International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5
Commission V Symposium, Newcastle upon Tyne, UK. 2010

A FUZZY CLUSTERING APPROACH FOR SUPERVISION OF BIOLOGICAL PROCESSES BY IMAGE PROCESSING

J.C Atine[a], SMILETEQUE.COM[b]

[a]jcatine@asr-web.com
[b]contact@smileteque.com
Residence Bel-Air Desrozires, Bat 14,
97170 Petit-Bourg.

KEY WORDS: Fuzzy logic, fuzzy decision tree, T-CAAR, bioprocess, diagnostic.

ABSTRACT:

We present in this paper a method of image segmentation (T-CAAR) based on a fuzzy logic tree. The construction of the classes corresponding to regions of interest has been made by an upset of membership function and a fusion operator into a tree structure. The comparison of this method with fuzzy c-mean, mean shift and watershed allows using an appropriate technique for biotechnological bioprocess concerning the biomass production. The automatized CAAR method incorporates spatial information which yield the result more accurate and more robust to noise. The data structure allows reducing the CPU time and affecting heavily the result.

1. INTRODUCTION

1.1 General Instructions

Before the underlying scientific principles were understood, fermentation has been known and practiced by mankind since prehistoric times. The next stage in its development was dominated by the success in the use of regulatory control mechanisms for the production of amino acids. The first breakthrough came in the late 1950s and early 1960s, when a number of Japanese researchers discovered that regulatory mutants were capable of overproducing amino acids. The higher value of final products like vitamins, baker's yeast, and antibiotics is produced by statistically non-stationary biochemical processes which need continually adaptive recipes for optimal performance. The microbiologists attempt to use computer science tools to understand and control the physiological states of micro-organisms. Different parameters directly link to microbial activity have been measured during a bio-process. We tried to discriminate alive cells from dead ones. For the present, some colorant is used to put in evidence the strong metabolic activity. This allows us to compute viability as the ratio of live cells to dead cells. Image processing allows separating the colored dead cells from not colored alive cells when a colorant is used.

A new method of color segmentation named T-CAAR is presented in this article. Our method constructs a fuzzy decision tree using Union and Find operators. The different groups are represented by a tree labeled by the root. Each node of the fuzzy tree allows us to cluster the pixels with the same spatial and colorimetric similarity. The T-CAAR algorithm is part of X-CAAR algorithms. X-CAAR algorithm is an intelligent structure of algorithm using different sampling method to treat data. The XCAAR Method can change his membership function during the treatment is desired by the user.

The fuzzy induction method we have adopted is based on a fuzzy decision tree that is an extension of binary decision trees. A fuzzy decision tree is a tree with fuzzy decisions functions. A 0-1 decision tree utilize 0-1 decision functions. The decision tree Algorithms operate in the same way to construct a downward tree, from the root to the leaf, according to a general method to get arborescent system of clustering. The choice criterion in the algorithm steps is the difference between objects (Olaru, 2003; Janikow, 1996). Decision trees are simple, clear and easy to do tools. The arborescent structure is similar to a rule based "if . . . then", this fact explains a decision obtained through a path.

Decision tree technique is a well-known method to take classification decisions in pattern recognition. Its principal property focuses on the fact that a broad number of classes could be maintained while the time of the final decision is minimized with small local decision. The "dividing to reign" techniques (divide and conquer strategy) is one of the great families of approaches to resolve problems like learning. First, it tries to identify partial problems to find a solution and then, it solves general problems combining these solutions. The decision trees learning algorithms are built on this principle.

Fuzzy decision tree takes numeric and symbolic data into account during the tree construction and clustering process. Reasoning using numeric-symbolic values, such as size is tall or middle size, is reasoning close to human thought and the use of fuzzy set theory lead a better comprehensibleness to decision tree during the process of numerical data. It has been showed that fuzzy set theory gives a better robustness when a new example is classified (Marsala, 1998). We choose different methods for comparison: thus fuzzy c-mean allows comparing with a fuzzy clustering method and meaning shift with a method based on estimation of probability density and watershed because it is a much known method used in cell's segmentation.

