State filtering and change detection using TBM conflict - Application to human action recognition in athletics videos

Emmanuel Ramasso, Michèle Rombaut, Denis Pellerin

State filtering and change detection using TBM conflict Application to human action recognition in athletics videos

Emmanuel Ramasso, Michèle Rombaut, Denis Pellerin

***Abstract*—In this paper, we propose a tool called *Temporal Credal Filter with Conflict-based Model Change* (TCF-CMC) to smooth belief functions on-line in Transferable Belief Model (TBM) framework. The TCF-CMC takes temporal aspects of belief functions into account and relies on conflict information explicitly modelled in TBM when combining beliefs. TBM fusion, in addition to uncertainty, takes into account imprecision and conflict inherent to features. The TCF-CMC takes part in a wider system for human action recognition in videos. The whole system is tested on 51 videos (11000 images) with moving camera and real conditions where the TCF-CMC improves running, jumping, falling and standing-up actions recognition in high jump, pole vault, long jump and triple jump activities. The TCF-CMC is also compared to hidden Markov models. Lastly, a TBM rules-based modelling is compared to Gaussian mixture.**

***Index Terms*—Transferable Belief Model, Belief State Filtering, Human Motion Analysis, Novelty Detection, Moving Camera.**

I. Introduction

HUMAN motion analysis [1] is usually based on probability theory [2] and this paper proposes a new approach based on Transferable Belief Model [3] (TBM). TBM is a more general framework because it relies on belief functions which generalize probability measures [4]. With TBM, a variety of knowledge can be represented from certain and precise up to total ignorance. Particularly, doubt and conflict between hypotheses are explicit. Actually, TBM is a sound framework which takes imprecision, uncertainty, inconsistency and reliability of features into account. Many applications of TBM exist but TBM-based human motion analysis in video is just in its infancy. Some methods based on Dempster-Shafer's evidence theory [5] have been proposed to classify human postures [6] and emotions [7]. However, these methods are static since they do not include temporal evolution of belief and features. As a solution, we have proposed the Temporal Credal Filter with Conflict-based Model Change (TCF-CMC) to take temporal aspects of belief functions into account. It has been applied to improve human actions recognition by smoothing belief and separating actions states [8]. The TCF-CMC takes part in a wider system for human action recognition utterly based on a TBM methodology [9].

In this paper, we propose two main contributions. Firstly, the TCF-CMC to which is added an unsupervised and on-line criterion used in both assessment of filtering and detection of new actions (novelty). TCF-CMC modelling is based on joint belief functions and includes parameter adaptation. Secondly, the complete system is tested on real videos of athletics jumps. The TCF-CMC is compared to probability-based hidden Markov models [10] and a thorough analysis allows to facilitate TCF-CMC's parameters setting.

The remainder of the paper is as follows. Section II describes related work. The proposed architecture for human action recognition as well as features and fusion process are presented Section III. The TCF-CMC is then described Section IV. The quality criterion which assesses the TCF-CMC filtering is presented Section V. The system is tested Section VI before concluding and providing future work.

E. Ramasso, M. Rombaut and D. Pellerin are with the GIPSA-lab, DIS team (ex LIS), Grenoble, France ({*ramasso,rombaut,pellerin*}*@lis.inpg.fr*).

II. Related work: probabilistic methods

Human motion analysis is of key of importance for many multimedia applications such as human-computer interaction, surveillance and video indexing and retrieval. The final goal is the recognition of human actions, activities and interactions [1]. This paper focuses on human action recognition. In the state-of-the-art, state-space and template matching approaches are generally used. The latter have been used in [11] with motion descriptors. In this paper, we focus on state-space methods. They rely mainly on hidden Markov model (HMM) [10] and dynamic Bayesian network (DBN) [12]. One model is generally learned for each activity and the recognition process (inference) determines the most likely one. DBNs, partially coupled HMM and multi-observation HMM are exploited in [13] for causality discovery and events modelling. In [14], a description and a comparison between DBN and HMM is proposed for sports video sequence interpretation. In [15], HMM are used for gesture recognition and on-line learning of gestures. Recent works have focused on drawbacks of the current approaches based on probabilities [16].

