

We chose to approach the Classification by analyzing histograms nonparametrically with emphasis on algorithmic and geometric approaches comparatively to statistical approaches. We justify this choice by the need first there was to provide solutions to the problem of dealing vectorially histograms of multicomponent images because of their size. The tools that we have developed, allowed us to better understand the mechanism of classification in a multidimensional space and to open channels to predict outcomes.

In Perspective, this work can be directly exploited to define new vectorial methods or strategies for the analysis of multidimensional histograms and their classification. The algorithms developed here can be used to serve statistical or spatio-colorimetric approaches for the segmentation of color and multicomponent images.

8. References

[1] R. Furferi, "Colour Classification Method for Recycled Melange Fabrics," *Journal of Applied Sciences*, Vol. 11, No. 2, 2011, pp. 236-246. doi:10.3923/jas.2011.236.246

[2] W. Yanqing, C. Deyun, S. Chaoxia and W. Peidong, "Vision-Based Road Detection by Monte Carlo Method," *Information Technology Journal*, Vol. 9, 2010, pp. 481-487.

[3] P. Vijayaprasad, M. N. Sulaiman, N. Mustapha and R. W. O. K. Rahmat, "Partial Fingerprint Recognition Using Support Vector Machine," *Information Technology Journal*, Vol. 9, No. 4, 2010, pp. 844-848. doi:10.3923/itj.2010.844.848

[4] L. Lixiong, W. Yuwei and W. Yuanquan, "A Novel Method for Segmentation of the Cardiac MR Images using Generalized DDGVF Snake Models with Shape Priors," *Information Technology Journal*, Vol. 8, No. 4, 2009, pp. 486-494. doi:10.3923/itj.2009.486.494

[5] R. Gonzalez and O. Wintz, "Digital Image Processing," 3rd Edition, Addison-Wesley Publishing Co., Massachusetts, 1991.

[6] A. Clement and B. Vigouroux, "A Compact Histogram for the Analysis of Multicomponent Images," *Proceedings of* 18*th Conference GRETSI on Signal Processing*, Vol. 1, 2001, pp. 305-307.

[7] S. Ouattara, A. Clément and F. Chapeau-Blondeau, "Fast Computation of Entropies and Mutual Information for Multispectral Images," *Proceeding of* 4*th International Conference on Informatics in Control*, *Automation and Robotics*, Angers, Vol. 1, May 2007, pp. 195-199.

[8] S. Ouattara and A. Clement, "Labelling of Compact Multidimensional Histograms for Analysis of Multicomponent Images," *Proceedings of the* 21*st Conference GRETSI on*

Copyright © 2011 SciRes.

the Image Processing, Troyes, September 2007, pp. 85-88.

[9] L. Busin, N. Vandenbroucke, L. Macaire and J.-G. Postaire, "Color Space Selection for Unsupervised Color Image Segmentation by Histogram Multithresholding," *Proceedings of IEEE International Conference on Image Processing* (*ICIP*'04), 2004, pp. 203-206.

[10] J. Hemming and T. Rath, "Computer-Vision-Based Weed Identification under Field Conditions Using Controlled Lighting," *Journal of Agricultural Engineering Research*, Vol. 78, No. 3, 2001, pp. 233-243. doi:10.1006/jaer.2000.0639

[11] G. Xuan and P. Fisher, "Maximum Likelihood Clustering Method Based on Color Features," *Proceedings of the First International Conference on Color in Graphics and Image*, Saint-Etienne, 2007, pp. 191-194.

[12] A. Clement and B. Vigouroux, "Unsupervised Segmentation of Scenes Containing Vegetation (Forsythia) and Soil by Hierarchical Analysis of Bidimensional Histograms," *Pattern Recognition Letters*, Vol. 24, No. 12, August 2003, pp. 1951-1957. doi:10.1016/S0167-8655(03)00034-5

[13] W.-Y. Wei, Z.-M. Li, G.-C. Zhang and G.-Q. Zhang, "Novel Color Microscopic Image Segmentation with Simultaneous Uneven Illumination Estimation Based on PCA," *Information Technology Journal*, Vol. 9, No. 8, 2010, pp. 1682-1685. doi:10.3923/itj.2010.1682.1685

[14] O. Lezoray and C. Charrier, "Color Image Segmentation using Morphological Clustering and Fusion with Automatic Scale Selection," *Pattern Recognition Letters*, Vol. 30, No. 4, March 2009, pp. 397-406. doi:10.1016/j.patrec.2008.11.005

[15] O. Lezoray, "Unsupervised 2D Multiband Histogram Clustering and Region for Color Image Segmentation," *Proceedings of the* 3*rd IEEE International Symposium on Signal Processing and Information Technology*, 2003, pp. 267-270.