2. THE LEARNING ALGORITHM LAMDA

We introduce here the membership function used in LAMDA. We use these function in T-CAAR method of X-CAAR algorithms.

The first function, named "fuzzy binomial" is the generalization to unity of binomial probability function. Let's take a variable x with two values associated to logical value 0 and 1.

With $\rho = prob[1]$ then $1-\rho = prob[0]$ and we can write $prob[x] = \rho^x (1-\rho)^{1-x}$.

In fuzzy logic, we consider that $x \in [0,1]$, can take all the real value between 0 and 1. We define the true fuzzy set by the membership function

$$\mu(x) = \rho^x (1-\rho)(1-x) \qquad (1)$$

We can see that this function depend on one parameter. In the binary case its estimation is the frequency of $x = 1$ such as for *N* sample can be written as $F_N \frac{1}{N}\sum_{i=1}^{N} x_i$. Likewise, we can define the estimation $\rho_N = \frac{1}{N}\sum_{i=1}^{N} x_i$. It is an average.

The second function considered is the non-normalized Gaussian:

$$\mu(x) = e^{-\frac{1}{2}(\frac{x-\rho}{\sigma})^2} \qquad (2)$$

Contrary to the first function, this one is defined for all $x \in \Re$. It is centered on the value $x = \rho$ which variability is represented by σ.

X-CAAR deals with various other functions as Mahalanobis, Max, Min, Euclidian distance. Each of the function in X-CAAR algorithm can work on RGB or XY component.

The presented function are used to obtain the relative membership degree we termed DAR in X-CAAR. They are parameterized by the average ρ and, in the Gaussian case by the variability parameter σ.

The fuzzy clustering algorithm for multivariate data analysis, LAMDA, has been experimented on various applications where lack of knowledge on classification criterion could be replaced by validation of results by an expert.

Utilization of adequacy and exigency concepts allows a dialog with an expert leading to an adjustment of the method parameters.

The method is robust facing the great number of variable and numeric data (Atine, 2004).

Each object is described by a set of n attributes or descriptors, and represented by vectors of n components. The set of those vectors will be called database. Qualitative or quantitative descriptors can be considered in LAMDA method. Qualitative descriptors take their values from a non-ordered set (color, sharp).

Quantitative ones take their values from a totally-ordered set (e.g. weight, temperature, etc). In this study, only quantitative values are considered.

To make possible a direct confrontation between classes and objects, it is necessary that the concepts are described with the same descriptor used for observations. Given an object x and a concept C; LAMDA computes for any descriptor *D* a matching degree between the value that *D* takes over x and the value that *D* takes over C. It is more appropriate to talk in our methodology about relative adequation degree (DAR) than "marginal adequacy degrees" in LAMDA. When the degrees are known, they are used by the system to compute the global adequacy degree GAD of object x to concept *C*.

LAMDA (Aguilar-Martin, 1990) can model the concept of maximum entropy, that is, total homogeneity or indistinctness. This concept is represented by a class named Non-Informative-Class (abbreviated NIC). All objects are accepted in the same way by NIC. So there is no distinction between objects. The context determines the existence of this class and implies that NIC is always present in the space of pre-established concepts.

In concept formation process, a class will be initialized any time.

Initially NIC is the concept with the highest adequacy degree to the object. Observations such that their maximum adequacy degree is taken by NIC, are not assigned to any of the significative pre-established concepts in the recognition concept process.

Then, NIC remains empty in the formation process, whereas NIC will be formed in the recognition process by those examples that cannot be described for the space of significative concepts, which is invariant in time. NIC plays the role of minimum threshold of meaning traditionally used in classic clustering algorithms. But this threshold is not fixed arbitrarily in the methodology because it is the context which determines it.

