Comparison of Classifiers for Human Activity Recognition

Óscar Pérez[1], Massimo Piccardi[2], Jesús García[1], and José M. Molina[1]**

[1] Departamento de Informática. Universidad Carlos III de Madrid.
Avenida de la Universidad Carlos III, 22 Colmenarejo 28270. Madrid. Spain.
`oscar.perez.concha@uc3m.es, jgherrer@inf.uc3m.es, molina@ia.uc3m.es`
[2] Faculty of Information Technology University of Technology, Sydney PO Box 123.
Broadway NSW 2007
`massimo@it.uts.edu.au`

Abstract. This work is a comparison of the classification performance in the human activity recognition field. We have selected seven types of classifiers (two Hidden Markov Models (HMM), a J.48 tree, two Bayesian classifiers, a classifier based on rules and a Neuro-Fuzzy system) and a set of manual labeled video sequences with four different human activities (inactive, active, walking and running). The features that identify activity patterns are extracted in a fashion that we only employ motion features based on the 2D centroid coordinates and the height and width of each person's blob. The results show that the classifiers reveals different performances according to the number of features employed and the set of classes to sort. The final conclusion allows us to say that the basic motion features are not enough to have a complete description of the problem and obtain a good classification. Moreover, the hand labelling introduces subjectivity and ill definition that we must take also into account.

1 Introduction

The topic of the human activity recognition is an open and challenging problem to solve by the research community. Usually, the analysis is carried out by extracting motion features [1] with which to compute the motion patterns to recognise a set of activities. For example, the works [2]-[4] show how to compute and then select the most relevant features for the final classification. There are other works that employ non motion features, like for instance the gait energy image [5] or silhouettes.

As it is pointed out in [6], many of the activity recognition works use body components features, like the position of head [7], hands and feet. The reality is that these features may not be found or located in many circumstances.

On the other hand, one of the main issues in this kind of works is the definition of the activities. They can be inferred from clusters of the extracted features [8] or

manually assigning action labels to the video sequences [9]. In this case, we have to take into account the subjectivity and ill-definition of the manual labelling.

Thus, this work proposes the comparison of different classifiers (HMM, J.48, PART, Bayes and Neuro-fuzzy) of activity recognition by using only motion features described and calculated by a sequence of displacements of the 2D centroid and the height and width of each person's blob. The CAVIAR [9] sequences are utilised in this paper to recognize the set of activities corresponding to 'Inactive' (IN), 'Active' (AC), 'Walking' (WK) and 'Running' (R).

The results show a good performance (above the 74% of correct classification) for all classifiers in the case of using only three clearly different classes: IN, WK and R. In the case of utilising the four activities, the classifiers present a low performance, except for the HMM which states have been adjusted by means of a Genetic Algorithm (GA) and the Naive Bayes classifier. The only exception is the Neuro-fuzzy classifier which can not exceed the 68% of correct classification in the best of the cases.

This work presents in Section 2 the selection and computation of the features for the future classification and the initial values of each classifier. Section 3 presents the steps followed to carry out the experiments and the results of them. Finally, Section 4 takes the conclusions inferred from the outcome of the experiments.

2 Selection of Features and Classifiers

2.1 Data Base and Extraction of Features

The data base used for our work is the one built for the CAVIAR project [9]. We selected four videos with the criterion of having the maximum number of activities: Fight_RunAway1.mpg, Fight_RunAway2.mpg, Fight_OneManDown.mpg and Fight_Chase.mpg. These videos were recorded by the CAVIAR team with a wide angle camera lens in the entrance lobby of the INRIA Labs at Grenoble, France. The sequences have half-resolution PAL standard (384 x 288 pixels, 25 frames per second) and were compressed using MPEG2.