III. System overview and features presentation

A. Architecture and features summary

The proposed human action recognition system [8] consists of four steps: features extraction, convertion into beliefs, fusion of beliefs according to rules and filtering by the TCF-CMC and assumes 1) the human is tracked by the camera, 2) trajectories of human's head, center of gravity and end of legs give sufficient information on actions and 3) a single human is moving. These assumptions are not very strong compared to the ones generally assumed [1] such as fixed view point, camera calibration and video quality. The system is generic enough to add new features and actions. Quality of filtering is tackled in this paper and used for action discovery.

Features used in this paper have been described in [9] and are roughly recalled hereafter. Numerical features are extracted

at each frame of the video. Three of them are computed by a camera motion estimator which are the horizontal ($P_{\rm hm}$) and vertical ($P_{\rm vm}$) motions and divergence ($P_{\rm div}$). A tracking algorithm provides human's head, center of gravity and end of legs position from which the variation of the center of gravity ($P_{\rm vcg}$), the angle between horizon and human axis ($P_{\rm swing}$) and a gait period ($P_{\rm gait}$) are derived. The feature vector is thus $\mathbf{x} = [P_{\rm hm}\ \ P_{\rm vm}\ \ P_{\rm div}\ \ P_{\rm vcg}\ \ P_{\rm swing}\ \ P_{\rm gait}]^T$.

B. Models of actions and fusion process

1) TBM background: We denote $\Omega_A = \{T_A, F_A\}$ the set of possible (symbolic) states called frame of discernment (FoD) and standing for *an action A is true* (T_A) or *false* (F_A). A basic belief assignment (BBA) $m_P^{\Omega_A}$, defined on the power set 2^{Ω_A} of Ω_A and depending on the value of feature P, is $m_P^{\Omega_A} : 2^{\Omega_A} \rightarrow [0,1]$, $X \rightarrow m_P^{\Omega_A}(X)$ and by construction $m_P^{\Omega_A}(\emptyset) = 0$ and $\sum_{X\subseteq\Omega_A} m_P^{\Omega_A}(X) = 1$. A value $m_P^{\Omega_A}(X)$ expresses a confidence in proposition $X \subseteq \Omega_A$ but does not imply any additionnal claims regarding subsets of X [4]. It is a fundamental difference with probability theory.

The fusion process is performed frame by frame for each action independently. Given two distinct BBAs $m_{P_1}^{\Omega_A}$ and $m_{P_2}^{\Omega_A}$, defined on the same FoD Ω_A, the combination is:

$$m_{P_1}^{\Omega_A} \textcircled{\triangle} m_{P_2}^{\Omega_A}(E) = \sum_{C\triangle D=E} m_{P_1}^{\Omega_A}(C).m_{P_2}^{\Omega_A}(D) \tag{1}$$

with $\triangle = \cap$ (resp. $\cup$) for the conjunctive (resp. disjunctive) rule. The resulting BBA, $m_{P_1,P_2,\ldots,P_K}^{\Omega_A}$, is defined on Ω_A.

2) Synthesizing belief functions from features: Two types of observation models are tested. The first one, based on fuzzy-sets, is described in [9] and consists of trapezoids providing belief on actions related to features values. Beliefs are combined according to predefined (fuzzy) rules. The combination is performed using Transferable Belief Model's rules [3]. The second one [17] consists in 1) estimating probabilistic models and 2) transforming and combining likelihoods to obtain a belief function. The reader may refer to [17] to be convinced about the contribution and usefulness of this methodology.

3) Integrating reliability of features: In TBM, reliability coefficient decreases belief [3] of sources (camera motion estimation and tracking) that do not work in good conditions thus is relevant for video analysis. Two coefficients are computed on-line and automatically from features: α_{dist} (tracking) is related to the distance between center of gravity and head and assumed quite constant between successive frames, and α_{sup} (camera) relies on the fact that silhouette's size is quite constant between successive frames.

IV. Temporal Credal Filter

The Temporal Credal Filter with Conflict-Based Model Change (TCF-CMC) has been proposed to smooth belief functions, making them consistent, and to separate belief state [8]. At frame f, the TCF-CMC robustly determines the state of an action A which can be either T_A (A is true) or F_A (A is false). The TCF-CMC process consists of three steps (Fig. 1): (i) Prediction, (ii) Fusion and (iii) Detection of model change.

