Generic System for Human-Computer Gesture Interaction: Applications on Sign Language Recognition and Robotic Soccer Refereeing

Paulo Trigueiros[1, 2, 4, 5], Fernando Ribeiro[1,4], Luis Paulo Reis[3,5]

[1]Escola de Eng. da Universidade do Minho, Departamento de Eletrónica Industrial (EEUM/DEI)
[2]Instituto Politécnico do Porto, Departamento de Informática (ISCAP/IPP),
[3]Escola de Eng. da Universidade do Minho, Departamento de Sistemas de Informação (EEUM/DSI)
[4]Centro ALGORITMI
[5]Laboratório de Inteligência Artificial e Ciência de Computadores (LIACC)
Emails:pjt@iscap.ipp.pt, fernando@dei.uminho.pt, lpreis@dsi.uminho.pt

Abstract - Hand gestures are a powerful way for human communication, with lots of potential applications in the area of human computer interaction. Vision-based hand gesture recognition techniques have many proven advantages compared with traditional devices, giving users a simpler and more natural way to communicate with electronic devices. This work proposes a generic system architecture based in computer vision and machine learning, able to be used with any interface for human-computer interaction. The proposed solution is mainly composed of three modules: a pre-processing and hand segmentation module, a static gesture interface module and a dynamic gesture interface module. The experiments showed that the core of vision-based interaction systems could be the same for all applications and thus facilitate the implementation. For hand posture recognition, a SVM (Support Vector Machine) model was trained and used, able to achieve a final accuracy of 99.4%. For dynamic gestures, an HMM (Hidden Markov Model) model was trained for each gesture that the system could recognize with a final average accuracy of 93.7%. The proposed solution as the advantage of being generic enough with the trained models able to work in real-time, allowing its application in a wide range of human-machine applications. To validate the proposed framework two applications were implemented. The first one is a real-time system able to interpret the Portuguese Sign Language. The second one is an online system able to help a robotic soccer game referee judge a game in real time.

Keywords - *Human-computer interaction; Gesture Recognition; Computer Vision; Machine Learning*

Introduction

Hand gestures are a powerful way for human communication, with lots of potential applications in the area of human computer interaction. Vision-based hand gesture recognition techniques have many proven advantages compared with traditional devices, giving users a simpler and more intuitive way of communication between a human and a computer. Using visual input in this context makes it possible to communicate remotely with computerized equipment, without the need for physical contact. For gesture-based applications, we need to model them in the spatial and temporal domains, where a hand posture is the static structure of the hand and a gesture is the dynamic movement of the hand. Being hand-pose one of the most important communication tools in human's daily life, and with the continuous advances of image and video processing techniques, research on human-machine interaction through gesture recognition led to the use of such technology in a very broad range of applications [1,2], of which some are here highlighted:

- Virtual reality: enable realistic manipulation of virtual objects using ones hands [3,4].
- Robotics and Tele-presence: gestures used to interact with and to control robots [5] are similar to fully-immersed virtual reality interactions, however the worlds are often real, presenting the operator with video feed from

cameras located on the robot. Here, for example, gestures can control a robot hand and arm movements to reach for and manipulate actual objects.

- Desktop and Tablet PC Applications: In desktop computing applications, gestures can provide an alternative interaction to mouse and keyboard [6-9]. Many gestures for desktop computing tasks involve manipulating graphics, or annotating and editing documents using pen-based gestures.
- Games: track a player's hand or body position to control movement and orientation of interactive game objects such as cars, or use gestures to control the movement of avatars in a virtual world, as for example the Play Station 2 Eye Toy [10] and the Microsoft Kinect [11].
- Sign Language: this is an important case of communicative gestures. Since sign languages are highly structural, they are very suitable as test-beds for vision-based algorithms [12-15].

There are areas where this trend is an asset, as for example in the application of these technologies in interfaces that can help people with physical disabilities, or areas where it is a complement to the normal way of communicating.

The main objective of this work is to describe and propose a generic system architecture based in computer vision and machine learning, able to be used with any interface for human-machine interaction. Computer vision based techniques have the advantage of being non-invasive and based on the way human beings perceive information from their surroundings [16].

To be able to implement such systems in a successful way, there are a number of requirements that they must satisfy [17], which are:

- Robustness: the system should be user independent and robust enough to a number of factors.
- Computational efficiency: algorithms and learning techniques should be the most effective possible and computational cost effective.
- Error tolerance: mistakes should be tolerated and accepted. The user should be able to repeat any command, instead of letting the system make wrong decisions.
- Scalability: the system must be easily adapted and configured so that it can serve a number of different applications.

The rest of the paper is as follows. Firstly, the related work is review in section 0. Section 0 introduces the system architecture and describes each one of the proposed modules. Application implementation is discussed in section 0. Experimental methodology and results are explained in section 0. Conclusions are drawn in section 0.

Related Work

Hand gestures, either static or dynamic, for human computer interaction in real time systems is an area of active research in computer vision and machine learning [18] and with many possible applications. However, vision-based hand gesture interfaces for real-time applications require fast and extremely robust hand detection, feature extraction and gesture recognition. Several approaches are normally used including Artificial Neural Networks (ANN), SVM and HMM's.