When an object is given to the system, the algorithm can be activated in two different modes. Recognition mode is used when concepts are already known. Otherwise, the algorithm makes a learning process that involves the formation of concepts or the modification of the existing concept. This learning process may be initialized by a set of pre-established concepts, or left without initial information, that is with NIC the only one initialized concept.

Let us assume that when object x is given to the algorithm, the existing concepts are $C_0, C_1, \ldots, C_k$, where C_0 is *NIC*.

• Step 0: For each component x_i of object x, get the extremum value $x_{\min}$ and $x_{\max}$ of the quantitative components.

Compute the normalized value x_i for one element x for the descriptor i using the formula:

$$x_i \leftarrow \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \qquad (3)$$

$i = 1 \ldots n$, where *n* is the number of attribute of x;

• Step 1: Compute the relative adequacy degrees of x_i to all concept C_j, which is DAR_{x_i/C_j}, $j = 0, \ldots, k$ et $i = 1, \ldots, n$;

• Step 2: Compute the global adequacy $GAD(x/C_j)$;

• Step 3: Search for the maximum, $\max_i GAD(x/C_j)$, computed degrees between all classes to deduce the possible owing class.

The fusion operation between the relative adequacy of concepts is made using a linear interpolation between union and intersection in fuzzy logic:

$$\alpha T + (1-\alpha) S \qquad (4)$$

T is a T-Norm and S is the dual T-Conorm (respectively intersection and union operators). α is exigency coefficient.

Each descriptor value contributes to the global adequacy of one object to one class through relative adequacy degree (DAR). The adjustable parameter is called "exigency". Maximum exigency corresponds to the conjunction AND associated to intersection.

Let's note that *DAR = 1* represents the total adequacy of one attribute to a class and *DAR = 0* represents the total inadequacy of one attribute. Respectively, if the combination of *DAR*, which is the global adequation degree, is equal to 1, then the object has a total adequacy to the class.

The best results are obtained, in our case, according the maximum to intersection. A maximum value is assigned to exigency degree $(\alpha = 1)$.

There exist nevertheless other connectives (or aggregation operators) (Waissman-Vilanova, 2000).

In pratic, 2 T-Norm are used: the Min and the Product, which provide respectively the dual T-Norms, the Max and the "probabilist sum" *(x + y − x * y).*

Analysis of step 1. Two cases may occur:

1. Recognition mode. Object x is placed in C_j . If C_j is NIC, then x is said "unrecognized";

2. Learning mode. There are two possibilities:

Let $\mu_j = \max_k GAD(x / C_k)$

• μ_j corresponds to NIC. This means that x is the first element of a new class C_{k+1} and the representation of this new class will depend on x . We use the following formula to initialize the new class:

$$\rho_{New} = \rho_{Nic} + \frac{(x_i - \rho_{NIC})}{N_0 + N + 1} \quad (5)$$

where ρ_{NIC} is determined by the context, *N* is the number of elements of the class ; and N_0 is a parameter that represent a virtual number of elements of NIC, this determine the fuzzy degree, the larger it is, the greater number of elements, different from x_i, will be accepted by C_j . $\mu_{NIC} = \mu_0$. N_0 is a fuzzy parameter and the larger it is, the greater number of elements will be accepted by C_j .

• μ_j does not correspond to NIC, in this case, x is placed in C and the representation of C will be modified to include x .

$$\rho_{New} = \rho_{Old} + \frac{1}{N_C + 1} * (x_i - \rho_{Old}) \quad (6)$$

where $N_C = N_0 + N$ where *N* is the number of elements in class *C* to which the object is assign at this step of the algorithm.

One of the drawbacks of LAMDA methodology is that it doesn't takes into account spatial position of data. This is caused by the order of treating the data by the algorithm (it is sequential).

We propose to study different algorithms allowing us to put in evidence the importance of the order of treating the data in the image and thus position of objects. It is evident that the mode of reading the data by the algorithm is link to the structure used.

