Tracking Objects in Color Image Sequences: Application to Football Images

Mourad Moussa, Ali Douik and Hassani Messaoud

Abstract—In this paper, we present a comparative study between two computer vision systems for objects recognition and tracking, these algorithms describe two different approach based on regions constituted by a set of pixels which parameterized objects in shot sequences. For the image segmentation and objects detection, the FCM technique is used, the overlapping between cluster's distribution is minimized by the use of suitable color space (other that the RGB one). The first technique takes into account a priori probabilities governing the computation of various clusters to track objects. A Parzen kernel method is described and allows identifying the players in each frame, we also show the importance of standard deviation value research of the Gaussian probability density function. Region matching is carried out by an algorithm that operates on the Mahalanobis distance between region descriptors in two subsequent frames and uses singular value decomposition to compute a set of correspondences satisfying both the principle of proximity and the principle of exclusion.

Keywords— Image segmentation, Objects tracking, Parzen window, Singular value decomposition, Target Recognition

I. INTRODUCTION

SOCCER is a very popular sport in the world and there is a great interest in better understanding its important fundaments if one wants to increase the performance of a team during a game, and better adapt the planning of the trainings. The movement of the players on the field, as a function of time, is useful information that can contribute for improving the performance of the players at different positions [1]. For tactical variations that a team can assume during a game, for example, the measured values may be associated to physiological variables as well as to technical and tactical information. Many studies were made about this issue, the first one concerned with the player's movement during the game by Reilly and Thomas [2] which employed audio recorders to register the estimated location of the players. Withers et al. [3] used a camera to analyze the movement of a unique soccer player. Mayhew and Wenger [4] also used a camera to track two players, each one filmed alternately for 7 min. They computed the time spent for each activity of these tracked players, such as walking, running, jogging, staying, as well as the frequency of the corresponding activities.

M. Moussa (corresponding author) is with the Department of Electrical Engineering, ATSI unit, National Engineers School of Monastir, Tunisia, e-mail: (Mourad.enim@yahoo.fr).

A. Douik is with ATSI unit, National Engineers School of Monastir, e-mail: (Ali.Douik@enim.rnu.tn).

H. Messaoud is the director of ATSI unit, National Engineers School of Monastir, e-mail: (Hassani.Messaoud@enim.rnu.tn).

Henning and Briehle [5] analyze the soccer player's movement by using a global positioning system (GPS). This kind of system locates the global position of the object by satellites which receive the signal emitted by a transmitter located on the earth surface. This methodology demands a device of 250 g to be carried by the tracked object from which the data are collected at a frequency of 1 Hz. A number of researchers have attempted to deal with object occlusion (and the resultant tracking problems) by attempting to track through occlusion. This can involve reasoning about object ordering along the camera optical axis, either using ground-plane information or simply reasoning about relative spatial ordering [6]. 3D or 2D object models may be used with a known camera-to ground-plane transformation to identify occlusion for a given set of object configurations. This can then be used to exclude subsets of image information from the model fitting process. However, complete occlusion leads to zero information from the image, and weak constraints on object position/configuration. [7] Takes a conservative approach of not tracking in uncertain situations, such as object occlusion. The ends of broken tracks are then joined based on spatio-temporal similarity measure. Another approach has been to use multiple cameras in an attempt to circumvent the occlusion problem. In [8], multiple football players are tracked using eight cameras. Each view is tracked separately and overlapping/occluding objects are associated with the same blob (as in our system). Each blob from each camera is projected to the ground-plane (using a camera calibration), and associated with one, or more, players using a 'closed-world' assumption, and a stochastic constraint optimisation procedure. Dynamic models such as the Kalman filter are often used to model the position of occluded objects, under the assumption of known dynamics (e.g. linear motion), when no visual information is available. In [9], the authors use both motion models and 2Ddynamic appearance models to track through occlusion. This work is interesting as it is a composite of a moving region 'blob' segmentation/extraction algorithm (that has no explicit notion of occlusion), and a model based tracker, that can track through partial and (short-term) complete occlusion using motion and appearance models.The paper is structured as follows: We start with an overview of work applying probabilistic approach for objects detection and tracking in computer vision. Section 3 describes algorithm allows to detect interest region from each frame. Section 4 summarizes the probability density estimation using Parzen windows to model two samples relative to each cluster. Section 5 presents conclusion and a perspective on future work.

