A predictive model for human activity recognition by observing actions and context

D.G. Romero[1]; A. Frizera[2]; A. D. Sappa[1,3]; B.X. Vintimilla[1]; T.F. Bastos[2]

[1]Escuela Superior Politécnica del Litoral, ESPOL,
Facultad de Ingeniería en Electricidad y Computación,
Campus Gustavo Galindo Km 30.5 Vía Perimetral, P.O. Box 09-01-5863,
Guayaquil, Ecuador
[2]Universidade Federal do Espírito Santo, Av. Fernando Ferrari 514, Campus
Goiabeiras, Vitória, ES, Brasil
[3]Computer Vision Center, Edifici O, Campus UAB,
08193 Bellaterra, Barcelona, Spain
dgromero@espol.edu.ec, anselmo@ele.ufes.br,
boris.vintimilla@espol.edu.ec, assapa@cvc.uab.es,tfbastos@ele.ufes.br

Abstract. This paper presents a novel model to estimate human activities — *a human activity is defined by a set of human actions*. The proposed approach is based on the usage of Recurrent Neural Networks (RNN) and Bayesian inference through the continuous monitoring of human actions and its surrounding environment. In the current work human activities are inferred considering not only visual analysis but also additional resources; external sources of information, such as context information, are incorporated to contribute to the activity estimation. The novelty of the proposed approach lies in the way the information is encoded, so that it can be later associated according to a predefined semantic structure. Hence, a pattern representing a given activity can be defined by a set of actions, plus contextual information or other kind of information that could be relevant to describe the activity. Experimental results with real data are provided showing the validity of the proposed approach.

1 Introduction

Human-Machine interaction is an attractive research field in the human assistance domain. One of the main goals is to develop methods for allowing automatic human activities recognition in order to guarantee an easy communication. Actually, the automatic human activity recognition is not just needed for human-machine interaction, but also for tasks related to human behavior understanding in the health context; suspicious behavior detection in the context of video surveillance; or activity registration in the context of pattern recognition can be benefited from it. This ever growing research field finds also applications in the assistance to elderly persons, where a continuous activity recognition is needed to help elder people in their everyday life [1].

The automatic recognition of human activities, expressions and intentions are multimodal problems; hence, the recognition process may use different sources of information to improve their results. Examples of such multimodal sources of information are facial or bodily expressions, voice or audio-visual signals, among other. Different studies has been published exploiting such a multimodality nature; for instance [2] introduce a learning framework to identify emotions from speech using a discriminant function based on Gaussian mixture models. The authors in [3] propose a multimodal approach for recognizing emotions from an ensemble of features. It involves face detection, followed by key-point identification and feature generation, which are finally used by the emotion recognition. Finally, in [4] the cross-modality information is explored in an audiovisual emotion recognition system. The authors study the relationship between acoustic features of the speaker and facial expressions of the interlocutor during dyadic interactions.

Due to the problem complexity, different taxonomies has been proposed in the literature related with motion recognition. One of the most widely accepted has been presented in [5]; it is also considered in the current work and consists of the following categories: *primitive action*, *action* and *activity*. A primitive action is an atomic movement that can be described at the level of the human body parts. An action consists of a set of primitive actions and describes a complex movement or full corporal expression. Finally, an activity consists of a set of consecutive actions performed by the person. Examples of the previous definitions are as follow: *i*) "putting the right arm in fron" is a primitive action; *ii*) "moving an object" is an action; *iii*) "playing chess" is an activity that involves primitive action, movement of objects and other actions related to the rules of the game.

During the last five decades different studies have been carried out to tackle problems related with the human activities inference. The main goal and motivation of current research in this field is to reach a human-machine interaction similar to human-human interaction. An interesting review can be found in [6]. Several approaches has been proposed to solve problems related with the activity understanding of human everyday life. For instance, in the pedestrian detection domain it is possible to empirically identify safe/unsafe activities of a person crossing a street. In this case, an automatic system can predict risky situations due to the recklessness of a pedestrian. In order to develop applications like the previous one, or some other that involves understanding of human actions, it is necessary to develop models. These models should cover the full description of these actions, considering the different variables that affect them (object, culture, environment, among other) together with their temporal and non-temporal surrounding framework. The current paper proposes a novel model to recognize human activities. It is based on RNN and Bayesian inference considering human movements together with contextual information. The manuscript is organized as follow: The proposed model is presented in Section 2. Experimental results and comparisons are provided in Section 3. Finally, conclusions are given in Section 4.

