Time-dependent interactive graphical models for human activity analysis

Marco Cristani and Vittorio Murino

Università degli Studi di Verona, Dipartimento di Informatica, Verona, Italy,
cristanm|murino@sci.univr.it

Abstract. Automatically monitoring and classifying human activities is one of the most challenging problems currently faced in machine learning. In this paper, we propose a statistical model aimed at modeling interactions among human subjects, in particular, conversational audio data is here analyzed. The proposed model, called Coupled Hidden Duration Semi Markov Model, takes inspiration from the large literature on Hidden Markov Models and its variants. The novelty introduced by the model is the capability of dealing with interacting state processes, where 1) states that characterize a single process exhibit different time durations, and 2) different processes involved in an interaction are not synchronized, i.e., their states do not begin/end at the same time instants. Comparative synthetical and real data experiments are presented, showing that the proposed model is able to tackle difficult interactive situations, not otherwise manageable by the state-of-the-art algorithms.

1 Introduction

Automatic activity recognition can be considered a recent growing area in automatic surveillance and monitoring research fields. In this context, the primary task is to classify multimodal patterns as simple human actions and, subsequently, starting from modeling individual actions, the last challenge is to actually cope with group activities, evidencing and quantifying the causal interactions among the individuals.
Dynamic Bayes Nets (DBN) (Jordan 1999) offer an elegant mathematical framework to combine the observations of the activities to be modeled (bottom-up) with complex behavioral priors (top-down), in order to provide expectations about the processes and dealing properly with the uncertainty.
A widely known DBN is the Hidden Markov Model (HMM) (Rabiner 1989), that models single Markov processes whose states are not directly visible. In the last years, several HMM extensions have been proposed, which can be roughly subdivided in two main families. First, there are models managing processes whose states are assumed to remain unchanged for some random time duration before their transitions, such as the

Hidden Semi Markov Model (HSMM) (Murphy 2002; Hongeng and Nevatia 2003).
The second family of HMM extensions copes with the interaction among Markov processes (Brand, Oliver, and Pentland 1997; Jordan 1999; Basu, Choudhury and Pentland 2001) in which various interaction aspects are considered. For instance, it is possible to extract the influence that *a whole process* exerts on another (Basu et al. 2001), or simply the inter-processes *state* conditional probabilities (Brand et al. 1997).

In this paper, we try to couple the two families of HMM extensions explained above, considering interactions among semi Markov processes. The modeling of such kinds of interactions is hard; actually, the interaction among "simple" Markov processes is evaluated at each time step, because at each time step every process makes a state transition, i.e., the processes are "transition" synchronized. In the case of interaction among semi Markov processes, the "transition" synchronization may be missed: roughly speaking, it may happen that a process continuously maintains a state while another interacting process performs several state transitions. Therefore, the computation of inter-chain conditional probabilities needs particular care in determining which are the states that condition or are conditioned by other states.

In order to deal with such a situation, we propose a novel framework, called Coupled Hidden Duration Semi Markov Model (CHD-SMM).
The aim of the paper is to use a CHD-SMM to capture and learn the nature of a *global process* (GP) formed by interacting semi Markov *individual processes* (IPs), highlighting aspects of the interaction present within.
In order to cope with complexity issues, the basic hypotheses over which a CHD-SMM can work are twofold. First, the whole global process modeled is formed by various *visible* semi Markov IPs. In other words, at each time step we can gather a shot of the different *visible* state labels assumed by every IP. The second hypothesis is that the IPs are interacting, in the sense that a transition of an IP towards one state is conditioned on the past story of all the other IPs.
In the experimental section, after reporting a synthetical example, we face the problem of recognizing high-level multiple audio activities, or conversational styles, where different human speech segments are modeled as IPs, whose visible states model periods of silence/speech of different duration. To the best of our knowledge, this is the first attempt to model such kinds of high-order interdependent dynamics, thus representing an improvement in the state of the art of the automatic monitoring literature.
The rest of the paper is organized as follows. In Section 2, the needed theoretical background is presented and Section 3 details the proposed model. Section 4 presents experimental results and concludes the paper.