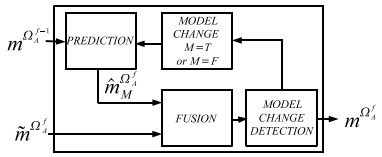

Fig. 1. TCF-CMC principle.

A. Prediction of the current BBA

1) BBA's evolution model: The BBA of the current TCF-CMC output $m^{\Omega_A^f}$ at frame f is supposed close to the previous one $m^{\Omega_A^{f-1}}$ because human motions are continuous. A model of evolution $\mathcal{M}$ is defined for each state of an action A with $\mathcal{M} \in \{\mathcal{T}, \mathcal{F}\}$ ($\mathcal{T}$ for T_A and $\mathcal{F}$ for F_A). It is expressed by a BBA defined on the joint space $\Omega_A^f \times \Omega_A^{f-1}$ with:

$$\begin{aligned} \text{Model } \mathcal{T} &\left\{ \begin{array}{l} m_{\mathcal{T}}^{\Omega_A^f\times\Omega_A^{f-1}}(T_A^f \cap T_A^{f-1}) = \gamma_{\mathcal{T}} \\ m_{\mathcal{T}}^{\Omega_A^f\times\Omega_A^{f-1}}(\Omega_A^f \times \Omega_A^{f-1}) = 1-\gamma_{\mathcal{T}} \end{array} \right. \\ \text{Model } \mathcal{F} &\left\{ \begin{array}{l} m_{\mathcal{F}}^{\Omega_A^f\times\Omega_A^{f-1}}(F_A^f \cap F_A^{f-1}) = \gamma_F \\ m_{\mathcal{F}}^{\Omega_A^f\times\Omega_A^{f-1}}(\Omega_A^f \times \Omega_A^{f-1}) = 1-\gamma_F \end{array} \right. \end{aligned} \tag{2}$$

where $\Omega_A^f \times \Omega_A^{f-1}$ is the cartesian product of FoDs.

2) Building the predicted BBA: The prediction $\hat{m}_{\mathcal{M}}^{\Omega_A^f}$ of the current BBA is computed by:

$$\hat{m}_{\mathcal{M}}^{\Omega_A^f} = (m_{\mathcal{M}}^{\Omega_A^f\times\Omega_A^{f-1}} \textcircled{\cup} m^{\Omega_A^{f-1}\uparrow\Omega_A^f\times\Omega_A^{f-1}})^{\downarrow\Omega_A^f} \tag{3}$$

The $\textcircled{\cup}$-rule emphasizes transitions between action states and never assigns more belief to an hypothesis than does the previous BBA [8]. Before combination by the $\textcircled{\cup}$-rule (Eq. 1), both BBAs must be defined on the same FoD, in this case $\Omega_A^f \times \Omega_A^{f-1}$, using the vacuous extension operator [18] ("$\uparrow$") which provides $m^{\Omega_A^{f-1}\uparrow\Omega_A^f\times\Omega_A^{f-1}}(C) = m^{\Omega_A^{f-1}}(B)$ if $C = B \times \Omega_A^f, B \subseteq \Omega_A^{f-1}$, and 0 otherwise. After combination the result is projected ("$\downarrow \Omega_A^f$") onto the current FoD Ω_A^f providing, $\forall B \subseteq \Omega_A^f$, $\hat{m}_{\mathcal{M}}^{\Omega_A^f\times\Omega_A^{f-1}\downarrow\Omega_A^f}(B) = \sum_C \hat{m}_{\mathcal{M}}^{\Omega_A^f\times\Omega_A^{f-1}}(C)$, with $C \subseteq \Omega_A^f \times \Omega_A^{f-1}, C \downarrow \Omega_A^f = B$.

B. Fusion of prediction and measure

Prediction $\hat{m}_{\mathcal{M}}^{\Omega_A^f}$ and measure $\tilde{m}^{\Omega_A^f}$ are combined by:

$$m^{\Omega_A^f} = \hat{m}_{\mathcal{M}}^{\Omega_A^f} \textcircled{\cap} \tilde{m}^{\Omega_A^f} \tag{4}$$

Conflict [3] denoted ϵ_f is quantified by the BBA value on emptyset i.e. $\epsilon_f = m^{\Omega_A^f}(\emptyset)$. A cumulative sum (CUSUM process) of ϵ_f is performed to detect model change.