II. RELATED WORKS

Recently, the usefulness of statistical and probabilistic concepts in computer vision has been recognized, with application in tasks like image segmentation and restoration, object recognition and tracking [10-30-31], and feature extraction. Several techniques have been proposed for motion segmentation, as image differencing, temporal analysis of grey-levels based on probabilistic models, robust motion estimation, or misalignment analysis based on the normal flow. In Ref. [11], body parts are segmented and tracked, using a body model to help resolving ambiguities and tracking failures. In the area of image registration [12], the alignment of two images that do not necessarily come from the same modality is done by maximizing the mutual information. This theoretically complicated and practically expensive step is elegantly performed with the stochastic optimization algorithm EMMA. The underlying pdf's are represented by Parzen window densities. The authors also show applications in the area of object tracking and photometric stereo. These techniques have parallels in principal component analysis and function learning. In [13], an information theoretic approach for feature extraction motivated by Fano's inequality for the error rate in classification is presented. This work also represents the continuous pdf's by Parzen window densities. It can be seen as a practical realization of a feature selection scheme based on the mutual information; in fact, it can also be found in textbooks on pattern recognition. Belief Filtering is used in tracking and identification, in this study the Peak Location Probability Function (PLPF) estimation has a role to determine the probability that a peak will be observed in a specific range bin location given that the observation was from some individual target class Ti. This probability is estimated from the peak locations of the training ensemble for each target class. A Parzen estimator with a normal kernel function along the range dimension is employed to estimate the PLPF. With this function, class probabilities are associated with peak locations. One approach is to use background subtraction or optical flow for detection of moving objects [14]. Detected objects are then tracked, and object classification is performed on the tracked object regions using simple descriptors of shape or motion.

The other approach, direct image-based detection of object classes, does not rely on object motion. Instead, each frame is scanned for regions having the characteristic appearance of objects of interest, such as vehicles [15] or pedestrians. These methods typically use combinations of simple image features such as edges, wavelets, or rectangular filters, and require training on large labelled datasets, especially for low-resolution far-field images. Methods based on background subtraction followed by object tracking do not suffer from the problem of false positives or scale selection, and have been demonstrated to run in real-time [16]. These methods may also be able to track robustly through temporary occlusion and clutter. While these methods are only applicable to static cameras, long-term scene analysis (i.e. more than a few hours) requires precisely such a setup. Statistical priors on object characteristics in a scene can greatly help classification. Contextual information (such as likely scales or positions of objects) may be manually specified by an operator for a given scene or learnt from examples [17] and used to prime object detection. However, to the best of our knowledge, no previous work has been done on learning local context features (such as position and direction of motion of objects) solely from long-term observation of moving object, to improve classification performance in a given scene.

III. OBJECT DETECTION

A. Introduction

Background subtraction is a very common method used for segmenting moving objects which consists of the difference between a set of images and its background model [29]. To consider problems in outdoor scenes such as changes of illumination, shadows, background objects, etc., these methods need to frequently update the background representation model. For this purpose, some statistical adaptive methods [18-32] have been discussed in the literature. These methods, which work well for relatively simple scenes with slow changes of illumination, update the background model for each frame by assuming its background as a Gaussian distribution. This approach can easily incorporate to the model objects that stop moving for a certain time. In some applications, however, it is desirable to keep these still or slow moving objects correctly tracked. In soccer games, for example, it is common to have players that stand still for many video frames.

The segmentation is usually accomplished in several stages: skin distribution modelling, which includes a training process, usually supervised, followed by several morphological operations for noise removal until no further segmentation is required to create an accurate representation model of skin colour.

State-of-the-art skin detection approaches do not usually include a justification of their colour space choice. This is probably because there is no colour space that can fully remove the changing lighting conditions, shadows and complex background containing surfaces and objects with skin-like colours, factors which can affect the accuracy of skin colour detection. However, a colour space more suited to discriminate skin from non skin colour values can make it easier to build an accurate classifier.