2 Fundamentals

2.1 Hidden Markov Models

The entities characterizing an Hidden Markov Model $\boldsymbol{\lambda}$ are: the set S of the N hidden states; the transition matrix $\mathbf{A} = \{a_{ij}\}$, where $a_{ij} = P(S_t = j|S_{t-1} = i), 1 \leq i, j \leq N$ with $a_{ij} \geq 0$, $\sum_{j=1}^{N} a_{ij} = 1$ and the variable S_t indicating the state at time t; the emission matrix $\mathbf{B} = \{P(O|j)\}$ indicating the probability that the state

j emits the symbol O, and the initial state probability distribution $\boldsymbol{\pi} = \{\pi_i\}$, where $\pi_i = P(S_1 = i)$, $1 \leq i \leq N$ with $\pi_i \geq 0$ and $\sum_{i=1}^{N} \pi_i = 1$. For convenience, we denote an HMM as a triplet $\boldsymbol{\lambda} = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$.
The learning of the HMM's parameters can be performed in two ways: as a direct derivation of the EM algorithm (the Baum-Welch procedure (Rabiner 1989)), or as a standard constrained optimization process (Zhong 2002).

2.2 HMM coupling architectures

The most intuitive structure of HMM coupling is represented by unstructured Cartesian product HMMs (Jordan, Ghahramani, Jaakkola, and Saul 1999), i.e. a group of HMMs in which the state of one model at time t depends on the states of all models (including itself) at time $t-1$, i.e.

$$P({}^{c}S_t|{}^{1}S_{t-1}, ..., {}^{C}S_{t-1}). \tag{1}$$

The state transition probability for C processes is described by a $(C+1)$ dimensional matrix leading to N^C free parameters (assuming a common number of hidden states N). In the Coupled HMMs (Brand et al. 1997), the joint conditional dependency of Eq.1 is substituted by the product of all marginal conditional probabilities. Another solution is the Influence Model (IM) (Basu et al. 2001): here the full transition probability is modeled with a linear combination of singular inter-chains transition probabilities, where the weights θ represent the influence among chains (*not among the states of the chains*). In formulae:

$$P({}^{c}S_t|{}^{1}S_{t-1},...,{}^{C}S_{t-1}) = \sum_{d=1}^{C} \theta_{(cd)} P({}^{c}S_t|{}^{d}S_{t-1}) \tag{2}$$

with $1 \leq c, d \leq C$, $\theta_{(cd)} \geq 0$, $\sum_{d=1}^{C} \theta_{(cd)} = 1$. In this case, the advantage is a good compromise between number of parameters needed ($CN^2 + C^2$, where the term CN^2 corresponds to the free parameters of the intra-chain transition tables, and C^2 for the influence coefficients) and expressivity of the model. In practice, the IM is able to model each interaction between pairs of chains, but is not able to model the joint effect of multiple chains together.

3 Coupled Hidden Duration Semi Markov Model

As pointed out in (Murphy 2002), Semi Markov processes can be approached as Markov processes whose *generalized* states (simply pointed out as *states* in this paper) are pairs of the form $S_{\langle t_k \rangle} =< S_{t_k}, D_{t_k} >$, where the variable S_{t_k} expresses the automaton *state label* occurring at time t_k, which is observed consecutively for D_{t_k} time instants, and $\langle t_k \rangle$ indicates the time interval $[t_k, t_k + D_{t_k}[$. In this paper, we model a GP composed by C interacting semi-Markov IPs. To do this, we suppose that the IPs are directly *visible*, i.e., an observation over an IP, at whatever instant t_k, permits deterministically to individuate its current state label S_{t_k}. To ease the reading, we describe

the case of $C = 2$ IPs, adding the apex $'$ to the quantities related to the second IP.
The joint probability of a GP states sequence is:

$$P\left(S_{\langle t_1\rangle}, S_{\langle t_2\rangle}, \ldots, S_{\langle t_M\rangle}, S'_{\langle t'_1\rangle}, S'_{\langle t'_2\rangle}, \ldots, S'_{\langle t'_{M'}\rangle}\right) \tag{3}$$

where, in general, the number of occurring states per chain can be different, i.e., $M \neq M'$, and $t_k \neq t'_k$, which means that no synchronization is present among state transitions. Anyway, we start the analysis supposing the states as perfectly synchronized (see Fig. 1(a)). In this case, we can rewrite Eq.3 dropping the apex $'$ from the time indexes, being them equal by definition. Subsequently, we first consider the factorization of the joint probability of a semi Markov process proposed in (Murphy 2002), which considers transition probability factors of the form $P(S_{\langle t_k\rangle}|S_{\langle t_{k-1}\rangle})$. Inspired by this, we rearrange Eq.3 as a product of coupled probability transition terms, assuming the form :

$$P(S_{\langle t_k\rangle}|S_{\langle t_{k-1}\rangle}S'_{\langle t_{k-1}\rangle}), \tag{4}$$