An ANN is a mathematical / computational model that attempts to simulate the structure of biological neural systems. They accept features as inputs and produce decisions as outputs [19]. Maung et al [18] applied it in a gesture recognition system for real-time gestures in unstrained environments. Vicen-Buéno et al. [20]

used it applied to the problem of traffic sign recognition. Bailador et al. [21] presented an approach to the problem of gesture recognition in real time using inexpensive accelerometers. Their approach was based on the idea of creating specialized signal predictors for each gesture class.

A SVM is a technique based on statistical learning theory, which works very well with high-dimensional data. The objective of this algorithm is to find the optimal separating hyper plane between two classes by maximizing the margin between them [22]. Faria [23,24] used it to classify robotic soccer formations and the classification of facial expressions, Ke [25] used it in the implementation of a real-time hand gesture recognition system for human robot interaction, Maldonado-Báscon [26] used it for the recognition of road-signs and Masaki [27] used it in conjunction with SOM (Self-Organizing Map) for the automatic learning of a gesture recognition mode. He first applies the SOM to divide the sample into phases and construct a state machine, and then he applies the SVM to learn the transition conditions between nodes. Almeida [28] proposed a classification approach to identify the team's formation in the robotic soccer domain for the two dimensional (2D) simulation league employing Data Mining classification techniques.

Trigueiros [29] has made a comparative study of four machine learning algorithms applied to two hand features datasets. In their study the datasets had a mixture of hand features. He has also made a comparative study of different image features for hand gesture machine learning [16] and proposed a vision-based system for the Portuguese sign language recognition [30], based in a SVM model with an accuracy of 99.2%, and a Vision-based Gesture Recognition System for Human Computer Interaction [31] based in machine learning algorithms and able to do real-time hand gesture recognition. Hidden Markov Models (HMMs) have been widely used in a successfully way in speech recognition and hand writing recognition [32], in various fields of engineering and also applied quite successfully to gesture recognition.

Oka [33] developed a gesture recognition system based on measured finger trajectories for an augmented desk interface system. They have used a Kalman filter for the prediction of multiple finger locations and an HMM for gesture recognition.

Perrin [34] described a finger tracking gesture recognition system based on a laser tracking mechanism which can be used in hand-held devices. They have used HMM for their gesture recognition system with an accuracy of 95% for a set of 5 gestures. Nguyen [35] described a hand gesture recognition system using a real-time tracking method with pseudo two-dimensional Hidden Markov Models. Chen [36] used it in combination with Fourier descriptors for hand gesture recognition using a real-time tracking method. Kelly [37] implemented an extension to the standard HMM model to develop a gesture threshold HMM (GT-HMM) framework which is specifically designed to identify inter gesture transition. Zafrulla [12] have investigated the potential of the Kinect depth-mapping camera for sign language recognition and verification for educational games for deaf children. They used 4-state HMMs to train each of the 19 signs defined in their study. Trigueiros [38] used HMM's applied to a Vision-based system capable of recognizing a set of referee commands for robotic soccer games.

Cooper [39] implemented an isolated sign recognition system using a 1st order Markov chain. In their model, signs are broken down in visemes (equivalent to phonemes in speech) and a bank of Markov chains are used to recognize the visemes as they are produced. Milosevic [40] implemented an HMM-based continuous gesture recognition algorithm, optimized for lower power, low cost microcontrollers without float point unit. The proposed solution is validated on a set of gestures performed with the Smart Micrel Cube (SMCube), which embeds a 3-axis accelerometer and an 8-bit microcontroller. They also explore a multiuser scenario where up to 4 people share the same device. Elmezain [41] proposed a system able to recognize both isolated and continuous gestures for Arabic numbers (0-9) in real-time. To handle isolated gestures, an HMM using Ergodic (it is possible to go from every state to every state), Left-Right (LR) and Left-Right Banded (LRB) topologies with different number of states was applied. The LRB in conjunction with the Forward algorithm presented the best performance with an average recognition rate of 98.94% and 95.7% for isolated and continuous gestures.

Proposed System Architecture

The design of any gesture recognition system essentially involves the following three aspects: (1) *data acquisition and pre-processing;* (2) *data representation or feature extraction* and (3) *classification or decision-making*. Taking this into account, a possible solution to be used in any vision-based hand gesture recognition system for human-machine interaction is represented in the diagram of Fig. 1.

The following sections will describe the Static gesture module and the Dynamic gesture module.