We have developed 3 supervised algorithms based on LAMDA principle, named X-CAAR. X-CAAR Algorithms are FLA-CAAR for flooding linear algorithm CAAR, FCA-CAAR for flooding chained algorithm CAAR, FRA-CAAR for flooding linear algorithm CAAR, T-CAAR for Tree based CAAR algorithm. These algorithms allow us to put in evidence the importance of the way of treatment of the data in X-CAAR methodology. In FLA-CAAR the user specifies the center and the algorithm read the data (using 4 neighborhoods) to the left as much as it is possible (if the GAD value allows recognition), after it change the direction, to the right, upward, and downward. The FRA algorithm read the data clockwise from initial points. In FCA algorithm the reading of the data can be represented by the way of a liquid for flooding a container from an initial point (or more points).

The structure of T-CAAR is most adapted to diffused images and allows us to obtain the more acceptable result.

3. T-CAAR CLUSTERING ALGORITHM

The notion of connections between the pixels is introduced by the notion of path and neighbor. A structure based on decision trees (Eyrolles; Fiorio, 1994) is used to represent the classified data.

This approach tries to classify a pixel according to some tests in each tree knot on the attributes which describe it. These tests are organized such as the response to one of them involves the next pixel test. The principle is to organize all the possible tests as a tree. A tree leaf indicates one of C classes (but to every class may correspond to several leaves) and to each node is associated a test (a selector) concerning an attribute, the R, G, B color values of pixels, their position or both of them.

In the data tree structure, only the father identifies the class and contains its parameters (the number of elements and the modality of its descriptors). Each node contains the parameters of the subclass he is the father. A forest is represented by a triplet weight (or size), father, modality; "Size" will keep the number of parts of the partition, so we will be able to delete non significant region.

The Structure we used is like: Structure SET = { ValueR, ValueG, ValueB, SET * Father, ValueX, ValueY } where ValueX and ValueY represent the barycentre of the current class.

The union cost is proportional to the height of the arborescence. Through "region growing" and tree representation method, fusions of regions are made. The "region growing" (Li, 2003) is often used in segmentation of aerial images (Bicego, 2003). First, we start from small regions, which are either pixels or points and we group them until we consider that we are in the optimal case.

Then, when a region becomes the son of another one it means that the root of the representative tree becomes the son of the element that identifies the father region. Region growing is often associated with union-find. Our operation find allows going up to the tree father and identifying the class. The Union allows to merge or not two elements. The Union operator uses the LAMDA fuzzy decision in equation 1.

In T-CAAR algorithm, if the image size is $N*M$, a number of $N*M$ classes are initialized at the beginning of the algorithm.

These classes contain one pixel, each of the pixels in the image.

Our classification process makes fusion of two classes that correspond to two different regions, into one class. Let's take an example of 2 classes *X* an *Y* which parameters are μ_X and μ_Y.

In T-CAAR, the number of pixels of each class is N_X and N_Y.

We define the notion of class weight in relation to the number of pixels contained in this class. If $N_X > N_Y$ the fusion of the classes will be made using the class *X* to compute the adequation degree. And if the next test $GAD(\mu_Y / \mu_X) > GAD(\mu_0)$ is satisfied then the union will be made. If $N_y > N_X$ the fusion of the classes will be made using the class *Y* to compute the adequation degree. And if the next test $GAD(\mu_x / \mu_y) > GAD(\mu_0)$ is satisfied then the union of object *X* to class *Y* will be done. μ_0 is the

parameter of the *NIC* and $GAD(\mu_Y / \mu_X)$ represents the GAD of the class X in comparison with a class of weaker weight, Y.
To construct the tree we use the update equation 6.