Such term considers the probability of being in state $S_{\langle t_k\rangle}$, given the previous own semi Markov state, and the semi Markov states assumed by the other synchronized IPs. In this case, evaluating Eq.4 is equivalent to the unstructured Cartesian HMM case, reaching a $O(N^C)$ space-complexity bound (see Fig. 1(b)). If the states are not synchronized (Fig. 1(c)) we can reach the exponential bound of $O((N^C)^L)$ (Fig. 1(c)), where L is the maximum state duration allowed. This happens when, during the interval $\langle t_{k-1}\rangle$ of length $D_{t_{k-1}}$, we have transitions in the other IP among different states at each time instant.
The situation can be better faced if we suppose that a transition among different states of one particular IP contributes to an evolution of the whole GP, introducing the concept of *implied transition*. Technically, we force a state transition of all the IPs whenever a single IP performs a transition between states having different label. This causes a state fragmentation, i.e., a state of an IP $S_{t_{k-1}}$ becomes two states $S_{\langle t_h\rangle}, S_{\langle t_{h-1}\rangle}$ if a transition between different states occurs at instant $t_h \in \langle t_{k-1}\rangle$ (Fig. 1(d)). Generalizing, in this framework the joint transition probability (4) can be fragmented in the following form:

$$\begin{aligned}
&P(S_{\langle t_k\rangle}|S_{\langle t_h\rangle}, S'_{\langle t_h\rangle}) \cdot \\
&P(S_{\langle t_h\rangle}|S_{\langle t_{h-1}\rangle}, S'_{\langle t_{h-1}\rangle})P(S'_{\langle t_h\rangle}|S_{\langle t_{h-1}\rangle}, S'_{\langle t_{h-1}\rangle}) \\
&\vdots \\
&P(S_{\langle t_{h-n+1}\rangle}|S_{\langle t_{h-n}\rangle}, S'_{\langle t_{h-n}\rangle})P(S'_{\langle t_{h-n+1}\rangle}|S_{\langle t_{h-n}\rangle}, S'_{\langle t_{h-n}\rangle})
\end{aligned} \tag{5}$$

where $\sum_{i=1}^{n}\langle t_{h-i}\rangle = \langle t_{k-1}\rangle$; it is worth to notice that the time indexes are equal for both the processes, being them common (See Fig. 1(d), and Fig. 2(a), 2(b)).
The effects of this factorization are: 1) an enrichment of the state space, due to the fact that we break each $S_{\langle t_k\rangle} = < S_{t_k}, D_{t_k} >$, forming generalized states with the same state label but different smaller durations. Therefore, if an original IP had N states we reach a state cardinality of $\tilde{M} > N$; 2) the complexity of the transition matrix goes down to