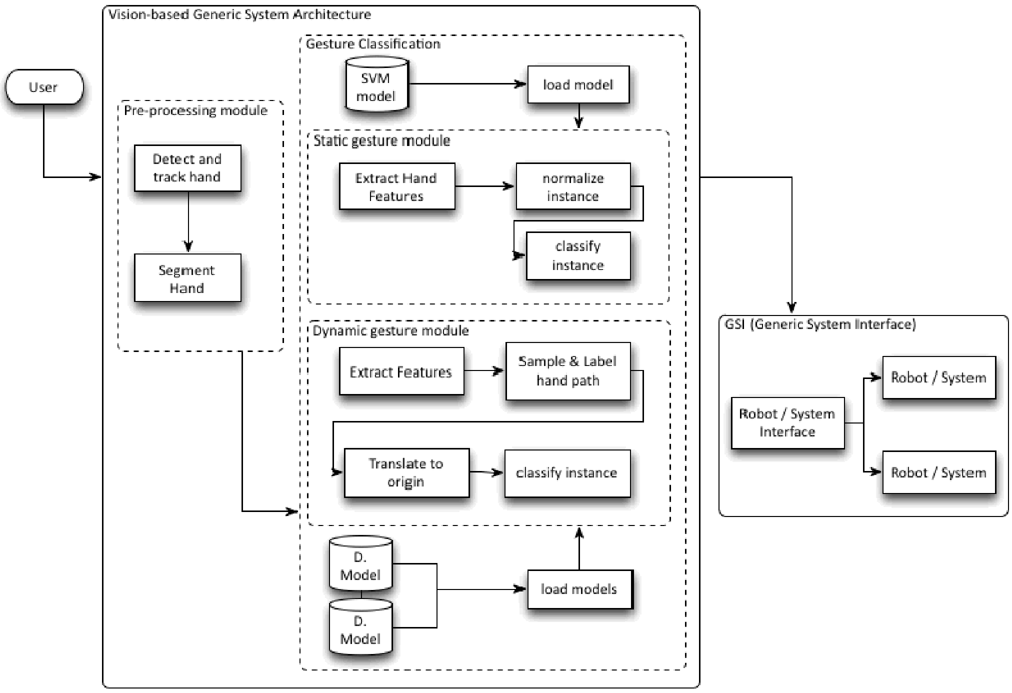

Fig. 1 Generic system architecture for a human-computer gesture interface

1.1 Static Gesture Module

For static gesture classification, hand segmentation and feature extraction is a crucial step in vision-based hand gesture recognition systems. This step is crucial to determine in the future whether given hands shape matches a given model, or which of the representative classes is the most similar. According to Wacs [42] proper feature selection, and their combination with sophisticated learning and recognition algorithms, can affect the success or failure of any existing and future work in the field of human computer interaction using hand gestures. The obtained segmented hand in the pre-processing module is used to extract hand features that are used later with classification algorithms [16]. The learned models for hand posture classification use feature vectors composed of centroid distance values. The centroid distance signature is a type of shape signature [43] expressed by the distance of the hand contour boundary points, from the hand centroid (x_c, yc) and calculated with the following equation:

$$d(i) = \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}, i = 0, \ldots, N-1 \tag{1}$$

This way, a one-dimensional function representing the hand shape is obtained. The number of equally spaced points N used in our implementation was 16. Due to the subtraction of centroid from the boundary coordinates, this operator is invariant to translation as shown by Tara [15]. All the vectors are normalized prior to training, by the z-normalization [44,45] as follows,

$$Z = \left(a_{ij} - \bar{a}\right)/\sigma \tag{2}$$

where $\bar{a}$ represents the mean of the instance i, and σ is the respective standard deviation, achieving this way scale invariance as desired.

The feature vectors thus obtained were used to train a multi-class SVM used in system implementations as shown in Fig. 2. The SVM is a pattern recognition technique in the area of supervised machine learning, which works very well with high-dimensional data. When more than two classes are present, there are several approaches that evolve around the 2-class case [46]. The one used in this system is the one-against-all, where c classifiers have to be designed. Each one of them is designed to separate one class from the rest.

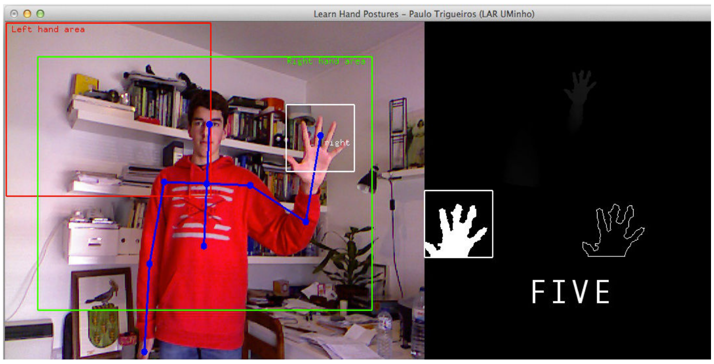

Fig. 2 Static gesture SVM classification

1.2 Dynamic gesture module

Dynamic gestures are time-varying processes, which show statistical variations, making HMMs a plausible choice for modelling the processes [47] [48]. So, for the recognition of dynamic gestures a HMM (Hidden Markov Model) model was trained for each possible gesture. A Markov Model is a typical model for a stochastic (i.e. random) sequence of a finite number of states [49]. A human gesture can be understood as a Hidden Markov Model where the true states of the model are hidden in the sense that they cannot be directly observed. HMMs have been widely used in a successfully way in speech recognition and hand writing

recognition [32]. In this system the 2D motion hand trajectory points are labelled according to the distance to the nearest centroid, based on Euclidean distance, and translated to origin resulting in a discrete feature vector like the one shown in Fig. 3. The feature vectors thus obtained are used to train the different HMMs and learn the model parameters: the initial state probability vector (Π), the state-transition probability matrix ($A=[a_{ij}]$) and the observable symbol probability matrix ($B=[b_j(m)]$). In the recognition phase an output score for the sample gesture is calculated for each model, given the likelihood that the corresponding model generated the underlying gesture. The model with the highest output score represents the recognized gesture. In our system a Left-Right (LR) HMM, like the one shown in Fig. 4, was used [50,44].
This kind of HMM has the states ordered in time so that as time increases, the state index increases or stays the same. This topology has been chosen, since it is perfectly suitable to model the kind of temporal gestures present in the system.