2 Proposed Approach

In the proposed approach the actions or human bodily expressions are considered as input data. They are obtained using a computer vision system like the one presented in [7]. Therefore, each identified action or human bodily expression are send to the inference engine; this inference engine is responsible for encoding, tracking and activity estimation according to an *actions-context-activity* association previously defined. The idea behind the proposed approach lies on the observation of human movements, which were already identified, together with the context information in order to predict possible interactions of the subject. In a general way, the proposed inference model can be represented as shown in Fig. 1.

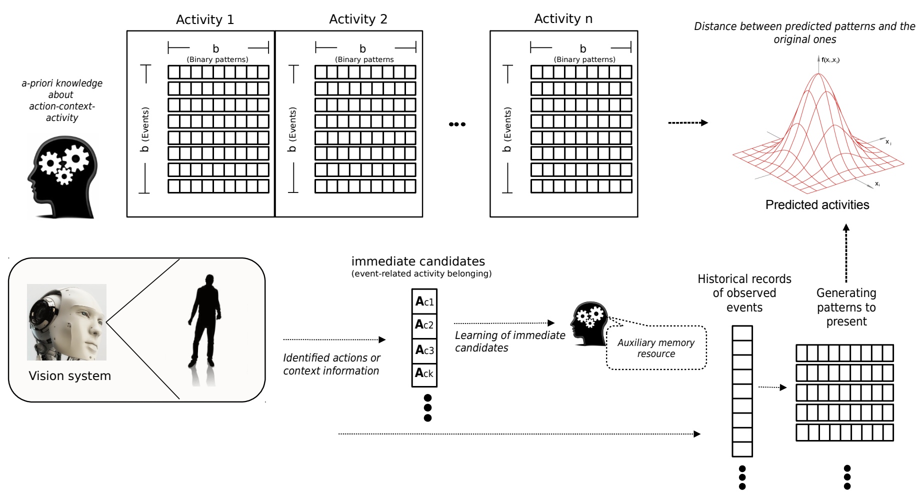


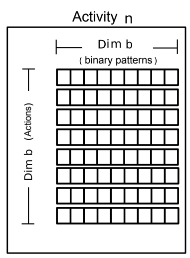
Fig. 1. General illustration of the proposed approach.

Human bodily expressions are important elements in the observation of human behavior. These expressions correspond to both movements of the body limbs (global level analysis) and facial expressions (detailed analysis). In order to reach the final goal, the information is encoded in such a way that it can be associated, later on, according to a predefined semantic structure. Hence, an encoded pattern representing an activity can be defined by a set of actions, contextual information or other kind of information that could be relevant to describe the given activity. Next, the three main elements of the proposed approach are described. First, the methodology to generate patterns that encodes the information is introduced in Section 2.1. Then, the inference engine based on recurrent neural network is presented in Section 2.2. Finally, the architecture used for tracking identified events is summarized in Section 2.3.

2.1 *Actions-Activity-Context* Patterns

One of the first problems to resolve after recognize actions and context is to abstract their representation. In the current work it's proposed a model that uses RNN as resources of *Associative Memory*, considering binary patterns and two *Hopfield* Neural Networks [8] [9] in a long and short term approach.

In this sense, let $Ac = \{ac_1^{(b)}, \ldots, ac_k^{(b)}\}$ be a set of actions, where k is the number of actions or context information that can be identified by the recognition system, and b is the set of bits, representing an associated event. Assuming a *Hopfield* recurrent neural network is used, then, the total number of neurons in the network will be $I = b^2$. In other words, an activity can be represented by a matrix $x_n^{(b \times b)}$, where $x_n \subset Ac$ and $n = \{1, 2, \ldots \lfloor 0.138I \rfloor\}$ (0.138 represents the theoretical storage capacity as will be defined in eq (3)). Figure 2($left$) illustrates the structure of a pattern of activity.