B. Fuzzy Segmentation

In this work, we use fuzzy algorithm to subtract the background pixels in soccer image sequences (Figure 1), the two main issues of the used technique are; (a) specify the choice of significant color system representation and prove that a system is not often suitable from an application to another, (b) overcome the overlapping between clusters by estimation of the significant color levels which allows to discriminate each clusters.
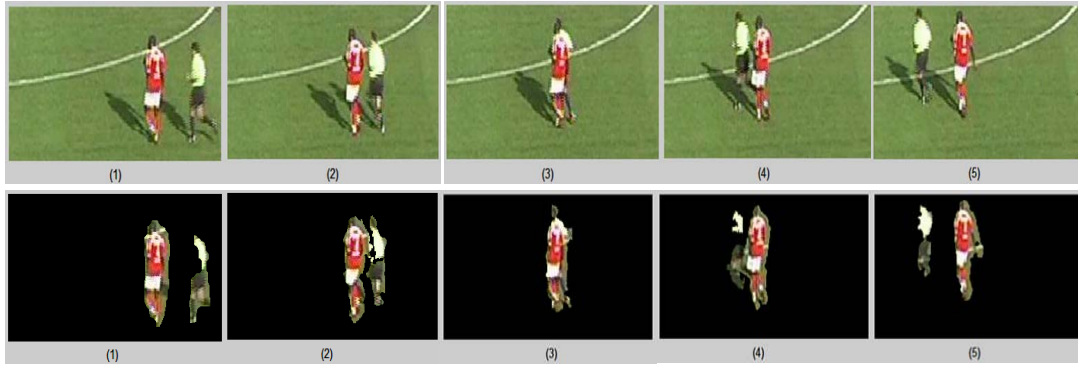
Fig. 1 Some selected frame describing the motin of two players in

The f-c-m technique is a classification algorithm founded on optimization of a quadratic criterion [27], where each class is represented by its gravity centre. It requires preliminary knowledge of class number and generates them by an iterative process minimizing an objective function J_m defined by (Eq.1), thus it allows obtaining fuzzy partition of image by giving to each pixel a membership degree to a given region.

$$J_m(U,C) = \sum_i \sum_k (U_{ik})^m \|x_i - c_k\|^2 \qquad (1)$$

Where m > 1 is a control parameter of the fuzzy degree, ck is the centre of class k, and U presents the matrix which regroups the membership degree values. Thanks to the play of fuzzy regions this algorithm is frequently used in many applications. It can be employed for pattern recognition or indexing images.

To execute this algorithm, we must follow the stages below:

1) Choose the number of class C: (for supervised algorithm the choice is a priori).

2) Initialize the partition matrix U as well as the centers ck with arbitrary way.

3) Evolve a matrix U and these centers according the following equations (Eq.2 and Eq. 3) this stage was considered as an update of the membership degrees and the centers.

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{dij}\right)\left(\frac{2}{m-1}\right)} \qquad (2)$$

$$c_k = \frac{\sum_{i}^{c} (u_{ik})^m x_i}{\sum_{i}^{c} (u_{ik})^m} \qquad (3)$$

4) Give a stopping condition allowing finishing execution algorithm. This test is defined by the difference between two consecutive values of J, and respecting a certain threshold *th*.

Classification by f–c–m algorithm present two problems:
- The number of clusters must be provided in advance (supervised Algorithms).

- Each region is characterized by a center, and the membership degrees are calculated with the Euclidean distance.