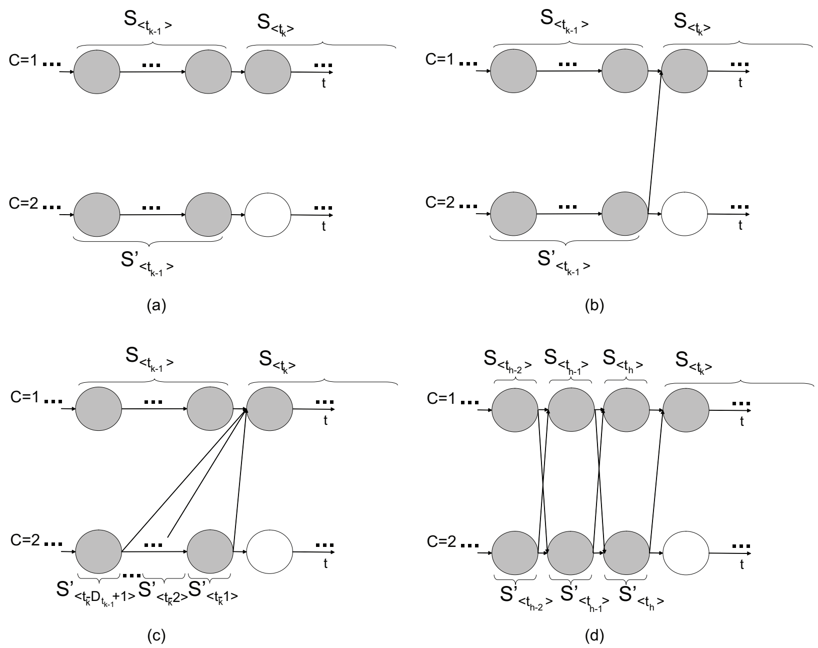

Fig. 1. Coupled semi Markov processes: (a) the states considered in Eq.3 are here perfectly synchronized; (b) the perfect synchronization permits to factorize into transition probabilities terms as written in Eq.4; (c) worst synchronization: the lower chain performs state transitions along all the time interval $\langle t_{k-1} \rangle$; (d) factorization obtained using the instants of implied transition: at each IP transition, every other IP performs also a state transition.

$O(\tilde{M}^C)$ instead of $O((N^C)^L)$. Nevertheless, the space complexity $O(\tilde{M}^C)$ results to be *unnecessarily* high, because several states segments with the same label but slightly different durations are present (see Fig. 2(a)).

Consequently, we decide to perform Gaussian clustering of all the existing durations (Fig. 2(c); in this way, starting from a set of fragmented states $\{< i, d >\}$ (indicating a possible pair of state label, state duration values) , we obtain a novel set of generalized (hidden) states that we call $\{< i, \tilde{d} >\}$, where $\tilde{d} \sim \mathcal{N}(\mu_w, \sigma_w)$ $w = 1 \ldots, W$, where W indicates the number of temporal clusters individuated. Therefore, the states of an IP are no more directly observable, being now the duration of a generalized state a quantity affected by uncertainty. Therefore, once we observe a sequence of state labels S_{t_k}, which is d elements long, we suppose it has been generated by a *hidden duration state* $< S_{t_k}, \tilde{D}_{t_k} >= \tilde{S}_{\langle t_k \rangle}$.
We call the model that permits to deal with such a modified IP as Hidden Duration Semi Markov Model (HD-SMM). This is a hybrid HMM, where the observed state label S_{t_k} indicates a state label which associated duration is modelled by a particular hidden state, that in this case we model with a Gaussian function (various forms of du-

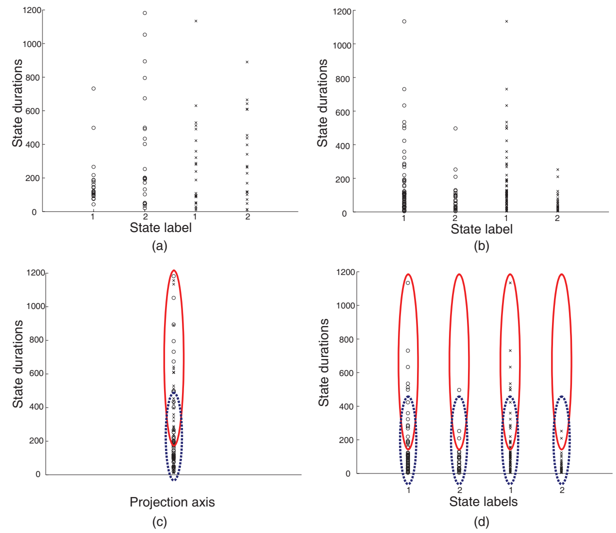

Fig. 2. Clustering of the durations, example with two (2-state labels) processes: (a) the first two columns represent the state labels of the first process (called 1 and 2) with the respective durations, before the time fragmentation. The third and fourth columns are related to the second process; (b) duration fragmentation: the number of smaller durations has grown; (c) projection on the temporal axis and clustering of the durations; (d) re-projection of the duration clusters on the state label axes, to ease the understanding.

rations are present in literature, starting from discrete mass functions, to Gaussian distributions, until the most recent Coaxian functions (Duong, Bui, Phung, and Venkatesh 2005). Here we use Gaussians as first explorative test, obtaining encouraging results as presented in Sec.4). The HD-SMM is indicated with $\boldsymbol{\lambda}^{\text{HD-SMM}} = \{\mathbf{A}_d, \mathbf{B}_d, \boldsymbol{\pi}\}$, where $\mathbf{A}_d$ indicates a $NW \times NW$ transition matrix, where N is the number of visible state labels, and W the number of temporal clusters individuated; the matrix $\mathbf{B}_d$ contains the Gaussian parameters that individuate different state durations. The analogy with the ordinary HMM machinery permit us to inherit all the classical inference and learning formulae.