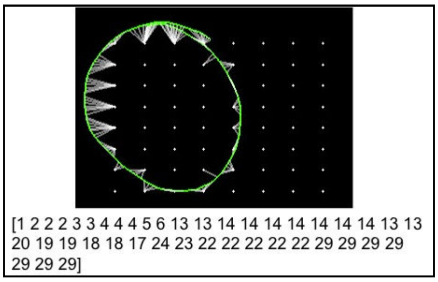

Fig. 3 Dynamic gesture path with extracted feature vector

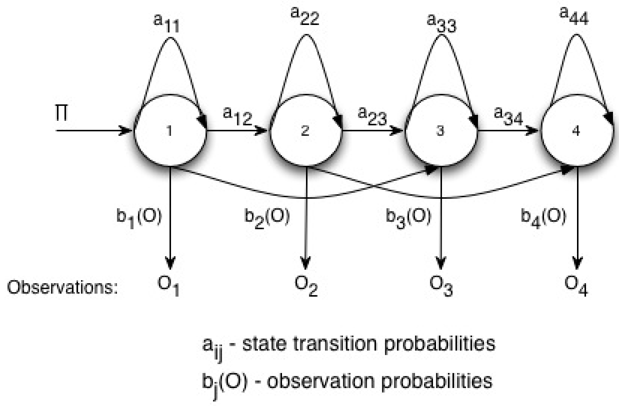

Fig. 4 A 4-state Left-Right HMM model

Applications

1.3 Portuguese Sign Langiage Recognition System

The Portuguese Sign Language Recognition Prototype is a real-time vision-based system whose purpose is to recognize the Portuguese Sign Language given in the alphabet of Fig. 5. The purpose of the prototype was to test the validity of a vision-based system for sign language recognition and at the same time, test and select hand features that could be used with machine learning algorithms allowing their application in any real-time sign language recognition systems. For that, the user must be positioned in front of the camera, doing the sign language gestures, that will be interpreted by the system and their classification will be displayed on the right side of the interface. The implemented solution uses only one camera, a Kinect camera [11], and is based on a set of assumptions, hereby defined:

- The user must be within a defined perimeter area, in front of the camera.
- The user must be within a defined distance range, due to camera limitations. The system-defined values are 0.7m for the near plane and 3m for the far plane.
- Hand pose is defined with a bare hand and not occluded by other objects.
- The system must be used indoor, since the selected camera does not work well under sun light conditions.

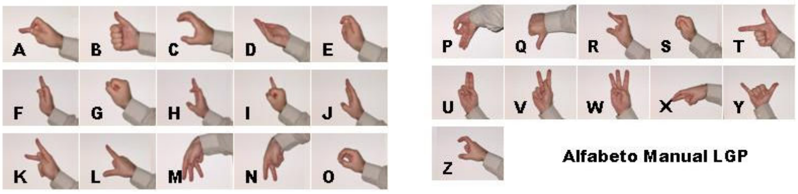

Fig. 5 Alphabet for the Portuguese Sign Language

1.4 Referee Command Language Interface System

To validate the proposed framework, an online system able to help a robotic soccer game referee judge a game in real time was implemented. It combines a vision-based hand gesture recognition system with a formal language definition, the *Referee CommLang*, into what was called the *Referee Command Language Interface System* (ReCLIS). The system builds a command based on system-interpreted static and dynamic referee gestures, and is able to send it to a computer interface, for robot control. The commands were defined in a new formal language described in section 1.4.1. With the proposed solution, there is the possibility of eliminating the assistant referee, thereby allowing a more natural game interface.

1.4.1 The Referee Command Language Definition

The *Referee CommLang* is a new and formal definition of all commands that the system is able to identify. As in [51], the language must represent all the possible gesture combinations and at the same time be simple in its syntax. The language was defined with BNF (Bakus Normal Form) [52]:

- Terminal symbols (keywords and operator symbols) are in a constant-width typeface.
- Choices are separated by vertical bars (|) and in greater-than and less-than symbols (< choice>).
- Optional elements are in square brackets ([optional]).
- Sets of values are in curly braces ({}).
- A syntax description is introduced with ::=.

The language has three types of commands: *Team commands, Player commands and Game commands*. This way, a language is defined to be a set of commands as follows:

<LANGUAGE> ::= {<COMMAND>}

<COMMAND> ::= <TEAM_COMMAND> | <GAME_COMMAND> |
<PLAYER_COMMAND>

<TEAM_COMMAND> ::= <KICK_OFF> | <CORNER> | <THROW_IN> | <GOAL_KICK> |
<FREE_KICK> | <PENALTY> | <GOAL> | <DROP_BALL>

<GAME_COMMAND> ::= <START> | <STOP> | <END_GAME> | <CANCEL> | <RESEND>

<PLAYER_COMMAND> ::= <SUBSTITUTION> | <PLAYER_IN> | <PLAYER_OUT> | <YELLOW_CARD> | <RED_CARD>

For the TEAM_COMMANDS there are several options according to the following definition:

<KICK_OFF> ::= KICK_OFF <TEAM_ID>

<CORNER> ::= CORNER <TEAM_ID>

<THROW_IN> ::= THROW_IN <TEAM_ID>

<GOAL_KICK> ::= GOAL_KICK <TEAM_ID>

<FREE_KICK> ::= FREE_KICK <TEAM_ID>

<PENALTY> ::= PENALTY <TEAM_ID>

<GOAL> ::= GOAL <TEAM_ID>

<DROP_BALL> ::= DROP_BALL

Finally, for the PLAYER_COMMAND, we have the following definitions.