The characteristic algorithm is showed below:

• We short the union pixels tests in order to test only adjacent regions, and to favor pixels whose attributes are close to each other. For example, let's take five objects A, B, C, D, F, see figure 1. Each point have the R, G, B feature. The neighbor objects are A, B, C, D according to figure 1. A shorted list containing the pair of connected elements according to the difference of the grey level _ is made (figure 1) ;

• The distance selected for the presented result is the Maximum of the difference between each plane R, G, B:
ε = max ($A_R - C_R$, $AG - C_G$, $AB - C_B$)

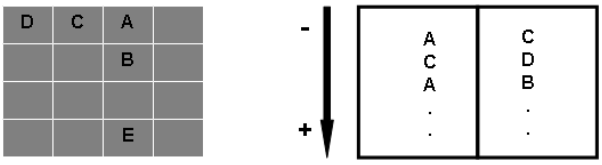


Figure 1: On the left we see the arrangement of the pixels in the plan. On the right, the shorted list. The smallest differences, between RGB plans, are placed towards the minus sign and the biggest differences are placed towards the plus sign.

The elements connected, in the resulted tree, at the left in figure 1 are close spatially and have the same theoretical color.
If A and B are close to each other, we take the decision relative to their own region. The construction of the region {A,B,C,D} is guided by the following plan, supposing that the possibility to merge A and C is stronger than the one to unite C and D and etc ;

• If the image size is N * M. The number of couples of pixels in all the image is $(N * M)^2 - N * M$. We use 4 and 8 neighborhood. In 4 neighborhood the number of pixels is $(N - 1) * M + N * (M - 1)$, and in 8 it is $(((N - 1) * M + N * (M - 1)) + (2 * (N - 1) * (M - 1)))$. In this method the union test with each pixel of the neighborhood is done supposing the first union test is superior to the last union test;

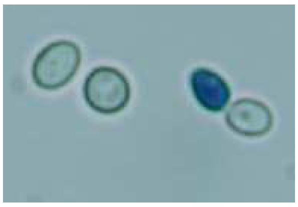
• In addition to this, to obtain the final region of our image, the merging test are organized such as the possible union of the first couple of elements superior to the one of the last couple of pixels owing to the image ;

• At the end of the classification, our system can allocate arbitrarily a color at each class or a color obtained from the mean of all the pixels belonging respectively to the region.

4. BIOLOGICAL IMAGE SEGMENTATION

The images have been obtained using a Nikon video-camera coupled to a Olympus microscope X40. The cell's sample is the result of a biochemical reaction between a colorant and the cells. If we have a low metabolic activity, the coloring agent penetrates into the cells. It is going to change the original color in a dark blue color. So dead cells (symbolized D) are colored blue. In the case of microscopic images, we have the diffusion of the incident light on the cell's membrane. Of course, our clustering method needs to take this problem into account, which means to tune some parameters. We showed that only fuzzy decision tree give satisfactory results. It allows to diagnostic the microbian cell's evolution into a bioreactor. Based on the knowledge of the acquisition, we have defined a triangular membership function for the image background. We use the flooding algorithm, FLA-CAAR, a part of the X-CAAR algorithm, associated with edge detection to easily distinguish the background from the cells. The background is labeled with the fill-color. The flooding algorithm operates along the largest found area for initialization, which is the background. Different filter are used ((Image - Filter * coefficient) + Image) for flooding in order to obtain a buffered region with the cells localization. The segmentation results on biological images are presented for an X-CAAR algorithm other than T-CAAR. The enhancement of the image allows us to increase the gray level between image background and the analyzed region of the image.
We can see now, the flooding step (FLA-CAAR).

```
FLA8(Im) { p : starting point of Im, p(x,y) 1)
FIND LEFT EDGE OF
COLOR AREA int LFillPts=x ;
//the location to check/fill on the left
ptr=p ;
//is the pointer to the current location
while(true)
{
        ptr[0]=fillcolor ; //fill with the color
        PixelsChecked[LFillPts,y]=true ;
        LFillPts-- ; //decrease counter
        ptr-=1;//decrease pointer
        if(LFillPts<=0) ||
        (pixel is in the area defined
        by a start color ; GAD > NIC^3)||
        (pixel have not been checked:
        PixelsChecked[LFillPts,y]=false)
        break ;
//exit loop if we're at edge
//of bitmap or color area
}
LFillLoc++ ; 1)FIND RIGHT EDGE OF COLOR AREA 2)
START THE LOOP
UPWARDS AND DOWNWARDS
//By making the same thing}
```



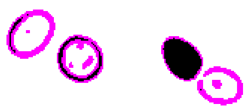
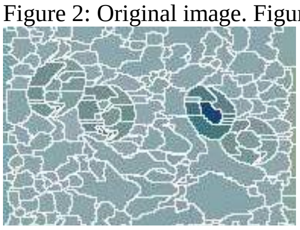


Figure 2: Original image. Figure 3: Fuzzy C-Means.



