



HAL
open science

Fuzzy Integral for Moving Object Detection

Fida El Baf, Thierry Bouwmans, Bertrand Vachon

► **To cite this version:**

Fida El Baf, Thierry Bouwmans, Bertrand Vachon. Fuzzy Integral for Moving Object Detection. FUZZ-IEEE 2008, Jun 2008, Hong-Kong, Hong Kong SAR China. pp.1729-1736, 10.1109/FUZZY.2008.4630604 . hal-00333086

HAL Id: hal-00333086

<https://hal.science/hal-00333086>

Submitted on 22 Oct 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fuzzy Integral for Moving Object Detection

Fida El Baf, Thierry Bouwmans, Bertrand Vachon

Abstract—Detection of moving objects is the first step in many applications using video sequences like video-surveillance, optical motion capture and multimedia application. The process mainly used is the background subtraction which one key step is the foreground detection. The goal is to classify pixels of the current image as foreground or background. Some critical situations as shadows, illumination variations can occur in the scene and generate a false classification of image pixels. To deal with the uncertainty in the classification issue, we propose to use the Choquet integral as aggregation operator. Experiments on different data sets in video surveillance have shown a robustness of the proposed method against some critical situations when fusing color and texture features. Different color spaces have been tested to improve the insensitivity of the detection to the illumination changes. Then, the algorithm has been compared with another fuzzy approach based on the Sugeno integral and has proved its robustness.

I. INTRODUCTION

ANALYSIS and understanding of video sequences is an active research field. Many applications in this research area (video surveillance [3], optical motion capture [4], multimedia application [2], video object segmentation [5], video coding [6]) need in the first step to detect the moving objects in the scene. So, the basic operation needed is the separation of the moving objects called foreground from the static information called the background. The process mainly used is the background subtraction. In the literature, many background subtraction methods can be found to be robust to the critical situations met in video sequence. These different methods are classified following the model used: Basic Background Modeling [8][9][10], Statistical Background Modeling [11][13][16] and Background Estimation [19][20][21]. In these different approaches, the features commonly used to handle critical situations are color, edge, stereo, motion and texture. Often, these features are used separately and the most used is the color one. The combination of several measuring features can strengthen the pixel's classification as background or foreground. In a general way, the Choquet and Sugeno integrals have been successfully applied widely in classification problems [23], in decision making [24] and also in data modelling [25] to aggregate different criteria. In the context of moving objects detection, these integrals seem to be good model candidates for fusing different measures from different features. Each integral has its particularity. The Choquet integral requires to interpret the scale as a continuum and the Sugeno integral allows to work with an ordinal scale. Recently, Zhang and Xu [1] have used texture feature and color features obtained from Ohta color space

to compute similarity measures between current and background pixels. Then, the measures are aggregated by applying the Sugeno integral. The assumption made by the authors reflects that the scale is ordinal. The moving objects are detected by thresholding the results of the Sugeno integral. In this work, the scheme used is based on Xu's algorithm. In the foreground detection, the values to be merged are the ratios of background pixel's features between a current image and the background image. The difference between these continuous values is real. In this context, the Choquet integral seems to be more suitable than Sugeno integral. So we propose to use the Choquet integral to aggregate color and texture features instead of the Sugeno integral. Then, the algorithm was improved by testing different color spaces which are more robust to shadows and illumination changes due to their geometrical characteristics. The rest of this paper is organized as follows: In Section 2, a brief review on background subtraction methods is given. Section 3 presents a brief overview of the proposed approach. Then, the features used for the foreground detection are described in Section 4. Fundamentals of fuzzy integrals are reminded in Section 5. After, we present in the Section 6 the application of the fuzzy integral for the foreground detection. Finally, a comparison of our algorithm with the method proposed by Zhang and Xu [1] is presented in Section 7, using video datasets from multimedia and video surveillance applications.

II. BACKGROUND SUBTRACTION: A BRIEF REVIEW

There are many background subtraction methods and the most recent surveys can be found in [3][7][35]. These different methods are commonly classified following the model used in the Background Modeling step. The simplest way to model the background is to acquire a background image which doesn't include any moving object. In some environments, the background isn't available and can always be changed under critical situations like illumination changes, objects being introduced or removed from the scene. So, the background representation model must be more robust and adaptive. The different background representation models can be classified in three classes:

- **Basic Background Modeling:** In this case, Background Representation is modeled using the average [8] or the median [9] or the histogram analysis over time [10]. Once the model is computed, the foreground detection is made as follows:

$$d(I_t(x, y) - B_{t-1}(x, y)) > T \quad (1)$$

Otherwise, pixels are classified as background. Where T is a constant threshold, $I_t(x, y)$ and $B_t(x, y)$ are

respectively the current and the background images at time t .