<SUBSTITUTION> ::= SUBSTITUTION <PLAYER_IN> <PLAYER_OUT>

<PLAYER_IN> ::= PLAYER_IN <TEAM_ID> <PLAYER_ID>

<PLAYER_OUT> ::= PLAYER_OUT <TEAM_ID> <PLAYER_ID>

<YELLOW_CARD> ::= YELLOW_CARD <TEAM_ID> <PLAYER_ID>

<RED_CARD> ::= RED_CARD <TEAM_ID> <PLAYER_ID>

<START> ::= START□

<STOP> ::= STOP□

<END_GAME> ::= END_GAME <PART_ID>

<CANCEL> ::= CANCEL□

<RESEND> ::= RESEND

<TEAM_ID > ::= CYAN | MAGENTA□

<PLAYER_ID> ::= PL1 | PL2 | PL3 | PL4 | PL5 | PL6 | PL7

<PART_ID> ::= 1ST | 2ND | EXTRA | PEN

1.4.2 Command classification

Since the ReCLIS uses a combination of dynamic and static gestures, modelling the command semantics became necessary. A *Finite State Machine* (FSM) is a usually employed technique to handle this situation [53,54] and thus implemented to control the transition between three possible states: DYNAMIC, STATIC and PAUSE as shown in Fig. 6.

A PAUSE state is used to control the transitions between gestures and somehow eliminate all unintentional actions between DYNAMIC/STATIC and STATIC/STATIC gestures. This state is entered every time a gesture is found, and exited after a predefined period of time or when a command sequence is identified.

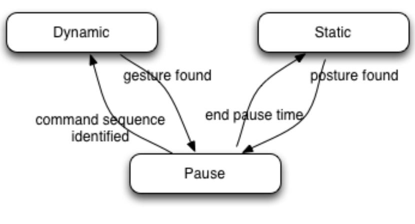

Fig. 6 The Finite State Machine (FSM) diagram

1.5 Implementation

The Human-Computer Interfaces (HCI) for both applications were developed using the C++ language, and the openFrameworks toolkit [55] with the OpenCV [56] and the OpenNI [57] addons, *ofxOpenCv* and *ofxOpenNI* respectively. OpenCV was used for some of the vision-based operations like extracting the hand blob contour, and OpenNI was responsible for the RGB and depth image acquisition. For SVM model training and gesture classification the open source Dlib library was used, a general-purpose cross-platform C++ library capable of SVM multiclass classification [58]. The resulting obtained model is compact and fast, able to be applied in any applications with real-time classification demands. In Fig. 7 one can see the Sign Language Prototype with two vowels correctly classified and displayed on the right side of the user interface.

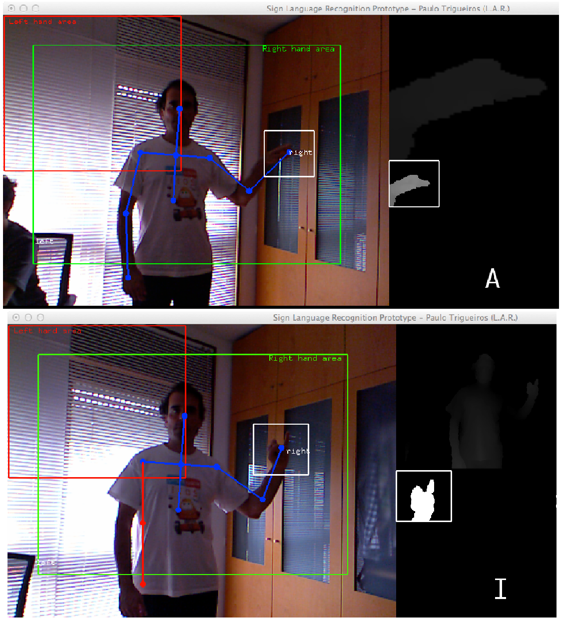

Fig. 7 Sign Language prototype interface wit two vowels correctly classified

For dynamic gesture model training and implementation an openFrameworks add-on implementation of the HMM algorithm for classification and recognition of numeric sequences was used. This implementation is a C++ porting of a MATLAB code from Kevin Murphy [59].
Fig. 8 shows the Referee Command Language Interface System with one command correctly classified.

Experimental Methodology and Results

The experimental methodology was divided into two parts: a hand posture database creation with the selected features and SVM model training and testing and a dynamic gesture database creation and HMM model training and testing.