Finally, connections among different HD-SMM lead to the Coupled Hidden Duration Semi Markov Model (CHD-SMM). In specific, HD-SMM are coupled together as occurs for normal HMMs, employing the coupling mechanism proposed in (Zhong 2002), that expresses the joint transition probability as convex combination of “IP to IP” influences coefficients. Rewriting the first term of Eq.5 we obtain:

$$\begin{aligned} &P(S_{\langle t_k \rangle} | S_{\langle t_{k-1} \rangle}, S'_{\langle t_{k-1} \rangle}) \\ &= \theta_{11} P(S_{\langle t_k \rangle} | S_{\langle t_{k-1} \rangle}) + \theta_{12} P(S_{\langle t_k \rangle} | S'_{\langle t_{k-1} \rangle}) \end{aligned} \tag{6}$$

and the same applies for the other elements of Eq.5. In conclusion, we remark that the learning of the CHD-SMM is performed in the same way as suggested in (Zhong 2002) for the Distance Coupled Hidden Markov Models: the training procedure is a constrained optimization process able to calculate separately transition parameters, influence factors and prior probabilities using Mean Field approximation, not reported here because deeply addressed in literature.

4 Experiments and discussion

4.1 A synthetic example

To evaluate the effectiveness of our model, we first show results on synthetic data. Data was generated by sampling 3 interacting 3-state semi Markov models, namely $c = 1, 2, 3$, whose states are named ${}^{(c)}S_{\langle t_k \rangle}$, with state label $S_{t_k} = 1, 2, 3$. The interaction has been exploited as follows: the process $c = 1$ is the "Leader" that evolves independently from the other processes with a flat transition table. All the state durations of the several models are modelled by different Gaussian pdfs, the mean values ranging from 10 to 70, with common standard deviation $\sigma = 2$.

The states of the leader are the *important* states, i.e. states that the other two processes, $c = 2, 3$, want to copy. When a transition of the leader occurs, a random time interval is extracted from an arbitrary Gaussian distribution with $\mathcal{N}(\mu_{\mathrm{r}} = 2, \sigma_{\mathrm{r}} = 0.4)$, that simulates a "reaction time" needed by processes 2 and 3 to notice the occurred change. After that, the 2 processes assume the same state of the leader. The length of the state sequences is 5000. The training stage of the CHD-SMM converged after 20-30 iterations.
Every state $S_{\langle t_k \rangle}$ of the semi Markov IPs, after the fragmentation caused by the presence of implied transitions and the subsequent clustering operation, turns into W new states, each one formed by the same state label S_{t_k}, and duration $D_{t_{k,w}},\ w = 1, \ldots, W$ modeled by a different Gaussian distribution. In this experiment we chose $W = 2$ in order to represent long and short state durations, obtaining $D_{k,1} = \mathcal{N}(4.7868, 1)$ and $D_{k,2} = \mathcal{N}(70.0996, 10)$. The choice of dividing each state sequence using exactly 2 states is motivated by the intuitive need to model "short" and "long" state durations. Choosing a larger number of states (up to 4), does not change the quality of the results.

The transition matrix and the influence matrix obtained after the training are meaningful, in the sense that they mirror precisely the process modeled. In the following (Fig. 3), we show some interesting intra and inter-chain transition matrices, where ${}^{ij}A$ indicates the transition state conditional probabilities that chain C_i exerts on C_j. The matrices have to be observed considering couples of successive rows: the first couple indicates the probabilities to depart from the first state, considered in its short (first row) and long (second row) duration, and so on for the other rows.

As one can notice (see coefficients in bold), matrix ${}^{11}A$ depicts correctly the modelled behavior of the leader. A short duration state, triggering the transition of the other two chains, is followed with high certainty by the same state, but with longer duration. After that, the choice of the next state is equally allocated to states two and three. Similar considerations can be done for the other two states.