- **Statistical Background Modeling:** Background Representation is modeled using a single Gaussian [11][12] or a Mixture of Gaussians [13][14][15] or a Kernel Density Estimation [16][17][18]. Statistical variables are used in the foreground detection to classify the pixels as foreground or background.
- **Background Estimation:** Background representation is estimated using a filter. For the foreground detection, any pixel of the current image that deviates significantly from its predicted value is declared foreground. This filter may be a Wiener filter [19], a Kalman filter [20] or a Tchebychev filter [21].

All these methods present the same following steps and issues:

- **Background Modeling** which describes the kind of model used to represent the background.
- **Background Initialization** which regards the initialization of the model.
- **Background Maintenance** which relies to the mechanism used for adapting the model to the changes occurred in the scene over time.
- **Foreground Detection** which consists in the classification of the pixel as a background or as a foreground pixel.
- **Choice of the picture's element** which is used in the previous steps. This element may be a pixel [13], a block [36][37] or a cluster [38].
- **Choice of the features** which characterize the picture's element. In the literature, there are five features commonly used: color features, edge features, stereo features, motion features and texture features. In [33], these features are classified as spectral features (color features), spatial features (edge features, texture features) and temporal features (motion features). These features have different properties which allow to handle differently the critical situations (illumination changes, motion changes, structure background changes).

Developing a background subtraction method, researchers must design each step and choose the features in relation to the critical situations they want to handle. In this article, we focus on the foreground detection and the use of color and texture features to increase robustness to illumination changes and shadows. The idea is to classify pixels as background or foreground using the fusion of similarity measures obtained from the features. The fusion is made with a fuzzy integral. We describe below the proposed system.

III. SYSTEM OVERVIEW

The first step of many video analysis systems is the segmentation of foreground objects from the background. This task is a crucial prerequisite for the effectiveness of the global system. A background subtraction algorithm should be able to cope with a number of the critical situations. In particular, it should deal with the presence of noise, continuous and

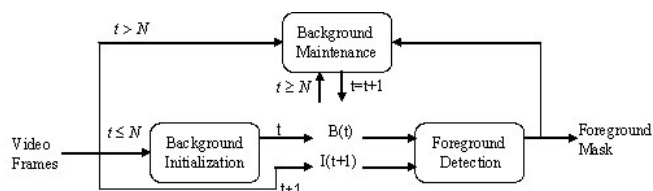


Fig. 1. System overview.

sudden light changes, temporal and permanent variation in background objects. These different critical situations can be handled in the different steps of the background subtraction: background representation, background initialization, background maintenance and foreground detection. The choice of the picture element size and of the features is essential too. In our approach, we have focussed on the foreground detection and the choice of the features but naturally the proposed pixel-wise foreground detection is a part of a complete background subtraction algorithm shown in Figure 1. The background initialization is made by using the average of the N first video frames where objects were present. An update rule of the background image is necessary to adapt well the system over time to some environmental changes. For this, a selective maintenance scheme is adopted as follows:

$$B_{t+1}(x, y) = (1 - \alpha) B_t(x, y) + \alpha I_{t+1}(x, y) \quad (2)$$

if (x, y) is background

$$B_{t+1}(x, y) = (1 - \beta) B_t(x, y) + \beta I_{t+1}(x, y) \quad (3)$$

if (x, y) is foreground

Here, the idea is to adapt very quickly a pixel classified as background and very slowly a pixel classified as foreground. Note that this background maintenance scheme allows the adaptation of the system to illumination changes but also the incorporation of motionless foreground objects. The learning

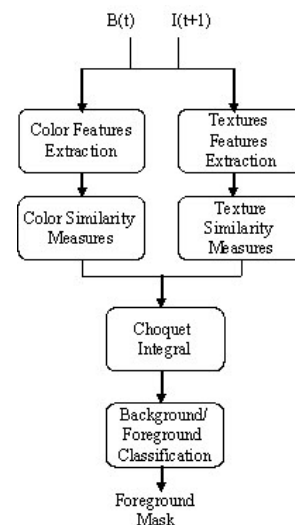


Fig. 2. Foreground detection process.

rate α determines the speed of the adaptation to illumination variations and the learning rate β controls the incorporation of motionless foreground objects. In Figure 2, the foreground detection process is presented in details. First, the color and the texture features are extracted from the background image B_t and the current image I_{t+1} . The similarity measures are computed for each feature which are then aggregated by the Choquet integral. The Background/Foreground classification is finally made by thresholding the Choquet integral's result. In the following sections, we describe the rationale for selecting and fusing the set of the adopted features.

IV. COLOR AND TEXTURE FEATURES

As seen before, the choice of the feature is essential. Intensity or color features are the main feature used because colours are often very discriminative features of objects, but they have several limitations in presence of some critical situations: illumination changes, camouflage and shadows. To solve these problems, some authors proposed to use other features like edge [28], texture [27] and stereo features [29], in addition to the color features. The stereo features deal with the camouflage but two cameras are needed. The edge handle the local illumination changes and the ghost leaved when waking foreground objects begin to move. The texture features are appropriate to illumination changes and to shadows, which are a main challenge in our work. In this context, we choose to add, to the color features, the Local Binary Pattern for texture proposed by [27]. In the following we discuss these two features.