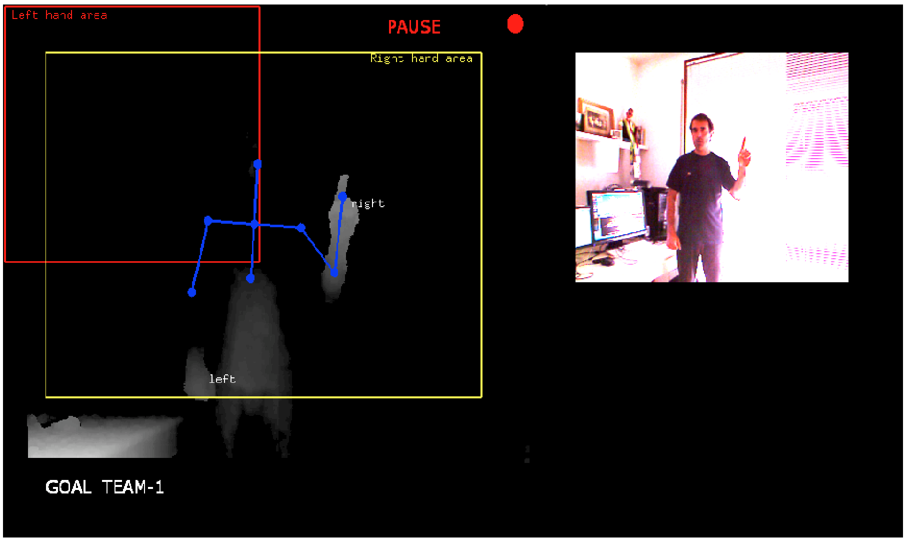

Fig. 8. The Referee Command Language Interface System.

The experiments were performed with a parameter optimization for the cost parameter C, and a 10-fold cross validation. A final accuracy of 99.4% was achieved, with a linear kernel and a C value equal to 1.
The obtained confusion matrix is shown in Table 1.

Table 1 Centroid distance features confusion matrix

		Actual class				
		1	2	3	4	5
Predicted class	1	**455**	0	0	2	0
	2	0	**394**	1	1	0
	3	0	0	**401**	1	0
	4	4	2	0	**382**	0
	5	0	0	1	0	**439**

For dynamic gesture recognition, an HMM model was trained off-line for each one of the 11 predefined gestures and the three parameters (the initial state probability vector, the state-transition probability matrix and the observable symbol probability matrix) were learned and saved. Once again four users were used to perform the predefined gestures and the extracted features were saved and used to train the models. The number of observation symbols (alphabet) and hidden states were learned by trial and error, and were defined to be 64 and 4 respectively. The test datasets obtained were analysed with the learned models and the final accuracy results obtained are represented in Table 2. So, for the dynamic gesture recognition an average accuracy of 93.7% was achieved.

Table 2 Hidden Markov Models accuracy for each gesture trained

Gesture	1	2	3	4	5	6	7	8	9	10	11
Accuracy	75%	100%	100%	100%	92%	88%	92%	100%	100%	96%	88%

Conclusions

This paper presented a system able to interpret dynamic and static gestures from a user with the goal of real-time human-computer interaction. Although the machine learning algorithms used are not the only solutions, they were selected based on obtained performance with the selected features. Thus, for hand posture classification a SVM model was learned from centroid distance features and a recognition rate of 99.4% was achieved. For dynamic gesture classification, a HMM model was learned for each gesture and a final average accuracy of 93.7% was achieved. We were able to test the system in real time situations, and it was possible to prove from the experiments that the trained models were able to recognize all the trained gestures, proving that this kind of models, as was already seen in other references, works very well for this type of problem. The experimental results also showed, that the proposed system was able to reliably recognize the pre-defined commands.
With the implemented applications, it was possible to prove that the core of vision-based interaction systems can be the same for all application, and that the proposed generic system architecture is a solid foundation for the development of hand gesture recognition systems that can be integrated in any human-machine interface application.

References

1. Mitra, S., Acharya, T.: Gesture recognition: A Survey. In: IEEE Transactions on Systems, Man and Cybernetics, vol. 37. vol. 3, pp. 311-324. IEEE, (2007)
2. Bourennane, S., Fossati, C.: Comparison of shape descriptors for hand posture recognition in video. Signal, Image and Video Processing **6**(1), 147-157 (2010). doi:10.1007/s11760-010-0176-6
3. Yoon, J.-H., Park, J.-S., Sung, M.Y.: Vision-Based bare-hand gesture interface for interactive augmented reality applications. In: 5th international conference on Entertainment Computing, Cambridge, UK, September 20-22 2006, pp. 386-389. Springer-Verlag, 2092520
4. Buchmann, V., Violich, S., Billinghurst, M., Cockburn, A.: FingARtips: gesture based direct manipulation in Augmented Reality. Paper presented at the 2nd international Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, Singapore, June 15-18
5. Trigueiros, P., Ribeiro, F., Lopes, G.: Vision-based hand segmentation techniques for human-robot interaction for real-time applications. In: Tavares, J.M., Jorge, R.M.N. (eds.) III ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, Olhão, Algarve, Portugal, 12-14 de Oubtubro 2011, pp. 31-35. Taylor and Francis, Publication
6. Vatavu, R.-D., Anthony, L., Wobbrock, J.O.: Gestures as point clouds: a $P recognizer for user interface prototypes. Paper presented at the 14th ACM international conference on Multimodal interaction, Santa Monica, California, USA,
7. Wobbrock, J.O., Wilson, A.D., Li, Y.: Gestures without libraries, toolkits or training: a $1 recognizer for user interface prototypes. Paper presented at the Proceedings of the 20th annual ACM symposium on User interface software and technology, Newport, Rhode Island, USA,
8. Li, Y.: Protractor: a fast and accurate gesture recognizer. Paper presented at the Conference on Human Factors in Computing Systems, Atlanta, Georgia, USA, April 10-15
9. Kratz, S., Rohs, M.: Protractor3D: a closed-form solution to rotation-invariant 3D gestures. Paper presented at the 16th International Conference on Intelligent User Interfaces, Palo Alto, CA, USA, February 13-16
10. Kim, T.: In-Depth: Eye To Eye - The History Of EyeToy. http://www.gamasutra.com/php-bin/news_index.php?story=20975 (2008). Accessed 29-03-2013 2013
11. Chowdhury, J.R.: Kinect Sensor for Xbox Gaming. IIT Kharagpur (2012)

12. Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., Presti, P.: American sign language recognition with the kinect. Paper presented at the 13th International Conference on Multimodal Interfaces, Alicante, Spain, November 14-18
13. Ong, S.C., Ranganath, S.: Automatic sign language analysis: a survey and the future beyond lexical meaning. IEEE Trans Pattern Anal Mach Intell **27**(6), 873-891 (2005). doi:10.1109/TPAMI.2005.112
14. Holt, G.A.t., Reinders, M.J.T., Hendriks, E.A., Ridder, H.d., Doorn, A.J.v.: Influence of handshape information on automatic sign language recognition. In: 8th International Conference on Gesture in Embodied Communication and Human-Computer Interaction, Bielefeld, Germany, February 25-27 2010, pp. 301-312. Springer-Verlag, 2127632
15. Tara, R.Y., Santosa, P.I., Adji, T.B.: Sign Language Recognition in Robot Teleoperation using Centroid Distance Fourier Descriptors. International Journal of Computer Applications **48**(2) (2012).
16. Trigueiros, P., Ribeiro, F., Reis, L.P.: A Comparative Study of different image features for hand gesture machine learning. In: 5th International Conference on Agents and Artificial Intelligence, Barcelona, Spain, 15-18 February 2013
17. Murthy, G.R.S., Jadon, R.S.: A Review of Vision Based Hand Gestures Recognition. International Journal of Information Technology and Knowledge Management **2**(2), 405-410 (2009).
18. Maung, T.H.H.: Real-Time Hand Tracking and Gesture Recognition System Using Neural Networks. Proceedings of World Academy of Science: Engineering & Technology **50**(Frebuary), 466-470 (2009).
19. Snyder, W.E., Qi, H.: Machine Vision. Cambridge University Press, (2004)
20. Vicen-Bueno, R., Gil-Pita, R., Jarabo-Amores, M.P., López-Ferreras, F.: Complexity Reduction in Neural Networks Appplied to Traffic Sign Recognition Tasks. In. (2004)
21. Bailador, G., Roggen, D., Tröster, G., Triviño, G.: Real time gesture recognition using continuous time recurrent neural networks. In: 2nd International Conference on Body Area Networks, Florence, Italy, June 11-13 2007, pp. 1-8. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 1460247
22. Ben-Hur, A., Weston, J.: A User's Guide to Support Vector Machines. In: Data Mining Techniques for the Life Sciences, vol. 609. pp. 223-239. Humana Press, (2008)
23. Faria, B.M., Lau, N., Reis, L.P.: Classification of Facial Expressions Using Data Mining and machine Learning Algorithms. In: aisti (ed.) 4ª Conferência Ibérica de Sistemas e Tecnologias de Informação, Póvoa de Varim, Portugal, 17 a 20 Junho de 2009 2009, pp. 197-206
24. Faria, B.M., Reis, L.P., Lau, N., Castillo, G.: Machine Learning Algorithms applied to the Classification of Robotic Soccer Formations and Opponent Teams. Paper presented at the IEEE Conference on Cybernetics and Intelligent Systems (CIS), Singapore, 28-30 June 2010
25. Ke, W., Li, W., Ruifeng, L., Lijun, Z.: Real-Time Hand Gesture Recognition for Service Robot. 976-979 (2010). doi:10.1109/icicta.2010.413
26. Maldonado-Báscon, S., Lafuente-Arroyo, S., Gil-Jiménez, P., Gómez-Moreno, H.: Road-Sign detection and Recognition Based on Support Vector Machines. In: IEEE Transactions on Intelligent Transportation Systems, vol. 8. vol. 2, pp. 264-278. (2007)
27. Oshita, M., Matsunaga, T.: Automatic learning of gesture recognition model using SOM and SVM. In: Bebis, G., Boyle, R., Parvin, B., Koracin, D., Chung, R. (eds.) International Conference on Advances in Visual Computing, Las Vegas, NV, USA, November 29 - December 1 2010, pp. 751-759. Springer-Verlag
28. Almeida, R., Reis, L.P., Jorge, A.M.: Analysis and Forecast of Team Formation in the Simulated Robotic Soccer Domain. Paper presented at the Proceedings of the 14th Portuguese Conference on Artificial Intelligence: Progress in Artificial Intelligence, Aveiro, Portugal,
29. Trigueiros, P., Ribeiro, F., Reis, L.P.: A comparison of machine learning algorithms applied to hand gesture recognition. In: 7th Iberian Conference on Information Systems and Technologies, Madrid, Spain, 20-23 July 2012, pp. 41-46
30. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision-based Sign Language Recognition System. In: World Conference on Information Systems and Technologies Madeira, Portugal, April 15-18 2014
31. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision-based Gesture Recognition System for Human-Computer Interaction. In: Tavares, J.M., Jorge, R.M.N. (eds.) IV ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing, Funchal. Madeira 2013. Taylor and Francis, Publication