The second matrix, ${}^{12}A$, expresses the high inter-chain dependency that the leader chain exerts on process no. 2, for both the durations of the states. Finally, matrix ${}^{21}A$

$$\begin{bmatrix} 0.01 & \mathbf{0.98} & 0.01 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{0.51} & 0.01 & \mathbf{0.47} & 0.01 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ \mathbf{0.43} & 0.03 & 0 & 0 & \mathbf{0.52} & 0.02 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ \mathbf{0.50} & 0.02 & \mathbf{0.46} & 0.02 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0.01 & \mathbf{0.98} & 0 & 0.01 & 0 & 0 \\ \mathbf{0.98} & 0 & 0 & 0.01 & 0 & 0.01 \\ 0 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 0.02 & \mathbf{0.95} & 0 & 0 & 0.03 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0.02 & 0 & 0.03 & \mathbf{0.95} & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0.01 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0.01 & 0.5 & 0.01 \\ 0 & 0.45 & 0 & 0 & 0 & 0.5 \\ 0.42 & 0.02 & 0 & 0 & 0.52 & 0.02 \\ 0.01 & 0.5 & 0 & 0.48 & 0 & 0 \\ 0.5 & 0.01 & 0.46 & 0.02 & 0 & 0 \end{bmatrix}$$

^{11}A $\quad$ ^{12}A $\quad$ ^{21}A

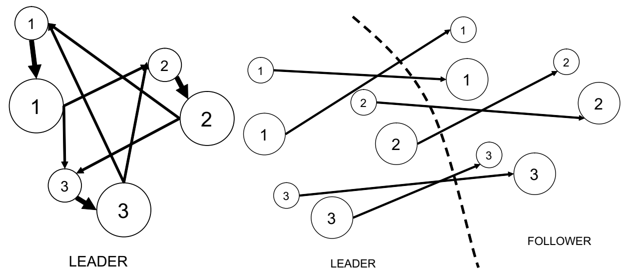

Fig. 3. Synthetical example: some learned transition matrices. In the matrix ^{11}A, the dashed boxes indicate the transition probabilities *from* the state label S_1 of the chain $c = 1$, both in the short-duration exemplar (first row) and in the longer one (second row). Below in the same column, the graphical depiction of ^{11}A is proposed, where the little (big) circles indicates the states with label i=1,2,3 and short (long) durations. Thick arrows represent high probabilities; some unimportant arrows are not reported for clarity. In ^{12}A the boxes highlight the strong conditional probabilities of each state of the leader on the follower. Below in the same column, the graphical depiction of ^{12}A is proposed; some unimportant arrows are not reported for clarity. The last matrix (^{21}A) shows that the leader does not follow the states assumed by the followers at all. The same results apply on Follower 3.

shows that the states of the follower are not able to probabilistically determine the states of the leader chain.
The obtained influence matrix is

$$\Theta_{\text{CHD-SMM}} = \begin{bmatrix} 1 & 0 & 0 \\ 0.99 & 0.008 & 0.002 \\ 0.99 & 0.001 & 0.009 \end{bmatrix} \tag{7}$$

that shows clearly the influence among the chains, where the coefficients θ_{ij} indicate the influence of the chain j over the chain i, with θ_{11} the auto-influence of the leader process on itself, and so on.

In order to compare our approach with the state of the art, we learn an Influence Model (IM) using directly the state sequence as if each state sequence would have been generated from Markov processes. This means that we have for each chain exactly three states, corresponding to the three different state labels and resulting in 3×3 transition matrices (In this case, we have exactly the same example as the one proposed in Basu et al. 2001). After the training, the resulting parameters are qualitatively different. As one can expect, the auto-transition probabilities dominates over the intra chain matrix, producing also high dependencies of the leader chain with respect to the other two processes. This can globally be visible in the influence matrix, that in this example assumes the values

$$\Theta_{\mathrm{IM}} = \begin{bmatrix} 0.83 & 0.04 & 0.13 \\ 0.22 & 0.78 & 0 \\ 0.23 & 0 & 0.77 \end{bmatrix} \tag{8}$$

in which the auto influence dominates and a comparable influence is present among different IPs, without an emerging leader.