A. Color Features

The selection of the color space, as color features, is one of the key factors for efficient color information extraction. In foreground detection, the most commonly used is the RGB space, because it is the one directly available from the sensor or the camera. The RGB color space has an important drawback ; their three components are dependent which increase its sensitivity to illumination changes. For example, if a background point is covered by the shadow, the three components values at this point could be affected because the brightness and the chromaticity information are not separated. A number of color space comparisons are presented in the literature [30][31][32]. After experimentally observing the effect of different color spaces on the segmentation result, the YCrCb was selected as the most appropriate color space. But first, let us define the different color spaces (the Ohta, HSV and YCrCb) tested which separate the luminance and the chrominance channels.

The axes of the Ohta space are the three largest eigenvectors of RGB space, found from the principal components analysis of a large selection of natural images. This color space is a linear transformation of RGB. The Equation (5) shows the relationship between RGB to the Ohta space:

$$\begin{aligned} I_1 &= (R + G + B) / 3 \\ I_2 &= (R - B) / 2 \text{ (or } (B - R) / 2) \\ I_3 &= (2G - R - B) / 4 \end{aligned} \quad (4)$$

HSV and YCrCb are closer to human interpretation of colours in the sense that brightness, or intensity, is separated from the base color. YCrCb uses cartesian coordinates to describe the base color while HSV uses polar coordinates. For HSV, the color information improves the discrimination between shadow and object, classifying as shadows those pixels having the approximately the same hue and saturation values compared to the background, but lower luminosity. The equations (6-7) below show the relationships between RGB and HSV, then YCrCb color spaces:

$$\begin{aligned} H &= 60(G - B) / \Delta \text{ if } \max(R, G, B) = R \\ H &= 60(B - R) / (\Delta + 120) \text{ if } \max(R, G, B) = G \\ H &= 60(R - G) / (\Delta + 240) \text{ if } \max(R, G, B) = B \\ S &= \Delta / \max(R, G, B) \\ V &= \max(R, G, B) \end{aligned} \quad (5)$$

where $\Delta = \max(R, G, B) - \min(R, G, B)$.

$$\begin{aligned} Y &= 0.25R + 0.504G + 0.098B + 16 \\ Cr &= 0.439R - 0.368G - 0.071B + 128 \\ Cb &= -0.148R - 0.291G + 0.439B + 128 \end{aligned} \quad (6)$$

For each color space, two components are chosen according to the relevant information which they contain so as to have the least sensitivity to illumination changes.

B. Texture Feature

The texture feature used is the Local Binary Pattern (LBP) which is developed by Heikkila and Pietikinen [27]. The LBP is invariant to monotonic changes in grey scale, which makes it robust against illumination changes. The operator labels the pixels of an image block by thresholding the neighbourhood of each pixel with the centre value and considering the result as a binary number:

$$LBP(x, y) = \sum_{i=0}^{N-1} s(g_i - g) 2^i$$

where g corresponds to the grey value of the center pixel (x, y) and g_i to the grey values of the N neighbourhood pixels. The function s is defined as follows:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

The original LBP operator worked with the 3×3 neighbourhood of a pixel.

Many fusion techniques can be used to fuse the color and the texture features. For this operation, we have chosen a fuzzy approach.

V. FUZZY INTEGRALS

The mathematical operator used for aggregation are multiples. In literature [22], we find the basic ones like the average, the median, the minimum and the maximum, as well as some generalizations like the Ordered Weighted

Average (OWA) having the minimum and the maximum as particular cases and the k-order statistics. Then, the family of fuzzy integrals has presented through its discret version a generalization of OWA or the weighted average using the Choquet integral, as well as the minimum and the maximum using the Sugeno integral. The advantage of fuzzy integrals is that they take into account the importance of the coalition of any subset of criteria.

In this section, we summarized briefly necessary concepts around fuzzy integrals (Sugeno and Choquet).

Let μ be a fuzzy measure on a finite set X of criteria and $h : X \rightarrow [0, 1]$ be a fuzzy subset of X .

Definition 1: The Sugeno integral of h with respect to μ is defined by:

$$S_\mu = \text{Max} (\text{Min} (h(x_{\sigma(i)}), \mu(A_{\sigma(i)}))) \quad (7)$$

where σ is a permutation of the indices such that $h_{\sigma(1)} \leq \dots \leq h_{\sigma(n)}$ and $A_{\sigma(i)} = \{\sigma(1), \dots, \sigma(i)\}$

Definition 2: The Choquet integral of h with respect to μ is defined by:

$$C_\mu = \sum_{i=0}^n h(x_{\sigma(i)}) (\mu(A_{\sigma(i)}) - \mu(A_{\sigma(i+1)})) \quad (8)$$

with the same notations as above.