The achieved algorithm allows overcoming ambiguity and clusters overlapping. A fast and robust solution based into the merge of elementary hyperspheric areas is given by the f-c-m algorithm. The results of computer vision applications are not always satisfactory in RGB levels. For this reason, we will introduce other methods to determine, among a set of chromatic levels used in image processing contribute to more separate the objects from a given frames [20]. Specify crucial levels among these commonly used information treatment therefore eliminate useless or redundant leading to refine segmentation results. The metric distance is the discrimination parameter able to compare resemblance and difference between two vectors (respectively two matrix), in this study Mahalanobis distance (Md) and Hausdorff distance (Hd) are used. The color space evaluated by (Md) is composed by the following three components Cb (YCbCr), Cr (YCbCr), a (Lab) on the one hand, the color space evaluated by (Hd) is constituted by Pr (YPbPr), H (LCH), a (Lab) on the other hand. The segmentation results shown in Fig. 1 proves that the chrominance component is more significant than the brightness, indeed the conversion RGB image to hybrid color space, led to a relevant segmentation.

IV. OBJECT TRACKING BY PARZEN WINDOW

A. Introduction

As it is explained in the preview section the tracking technique is based on pixel classification. Players were characterized by color pixel, the two sets of pixel S_i^n and S_j^m relative to cluster1 and cluster2 are represented in 2 D

space constituted by R-G component (Red-Gray). Many studies incorporate color information and morphologic features to track objects in computer vision application, for ball tracking, HSV colour space is used to make color based probabilistic tracker [19], Hue and Saturation values are relatively stable indicators of skin colour regardless of the luminance range, and that the choice of alternative colour spaces such as normalised RGB does not make a significant difference, a simple Bayesian posterior map is used to identify the most probable locations of skin in each frame. In our case, grey and red levels are a significant color space allows representing the various clusters with a minimum overlapped degree.

B. Probability density estimation using Parzen Window

The parameters of the distribution functions need to be learned from training data. Probability density function of each distribution is estimated using a nonparametric method. The membership posterior probability of each player to suitable cluster is calculated from the given conditional probabilities and prior probabilities using equation (4), where $P(C_1)$ is the prior probability of the cluster1 and $P(C_2)$ is the prior probability of the cluster2, which were both set at 0.5.

$$p(C_1 / S_i^n) = \frac{p(S_i^n / C_1)P(C_1)}{p(S_i^n / C_1)P(C_1) + p(S_j^m / C_2)P(C_2)} \quad (4)$$

The distribution parameters, $p(S_i^n / C_1)$ and $p(S_j^m / C_1)$, are learned from the training data using a non-parametric probability estimation method. The color features of each class are then used to model the probability density functions for each distribution. In general, supervised learning is based on the assumption that the forms of the underlying density functions are known. However, since features that do not follow known distribution and parametric densities are multimodal densities rather than unimodale, the probability density functions of each class are estimated using a non-parametric method. The general formulation of the nonparametric density estimation function is a Parzen window. However, the Gaussian probability density function is a popular kernel for Parzen-window density estimation [21]. Using Equation (5), the Parzen-window estimate with the Gaussian kernel becomes

$$p(x) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{(h\sqrt{2\pi})^d} \exp\left(-\frac{\|x - x_n\|^2}{2h^2}\right) \quad (5)$$

Where h is the standard deviation of the Gaussian probability density function along each dimension d. If h is too large, the density estimate $p(x)$ is very smooth, while if h is too small, the estimate $p(x)$ is just a superposition of n sharp pulses centered at the training samples. In practice, h should be determined by an acceptable compromise, since the number of training samples is always limited. A smooth parameter h is typically chosen based on the number of available observations n.

In our case, given that the pixel's distribution for objects which will be tracked in the R-G space, the theoretical frontier of decision (presented by a green line) allows to discriminate the two sets of pixel's intensities defined by a linear function (y = ax + b). As for estimated frontier of decision (presented by Magenta curve), we are estimated Gaussian density using a Parzen windows having a 0.1 as width (Figure 2).
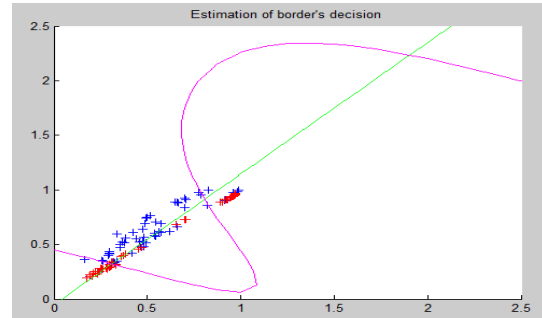

Fig. 2 Representation of theoretical and estimated decision's frontier

In this paper, the value of h was set at 0.12 based on experiments given by table 1.

