Automatic Analysis of Non-verbal Communication

Víctor Ponce, Sergio Escalera, Xavier Baró and Petia Radeva

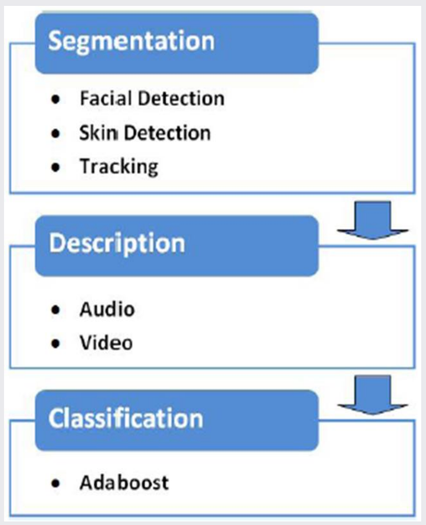
Computer Vision Center, Universitat Autònoma de Barcelona,
08193 Cerdanyola, Spain
Dept. Matemàtica Aplicada i Anàlisi, UB, Gran Via 585,
08007, Barcelona, Spain
v88ponce@gmail.com, {sergio,xevi,petia}@maia.ub.es

Centre de Visió per Computador

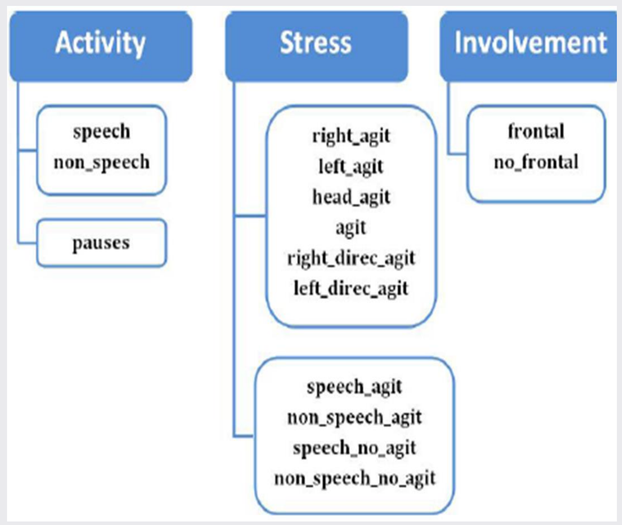
Abstract

Oral expression and communication is one of the most important competencies for personal, academic, professional and civic life [3]. According to the American Society of Personnel Administrators [1], it is considered that a good oral communication skill is important for obtaining a job, and for a good efficiency at work [2]. The main objective of this project is to obtain a Software tool that is able to obtain a series of features of a subject from automatic audiovisual analysis. The extraction of the features obtained from the oral and nonverbal language is something of particular interest in the analysis of psychological factors that a subject presents. This analysis is useful to improve the quality of oral communication: presentations, job interviews, etc. This is the ultimate goal of the project. The system has been applied to 15 end career project videos and presentations of fourth course students. It has been created a version that analyzes a recording and other that makes it in real-time via WebCam.

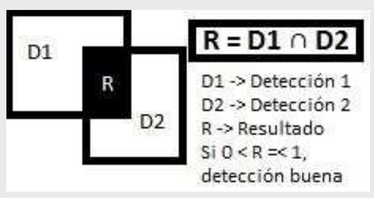
1. METHODOLOGY



Schema of the system for analysis



Descriptors: Group of features

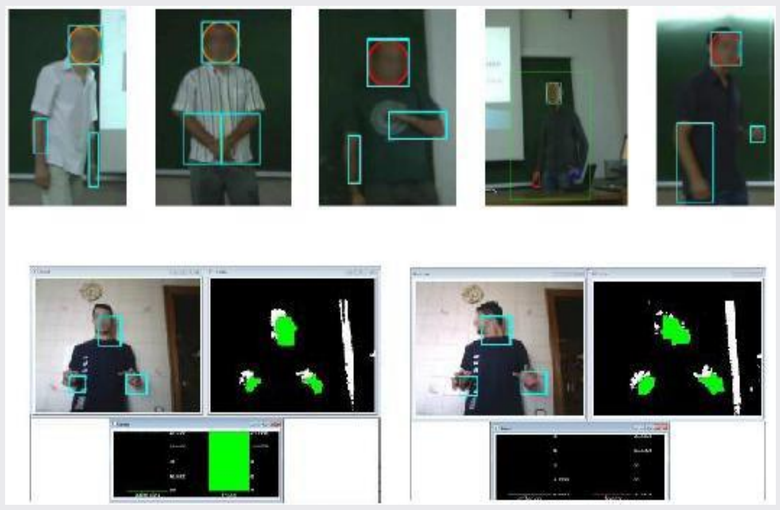


Facial detections intersections

Color model sample taken

Examples of detected regions for different students.

The regions of interest have been normalized regarding the detected facial area, in order to make comparable the values of the characteristics obtained by all students.



The two examples on bottom relate to a case of frontal look and the opposite case.

GUI : Detected regions on the left, segmentation on the right and statistics bars on the bottom center.

2. VALIDATION

The data analyzed consist on 30 videos recorded in presentations of students. Using Adaboost, we train a classifier which combines different simple decisions to obtain a strong hypothesis of the conversation. This method not only makes a selection of the most relevant hypothesis, but also provides a rule combination based on a weighted sum of the characteristics.

This method is used in two evaluation parts of the system, in order to find a classifier which separates between two main groups of conversations, those of higher ”quality” from those of fewer ”quality”. Moreover, it has also been used to analyze the order in which the characteristics are selected from higher to less relevance (ranking).

3. RESULTS

Evaluations:

Característica	Valor
Agit_cab	↑
Agit_direc_der	↑
No_Habla_No_Agit	↓
Agit_direc_izq	↑
Agit_izq	↑
Agit_der	↑
Habla	↑
Frontal	↑

Ranking of features chosed by Adaboost in order to split the best marks of the worst. At right: high or low values for discriminate the best grades.

Característica	Valor
Agit_cab	↑
Agit_direc_izq	↑
Agit_der	↓
Frontal	↑
Agit_izq	↑
Habla	↑
No_Habla_No_Agit	↑
No_Frontal	↓

Ranking of features chosed by Adaboost in order to split the best marks of the worst, by presentations and observers notes.

4. CONCLUSIONS

We presented a tool for automatic analysis of oral and gestural communication of students in public presentations. The system is able to automatically detect the regions corresponding to face, hands and arms, extracting a set of features that are analyzed by statistical classifiers. Results obtained on 30 videos showed the viability and usability of the system to obtain assessments of oral and gestural expression of the students, offering a ”feedback” that can be useful to improve the quality of their presentations. The most immediate future work is to increase the discretization of the presentations score, increasing from two to N ”quality” categories, in order to obtain a more accurate description of oral and gestural communication. We also want to include more accurate features for agitation and speech in order to differentiate between nervousness or involvement situations. These situations can be attacked directly by combining characteristics instead of individual indicators. For instance: The student speaks continuously but he agitates without paying attention to the public.

[1] D.B. Curtis, J. L. Winsor, and R.D. Stephens. National preferences in business and communication education. Communication Education, Vol. 38 (1), pp. 6-14. 1989.

[2] J. L. Winsor, D.B. Curtis, and R.D. Stephens. National preferences in business and communication education: A survey update. Journal of the association of Comunication Administration, Vol. 3, pp. 170-179. 1997.

[3] T. Allen, Charting a communicative pathway: Using assessment to guide curriculum development in a re-vitalized general education plan.. Communicative Education, 51(1) 26-39. 2002.