P1	e	e	f	9	c	f	9	f
P2	f	9	c	e	9	8	9	7
P3	6	c	6	f	9	e	9	9
P4	f	6	f	9	8	7	9	e
P5	e	9	f	f	f	5	9	9

Fig. 2. Activity pattern: ($left$) activity pattern structure; ($right$) activity patterns (P_i) with their representative events associated to them.

As it was mentioned in the introduction, one of the goals of current approach is to support the recognition of unsorted events. The previous statement means that it could exist a semantic similarity between a set of events, which have been observed in a different order. Additionally, this provides to the system the capability to estimate an activity even in those cases where the observed events are related with some other activity already known.

In this context, the neural network takes into account representative observations within a given activity, and then the probability of the candidate activities could change as soon as a new event occurs. In the "retrieval" phase an observation will induce to a known state, converging in each new observation to some trained pattern. The resulting patterns could have variations that can be measured using the *Hamming distance* [11] at each new observation presented to the network.

In the current work, the inputs and outputs of the neural network correspond to binary patterns ($x_n^{(I)}$), which represent a combination of observable events (each event is encoded using 8 bits). Initially, these combinations are randomly defined using *Linux Pseudo-Random Number Generator* (LRNG) [12] [13]. The initial combinations, randomly generated, are used to define a set of observable events k, without semantic relationships between them. Hence, in order to have a balance between the resulting patterns (*unbias*), they are discriminated so that $P(x_i(j) = 1) = P(x_i(j) = -1) = 1/2$, where $j = \{1, 2, \ldots, I\}$.

Unfortunately, the process presented above for the random generation of patterns and posterior discrimination, do not provide enough information regarding the correlation between them. For this reason, each activity (semantic relationship between events) will be assigned to the less correlated patterns, which are selected by means of a *cross-correlation* criterion based on the *Pearson's* correlation (PPMCC or PCC) [14, 15]. The *Pearson's* correlation coefficient (r_{xy}) is defined as the covariance of two variables divided by the product of their standard deviations; it is expressed as:

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i y_i}{\sqrt{n \sum x_i^2 - (\sum x_i)^2} \sqrt{n \sum y_i^2 - (\sum y_i)^2}}. \tag{1}$$

The correlation coefficient spans $[-1, 1]$, where 1 means that the linear equation perfectly describes the relation between X and Y, with all the data lying on a straight line where X and Y increase. On the contrary, -1 means that all the data lie on the straight line where X and Y decrease. Hence, the 0 value represents the lack of linear correlation between the variables. Figure 2(*right*) shows five activity patterns with their corresponding associated events (i.e., actions or context), where each alphanumeric character represents a binary code according with the 8-bits ASCII table.

Once all the patterns have been defined, a semantic annotation is given to them according to a specific application (e.g., events representing actions or context information and patterns representing activities). Table 1 shows the results of the selection process based on the *Pearson's* correlation.

Table 1. Cross-correlation using the Pearson's correlation.

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
P1	1	0.25	0.188	0.5	0.312	0.5	0.375	0.562	0.312	0.438
P2	0.25	1	0.375	0.438	0.438	0.125	0.562	0.312	0.25	0.5
P3	0.188	0.375	1	0.312	0.438	0.438	0.438	0.25	0.562	0.125
P4	0.5	0.438	0.312	1	0.312	0.438	0.375	0.5	0.562	0.438
P5	0.312	0.438	0.438	0.312	1	0.375	0.625	0.188	0.188	0.562
P6	0.5	0.125	0.438	0.438	0.375	1	0.312	0.312	0.688	0.188
P7	0.375	0.562	0.438	0.375	0.625	0.312	1	0.062	0.25	0.5
P8	0.562	0.312	0.25	0.5	0.188	0.312	0.062	1	0.5	0.375
P9	0.312	0.25	0.562	0.562	0.188	0.688	0.25	0.5	1	0.062
P10	0.438	0.5	0.125	0.438	0.562	0.188	0.5	0.375	0.062	1

2.2 Recurrent Neural Network based approach

The *Hopfield* neural network has two main applications. In the first one they are used as *associative memories*, while in the second one they are used to solve *optimization problems*. On the other hand, the inference process in the context of recurrent neural networks could be faced in two ways: the first one is by considering the internal probabilistic constitution of the neural network, while the second one, is focused on the probabilistic meaning of the results, being the latter one the approach considered in the current work.