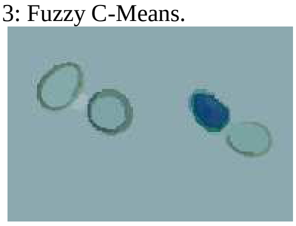


Figure 4: Watershed. Figure 5: Mean-Shift.

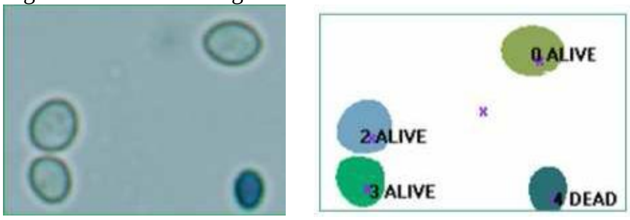


Figure 6: T-CAAR result. The results are presented with a random color per each class for a better observation.

5. COMPARISONS AND EXPERIMENTS

This new method, T-CAAR, has been compared with traditional algorithm: Mean-Shift, Fuzzy C-Mean, and Watershed. We can see visually the result on biological images from figure 2 to 5.
When we watch the cells through the microscope, a diffusion effect appears inside the cell. We have likewise the light problem around the edge of this cell which can be discontinuous.

Center_x	Center_y	Surface	Longest axis	State
143	24	1137	42	A
28	77	1010	40	A
27	109	964	38	A
152	117	695	33	D

Table 1: Data result obtained from the algorithm T-CAAR for the cells. A: alive ; D : dead. Unit for surface is pixels.

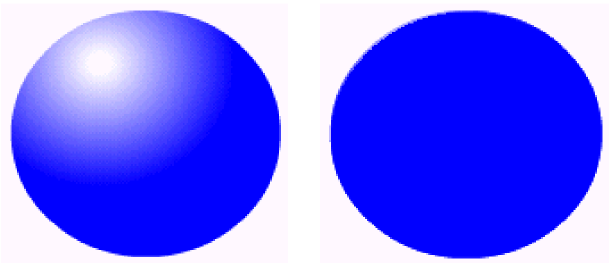


Figure 7: On the left: the reference images which is a 3D sphere. On the right: a disc, the hoped result.

The diffusion effect which is present on the cell has been simulated on a sphere under the light. This object is modeled in 3 dimensions in order to measure the data in these images initially without noise. The owing classes of the pixels are known and we add diffusion to the sphere. In addition to that, the region inside the cell is statistically equivalent to the background.
The goal to reach by the classification of the pixels is to make disappear the diffusion task on the object in order to get out from the clustering process a full sphere we can see like a disc. The object surface is the evaluation criterion chosen and the error is computed on the shape of the obtained surface. The image we want to have is represented at figure 7. Concerning the cell, we should obtained a full cell (the form represented the cell) to measure morphological data.

Table 2 shows the results obtained for the algorithms T-CAAR (TL),Mean-Shift (MS), Fuzzy C-Mean (FCM), LAMDA,Water-Shed, Modified WaterShed (Watershed + k nearest neighbor), XCAAR (FLA-CAAR, FCA-CAAR and FRA-CAAR). We vary a random noise in the image. The ration of correct clustering varies from 0 to 1; 1 is obtained for the better segmented image and 0 for an invalid recognition of the disc in the image. CF supposes the background knowledge.