TABLE I
INTEREST VALUES AMONG 50 OF STANDARD DEVIATION AND CORRESPONDING ERROR

h	0.09	0.1	0.11	0.12	0.13	0.14
Error%	33.66	31.16	30.83	30.33	31.33	33.16

Figure 3 shows the results for some examples of Parzen windows values to estimate probability density functions.
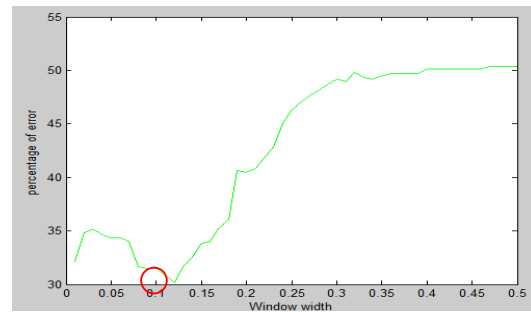

Fig. 3 Evaluation of error according Parzen windows values

Each value leads to evaluate the percentage of error between original and estimated memberships probability. Finally, according Table 1 and Figure 3 the optimal value of h is at 0.12.

V. OBJECT TRACKING BY REGION MATCHING

A. Review Stage

The tracking starts with an empty edge set and proceeds by adding edges. Ideally, an edge links two nodes from consecutive layers if they represent the same MO (or part of it) at different time instants. A *track* is a chain of at least two nodes belonging to consecutive frames. Every node in a track,

other than the first and the last, has degree two. Any node with a degree of two, one or zero is, respectively, referred as *internal*, *external*, or *isolated* node. An isolated node does not belong to any track. The *length* of a track is the number of nodes belonging to that track, the *area* of a track is the median area of the blobs corresponding to the nodes of that track. The goal is to find one track for each MO, thereby identifying blobs as objects [28].

B. Object Matching

In a first phase, tracks are initialized by matching nodes from one layer to the next. A dissimilarity (distance) measure between blobs is defined taking into account the appearance (shape and colour) of the blob and its position. In particular, each blob is described by a feature-vector **b** composed of solidity, eccentricity, orientation, area, dimensions of the bounding box, average colour, contrast (standard deviation of the colour) and position of the centroid in the mosaic reference frame [22].

The dissimilarity of blobs I_i^n and J_j^m is made by Mahalanobis distance between the respective feature vectors [23]:

$$d_{ij} = (b_i - b_j)^T (V_I + V_J)^{-1} (b_i - b_j) \qquad (6)$$

Where V_I and V_J are the covariance matrices of the feature vectors in images *I* and *J*, respectively. In practice, V is a diagonal matrix containing normalizing weights for each feature-vector element. Matching is carried out with a technique introduced by [24] and elaborated upon by Pilu (henceforth referred to as "SL matching"), where the singular value decomposition (SVD) of a suitable matrix is used for associating features of two images.

Let $\{I_i^n\}_{1...k}$ and $\{J_j^m\}_{1...l}$ be the two sets of blobs which are to be put in one-to-one correspondence. The first stage is to build a *proximity matrix* **P** of the two sets of features: $P_{ij} = e^{-d_{ij}/2}$ The next stage is to perform the SVD of **P:**

$$P = USV^T \qquad (7)$$

Where **U** and **V** are orthogonal and **S** is a non-negative $m \times n$ diagonal matrix. Finally, **S** is converted into a new $m \times n$ matrix **D** by replacing every diagonal element S_{ii} with 1, thus obtaining another matrix $Q = UDV^T$ of the same shape as the original proximity matrix and whose rows are mutually orthogonal. The element Q_{ij} indicates the extent of pairing between the blobs *Ii* and *Jj*. This matrix incorporates the principle of proximity (that favours a match with the closest feature) by construction of **P** and the principle of exclusion (that prohibits many-to-one correspondences) by virtue of its orthogonality.