Among all the measurements provided as ground truth, we selected the 2D centroid position *(x,y)*, height and width of each person's surrounding box *(h,w)*. Then, we compute a set of 40 features that are divided into two groups:

1. Velocity and Speed for a frame window (*f*) of 3, 5, 15 and 25 frames:
   – velocity for the x-axis and y-axis:

$$v_f^x = (x_i - x_{i-f}) \qquad (1)$$

$$v_f^y = (y_i - y_{i-f}) \qquad (2)$$

   – speed:

$$speed_f = \sqrt{(x_i - x_{i-f})^2 + (y_i - y_{i-f})^2} \qquad (3)$$

– mean speed:

$$mean\_speed_f = \sum_{j=0}^{f-1} \sqrt{(x_{i-j} - x_{i-(j+1)})^2 + (y_{i-j} - y_{i-(j+1)})^2} \quad (4)$$

2. Height, Width and Area for a frame window (f) of 3, 5, 15 and 25 frames:
   – difference of height, width and area:

$$diff\_height_f = (h_i - h_{i-f}) \quad (5)$$

$$diff\_width_f = (w_i - w_{i-f}) \quad (6)$$

$$diff\_area_f = |diff\_height_f \cdot diff\_width_f| \quad (7)$$

   – mean difference of height, width and area:

$$mean\_diff\_height_f = \sum_{j=0}^{f-1}(h_{i-j} - h_{i-(j+1)}) \quad (8)$$

$$mean\_diff\_width_f = \sum_{j=0}^{f-1}(w_{i-j} - w_{i-(j+1)}) \quad (9)$$

$$mean\_diff\_area_f = |mean\_diff\_height_f \cdot mean\_diff\_width_f| \quad (10)$$

where i is the current frame.

The next step was to extract the most relevant motion features for the activity recognition. Our approach consists of using a wrapper method that produces empirical evaluations for a classifier, in our case J.48, and a Genetic Algorithm as a search method for the best solution. To carry out this wrapper we employed WEKA 3.5.2 [10].

2.2 Election of Classifiers

Subsequently, we had to choose the set of classifiers to compare. First of all, we decided to select HMM (Hidden Markov Models) since they classify each scene as a function of the future, actual and previous frames and they have been used efficiently for this task in previous works. As it was suggested in [11], the Baum-Welch method and a Genetic Algorithm (GA) are employing to adjust the parameters of the mixtures of gaussians that in our work define each state of the HMM. In addition, we selected a J.48 tree, two Bayesian classifiers (Bayesian Network and Naive Simple), a classifier based on rules (PART)(all of them included in the WEKA software [10]) and finally a Neuro-Fuzzy classifier [12] so that we have a wide variety of different methods for classification.

The parameters for the different classifiers are adjusted as follows:

– HMM (Baum-Welch):

- $\mu_0 = (0, 1, 2, 7.0)$ for 4 activities and $\mu_0 = (0, 1, 7.0)$ for 3 activities.
  - The σ_0 and the transition matrix A are randomly selected in an interval of [-2..2].
  - Prior probability: $prior_0 = $ (IN=1.0, WK=0.0, AC=0.0, R=0.0).
- HMM (GA):
  - μ_0, σ_0, the transition matrix A and $prior_0$ are chosen randomly.
  - The typical variables of the GA are initialised as in [11].
- Neuro-fuzzy:
  - Number of variables in each fuzzy set: 2.
  - Form of the set: triangular.
  - Size of the rule base: automatically determine.
  - Rule learning Procedure: best per class
  - Learning Rate: 0.01
  - Maximum Number of epochs: 500
  - Minimum Number of epochs: 0
  - Number of Epochs after optimum: 100
  - Admissible Classification errors: 0
- Rest of classifiers: The default parameters of WEKA.

Then, we divided the data into two groups, one for training and other one for validation. Hence, the first and the third videos are used as training data, whereas the second and the forth constitute the validation data. We split up the data in this fashion so that every group have a great variety of data.