An interesting interpretation of the fuzzy integrals arises in the context of the source fusion. The measure μ can be viewed as the factor which describes the relevance of the sources of information where h denotes the values the criteria have reported. The fuzzy integrals then aggregates nonlinearly the outcomes of all criteria. The Choquet integral is adapted for cardinal aggregation while Sugeno integral is more suitable for ordinal aggregation. More details can be found in [23][24][25][26].

In fusion of different criteria or sources, the fuzzy measures take on an interesting interpretation. A pixel can be evaluated based on criteria or sources providing information about the state of the pixel whether pixel corresponds to background or foreground. The more criteria provide information about the pixel, the more relevant the decision of pixel's state. Let $X = \{x_1, x_2, x_3\}$, with each criterion, we associate a fuzzy measure, $\mu(x_1) = \mu(\{x_1\})$, $\mu(x_2) = \mu(\{x_2\})$ and $\mu(x_3) = \mu(\{x_3\})$ such that the higher the $\mu(x_i)$, the more important the corresponding criterion in the decision. To compute the fuzzy measure of the union of any two disjoint sets whose fuzzy measures are given, we use an operational version proposed by Sugeno which called λ -fuzzy measure. To avoid excessive notation, let denote this measure by μ_λ -fuzzy measure, where λ is a paramater of the fuzzy measure used to describe an interaction between the criteria that are combined. Its value can be determined through the boundary condition, i.e. $\mu(X) = \mu(\{x_1, x_2, x_3\}) = 1$. The fuzzy density values over a given set $K \subset X$ is computed as:

$$\mu_\lambda(K) = \frac{1}{\lambda} \left[\prod_{x_i \in K} (1 + \lambda \mu_\lambda(x_i)) - 1 \right] \quad (9)$$

In the following section, we describe the use of the Choquet integral in the context of foreground detection.

VI. FUZZY INTEGRAL FOR FOREGROUND DETECTION

Foreground detection is based on a comparison between current and background images. In general, a simple subtraction is made between these two images to detect regions corresponding to foreground. Another way to establish this comparison consists in defining a similarity measure between pixels in current and background images. In this case, pixels corresponding to background should be similar in the two images while pixels corresponding to foreground should not be similar. In general, the most used features are color but texture feature can be a further tool to gain more robustness against illumination changes. So, we propose, in the following subsections, to compute similarity for color and texture features. Once these measures are computed, they will be aggregated by a Choquet integral.

A. Color Similarity Measures

In the following, we describe the similarity measure in a general way, i.e the color features may be any color space with three components noted I_1 , I_2 and I_3 . Then, the color similarity measure $S_k^C(x, y)$ at the pixel (x, y) is computed as in [1]:

$$S_k^C(x, y) = \begin{cases} \frac{I_k^C(x, y)}{I_k^B(x, y)} & \text{if } I_k^C(x, y) < I_k^B(x, y) \\ 1 & \text{if } I_k^C(x, y) = I_k^B(x, y) \\ \frac{I_k^B(x, y)}{I_k^C(x, y)} & \text{if } I_k^C(x, y) > I_k^B(x, y) \end{cases}$$

where $k \in \{1, 2, 3\}$ is one of the three color features, B and C represent respectively the background and the current images at time t . B can be obtained using any of the background modelling method. Note that $S_k^C(x, y)$ is between 0 and 1. Furthermore, $S_k^C(x, y)$ is closed to one if $I_k^C(x, y)$ and $I_k^B(x, y)$ are very similar.

B. Texture Similarity Measure

The texture similarity measure $S^T(x, y)$ at the pixel (x, y) is computed as follows:

$$S^T(x, y) = \begin{cases} \frac{L^C(x, y)}{L^B(x, y)} & \text{if } L^C(x, y) < L^B(x, y) \\ 1 & \text{if } L^C(x, y) = L^B(x, y) \\ \frac{L^B(x, y)}{L^C(x, y)} & \text{if } L^C(x, y) > L^B(x, y) \end{cases}$$

where $L^B(x, y)$ and $L^C(x, y)$ are respectively the texture LBP of pixel (x, y) in the background and current images at time t . Note that $S^T(x, y)$ is between 0 and 1. Furthermore, $S^T(x, y)$ is close to one if $L^B(x, y)$ and $L^C(x, y)$ are very similar.