If Q_{ij} is both the largest element in its row and in its column, then I_i^n and J_j^m are regarded as corresponding with each other, provided that their Mahalanobis distance is below a certain threshold. When Q_{ij} is the greatest element in row *i* but not the greatest in column *j*, "then we may regard I_i as competing unsuccessfully for partnership with J_j" [24], and a matching is not established. The use of Mahalanobis distance is customary in data association [25], but it is often used in a nearest-neighbour scheme (proximity principle). The approach described above extends it by introducing also the exclusion principle. On the other hand, our proximity matrix **P** generalises the solution proposed by Ref. [26], because using Mahalanobis distance in a feature space allows taking into account both appearance and spatial position (and possibly other features) in a consistent way. This matching across the sequence produces tracks (figure 4). Many of them are due to noise, and only a few correspond to MOs. A track may represent only a *part* of an object, in the case of occlusion with a static element or because of over-segmentation.

C. Tracks building

If blobs are allowed to split (for the reasons described before) and merge (because the projections of the objects in the image overlap or because they physically touch each other) their descriptors change significantly, hence SL matching is likely to fail.
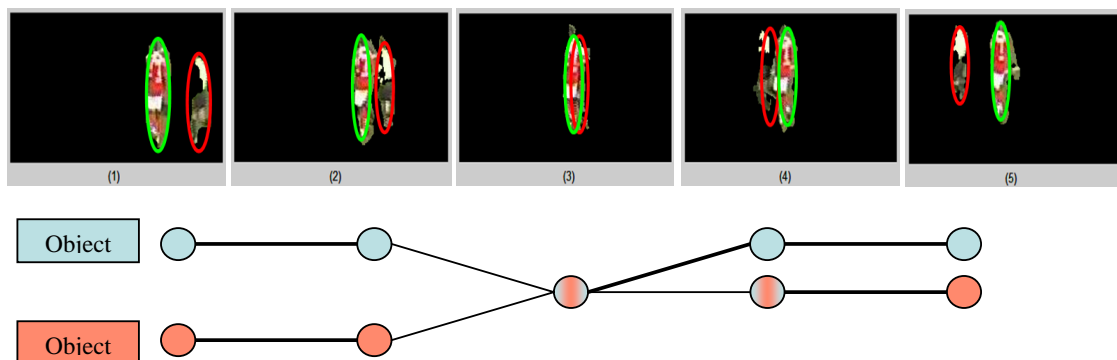

Fig. 4 Overlapping objects. When two objects merge, SL matching (bold line) typically fails in the merge section. The connected component is created by the edges added by template matching (thin line).

This also happens when objects enter/exit the view frustum or appear/disappear behind large occlusors. Moreover, the outputs of the blob matching phase are tracks, which are not suited to represent objects in split or merge situations, where nodes should be allowed to have degree greater than two. A template matching, which is likely to succeed where the SL matching failed, is employed in order to:

− Connect tracks representing fragments of the same MO.
− Connect tracks representing overlapping MOs (figure 4).
− Prolong tracks representing MOs.

The template matching procedure uses colour information and sum of squared differences (SSD) metric. Each external node is template-matched against the nodes (blobs) of the adjacent layer contained in a suitable search window, including internal nodes and those corresponding to a bad track. If one node is isolated we are prolonging the track. If both nodes are external we are chaining the two tracks. If one node is internal, we are increasing its degree above two, thereby catering for splitting and merging situations.

VI. CONCLUSION

In this paper, we describe the F-C-M algorithm, which allows detecting objects from colour image issued from soccer game. Indeed, this method incorporates color information by the use of another colour space that the RGB one, and fuzzy reasoning. Moreover, algorithms used in this paper are based on region approach, the first technique represents the sets of pixels corresponding to objects which will be tracked in (R-G) space and the probability density function of each distribution is estimated using nonparametric method. The crucial role played by optimal value of standard deviation leading to give enough precision to memberships probability of each cluster. The second technique use the features relative to objects in shot's sequence and then dissimilarity are computed by Mahalanobis distance. Region matching is achieved by the use of Singular Values Decomposition to associate features of two images, two stages were led; the first one is to build a proximity matrix of the two sets of features, the next stage is to perform this matrix. Finally, through this paper we can say that, the incorporation of color information and morphological features and also optimisation of statistical parameters lead to efficient results.