C. Detection of model change by a CUSUM process

The CUSUM of ϵ_f is given at f by $\mathbf{CS}(f)$:

$$\mathbf{CS}(f) \leftarrow \lambda \times \mathbf{CS}(f-1) + \epsilon_f \tag{5}$$

where $\lambda \in [0,1]$ is a fader coping with low conflict during a long time and forgetting gradually past event. The CUSUM is robust because integrates conflict along time. When $\mathbf{CS}(f) > \mathcal{T}_w$ (warning threshold), the frame number f_w is stored and the model is kept until $\mathbf{CS}(f) > \mathcal{T}_s$ (stop threshold), at frame f_s,

where the model is changed. The new model is applied from f_s and belief on $[f_s - \mathcal{W}, f_s]$ is compelled to be vacuous (all belief on the whole FoD) to emphasize action states transition.

D. TCF-CMC output

When a conflict appears between prediction $\hat{m}^{\Omega_A^f}_{\mathcal{M}}$ and measure $\tilde{m}^{\Omega_A^f}$, i.e. when $0 < \mathbf{CS}(f) < \mathcal{T}_s$, the prediction is kept instead of an erroneous measurement to avoid propagating conflict which is absorptive by the $\textcircled{\cap}$-rule:

$$m^{\Omega_A^f} = \hat{m}^{\Omega_A^f}_{\mathcal{M}} \textcircled{\cap} \tilde{m}^{\Omega_A^f} \text{ if } \epsilon_f \leq \delta_\emptyset \text{ , and } \hat{m}^{\Omega_A^f}_{\mathcal{M}} \text{ otherwise} \quad (6)$$

where $\delta_\emptyset$ is a threshold reflecting a tolerance to the conflict adaptively computed using the mean of conflict over a window (size N) of a few frames: $\delta_\emptyset = 1/N \cdot \sum_{f_i=(f-N-1)}^{f} \epsilon_{f_i}$.

V. On-line unsupervised quality criterion

A. Local Quality recognition performance

An unsupervised quality criterion is computed to assess, on-line, the quality of the filtering and, as well, the recognition by TCF-CMC. The criterion has the particularity to be computed automatically without any additionnal knowledge. We denote $LQ^{f_s:f}_{i,j}[\mathcal{M}](s)$ the Local Quality recognition performance criterion of state $s \in \{T_A, F_A\}$ of action A_i within activity S_j in interval of frames $[f_s, f]$ (between the start frame f_s and the current frame f) given the model $\mathcal{M}$. It is computed as an on-line weighted mean by:

$$\begin{aligned} LQ^{fs:f}_{i,j}[\mathcal{M}](s) &= \left(1 - \frac{1}{f - fs}\right) \cdot LQ^{fs:(f-1)}_{i,j}[\mathcal{M}](s) \quad (7) \\ &+ \frac{m^{\Omega_A^f}(s)}{f - f_s} \cdot (1 - \epsilon_f) \end{aligned}$$

Eq. 7 embeds past events and innovation. Conflict weighs the current belief. The quality LQ is high when both the conflict is low and the belief is high.

This criterion is used for *false alarm detection* as follows:

$$\text{if } LQ^{f_s:f}_{i,j}[\mathcal{M}](s) < \delta_{\text{FA}} \text{ then FALSE ALARM} \quad (8)$$

where δ_{FA} is the minimal quality value required to validate an action. When a false alarm appears, the interval of frames $[f_s, f]$ is rescanned with the other model and with a CUSUM always set to 0 (not active).

B. On novelty detection

Adaptive systems are of great interest in most of applications related to Computer Vision. We propose to use the quality criterion in order to detect new actions and to correct features models. The procedure is simple: we consider a pool of N possible actions. For each of them we have set up the features models (gaussian or trapezoids). Given observations, we perform the filtering based on the TCF-CMC. When all actions quality criteria are low, we are facing a new action. The decision rule to detect and create a new action is thus:

$$\text{if } \forall i, \forall j, LQ^{f_s:f}_{i,j}[\mathcal{T}](s) < \delta_{\text{FA}} \text{ then NEW ACTION} \quad (9)$$

When a new action is detected on a given interval of frames $[f_s, f]$, then the features models are automatically computed. For that, a feature vector containing the values of features in this interval is created and could be analyzed by a learning procedure, such as EM, in order to compute the models corresponding to the new action.