C. Aggregation of color and texture similarity measures by the Choquet Integral

As defined above, the computed measures are obtained by dividing the intensity values in background and current images with endpoints denoted by 0 and 1. Where 0 means that the pixels at the same location in background and current

images respectively are not similar and 1 means that these pixels are similar i.e. pixel corresponding to background. In such a case, the scale is continuum and is constructed as a cardinal one where the distances or the differences between values can be defined. For example the distance between 0.1 and 0.2 is the same than the distance between 0.8 and 0.9, because numbers have a real meaning. While in the case of an ordinal scale, the numbers correspond to modalities when an order relation on the scale should be defined. A typical example of this former when we define a scale [a, b, c, d, e] to evaluate the level of some students, where "a" corresponds to "excellent" and "e" to "very bad". So that, the difference between "b" (very good) and "c" (good) is not necessary the same as the difference between "c" (good) and "d" (bad). Hence, operations other than comparison on a cardinal scale can be allowed like standard arithmetic operations, typically addition and multiplication. In this sense, the Choquet integral is considered as more suitable than the Sugeno integral because of its ability to aggregate well features on a cardinal scale and to use such arithmetic operations. So, for each pixel, color and texture similarity measures are computed as explained in section 4 from the background and the current frame. We define the set of criteria $X = \{x_1, x_2, x_3\}$ with, $(x_1, x_2) =$ two components color features of the chosen color space (i.e. Ohta, HSV, YCrCb etc) and $x_3 =$ texture feature obtained by the LBP.

For each x_i , let $\mu(x_i)$ be the degree of importance of the feature x_i in the decision whether pixel corresponds to background or foreground. The fuzzy functions $h(x_i)$ are defined in $[0, 1]$ so that, $h(x_1) = S_1^C(x, y)$, $h(x_2) = S_2^C(x, y)$ and $h(x_3) = S^T(x, y)$. To compute the value of Choquet integral for each pixel, we need firstly to rearrange the features x_i in the set X with respect to the order: $h(x_1) \geq h(x_2) \geq h(x_3)$.

The pixel at position (x, y) is considered as foreground if its Choquet integral value is less than a certain constant threshold Th :

if $C_\mu(x, y) < Th$ then (x, y) is foreground.

which means that pixels at the same position in the background and the current images are not similar. Th is a constant value depending on each video data set.

VII. EXPERIMENTAL RESULTS

We have applied our algorithm to different datasets: the first one is our Aqu@theque dataset used in a multimedia application [2], where the output images are 384×288 pixels. The second one is the VS-PETS 2003¹ used in video sport application with image's size is 720×576 pixels. The third and the fourth ones are PETS 2000² and the PETS 2006³ dataset applied in video surveillance. The output images of these two last datasets are respectively 768×576 and

720×576 pixels. For each datasets, we provide a comparison with another approach based on Sugeno integral [1]. The results are obtained without post processing and the threshold for each algorithm is optimized to give the best results.

A. Aqu@theque dataset

This dataset contains several video sequences presenting fishes in tank. The goal of the application Aqu@theque [2] is to detect fishes and identify them. In these aquatic video sequences, there are many critical situations. For example, there are illumination changes owed to the ambient light, the spotlights which light the tank from the inside and from the outside, the movement of the water due to fish and the continuous renewal of the water. These illumination changes can be local or global following their origin. Furthermore, the constitution of the aquarium (rocks, algae) and the texture of fishes amplify the consequences of the brilliant variation. Figure 3 shows the experiments made on one sequence. In table I, we show the fuzzy density values that we have tested. The best results are obtained with $\{0.53, 0.034, 0.13\}$. It is noticed that the results obtained using the proposed method are better than using the method proposed by [1] with the same color space, i.e. Ohta. The results obtained with the Choquet integral using other color spaces, i.e. the HSV and YCrCb confirmed that optimum results are obtained using Choquet integral with the YCrCb color features.

The quantitative evaluation has been done firstly using the similarity measure derived by Li [33]. Let A be a detected region and B be the corresponding ground truth, the similarity between A and B can be defined as:

$$S(A, B) = \frac{A \cap B}{A \cup B} \quad (10)$$

If A and B are the same, $S(A, B)$ approaches 1, otherwise 0 i.e. A and B have the least similarity. The ground truth are marked manually. Table II shows the similarity value

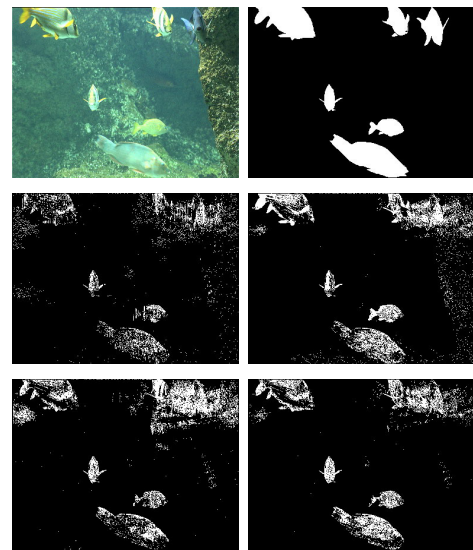


Fig. 3. First row: The current image, the ground truth. Second row: Sugeno-Ohta, Choquet-Ohta. Third row: Choquet-HSV and Choquet-YCrCb.