Analysis of the Result on Test Images

The real images treated are noisy, so we submit the algorithm to noise in order to put in evidence the algorithm which give the better result.
The noise added to the image is points of random color, added randomly in the image. It is the percentage of points of equiprobable random color.
In the resulting image of the clustering, we want the diffusion task to disappear because it is the most important phenomenon observed on the cell.
Here is the result of each algorithm. The watershed gives several classes. It gives bad result relatively to noise and when the image is too smooth. Watershed gives non-exploitable result. We try to merge the region obtained by the watershed using a k nearest neighbor but it isn't more interesting.

Noise	0	2	4	8	16	32
Algorithms	Ration of correct classification					
T-CAAR/RGB	0.94	0.90	0.86	0.86	0.82	0.71
T-CAAR/HSL	**0.98**	**0.98**	**0.98**	**0.98**	**0.98**	**0.97**
FCM/RGB	0.62	0.66	0.66	0.64	0.61	0.55
L/HSL CF	**0.98**	**0.99**	**0.97**	0.94	0.87	0.73
W		0.59	0.29	0.27	0.31	0.38
M-W		0.59	0.48	0.37	0.37	0.32
FLA-CAAR 8N/RGB	0.83	0.81	0.76	0.68	0.74	0.50
FCA-CAAR 8N/RGB	0.66	0.65	0.63	0.62	0.56	0.47
FRA-CAAR 8N/RGB	0.66	0.65	0.63	0.61	0.55	0.51

Table 2: Result obtained using different percentage of noise added to the images. The results are given for the following algorithms T-CAAR, Fuzzy C-Mean (FCM), LAMDA (L), WaterShed (W), FLA-CAAR, FCA-CAAR. The proportion of correct classification varies for 0 to 1; 1 is obtained for the best result and 0 for an invalid recognition of the disc. CF supposes the background knowledge. N is the neighborhood used. M-W is a watershed combined to nearest neighbor algorithm.

Extracting the background before applying LAMDA, the method seems to work but another problem is due to spatial integrity of the data. This algorithm merges two clusters spatially different.
Comparatively if we run T-CAAR with background knowledge, the ball classification is up to 1 even if noise reaches 32.
The percentage of error for the FCM is due to the number of classes fixed in advance by the user. The algorithm can merge noise with a part of the sphere.
Results obtained with classical "Mean Shift" (MS) on the sphere are correct for an image without noise. This algorithm is sensitive to noise and product and over segmentation. We try to overcome the problem of this segmentation reconstructing the neighborhood map and using the region merging technique.
Using a noise of 2%, the MS realize an acceptable recognition, but the result present graze in the sphere and the obtained area is greater than the original model. The combined kppv-mean (nearest neighbor-mean) does not present marks in the sphere

but the area remains greater. This problem reacts on the computed ratio.
Globally, T-CAAR is less sensitive to noise; it leads to a good recognition even if there is more, noise in the image. The result is better when we take in account the knowledge of the background and working in HSL space. We decrease the light contribution by 0.5 in T-CAAR HSL. We can see that the result depend on the color space.
Varying the diffusion intensity, on the 3D object, we can observe equivalent result to the presented result for T-CAAR.
The processes of merging region obtained from the Watershed and the MeanShift result, allows overcoming over-segmentation.
Some cell parts can disappear because of the end merging process because these part are quite similar to the background.
Figure 8 shows the result obtained with the flooding LAMDA methods (X-CAAR) on cells : in input , 7 centers of clusters linked to 7 process in memory of the computer. This process can identify 7 clusters. The edges of the cells, due to the diffusion problem remains. They don’t belong to any of the pre-establish concept.

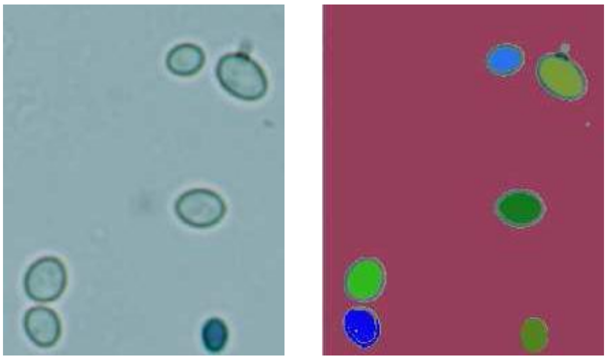


Figure 8: Results obtained with “Flooding-CAAR” on cell images.