VI. Experiments

The system is tested for action recognition in athletics jumps. The goal is to recognize 3 actions: running, falling and jumping (plus standing-up in some cases), within 4 different athletics jumps (activities): high jump, pole vault, triple jump and long jump. Fig. 2 depicts the first three tests and the fourth one concerns the analysis of the TCF-CMC parameters.

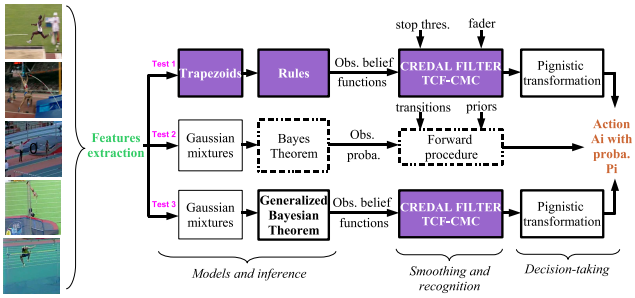

Fig. 2. Tests: dotted-lines boxes concern HMM and probabilities, bolded-line for TBM (belief). Both used observations (obs.) obtained from features. Probabilities are used to bet on actions. Left: some images from videos.

A. Database description and evaluation criteria

The database used for testing is made of 51 videos acquired by a moving camera and with several unknown view angles (original conditions [1]). There are 25 pole vaults (3000 frames), 15 high jumps (2300 frames), 12 triple jumps (3200 frames) and 11 long jumps (2200 frames). The videos are from various sources (TV, DVD, VHS) compressed using divx encoder in 25 fps and 352x288 size. About a half of the database are indoor and another half outdoor meetings. In addition to camera motion, illumination change, view variation and other moving people, the challenge of the tests concerns the fact that each video represent one jump and that each jump is made of actions. Therefore, actions are *not* pre-segmented as usually done in experiments. We assume that the system has to be able to detect actions separately within activities.

The database was manually annotated with states of actions true and false. Evaluation relies on recall ($\mathcal{R}$) and precision ($\mathcal{P}$) indexes with $\mathcal{R} = \frac{C \cap R}{C}$, $\mathcal{P} = \frac{C \cap R}{R}$, where C is the reference set obtained by expert annotations, R is the set of retrieved frames provided by the recognition module and $C \cap R$ is the number of correctly retrieved frames. In order to obtain only one criterion, the F_1-measure, well-known in database management, is used: $F_1 = \frac{2 \times \mathcal{R} \times \mathcal{P}}{\mathcal{R} + \mathcal{P}}$.

B. Test 1: Improvement of belief on actions by the TCF-CMC

The goal is to improve belief on actions and separate action states (true or false). Belief on each action is provided by discriminant rules and trapezoids [9] in order to distinguish all types of actions. The same TCF-CMC setting was done

for high jump, pole vault and long jump. For information we set: $\lambda = 0.9$, $\gamma_{\mathcal{T}} = \gamma_{\mathcal{F}} = 0.9$, $\mathcal{T}_s = 2.9$, $\mathcal{T}_w = 1$, $\mathcal{W} = 5$, $\delta_{FA} = 60\%$. Concerning triple jump, actions are shorter thus $\mathcal{T}_s$ was decreased a little. In order to bet on action state, we use the pignistic probability (BetP) [3] defined as $\mathrm{BetP}(T_A) = (m(T_A)+0.5*m(T_A \cup F_A))/(1-m(\emptyset))$. If $\mathrm{BetP}(T_A) > 0.51$ then A is considered as true.