¹<http://ftp.pets.rdg.ac.uk/VS-PETS/TESTING/CAMERA3>

²<http://ftp.pets.rdg.ac.uk/PETS2000>

³<http://www.cvg.rdg.ac.uk/PETS2006/data.html>

TABLE I
FUZZY MEASURE VALUES

$\{x_1\}$	$\{x_2\}$	$\{x_3\}$	$\{x_1, x_2\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\{X\}$
0.6	0.3	0.1	0.9	0.7	0.4	1
0.5	0.4	0.1	0.9	0.6	0.5	1
0.5	0.3	0.2	0.8	0.7	0.5	1
0.50	0.39	0.11	0.89	0.61	0.5	1
0.53	0.34	0.13	0.87	0.66	0.47	1

TABLE II
SIMILARITY MEASURE

Integral Color Space	Sugeno Ohta	Choquet Ohta	Choquet HSV	Choquet YCrCb
$S(A, B)$	0.27	0.44	0.34	0.46

obtained for the previous experiments. It is well identified that optimum results are obtained by the Choquet integral. Furthermore, the Ohta and the YCrCb spaces give almost similar results ($S_{Ohta} = 0.44, S_{YCrCb} = 0.46$), when the HSV space registers ($S_{HSV} = 0.34$). When observing the effect of YCrCb and Ohta spaces on the images, we have noticed that the YCrCb is slightly better than the Ohta space. To see the progression of the performance of each algorithm, we use the ROC curves [34]. For that, we compute the false positive rate (FPR) and the true positive rate (TPR) as follows:

$$FPR = \frac{FP}{FP + TN} \quad ; \quad TPR = \frac{TP}{TP + FN}$$

where TP is the total of true positives, TN the total of true negatives, FP the total of false positives and FN the total of false negatives. The FPR is the proportion of background pixels that were erroneously reported as being moving object pixels. And the TPR is the proportion of moving object pixels that were correctly classified among all positive samples. The Figure 4 represents the ROC curves for the Sugeno and the Choquet integrals with the Ohta color space. These curves confirm that the Choquet integral outperforms the Sugeno one using the Ohta space. Then, we have compared the previous results with other color spaces. The Figure 5 shows the ROC curves for the Choquet integral with the Ohta, HSV and YCrCb color spaces. Once again, The curves confirm the previous conclusion. Indeed, the Area Under Curve (AUC) are almost similar for YCrCb and Ohta spaces.

Thus, in the following, we present the results for other datasets using only the YCrCb space.

B. VS-PETS dataset

The dataset is formed by outdoor scenes (soccer video sequence). Figure 6 shows the results obtained with the method proposed by [1] and with the Choquet integral using the YCrCb color space. The silhouettes are better detected and the illumination variations on the white border are less detected using our method.

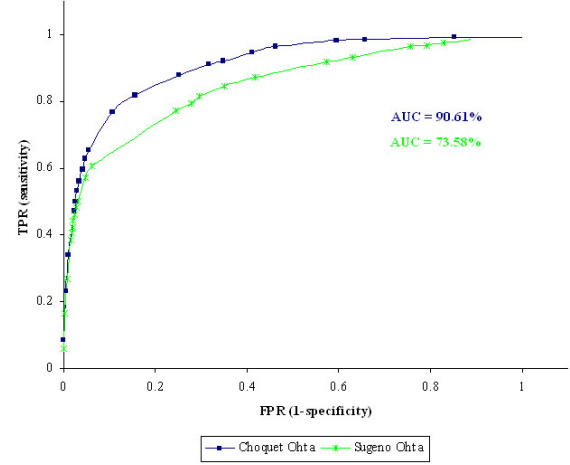


Fig. 4. ROC Curve : Comparison of the two detection algorithms using respectively the Sugeno and the Choquet integrals in Ohta color space.

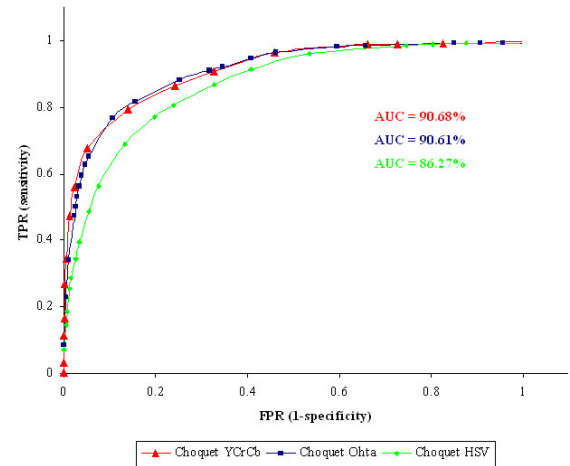


Fig. 5. ROC Curve : Evaluation of the effect of different color spaces to the detection algorithm using the Choquet integral.

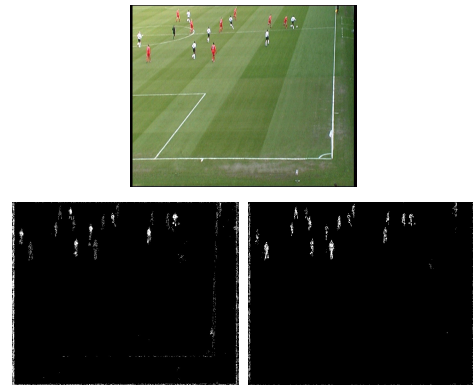


Fig. 6. First row: The current image. Second row: Sugeno-Ohta, Choquet-YCrCb.