4.2 Conversational style classification

In this experiment, we show the ability of the CHD-SMM to model and classify different styles of conversations. We have 2 subjects, A and B, involved in several conversation sessions that hold for around 5 minutes each. Our class-library is formed by different conversation moods, labeled as follows:
1) *Flat discussion*: A and B are discussing calmly together.
2) *Successful interrogation*: A asks questions, and B answers promptly.
3) *Unsuccessful interrogation*: A asks questions but B does not respond promptly, producing longer periods of silence before (eventually) answering.
4) *Fight*: A and B are arguing.
The training data set is formed by 20 conversations sessions, for each conversation mood. Each session lasts 5 minutes circa, and is performed by actors. We ask them to improvise the fourth situation written above. To validate the semantic content of the audio sequences, we ask 10 testers to manually classify all the sequences, using the labels listed above. Subsequently, we perform a voice-unvoice test on the sample sequences, in order to obtain data-sets of voice/silence values. Therefore, the mood classes are modeled by CHD-SMM trained with the labeled data-sets. After the fragmentation of the data set, we perform Gaussian clustering using 2 clusters, in order to model "long" and "short" segments of state labels, for each state label. The learning time has been 10 sec. for each sequence, with 20-30 iterations before to converge.
In order to get better insight into the proposed method, we compare the CHD-SMM with the Influence Model, using as training sequences the same used for the CHD-SMM. Moreover, we explicitly model the turn taking dynamics of the conversation, by transforming the original training sequences. In practice, we reduce the data set, by maintaining for all the IPs only the states labels occurring in the instant of implied transition. With such data, we train an Influence Model similar to that proposed in (Basu et al. 2001), that we call Turn Taking Influence Model (TTIM).

The classification is performed using a Maximum Likelihood classifier, and the classification accuracy has been estimated using the Leave One Out (LOO) scheme (Duda, Hart and Stork 2001). The results of the classification are visible in Tab. 1. One can notice that the classification results based on the Influence Model show that, even with not-so-high accuracy, the conversation style "Successful Interrogation" (situation 2) and "Unsuccessful Interrogation" (situation 3) are better classified, compared to the other two situations. This is due to the different auto-transition probabilities (relative to the long silence periods for situation 3 and long speech periods for situation 2),

Test	IM	TTIM	CHD-SMM
1)	76%	85%	91%
2)	85%	65%	98%
3)	84%	70%	93%
4)	75%	90%	92%

Table 1. LOO classification accuracies for the four different problems.

which likely help to get a good discrimination. Anyway, the strong auto-transition probabilities overwhelm the transition probabilities among different states of the other two conversation situations 1 and 4, producing low classification accuracies. This is actually a problem because auto-transitions smooth away the turn-taking dynamics that strongly characterizes a conversation style. This effect is also visible by observing the resulting influence matrix of each sequence, mainly exhibiting auto-influences.

In the TTIM based classification, situations 2 and 3 show low accuracy, principally because the turn taking data does not represent any modeling of duration.

Globally, both the TTIM and the IM schemes exhibit worse performances with respect to our approach, that instead is able to model both a rough idea of duration of the states and exploiting well the inter-chain influences.

One drawback of the proposed method is that the exact duration of the states of each single IP is lost, due to the implied fragmentation of the state segments. In any case, the good classification performances encourage us to perform further testing of the model.

References

S. Basu, T. Choudhury and A. Pentland, "Towards measuring human interactions in conversational settings," in *IEEE Int'l Workshop on Cues in Communication (CUES 2001)*, Hawaii, CA, 2001.

M. Brand, N. Oliver, and S. Pentland, "Coupled hidden markov models for complex action recognition," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1997.

R. Duda, P. Hart, and D. Stork, *Pattern Classification*. John Wiley and Sons, 2001.

T. Duong, H. Bui, D. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-markov model," in *CVPR (1)*, 2005, pp. 838–845.

M. I. Jordan, *Learning in graphical models*. MIT Press, 1999.

M. I. Jordan, Z. Ghahramani, T. Jaakkola, and L. Saul, "An introduction to variational methods for graphical models," *Machine Learning*, vol. 37, no. 2, 1999, pp. 183–233.

K. Murphy, "Hidden semi-markov models (hsmm's)," November 2002, available at `www.ai.mit.edu/~murphyk`.

S. Hongeng and R. Nevatia, "Large-scale event detection using semi-hidden markov models," in *Proceedings of the IEEE International Conference on Computer Vision*, Nice, France, October 2003, pp. 1–8.

L. Rabiner, "A tutorial on Hidden Markov Models and selected applications in speech recognition," *Proc. of IEEE*, vol. 77, no. 2, 1989, pp. 257–286.

S. Zhong and J. Ghosh, "Hmms and coupled hmms for multi-channel eeg classification," in *Proc. IEEE Int. Joint Conf. on Neural Networks*, 2002, pp. 1154–1159.