3 Experiments

3.1 Wrapper

As we explained before, the first step is the selection of features by means of a wrapper that uses a J.48 and a GA as a classifier and search method respectively. The outcome of this filter is the importance weight of each feature in the classification process. We decided to extract and choose the features like that: a first group with the features above 69% of importance and a more restristed one with the features above 99% of weight. The next table show the selected features and their correspondent weight in pertecentage:

The histograms of these features show the difficulty of this problem due to the subjectivity in the manual process carried out when labeling the activities. The challenging goal of the classifiers is to sort the classes taking these features with such a low level of separability.

3.2 Training Classifiers

The following step was to employ the training and validation data in the different classifiers. In a first approach, we train the classifiers for three activities (inactive, walking and running) as a reduce case of the general one and subsequently the training is carried out with the four activities.

Thus, we carried out four rounds for the comparison of performance:

Table 1. Selected Features above 69% of weight

v_3^x	100%
v_3^y	100%
v_5^y	80%
v_{15}^x	100%
v_{25}^x	80%
v_{25}^y	90%
$speed_3$	100%
$speed_5$	70%
$speed_{15}$	90%
$speed_{25}$	90%
$diff\_height_{15}$	80%
$diff\_height_{25}$	70%
$mean\_diff\_area_{25}$	(90%)

1. 3 activities (IN, WK and R) and 13 features (the ones above the the 69% of weight)
2. 3 activities (IN, WK and R) and 4 features (the ones above the the 99% of weight)
3. 4 activities (IN, WK, AC and R) and 13 features (the ones above the the 69% of weight)
4. 4 activities (IN, WK, AC and R) and 4 features (the ones above the the 99% of weight)

The next tables show the percentage of correct classifications for each classifier in each trial.

Table 2. Results for 13 features and 3 activities (inactive, walking, running)

	HMM (Baum-Welch)	HMM (GA)	J.48	Bayes Net	Naive Bayes	PART	Neuro-Fuzzy
Training	63.90%	70%	99.44%	92%	90.63%	99.47%	87.91%
Validation	48.81%	64%	74.79%	82.82%	80.13%	76.29%	67.74%

Table 3. Results for 4 features and 3 activities (inactive, walking, running)

	HMM (Baum-Welch)	HMM (GA)	J.48	Bayes Net	Naive Bayes	PART	Neuro-Fuzzy
Training	86.49%	89.92%	90.93%	87.00%	88.81%	91.34%	64.79%
Validation	76.89%	88.61%	77.74%	75.76%	79.43%	78.38%	49.40%

It is interesting to show some of the confusion matrix to check how the classification is carried out. We can see below the confusion matrix corresponding to the HMM adjusted with Baum-Welch and GA and the Naive Bayes for the
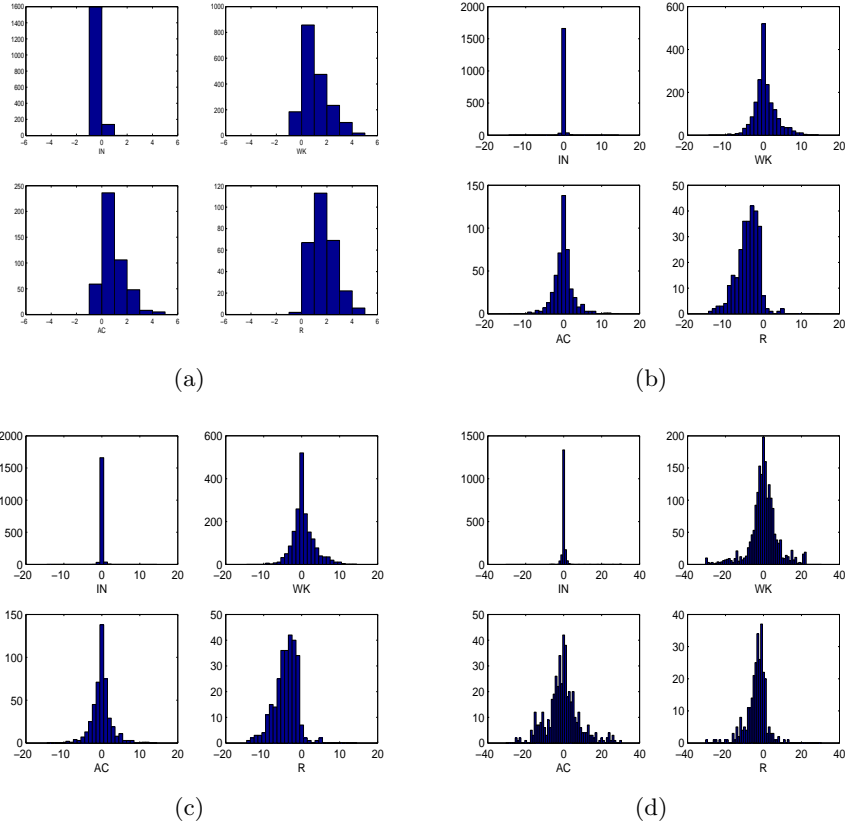
Fig. 1. *Speed_*3 (a) v_3^x (b), v_3^y (c) and $diff_height_$15 (d) for inactive (IN), walking (WK), active (AC) and running (R)