According to [16], there is a difference between the theoretical storage capacity and the experimental one, both using the Hebbian learning rule as well as the pseudo-inverse rule (also referred to as projection learning rule) [17]; the main reasons for these differences [9] are:

- The high correlation between training patterns; the correlation between patterns reduces the performance of the network.
- The storage capacity of Hopfield networks is related to the sparsity of training patterns.
- The *global inibition* is another factor that can affect the storage capacity.

The number of patterns to be stored can be increased if a small error is accepted. In the current work a storage of $N = 8$ patterns, of $I = 64$ bits, is considered. Hence, the theoretical error due to bit swapping in the first iteration is about 0.2%; this error will be larger if next iterations are considered.

Regarding the rate between the number of stored patterns and the number of neurons N/I, [18] states that the maximum number of patters to be stored, with an acceptable maximum error value, is defined by the following expression (this statement is shared by [19]):

$$N_{max} \simeq \frac{I}{4lnI + 2ln(1/\epsilon)}, \tag{2}$$

actually, in [19] a critical limit is presented, which considers the abrupt drop in the performance of the network with the increase of the ratio N/I; this critical limit is defined as:

$$N_{crit} = 0.138I. \tag{3}$$

As presented above, the balanced combination of each pattern and the selection of those with the lowest correlation are intended to reach a fair initial distribution of the probabilities of each state. Later on, during the training of the network, these probabilities are biased toward the most active state. In other words, if in the activity "$activ_1$" the event "a_1" is performed more than one time, this event will have a higher weight in the given activity, which will be reflected in the neural network. In the same way, activities that have been trained using the proposed model, will be tolerant to the presence of events that may belong to other activities, or events identified by the system (computer vision or other systems).

2.3 Tracking of Identified Events

Initially, all the set of actions-activity-context patterns (x_n) are stored in the main associative memory block (long-term), considering the limit of n defined in the previous section (see Fig. 1). Then, as soon as a new event is identified, the second block of memory (shot-term) stores all the patterns x_n that contain a new action, these selected patterns are referred immediate candidates

$$X_{sel} = \{\forall x_n \in X | ac_k \in x_n, n = \{1, 2, \ldots, \lfloor 0.138I \rfloor\}\}. \tag{4}$$

The new set of patterns X_{sel} ($X_{sel} \subset X$) is temporally stored, so that the recurrent network responsible of store these patterns will behave like a short-term memory. In this way, the patterns in time $t-1$ could be "forgotten" when a new event appears (Fig. 3).

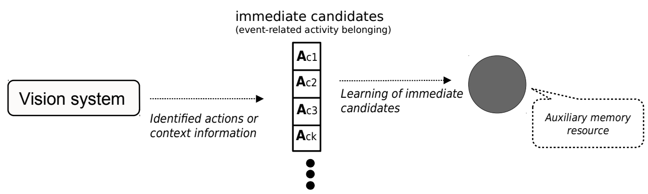


Fig. 3. Temporal storing of immediate candidates.

Considering the constraints in the storage capacity of activity patterns, the probability of a candidate activity being one of the learned activities, given the appearance of a new action, can be determined using the Bayes theorem (assuming that $ac_k \in X$):

$$P(x_n | ac_k) = \frac{P(x_n) P(ac_k | x_n)}{P(ac_k)}. \tag{5}$$

The aim of the short-term resource is to provide the system a knowledge base to be used during the appearance of the first actions.

After storing (memorizing) the immediate candidates, the observed event is kept in a historical register $hist(ac_j)$; for each new observed event, a new set of patterns H to be presented to the network is created from $hist$, fulfilling $x_n(i) = H(ac_j)$, for $\forall i \in \Re| \leq b$ y $\forall j \in \Re | j \leq k$.