Tab. I (col. T_1 and T_2) shows an important gain obtained thanks to the TCF-CMC which smooths belief and detects some false alarms. Illustrations depicted Fig. 3 (right) show that actions belief are smooth and consistent. Moreover, they shed light on the need to use temporal contraints between actions in order to avoid overlapping (see running and jumping). Fig. 4 (lines 1 and 2) illustrates evolution of CUSUM and conflict which are generally high when an action is constrained to be in one state, e.g. true, while data expressed the contrary, e.g. false (when stop threshold is not reached). The fader allows to smooth the CUSUM and to forget past event thus avoiding some false alarms (see line 2, col 3). Lines 3 and 4 show the LQ recognition performance criterion (Eq. 7) used for false alarms detection (Eq. 8). If the LQ of two different actions are high then actions are overlapping thus LQ allows to detect position of transitions. When all LQ are low, a new action is detected. For instance interval $[82, 87]$ is such a case (since interval size is short, it is not necessary to learn another action but this information can be exploited to adapt the models by adding more variance in trapezoids/Gaussian mixture models (MoG)). Action recognition is of good quality if LQ is close to 1 during both true and false states (e.g. jumping and running are good but not falling).

C. Test 2: TCF-CMC with trapezoids and rules vs. HMM

The goal is to compare HMM and TCF-CMC for the recognition of the three actions plus another one called *standing-up* (sometimes appears in some activities). HMM setting: 16 models of HMM are built (one for each action within each activity), each one is made of 2 states (true/false), observations are continuous, modelled by MoG[1] and assumed independent conditioned on the state, Baulm-Welch algorithm [10] is used to learn transition matrices, priors and MoG mixing coefficients. Learning is performed on $50-75\%$ (according to the jump) of the database and tests are done on the remainder. Then, $50 - 75\%$ is again randomly selected, the test is rerun and we compute the mean on both results.

Assessment: we apply the forward algorithm [10] to smooth probabilities on states. At each frame, an action is true if its probability is greater than 0.51 (as for the TCF-CMC assessment). The forward procedure assesses on-line the recognition as the TCF-CMC does, so the comparison is relevant. Using references, we compute F_1-measures and results are given in Tab. I, col. T_2 and T_3. Results are of good quality, close to the TCF-CMC results. The great difference is the learning step since 75% of the learning set was necessary to learn MoGs however the main advantage is that learning is automatic. Obviously, integrating explicit duration, increasing the number of states or the number of components might improve the results but these results show that the TBM-based approach is efficient. Note that standing up action is detected which is a difficult task because it is changing and short. When sufficient data are available, EM can learn associated models however, no rules were provided for the trapezoids-based approach.

[1]Nb. of component is set up automatically using [19] and equals 2 to 3 according to jumps. HMM implementation is embedded in BNT toolbox [12].

D. Test 3: TCF-CMC with Gaussian mixtures vs. HMM

We compare HMM to the TCF-CMC when observations models are given by MoG for both of them. Likelihoods provided MoG are transformed into belief functions using the method proposed by Smets in [18] and exploited for target identification in [17]. The method is based on the Generalized Likelihood Principle and on the Generalized Bayesian Theorem (GBT) [18]. Comparison results are gathered in columns T_3 and T_4 in Tab. I and shed light on importance of doubt and conflict: TBM modelling (using GBT) and TCF-CMC process increase the recognition rate (as in [17]). This is explained by three important characteristics of the proposed system: doubt and conflict are explicitly modelled, and the TCF-CMC plays a crucial role by separating actions states.

E. Test 4: TCF-CMC sensitivity to the parameters

In this section, the TCF-CMC sensitivity is analyzed according to the two most important parameters which are $\mathcal{T}_s$ (stop threshold of CUSUM) and λ (fader). We study the sensitivity according to trapezoids-based and MoG-based feature models. To measure the sensivity, we apply the TCF-CMC on 10 videos sequences of high jump (66% of the database). We compute the F_1-measure for three actions. Fig. 5 pictorially describes the results. Concerning trapezoids (top), several common values can be chosen for the three actions (thus the setting is easy): for $\lambda = 0.9$ and $\mathcal{T}_s = 2.9$ we obtain a performance $> 72\%$ for the three actions (this setting was used in test T_2). Sensitivity seems more important for jumping action while running and falling actions are more robust. Concerning MoG (bottom), jumping action is still more sensitive to parameters changing than running and falling actions. Compared to trapezoids and rules, the three frontiers of decision (for each action) are quite different using MoG whereas they are quite similar using the former (see gradual changes in gray levels). Furthermore, the recognition rate surface (in gray level) is smoother with trapezoids than with MoG but locally better with the latter. Note on complexity: for K actions, T images and n features, complexity is $\mathcal{O}(T \cdot K \cdot 2^n)$.