case of 4 features and 3 activities. Furthermore, we included the confusion matrix of the HMM (GA) and Naive Bayes for 4 features and 4 activities.

4 Conclusions

The outcome of the classification of 3 activities shows that the results are very similar for 4 and 13 features in the case of the J.48, PART and the two Bayesian classifiers, being above the 74% of correct classification in the validation set of data. Nevertheless, the classification differs a lot for the HMMs, which present a good performance in the training for 4 features and a poor peformance in the case of 13 features. This can be due to the adjustment of the HMM's states by means of a mixture of Gaussians, that is, the more dimensions the more difficult is for the mixture of Gaussian to model the probability landscape of each activity. In the other hand, the Neuro-fuzzy classifier has the opposite behaviour and it

Table 4. Results for 13 features and 4 activities (inactive, walking, running)

	HMM (Baum-Welch)	HMM (GA)	J.48	Bayes Net	Naive Bayes	PART	Neuro-Fuzzy
Training	54.11%	56.27%	98.87%	83.03%	70.53%	98.71%	52.13%
Validation	32.68%	39.32%	54.56%	50.71%	39.25%	55.68%	26.03%

Table 5. Results for 4 features and 4 activities (inactive, walking, running)

	HMM (Baum-Welch)	HMM (GA)	J.48	Bayes Net	Naive Bayes	PART	Neuro-Fuzzy
Training	63.23%	80.57%	85.84%	77.62%	79.30%	86.21%	61.55%
Validation	56.53%	78.31%	60.00%	65.01%	73.15%	62.14%	46.90%

is able to build a more robust classifier by taking a high number of features. In the case of 4 features, the performance for all classifiers is lower than in the previous case. The reason why this happens is that it is very difficult to distinguish between active and walking with the only information of centroid and surrounding boxes. We can highlight the good working of the Naive Bayes and HMM (GA), with 73.15% and 78.31%.

Furthermore, we can infer very interesting conclusions by analysing the confusion matrix for each classifier. For example, the good result obtained with the HMMs (Baum-Welch and GA) in the case of 4 features and only 3 activities informs that the running case is ignored, and the total performance is still good since the running samples are much less than the rest classes. In general, this phenomenon happened for all the classifiers as we can observe in the tables below.

In order to avoid this, we took several experiments where the training data had the same number of samples for each activity and surprisingly the results did not improve. This can be explained as the difficulty in distinguishing the activities due to the ill definition of the labels in the ground truth. Thus, the good results are partly obtained for the good classification of the most common class.

This can be proof by seeing the confusion matrix for 4 activities. We can check that the active (AC) and running (R) activities are misclassified almost always.