C. PETS 2000 and 2006 dataset

The algorithm is tested also on PETS 2000 and 2006 benchmark data indoor and outdoor sequences in video surveillance context. The goal is to detect moving persons and/or vehicles. Once again the use of Choquet integral with YCrCb color space shows a robustness to illumination changes and shadows, as we can see in Figure 7-8.

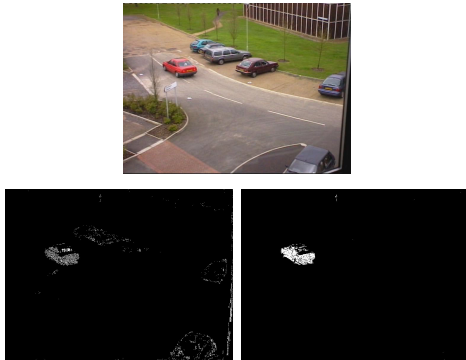


Fig. 7. First row: The current image. Second row: Sugeno-Ohta, Choquet-YCrCb.



Fig. 8. First row: The current image. Second row: Sugeno-Ohta, Choquet-YCrCb.

VIII. CONCLUSION

In this paper, we have presented a foreground detection method using the Choquet integral for fusing color and textures features. Experiments in multimedia and video surveillance datasets show that the Choquet integral gives better results than the use of the Sugeno integral proposed by Zhang and Xu [1]. YCrCb and Ohta spaces provide similar results. Furthermore YCrCb is slightly better than the Ohta space. The proposed algorithm is more robust to shadows and illumination changes than the method proposed by Xu. Further research consists in fusing other features like edge or motion features and learning the fuzzy densities.

ACKNOWLEDGEMENTS

We thank Dan Jurca and Hoel Le Capitaine for their contributions and the professor Radu Vasu who is the responsible for research and international programmes at the University of Timisoara.

REFERENCES

- [1] H. Zhang, D. Xu, "Fusing Color and Texture Features for Background Model," *Third International Conference on Fuzzy Systems and Knowledge Discovery*, vol. 4223, no. 7, pp. 887-893, Sept. 2006.
- [2] F.El Baf, T. Bouwmans, "Comparison of Background Subtraction Methods for a Multimedia Learning Space," *International Conference on Signal Processing and Multimedia*, Jul. 2007.
- [3] S. Cheung, C. Kamath, "Robust Background Subtraction With Foreground Validation for Urban Traffic Video," *Journal of Applied Signal Processing*, Sept. 2005.
- [4] J. Carranza, M. Magnor, "Free-Viewpoint Video of Human Actors," *ACM Transactions on Graphics*, vol. 22, no. 3, pp. 569-577, 2003.
- [5] A. Colombari, A. Fusiello, V. Murino, "Video Objects Segmentation by Robust Background Modeling," *ICIAP 2007*, pages 155-164, 2007.
- [6] J. Zhang, "Region of Interest Segmentation and Video Coding in Dynamic Video Backgrounds," *PhD Thesis in Electrical Engineering, University of Massachusetts Dartmouth*, May 2007.
- [7] S. Elhabian, K. El-Sayed, S. Ahmed, "Moving Object Detection in Spatial Domain using Background Removal Techniques - State-of-Art," *Recent Patents on Computer Science*, vol. 1, no. 1, pp. 32-54, Jan 2008.
- [8] B. Lee, M. Hedley "Background Estimation for Video Surveillance," *Image and Vision Computing New Zealand*, pp. 315-320, 2002.
- [9] N. McFarlane, C. Schofield, "Segmentation and tracking of piglets in images," *British Machine Vision and Applications*, pp. 187-193, 1995.
- [10] J. Zheng, Y. Wang, "Extracting Roadway Background Image: A mode based approach," *Transportation Research Board*, 2006.
- [11] C. Wren, A. Azarbayejani, "Pfunder : Real-Time Tracking of the Human Body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780 -785, Jul. 1997.
- [12] M. Zhao, N. Li, C. Chen, "Robust automatic video object segmentation technique," *IEEE International Conference on Image Processing, ICIP 2002*, Sept. 2002.
- [13] C. Stauffer, "Adaptive background mixture models for real-time tracking," *Proceedings IEEE Conference on Computer Vision and Pattern Recognition*, pp. 246-252, 1999.
- [14] J. Zhang, C. Chen, "Moving Objects Detection and Segmentation in Dynamic Video Backgrounds," *Conference on Technologies for Homeland Security*, pp. 64-69, Woburn, USA, May 2007.
- [15] R. Tan, H. Huo, J. Qian, T. Fang, "Traffic Video Segmentation using Adaptive-K Gaussian Mixture Model," *The International Workshop on Intelligent Computing, IWICPAS 2006*, China, August 2006.
- [16] A. Elgammal, L. Davis, "Non-parametric Model for Background Subtraction," *6th European Conference on Computer Vision*, Jun. 2000.
- [17] C. Ianasi, V. Gui, C. Toma, D. Pescaru, "A Fast Algorithm for Background Tracking in Video Surveillance, Using Nonparametric Kernel Density Estimation," *Facta Universitatis, Series : Electronics and Energetics*, vol. 18, no. 1, pp. 127-144, Apr. 2005.
- [18] A. Tavakkoli, M. Nicolescu, G. Bebis, "Robust Recursive Learning for Foreground Region Detection in Videos with Quasi-Stationary Backgrounds," *Proceedings of the International Conference on Pattern Recognition, ICPR 2006*, vol. 1, pp. 315-318, Hong Kong, August 2006.
- [19] K. Toyama, J. Krumm, "Wallflower: Principles and Practice of Background Maintenance," *International Conference on Computer Vision*, Sept. 1999.
- [20] S. Messelodi, C. Modena, "A Kalman filter based background updating algorithm robust to sharp illumination changes," *International Conference on Image Analysis and Processing*, Sept. 2005.
- [21] R. Chang, T. Ghandi, M. Trivedi, "Vision modules for a multi sensory bridge monitoring approach," *IEEE Conference on Intelligent Transportation Systems*, Oct 2004.
- [22] M. Detryniecki, "Fundamentals On Aggregation Operators," *AGOP*, Berkeley, 2001.
- [23] H. Tahani, J. Keller, "Information fusion in computer vision using the fuzzy integral," *TTransaction on Systems, Man, and Cybernetics*, vol. 20, no. 3, pp. 733-741, 1990.
- [24] Y. Narukawa, T. Murofushi, "Decision Modelling Using the Choquet Integral," *Modeling Decisions for Artificial Intelligence*, vol. 3131, pp. 183-193, 2004.
- [25] M. Sugeno, S. Kwon, "A new approach to time series modeling with fuzzy measures and the Choquet integral," *4th IEEE International Conference on Fuzzy Systems*, pp. 799-804, Mar. 1995.
- [26] M. Grabish, "Fuzy integral in multicriteria decision making," *Fuzzy Sets and Systems*, pp. 69-279, 1995.