[16] C. G. Looney, "Fuzzy Connectivity Clustering with Radial Basis Kernel Functions," *Fuzzy Sets and Systems*, Vol. 160, No. 13, 2009, pp. 1868-1885. doi:10.1016/j.fss.2008.12.010

[17] O. Nempont, J. Atif, E. Angelini and I. Bloch, "A New Fuzzy Connectivity Measure for Fuzzy Sets: And Associated Fuzzy Attribute Openings," *Journal of Mathematical Imaging and Vision*, Vol. 34, No. 2, 2009, pp. 107-136. doi:10.1007/s10851-009-0136-3

[18] D. Al-Bashish, M. Braik and S. Bani-Ahmad, "Detection and Classification of Leaf Diseases Using K-Means-Based Segmentation and Neural-Networks-Based Classification," *Information Technology Journal*, Vol. 10, No. 2, 2011, pp. 267-275. doi:10.3923/itj.2011.267.275

[19] S. Dehuri, C. Mohapatra, A. Ghosh and R. Mall, "A comparative Study of Clustering Algorithms," *Information Technology Journal*, Vol. 5, 2006, pp. 551-559. doi:10.3923/itj.2006.551.559

[20] H. Zhang, J. E. Fritts and S. A. Goldman, "Image Segmentation Evaluation: A Survey of Unsupervised Methods," *Computer Vision and Image Understanding*, Vol. 110, No. 2, 2008, 260-280. doi:10.1016/j.cviu.2007.08.003

[21] O. Kubassova, M. Boesen and H. Bliddal, "General Framework for Unsupervised Evaluation of Quality of Segmentation Results," 15*th IEEE International Conference on Image Processing* (*ICIP*'08), 2008, pp. 3036-3039.

[22] S. Chabrier, B. Emile, C. Rosenberger and H. Laurent, "Unsupervised Performance Evaluation of Image Segmentation," *EURASIP Journal on Applied Signal Processing*, 2006, pp. 1-12. doi:10.1155/ASP/2006/96306

[23] C. P. Juan, E. S. David and F. R. Francisco, "Image Segmentation Based on Merging of Sub-Optimal Segmentations," *Pattern Recognition Letters*, Vol. 27, No. 10, 2006, pp. 1105-1116.

[24] D. Dubois and H. Prade, "Fuzzy Sets and Systems — Theory and Applications," Academic Press, New York, 1980.

[25] I. Bloch, "Fuzzy Spatial Relationship for Image Processing and Interpretation: A Review," *Image and Vision Computing*," Vol. 23, No. 2, February 2005, pp. 89-110. doi:10.1016/j.imavis.2004.06.013

[26] C. Demko and E. Zahzah, "Image Understanding Using Fuzzy Isomorphism of Fuzzy Structures," *Proceedings of International Joint Conference of the Fourth IEEE International Conference on Fuzzy Systems and the Second International Fuzzy Engineering Symposium*, Yokohama, Vol. 3, 1995, pp. 1665-1672.

[27] B. M. Carvalho, G. T. Herman and T. Y. Kong, "Simultaneous Fuzzy Segmentation of Multiple Objects," *Discrete Applied Mathematics*, Vol. 151, No. 1-3, 2005, pp. 55-77. doi:10.1016/j.dam.2005.02.031

[28] P. K. Saha, J. K. Udupa and D. Odhner, "Scale-Based Fuzzy Connected Image Segmentation: Theory, Algorithms, and Validation," *Computer Vision and Image Understanding*, Vol. 77, No. 2, 2000, pp. 145-174. doi:10.1006/cviu.1999.0813

[29] J. Freixenet, X. Munoz, D. Raba, J. Marti and X. Cufi, "Yet Another Survey on Image Segmentation: Region and Boundary Information Integration," *Lecture Notes in Computer Science*, Vol. 2352, 2002, pp. 21-25.

[30] D. Coquinand and Ph. Bolon, "Application of Baddeley's Distance to Dissimilarity Measurement between Gray Scale Images," *Pattern Recognition Letters*, Vol. 22, No. 14, December 2001, pp. 1483-1502. doi:10.1016/S0167-8655(01)00104-0

[31] L. Vinet, "Segmentation and Mapping of Areas of Stereoscopic Pairs of Images," Ph.D. Dissertation, University of Paris IX Dauphine, Paris, 1991.

[32] D. P. Huttenlocher and W. J. Rucklidge, "A Multi-Resolution Technique for Comparing Images Using the Hausdorff Distance," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, New York, 1993, pp. 705-706. doi:10.1109/CVPR.1993.341019

[33] M. D. Levine and A. M. Nazif, "Dynamic Measurement of Computer Generated Image Segmentations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 7, No. 2, March 1985, pp. 155-164.

Copyright © 2011 SciRes.

doi:10.1109/TPAMI.1985.4767640

[34] M. Borsotti, P. Campadelli and R. Schettini, "Quantitative Evaluation of Color Image Segmentation Results," *Pattern Recognition Letter*, Vol. 19, No. 8, 1998, pp. 741-747. doi:10.1016/S0167-8655(98)00052-X

[35] R. Zeboudj, "Automatic Thresholding, Contrast and Contours: The Pre-Treatment with the Image Analysis," Ph.D. Dissertation, University of Saint Etienne, Saint Etienne, 1988.

[36] C. Rosenberger, "Adaptative Evaluation of Image Segmentation Results," 18*th International Conference on Pattern Recognition*, Vol. 2, 2006, pp. 399-402. doi:10.1109/ICPR.2006.214

Copyright © 2011 SciRes.