References

1. J.C. Nascimento, M. A. T. Figueiredo and J. S. Marques: Segmentation and Classification of Human Activities, HAREM 2005: International Workshop on Human Activity Recognition and Modelling, Oxford, UK, September 2005
2. Pedro Canotilho Ribeiro and José Santos-Victor. Human Activity Recognition from Video: modeling, feature, selection and classification arquitecture, HAREM 2005: International Workshop on Human Activity Recognition and Modelling, Oxford, UK, September 2005

8

Table 6. Results for 4 features and 3 activities (inactive, walking, running) for the HMM (Baum-Welch) and 76.89% of correct classifications

HMM (Baum-Welch)	Predicted		
Actual	IN	WK	R
IN	619	14	0
WK	437	1440	0
R	1	167	0

Table 7. Results for 4 features and 3 activities (inactive, walking, running) for the HMM (GA) and 88.61% of correct classifications

HMM (Baum-Welch)	Predicted		
Actual	IN	WK	R
IN	599	34	0
WK	95	1774	8
R	0	168	0

3. Sebastian Brännström. Extraction, Evaluation and Selection of Motion Features for Human Activity Recognition Purposes, Master's Thesis in Computer Science at the School of Engineering Physics Royal Institute of Technology 2006
4. Osama Masoud and Nikos Papanikolopoulos. A Method For Human Action Recognition, Department of Computer Science and Engineering University of Minnesota 2003
5. Xiaotao Zou and Bir Bhanu. Human Activity Classification Based on Gait Energy Image and Coevolutionary Genetic Programming Pattern Recognition, 2006. ICPR 2006. 18th International Conference on Volume 3, 20-24 Aug. 2006 Page(s):556 - 559
6. M. Leo, T. D'Orazio, I. Gnoni, P. Spagnolo and A. Distante: Complex Human Activity Recognition for Monitoring Wide Outdoor Environments, Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on Volume 4, 23-26 Aug. 2004 Page(s):913 - 916 Vol.4
7. Paul E. Rybski and Manuela M. Veloso: Human Activity Recognition from Video: modeling, feature selection and classification architecture, HAREM 2005 - International Workshop on Human Activity Recognition and Modeling. Oxford, UK, September 2005
8. Zhou Feng and Tat-Jen Cham. Video-based Human Action Classi.cation with Ambiguous Correspondences, Computer Vision and Pattern Recognition, 2005 IEEE Computer Society. 20-26 June 2005 Volume: 3, On page(s): 82- 82
9. "CAVIAR PROJECT", http://homepages.inf.ed.ac.uk/rbf/CAVIAR
10. "WEKA", http://www.cs.waikato.ac.nz/ml/weka/
11. Óscar Pérez, Massimo Piccardi, Jesús García, M.A. Patricio, J.M. Molina. Comparison between Genetic Algorithms and the Baum-Welch Algorithm in Learning HMMs for Human Activity Classification Proceedings of EvoIASP2007: Ninth European Workshop on Evolutionary Computation in Image Analysis and Signal Processing. Valencia, Spain, 11-13 April 2007
12. "NEFCLASS", http://fuzzy.cs.uni-magdeburg.de/nefclass/

Table 8. Results for 4 features and 3 activities (inactive, walking, running) for the Naive Bayes and 79.43% of correct classifications

HMM (Baum-Welch)		IN	WK	R
Actual		IN	WK	R
	IN	572	61	0
	WK	803	1553	119
	R	0	164	0

Table 9. Results for 4 features and 4 activities (inactive, walking, active, running) for the HMM (GA) and 78.31% of correct classifications

HMM (Baum-Welch)		IN	WK	AC	R
Actual		IN	WK	AC	R
	IN	600	99	0	0
	WK	190	1497	12	8
	AC	0	170	0	0
	R	0	168	0	0

Table 10. Results for 4 features and 4 activities (inactive, walking, active, running) for the Naive Bayes and 73.15% of correct classifications

HMM (Baum-Welch)		IN	WK	AC	R
Actual		IN	WK	AC	R
	IN	576	57	0	0
	WK	202	1377	0	128
	AC	15	155	0	0
	R	2	160	0	0