32. Rabiner, L.R.: A tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE **77**(2), 257-286 (1989). doi:10.1109/5.18626
33. Oka, K., Sato, Y., Koike, H.: Real-time fingertip tracking and gesture recognition. IEEE Computer Graphics and Applications **22**(6), 64-71 (2002). doi:10.1109/mcg.2002.1046630
34. Perrin, S., Cassinelli, A., Ishikawa, M.: Gesture recognition using laser-based tracking system. In: Sixth IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, South Korea, May 17-19 2004, pp. 541-546
35. Binh, N.D., Shuichi, E., Ejima, T.: Real-Time Hand Tracking and Gesture Recognition System. In: Proceedings of International Conference on Graphics, Vision and Image Cairo - Egypt, December 2005 2005, pp. 362--368
36. Chen, F.-S., Fu, C.-M., Huang, C.-L.: Hand gesture recognition using a real-time tracking method and hidden Markov models. Image and Vision Computing **21**(8), 745-758 (2003). doi:10.1016/s0262-8856(03)00070-2
37. Kelly, D., McDonald, J., Markham, C.: Recognition of Spatiotemporal Gestures in Sign Language Using Gesture Threshold HMMs. In: Wang, L., Zhao, G., Cheng, L., Pietik√§inen, M. (eds.) Machine Learning for Vision-Based Motion Analysis. Advances in Pattern Recognition, pp. 307-348. Springer London, (2011)
38. Trigueiros, P., Ribeiro, F., Reis, L.P.: Vision Based Referee Sign Language Recognition System for the RoboCup MSL League. In: 17th annual RoboCup International Symposium, Eindhoven, Holland, 1 July 2013 2013
39. Helen, C., Richard, B.: Large lexicon detection of sign language. In: 11th International Conference on Human-Computer Interaction, Rio de Janeiro, Brazil 2007, pp. 88-97. Springer-Verlag
40. Milosevic, B., Farella, E., Benini, L.: Continuous Gesture Recognition for Resource Constrained Smart Objects. In: IARIA (ed.) The Fourth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies, Florence, Italy, October 25-30 2010, pp. 391-396
41. Elmezain, M., Al-Hamadi, A., Appenrodt, J., Michaelis, B.: A Hidden Markov Model-based Continuous Gesture Recognition System for Hand Motion Trajectory. In: 19th International Conference on Pattern Recognition, Tampa, Florida, USA, December 8-11 2008, pp. 1-4
42. Wachs, J.P., Stern, H., Edan, Y.: Cluster Labeling and Parameter Estimation for the Automated Setup of a Hand-Gesture Recognition System. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans **35**(6), 932-944 (2005). doi:10.1109/tsmca.2005.851332
43. Trigueiros, P.: Hand Gesture Recognition System based in Computer Vision and Machine Learning: Applications on Human-Machine Interaction. University of Minho (2014)
44. Alpaydin, E.: Introduction to Machine Learning. MIT Press, (2004)
45. Montgomery, D.C., Runger, G.C.: Applied Statistics and Probability for Engineers. Wiley, (1994)
46. Theodoridis, S., Koutroumbas, K.: An Introduction to Pattern Recognition: A Matlab Approach. Academic Press, (2010)
47. Rabiner, L.R., Juang, B.H.: An introduction to hidden Markov models. IEEE ASSp Magazine (1986).
48. Wu, Y., Huang, T.S.: Vision-Based Gesture Recognition: A Review. Paper presented at the Proceedings of the International Gesture Workshop on Gesture-Based Communication in Human-Computer Interaction,
49. Fink, G.A.: Markov Models for Pattern recognition - From Theory to Applications. Springer, (2008)
50. Camastra, F., Vinciarelli, A.: Machine Learning for Audio, Image and Video Analysis. Springer, (2008)
51. Reis, L.P., Lau, N.: COACH UNILANG - A Standard Language for Coaching a (Robo) Soccer Team. In: Birk, A., Coradeschi, S., Tadokoro, S. (eds.) RoboCup 2001: Robot Soccer World Cup V, vol. 2377. Lecture Notes in Computer Science, pp. 183-192. Springer Berlin Heidelberg, (2002)
52. Backus, J.W., Bauer, F.L., Green, J., Katz, C., McCarthy, J., Perlis, A.J., Rutishauser, H., Samelson, K., Vauquois, B., Wegstein, J.H., Wijngaarden, A.v., Woodger, M.: Revised Report on the Algorithmic Language ALGOL 60. In: Communications of the ACM, vol. 6. vol. 1, pp. 1-17. ACM, (1960)
53. Buckland, M.: Programming Game AI by Example. Wordware Publishing, Inc., (2005)
54. Millington, I., Funge, J.: Artificial Intelligence for Games, second edition ed. Elsevier, (2009)
55. Lieberman, Z., Watson, T., Castro, A.: openFrameworks. http://www.openframeworks.cc/ (2004). 2011

56. Bradski, G., Kaehler, A.: Learning OpenCV: Computer Vision with the OpenCV Library, 1st ed. O'Reilly Media, (2008)
57. OpenNI: The standard framework for 3D sensing. http://www.openni.org/ (2013).
58. King, D.E.: Dlib-ml: A Machine Learning Toolkit. Journal of Machine Learning Research **10**, 1755-1758 (2009).
59. Murphy, K.: Hidden Markov Model (HMM) Toolbox for Matlab. http://www.cs.ubc.ca/~murphyk/Software/HMM/hmm.html (1998). 2012