VII. Conclusion and future work

The proposed Temporal Credal Filter with Conflict-Based Model Change (TCF-CMC), defined in Transferable Belief Model (TBM) framework, worked well in smoothing belief functions on-line. The TCF-CMC relies on conflict emphasized in the fusion process using TBM. Beliefs can be provided by a combination of features, independantly from the models used for the convertion (trapezoids, gaussians or others). A new criterion is also proposed for novelty detection. The TCF-CMC improved human action recognition in 51 real videos

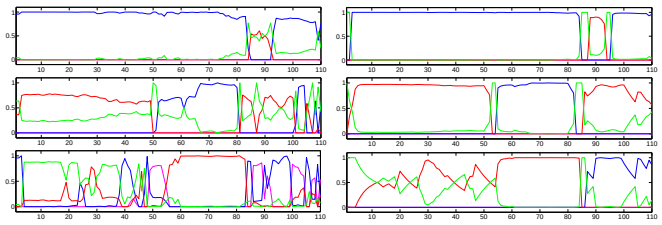

Fig. 3. Evolution of belief ($\in [0,1]$) of actions running (1^{st} line), jumping (2^{nd} line) and falling (3^{rd} line) in a high jump before (left) and after (right) filtering by TCF-CMC. Legend: States true (T_A, in blue), false (F_A, false), ignorance ($T_A \cup F_A$, in green) and conflict ($\emptyset$, in magenta).

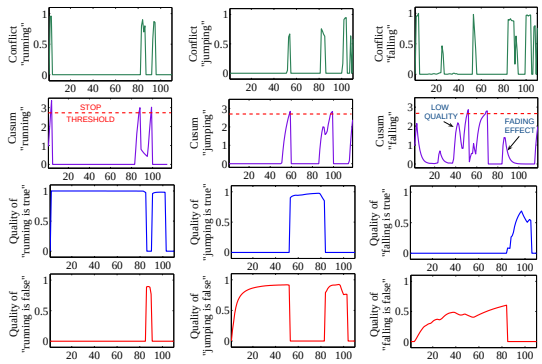

Fig. 4. Video of high jump, the same as Fig. 3. Lines 1 and 2: conflict and CUSUM evolution during filtering. Lines 3 and 4: Local Quality recognition performance evolution for each state (true and false) and each action.

of athletics jumps and good performance was obtained compared to usual hidden Markov models. Parameters setting is also easy. Fascinating future researches are actions sequences (activity) recognition, pattern discovery and adaptation using TBM and notably the TCF-CMC, conflict information and the quality criterion proposed in this paper.

ACKNOWLEDGEMENT

We thank C. Panagiotakis and G. Tziritas (Depart. of Computer Science, Univ. of Crete, P.O. Box 2208, Heraklion, Greece) for data and codes exchanges. This work was supported partially by SIMILAR European excellence network.

TABLE I

F_1-MEASURE (IN %) AFTER TRAPEZOIDS AND RULES DETECTION WITHOUT TCF-CMC (T_1) AND WITH TCF-CMC (T_2), AFTER HMM WITH MOG (T_3) AND AFTER TCF-CMC WITH MOG (T_4).

HIGH JUMP	T_1	T_2	T_3	T_4
running	69%	79%	81%	88%
jumping	61%	71%	67%	74%
falling	69%	76%	76%	81%
standing up	–	–	53%	58%
POLE VAULT	T_1	T_2	T_3	T_4
running	71%	79%	75%	79%
jumping	61%	70%	63%	69%
falling	60%	74%	64%	68%
standing up	–	–	43%	53%
TRIPLE JUMP	T_1	T_2	T_3	T_4
running	60%	72%	83%	92%
jumping	60%	65%	63%	68%
falling	55%	61%	56%	60%
standing up	–	–	58%	69%
LONG JUMP	T_1	T_2	T_3	T_4
running	78%	82%	86%	95%
jumping	56%	63%	59%	72%
falling	66%	70%	65%	67%
standing up	–	–	43%	56%