REFERENCES

[1] J. Allen, R. Butterly, M. Welch and R. Wood, "The physical and physiological value of 5-a-side soccer training to 11-a-side match play". J. Hum. Move. Stud., vol. 34, pp.1-11 1998.
[2] T. Reilly, and V. Thomas, "A motion analysis of work-rate in different positional roles in professional football match play". J. Human Move. Stud., vol. 2, pp.87-97, 1976.
[3] R.T. Withers, Z. Maricic, S. Wasilewski and L. Kelly, "Match analyses of Australian professional soccer players". J. Human Move. Stud., Vol. 1, No. 2, pp. 205-222, 1982.
[4] S.R. Mayhew, and H.A. Wenger, "Time-motion analysis of professional soccer", J. Human Move. Stud., vol. 11, pp. 49-52, 1985.
[5] E. Hennig, and R. Briehle, "Game analysis by gps satellite tracking of soccer players". In Proc. XI Congress of Canadian Society of Biomechanics, Canada, July, 2000, pp.44.
[6] A. Elgammal, and L. Davis, "Probabilistic framework for segmenting people under occlusion", in Proceedings of the Eighth International Conference on Computer Vision IEEE, pp.9-12, 2001.
[7] C. Stauffer, "Estimating tracking sources and sinks", in: Proceedings of the IEEE Workshop on Event Mining, 2003.
[8] M. Xu, J. Orwell, L. Lowey and D. Thirde, "Architecture and algorithms for tracking football players with multiple cameras'", IEE Proceedings: Vision, Image and Signal Processing, Vol. 152, No. 2, pp. 232-241, 2005.
[9] I. Haritaoglu, D. Harwood, and L. Davis, "Who? when? where? what? a real time system for detecting and tracking people", in Proceedings of the International Conference on Automatic Face and Gesture Recognition, pp. 222-227, 1998.
[10] L. Bretzner, I. Laptev and T. Lindeberg, "Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering', in Proc. 5th IEEE Int. Conf. Automatic Face and Gesture Recognition, pp. 423-428, 2002.
[11] S. Park, and J. Aggarwal, "Segmentation and tracking of interacting human body parts under occlusion and shadowing", in: IEEE Workshop on Motion and Video Computing, pp. 105-111, 2002.
[12] P. Viola, and W.M. Wells, "Alignment by Maximization of Mutual Information", International Journal on Computer Vision, Vol. 24, No. 2, pp.137-154, 1997.
[13] J. Fisher, and J.C. Principe, "A Nonparametric Method for Information Theoretic Feature Extraction", Proc. Defense Advance Research Projects Agency (DARPA) Image Understanding Workshop, 1997.
[14] O. Javed, and M. Shah, "Tracking and object classification for automated surveillance", In Proceedings ECCV, 2002.
[15] Z. Kim, and J. Malik, "Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking", In Proceedings ICCV, 2003.
[16] C. Stauffer, and E. Grimson, "Learning patterns of activity using real-time tracking", IEEE Trans. PAMI, Vol. 22, No. 8, pp.747-757, 2000.
[17] A. Torralba, and P. Sinha, "Statistical context priming for object detection", In Proc. ICCV, pp. 763-770, 2001.
[18] S.J. McKenna, S. Jabri, Z. Duric, M, Rosenfeld, and A.H. Wechsler, "Tracking groups of people", Comput. Vision Image Understand, vol. 80, pp. 42-56, 2000.
[19] S. Lu, G. Tsechpenakis, and D. N. Metaxas, "Blob analysis of the head and hands: A method for deception detection", In 38th Hawaii International Conference on System Sciences, 2005.
[20] M. Sivabalakrishnan and D. Manjula. "Human Tracking and Segmentation using Color Space Conversion", International Journal of Computer Science Issues, Vol. 7, No. 5, pp. 284-288, 2010.
[21] K. Huang, J. Lin, J.A. Gajnak, and R.F. Murphy, "Image Content-based Retrieval and Automated Interpretation of Fluorescence Microscope Images via the Protein Subcellular Location Image Database", IEEE International Symposium on Biomedical Imaging, pp.325-328, 2002.
[22] A. Buzzanca, G. Castellano and A.M. Fanelli, "Feature-Driven Classification of Musical Styles". World Academy of Science, Engineering and Technology, vol. 57, pp. 471-475, 2009.
[23] W. S. Qureshi and N. Alvi. "Object Tracking using MACH filter and Optical Flow in Cluttered Scenes and Variable Lighting Conditions". World Academy of Science, Engineering and Technology, Vol. 60, pp. 709-712, 2009.
[24] G. Scott, and H. Longuet-Higgins, "An algorithm for associating the features of two images", in: Proceedings of the Royal Society of London B, Vol. 244, pp.21-26, 1991.
[25] I. Cox, "A review of statistical data association techniques for motion correspondence", Int. J. Comput. Vision, Vol. 10, No. 1, pp. 53-66, 1993.
[26] M. Pilu, "A direct method for stereo correspondence based on singular value decomposition", in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.261-266, 1997.
[27] G. Padmavathi, M. Muthukumar and S. K. Thakur. "Non linear Image segmentation using fuzzy c means clustering method with thresholding for underwater images", International Journal of Computer Science Issues, Vol. 7, No. 3, pp. 35-40, 2010.
[28] Colombari, A., Fusiello, A. and Murino V. "Segmentation and tracking of multiple video objects", Journal of pattern recognition, vol. 40, pp. 1307 - 1317, 2007.