- [27] M. Heikkilä, M. Pietikinen, "A texture-based method for modeling the background and detecting moving objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 657-662, 2006.
- [28] O. Javed, K. Shafique, M. Shah, "A Hierarchical Approach to Robust Background Subtraction using Color and Gradient Information," *IEEE Workshop on Motion and Video Computing, WMVC 2002*, Orlando, Dec. 2002.
- [29] M. Harville, G. Gordon, J. Woodfill, "Adaptive background subtraction using color and depth," *In Proceedings of the IEEE International Conference on Image Processing*, Thessaloniki, Greece, Oct. 2001.
- [30] F. Kristensen, P. Nilsson, V. wall, "Background Segmentation Beyond RGB," *ACCV 2006*, pp. 602-612, Hyderabad, Indian, 2006.
- [31] H. Ribeiro, A. Gonzaga, "Hand Image Segmentation in Video Sequence by GMM: a comparative analysis," *XIX Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI 2006*, pp. 357-364, Manaus, Brazil, 2006.
- [32] S. Kanprachar, S. Tangkawanit, "Performance of RGB and HSV color systems in object detection applications under different illumination intensities," *International Multi Conference of Engineers and Computer Scientists*, vol. 2, pp. 1943-1948, Kowloon, China, March 2007.
- [33] L. Li, W. Huang, "Statistical Modeling of Complex Background for Foreground Object Detection," *IEEE transaction image processing*, vol. 13, no. 11 pp. 81459-1472, Nov. 2004.
- [34] J. Davis, M. Goadrich, "The Relationship between Precision-Recall and ROC Curves," *23rd International Conference on Machine Learning, ICML 2006*, vol. 148, pp: 233-240, Pittsburgh, Pennsylvania, 2006.
- [35] M. Piccardi, "Background subtraction techniques: a review," *Proceedings of the International Conference on Systems, Man and Cybernetics*, pp. 3199-3104, Oct. 2004.
- [36] X. Fang, W. Xiong, B. Hu, L. Wang, "A Moving Object Detection Algorithm Based on Color Information," *International Symposium on Instrumentation Science and Technology*, vol. 48, pp. 384-387, 2006.
- [37] D. Pokrajac, L. Latecki, "Spatiotemporal Blocks-Based Moving Objects Identification and Tracking," *IEEE Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, pp. 70-77, Oct. 2003.
- [38] H. Bhaskar, L. Mihaylova, S. Maskell, "Automatic Target Detection Based on Background Modeling Using Adaptive Cluster Density Estimation," *LNCS from the 3rd German Workshop on Sensor Data Fusion: Trends, Solutions, Applications*, Germany, Sep. 2007.