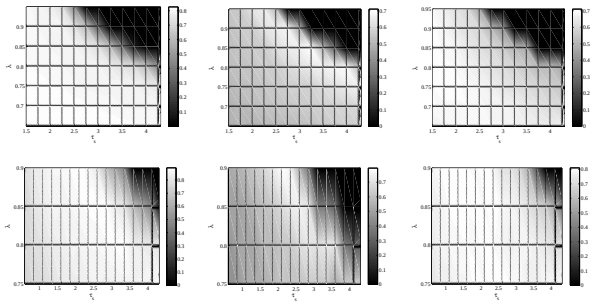

Fig. 5. TCF-CMC sensivity using trapezoids+rules (top) and MoG (bottom): F_1-mesure vs cusum's stop threshold and fader for running (1^{st} col.), jumping (2^{nd} col.) and falling (3^{rd} col.) actions (mean on 10 highjumps).

REFERENCES

[1] T. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *CVIU*, vol. 81, pp. 231–268, 2001.
[2] S. Hongeng, R. Nevatia, and F. Bremond, "Video-based event recognition and probabilistic recognition methods," *CVIU*, vol. 96, pp. 129–162, 2004.
[3] P. Smets and R. Kennes, "The Transferable Belief Model," *Artificial Intelligence*, vol. 66, pp. 191–234, 1994.
[4] G. Klir and M. Wierman, *Uncertainty-based information. Elements of generalized information theory, 2nd edition*, Physica-Verlag, 1999.
[5] P. Smets, "What is dempster-shafer's model?" in *Advances in the Dempster-Shafer theory of evidence*, 1994, pp. 5–34.
[6] V. Girondel, A. Caplier, L. Bonnaud, and M. Rombaut, "Belief theory-based classifiers comparison for static human body postures recognition in video," *Signal Processing*, vol. 2, pp. 29–33, 2005.
[7] Z. Hammal, A. Caplier, and M. Rombaut, "Belief theory applied to facial expressions classification," in *Int. Conf. on Advances in Pattern Recognition*, Bath, United Kingdom, 2005.
[8] E. Ramasso, M. Rombaut, and D. Pellerin, "A Temporal Belief Filter improving human action recognition in videos," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. 2, 2006, pp. 141–144.
[9] E. Ramasso, D. Pellerin, C. Panagiotakis, M. Rombaut, G. Tziritas, and W. Lim, "Spatio-temporal information fusion for human action recognition in videos," in *European Signal Processing Conf.*, 2005.
[10] L. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, pp. 257–285, 1989.
[11] A. Efros, A. Berg, G. Mori, and J. Malik, "Recognizing action at a distance," in *IEEE ICCV*, France, 2003, pp. 726–733.
[12] K. P. Murphy, "Dynamic Bayesian networks: Representation, inference and learning," Ph.D. dissertation, UC Berkeley (CSD), 2002.
[13] T. Xiang and S. Gong, "Discovering Bayesian causality among visual events in a complex outdoor scene," in *IEEE AVSS*, 2003, pp. 177–182.
[14] Y. Luo, T. Wu, and J. Hwang, "Object-based analysis and interpretation of human motion in sports video sequences by dynamic Bayesian networks," *CVIU*, vol. 92, pp. 196–216, 2003.
[15] A. Wilson and A. Bobick, "Hidden Markov models for modeling and recognizing gesture under variation," *Int. Jour. of Pattern Recognition and Artificial Intelligence*, vol. 15, pp. 123–160, 2001.
[16] M. Shah, "Understanding human behavior from motion imagery," *Machine Vision and Applications*, vol. 14, pp. 210–214, 2003.
[17] F. Delmotte and P. Smets, "Target identification based on the Transferable Belief Model interpretation of Dempster-Shafer model," *IEEE Trans. on SMC*, vol. 34, 2004.
[18] P. Smets, "Beliefs functions: The Disjunctive Rule of Combination and the Generalized Bayesian Theorem," *IJAR*, vol. 9, pp. 1–35, 1993.
[19] M. Figueiredo and A. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. on PAMI*, vol. 24, pp. 381–396, 2002.