In this way, at least one of the patterns from H will be an *incomplete* and/or partially wrong version of x_n. Hence, for every new observation, considering an 64-bits activity pattern (8 observable events of 8 bits), it will exist a pattern to be presented to the network with an overlap of at least 1/8 with some of the patterns initially learned. Once the neural network returns the estimated activity (with an acceptance rate based on the mean and variance of Hamming distance values between the stored patterns in long-term memory x_n and those retrieved from it x_{des}), the total content in the historical register $hist$ is deleted; starting again with a new and continuous loop of analysis. In a particular situation, if the information presented to the long-term memory is not enough, the patterns in H will be presented to the short-term memory, the probability of retrieve a correct patterns could be obtained from the equation (5) when $ac_k \in X$. Although the short term memory tends to give a strong conclusion (due to the reduced number of stored patterns), it is not reliable. For instance, it could give an answer even in the case when a observed event doesn't belong to some learned patterns, in this particular case, the obtained conclusion would be wrong (false positive).

The proposed model allows the estimation of an activity from $1/b$ of the actions that define it, considering the theoretical limit of storage capacity of Hopfield neural networks (assuming a correlation between patterns ($r_{x,y} \to 0$)), fixed as $N_{max} = \lfloor 0.138I \rfloor$.

3 Experimental Results

The proposed approach has been evaluated with the *Hollywood* data set [22]. It contains annotations of actions performed in scenes of different films. Figure 4 shows some frames of this data set; they correspond to the following actions: "AnswerPhone"; "GetOutCar" and "Kiss".

Fig. 4. Frames from the Hollywood data set, corresponding to the following actions: ($left$) AnswerPhone; ($middle$) GetOutCar; ($right$) Kiss.

Table 2 shows the annotations of identified actions in different time intervals for different films (they are provided by [22] and are used as ground truth). In the current work these annotations have been used to estimate the titles of the video sequences during the actions observation.

Table 2. Annotated actions, provided by [22], of different films.

Video Sequence	Time	Annotation	Video Sequence	Time	Annotation
"As Good As It Gets – 00259.avi"	(10-125)	$< SitDown >$	"LOR-FellowshipOfTheRing – 01181.avi"	(10-96)	$< StandUp >$
"As Good As It Gets – 01311.avi"	(1-180)	$< StandUp >$	"LOR-FellowshipOfTheRing – 01181.avi"	(1796-1875)	$< SitUp >$
"As Good As It Gets – 01400.avi"	(620-730)	$< Kiss >$	"LOR-FellowshipOfTheRing – 01286.avi"	(1-80)	$< SitUp >$
"As Good As It Gets – 01619.avi"	(1-50)	$< Kiss >$	"LOR-FellowshipOfTheRing – 01494.avi"	(25-232)	$< SitUp >$
"As Good As It Gets – 01834.avi"	(150-450)	$< SitDown >$	"LOR-FellowshipOfTheRing – 01712.avi"	(246-372	$< Kiss >$
"As Good As It Gets – 01935.avi"	(1-209)	$< StandUp >$	"LOR-FellowshipOfTheRing – 02501.avi"	(65-93)	$< SitUp >$
"As Good As It Gets – 02002.avi"	(1-664)	$< Kiss >$	"LOR-FellowshipOfTheRing – 02707.avi"	(1-410)	$< HugPerson >$
"Erin Brockovich – 00816.avi"	(1-50)	$< StandUp >$	"Dead Poets Society – 00068.avi"	(18-55)	$< SitDown >$
"Erin Brockovich – 01233.avi"	(450-530)	$< StandUp >$	"Dead Poets Society – 00148.avi"	(163-233)	$< HandShake >$
"Erin Brockovich – 01768.avi"	(43-141)	$< AnswerPhone >$	"Dead Poets Society – 00205.avi"	(1-89)	$< HandShake >$
"Erin Brockovich – 01916.avi"	(1-91)	$< StandUp >$	"Dead Poets Society – 01587.avi"	(1-377)	$< SitDown >$
"Erin Brockovich – 02110.avi"	(1-180)	$< SitDown >$	"Dead Poets Society – 01587.avi"	(1-377)	$< Kiss >$
"Erin Brockovich – 02137.avi"	(1-180)	$< StandUp >$	"Dead Poets Society – 01741.avi"	(144-195)	$< AnswerPhone >$
"Erin Brockovich – 02262.avi"	(130-241)	$< SitDown >$	"Dead Poets Society – 02590.avi"	(148-230)	$< SitUp >$
"Pianist, The – 01525.avi"	(1-141)	$< HandShake >$			
"Pianist, The – 01285.avi"	(1-150)	$< StandUp >$	"Pianist, The – 01285.avi"	(360-445)	$< GetOutCar >$
"Pianist, The – 00926.avi"	(100-367)	$< HugPerson >$	"Pianist, The – 01255.avi"	(450-716)	$< HugPerson >$
"Pianist, The – 01334.avi"	(350-530)	$< SitDown >$	"Pianist, The – 01285.avi"	(925-995)	$< HandShake >$