[29] G.Padmavathi1, P. Subashini and A. Sumi. "Empirical Evaluation of Suitable Segmentation Algorithms for IR Images", International Journal of Computer Science Issues, Vol. 7, No. 4, pp. 22-29, 2010.

[30] H. H. Huynh, J. Meunier, J.Sequeira, and M. Daniel, "Real time detection, tracking and recognition of medication intake". World Academy of Science, Engineering and Technology, vol. 60, pp. 280-287, 2009.

[31] H. K. Kim, S. H. Park, D. H. Kim, and S. J. Ko, "Probabilistic Center Voting Method for Subsequent Object Tracking and Segmentation". World Academy of Science, Engineering and Technology, vol. 59, pp.450-454, 2009.

[32] M. Sujaritha and S. Annadurai, "Color Image Segmentation using Adaptive Spatial Gaussian Mixture Model". World Academy of Science, Engineering and Technology, vol. 61, pp. 744-748, 2010.

BIOGRAPHIES

Mourad Moussa was born in Ben Arous, Tunisia in 1977. He received the Master degree in Automatic from the "Ecole Nationale d'Ingnieurs de Monastir" in 2006 and the Ph.D. degree in Electrical Engeneering from the "Ecole Nationale d'Ingnieurs de Monastir" in 2011. His research is related to Image Processing.

Ali Douik was born in Tunis, Tunisia. He received the Master degree from the "Ecole Normale Supérieur de l'Enseignement Technique de Tunis" in 1990 and the Ph.D. degree in Electrical Engeneering from the "Ecole Supérieure des Sciences et Technique de Tunis" in 1996. He is currently an Associate Professor in the "Ecole Nationale d'Ingénieurs de Monastir". His research is related to Automatic Control, robust control, evolutionary optimization control and Image Processing.

Hassani Messaoud was born in Mahdia, Tunisia. He received the Master degree from the "Ecole Normale Supérieur de l'Enseignement Technique de Tunis", in 1985 and the Ph.D. degree in Automatic Control from the "Ecole Centrale de Lille, France", in 1993. In 2001, he received the ability degree from the "Ecole Nationale d'Ingénieurs de Tunis". He is presently Professor in the "Ecole Nationale d'Ingénieurs de Monastir"and Director of the laboratory ATSI. He has been the supervisor of several PhD thesis and is author or co-author of several Journal articles. His research is related to Automatic Control and Signal Processing.