The remain method for computing cell viability is T-CAAR, because it gives the better result for our images.
Table 1 shows the result obtained with T-CAAR using R, G, B attribute and spatial position by the structure. The images associated to the table 1 are showed in figure 6. Remember that the images have been obtained using a video camera coupled to a microscope.
The cell sample is the result of a biochemical reaction between a colorant and the cells. If we have a weak metabolic activity, coloring agent get into the cells. This action changes the original color of the dead cells to a dark blue color. The dead cells membrane cannot protect it and let the blue solution get in. Alive cells stay as they are. The resulting clustering allows to operate the diagnosis of the microbial cells in the fermenter.
Among the aggregation function used we have Gauss and Mahalanobis.
Coarsely, a classical function gives less precise results.

6. FUTUREWORK AND IMPROVEMENT

We present works on viability of cells and we succeed. We should now search for a method and a function of adequation which allows us to put in evidence scars on cells to study splitting. This task is different because scar is small and can be confused with noise.

7. CONCLUSION

We present in this paper a clustering algorithm based on fuzzy region merging. This new method has been compared with traditional algorithm: Mean-Shift, Fuzzy C-Mean, and Watershed.
The ability to work on the low level image definition has been proved. For example, biological systems analysis by microscope.
These good results are due to the update of the memberships function when an instance is associate to a class and the possibility to follow a particular order in pixels merging by neighborhood.

REFERENCES

Aguilar-Martin J., Jarachi F., Chan M., 1990, “Partitioned identification techniques from poisson observations : application to cerebral blood flow estimation,” Report LAAS No 89244, 11th IFAC World Congress, Tallinn (URSS), pp. 24-27, 13-17.

Atine J-C, 2004, “People action recognition in images sequences using a 3D articulated model”, dans proc. de International Conference on Image Analysis and Recognition.

Bicego M., Dalfini S.,Murino V., 2003, “Extraction of geographical entities from aerial images,” Proc. of IEEE-ISPRSWorkshop on Remote Sensing and Data fusion over Urban Areas URBAN ’03, Berlin, Germany, 125-128.
Comaniciu D., Meer P., 2002, “Mean Shift : A Robust Approachtoward Feature Space Analysis,” IEEE-PAMI, vol. 24, 603619.

Comaniciu D., Meer P., 1999, “Mean Shift Analysis and Applications,” Proc. of ICCV, pp. 1197-1203.

Eyrolles Ed.,“Apprentissage artificiel, concepts et algorithmes” Ed. , pp. 334-362.

Fiorio C., Gusted J., 1994, “Two linear time Union-Find strategies for image processing.,” LIRMM.

Janikow C. Z., 1996, “Exemplar Learning in Fuzzy Decision Trees,” Department of Mathematics and Computer Science, University of Missouri - St. Louis.

Li M., Ishwar K. Sethi, Li D. and Dimitrova N., 2003, “Region Growing Using Online Learning.,” CISST pp. 73-76.

Marsala C., 1998, “Construction d’arbres de d´ecision flous : le système Slammbô,” Rencontre Francophone sur la logique floue et ses applications.

Olaru C., Wehenkel L., 2003,“A complete fuzzy decision tree technique,” Fuzzy Set ans Systems 138, pp. 221-254.

Wang H. and Suter D., 2003, “Color Image Segmentation Using Global Information and Local Homogeneity,” Proc. VIIth Digital Image Computing: Techniques and Applications, Sydney, 10-12.

Waissman-Vilanova J., 2000, “Construction d’un modele comportemental pour la supervision de procedes : application `a une station de traitement des eaux”, Rapport LAAS no 00601.