Each actions have been encoded and selected according to the lowest *Pearson's* correlation. Table 3 shows patterns (in ASCII format) corresponding to the annotated actions according to the video sequences presented in Table 2. Each set of patterns defines an activity representing the title of the film. The annotated actions corresponds to: f=$< SitDown >$; Y=$< StandUP >$; R=$< Kiss >$; K=$< AnswerPhone >$; q=$< HandShake >$; $|$=$< SitUP >$ and g=$< HugPerson >$. The Figure 5 depicts the codified patterns corresponding to the defined activities in Table 3.

Table 3. Titles (*activities*) in *Hollywood* data set [22] with their corresponding codified annotations (*actions*).

Title	**Codified Annotations**
As Good As It Gets	$fYRRfYR$
Erin Brockovich	$YYKYfYf$
Pianist, The	$ggYXqfq$
LOR-FellowshipOfTheRing	$Y\vert\vert\vert R\vert g$
Dead Poets Society	$fqqYRK\vert$

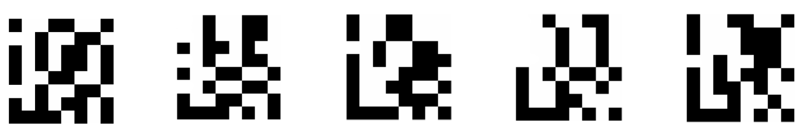


Fig. 5. Activities defined in Table 3 (left to right corresponds to films from top to bottom).

As soon as a new action is observed, the different stages of the proposed model are performed: selection of possible activities, generation of candidate patterns, presentation of candidates to the network; and finally, selection of the highest probability answer. Table 4 presents the results obtained with the title of the "As Good As It Gets" film is estimated by the model. Each row in the "Obs. Actions" column corresponds to the set of actions that have been identified in different pieces of the film. The inference process is shown to the extent that actions are observed. It can be appreciated that in the second observation the obtained conclusion corresponds with the activity that has the most representative action (see the activities defined in Table 3), considering all the previous actions associated to the activities. In this example, the correct result can be obtained just when a new action is observed. This fact is also appreciated in Table 5, where an action that do not belong to the activity of interest (AOI) is observed; a similar trend is observed, but in this case due to the previous wrong information, showing the influence of the learning process (associative reinforcement between events and activities) in the results given by the model.

Table 4. Inference of the title of films according to observed actions.

		Films titles			
Obs.	Obs. Actions	Good... (AOI)	Pianist...	Poets...	Erin...
1	f	**35.29%**	–	17.65%	17.65%
2	fY	70.59%	17.65%	17.65%	**82.35%**
3	fYR	**100%**	17.65%	29.41%	82.35%
4	fYRR	**100%**	17.65%	29.41%	82.35%

Table 5. Inference of films titles according to observed actions and by inserting a wrong observation (q).

		Films titles			
Obs.	Obs. Actions	Good... (AOI)	Pianist...	Poets...	Erin...
1	f	**40.0%**	–	30%	30%
2	fq	14.29%	7.14%	**21.43%**	–
3	fqY	**72.22%**	27.78%	44.44%	22.22%
4	fqYR	**100%**	82.35%	58.82%	82.35%

4 Conclusions

The current work presents an approach for events tracking captured by a vision system and other sensors that can provide contextual information towards the automatic inference of human activities. The usage of states in a recurrent neural network for representing human actions and context information has been stated. Additionally, it proposes a cyclic model for tracking actions and inferring activities, which consider predominant actions that influence it. In this way, it was possible to incorporate two common properties present in everyday life: first, the fact that different persons can do different set of actions during the performance of a given activity; second, the inclusion of contextual information, which affect the decision criteria during the automatic inference of activities as an association of movements or corporal expressions. The model proposed for the activity inference has been described supporting the selected options.

The results obtained with the *Hopfield* network opens the possibility to incorporate other recurrent neural networks working as associative memory within the proposed cyclic model. Finally, the state of the art on recurrent neural networks, show the potential of these techniques for applications related with associative memory. The current computational capability represents an appealing factor for implementing techniques as the ones proposed in the current work.

References

1. Zhang, M., Sawchuk, A.A.: Manifold learning and recognition of human activity using body-area sensors. In: Machine Learning and Applications and Workshops (ICMLA). Volume 2. (2011) 7–13
2. Yun, S., Yoo, C.: Loss-scaled large-margin gaussian mixture models for speech emotion classification. Audio, Speech, and Language Processing, IEEE Transactions on **20** (2012) 585–598
3. Tariq, U., Lin, K.H., Li, Z., Zhou, X., Wang, Z., Le, V., Huang, T., Lv, X., Han, T.: Recognizing emotions from an ensemble of features. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on **42** (2012) 1017–1026
4. Mariooryad, S., Busso, C.: Exploring cross-modality affective reactions for audiovisual emotion recognition. Affective Computing, IEEE Transactions on **4** (2013) 183–196
5. Moeslund, T.B., Hilton, A., Kruger, V.: A survey of advances in vision-based human motion capture and analysis. Computer Vision and Image Understanding **104** (2006) 90–126
6. Saygin, A.P., Cicekli, I., Akman, V.: Turing test: 50 years later. Minds and Machines **10** (2000) 12–20
7. López, D.R., Neto, A.F., Bastos, T.F.: Reconocimiento en-línea de acciones humanas basado en patrones de rwe aplicado en ventanas dinámicas de momentos invariantes. Revista Iberoamericana de Automática e Informática Industrial RIAI **11** (2014) 202–211
8. Hopfield, J.J.: Neural networks and physical systems with emergent collective computational abilities. Proceedings of the national academy of sciences **79** (1982) 2554–2558
9. Trappenberg, T.: Fundamentals of computational neuroscience. Oxford University Press (2010)
10. Stroffek, J., Marsalek, P.: Short-term potentiation effect on pattern recall in sparsely coded neural network. Neurocomputing **77** (2012) 108– 113
11. Mandic, D.P., Chambers, J.: Recurrent neural networks for prediction: Learning algorithms, architectures and stability. John Wiley & Sons, Inc. (2001)
12. Gutterman, Z., Pinkas, B., Reinman, T.: Analysis of the linux random number generator. In: Security and Privacy, 2006 IEEE Symposium on, IEEE (2006) 15–pp
13. Vadhan, S.P.: Pseudorandomness. Now (2012)
14. Goh, K.I., Cusick, M.E., Valle, D., Childs, B., Vidal, M., Barabasi, A.L.: The human disease network. Proceedings of the National Academy of Sciences **104** (2007) 8685–8690
15. Nikolić, D., Mureşan, R.C., Feng, W., Singer, W.: Scaled correlation analysis: a better way to compute a cross-correlogram. European Journal of Neuroscience **35** (2012) 742–762
16. Wu, Y., Hu, J., Wu, W., Zhou, Y., Du, K.L.: Storage capacity of the hopfield network associative memory. In: Intelligent Computation Technology and Automation (ICICTA), 2012 Fifth International Conference on. (2012) 330–336
17. Valiant, L.G.: Projection learning. Machine Learning **37** (1999) 115–130
18. MacKay, D.J.: Information theory, inference and learning algorithms. Cambridge university press (2003)
19. McEliece, R., Posner, E.C., Rodemich, E.R., Venkatesh, S.: The capacity of the hopfield associative memory. Information Theory, IEEE Transactions on **33** (1987) 461–482

20. Wegner, P.: A technique for counting ones in a binary computer. Communications of the ACM **3** (1960) 322
21. Yao, Z., Gripon, V., Rabbat, M.G.: A massively parallel associative memory based on sparse neural networks. arXiv preprint arXiv:1303.7032 (2013)
22. Marszalek, M., Laptev, I., Schmid, C.: Actions in context. In: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE (2009) 2929–2936