



A Voice User Interface for football event tagging applications



Sivio Barra¹, Alessandro Carcangiu², Salvatore Carta³, Alessandro Sebastian Podda³ and Daniele Riboni³

¹DIETI, University of Naples 'Federico II', 80138 Naples, Italy

²FootureLab, 07100 Sassari, Italy

³Department of Mathematics and Computer Science, University of Cagliari, 09124 Cagliari, Italy



Abstract

- Manual event tagging may be a very **long and stressful activity**, due the monotonous operations involved;
- For online video tagging, as for football matches, events to tag often consist of **many thousands of actions**;
- To overcome these issues, we enhanced the GUI of a football match tagging application, and **integrated it with a VUI**;
- Empirical tests revealed the efficiency and the benefits brought by our solution, with **reduced tagging time and error rate**.

Introduction

Graphic User Interfaces (GUI) have drastically improved the interactions between users and machines, so becoming a *must* for almost the totality of the developed software applications. Albeit some of them are extremely customised on the typology of the interacting users [1], they still present some drawbacks:

- individuals with physical disabilities may find it difficult to access a classic GUI application designed (as in the specific case) for a standard usage with mouse and keyboard [2];
- protracted and repetitive actions may be demotivating and stressful for the users, causing them to make mistakes [3];

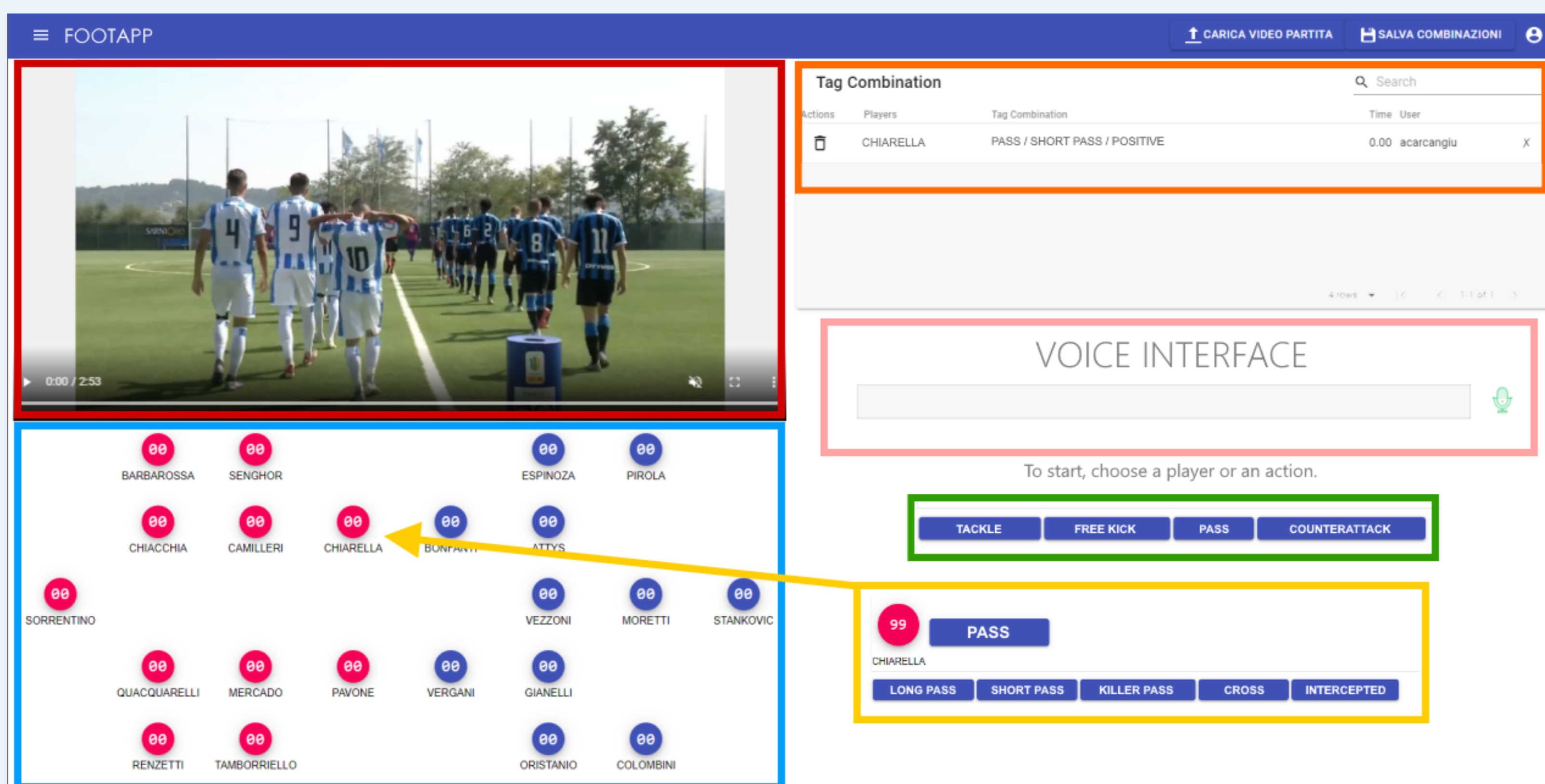
In this poster, we describe a voice user interface integrated into a GUI, specifically enhanced to handle specific interaction patterns which can be harder to address, aiming at fastening and improving usability in tagging events in football matches: the application has been called **FootAPP** (Football event Tagging Applications).

Football Match Event Tagging

Football event tagging aims at labelling all the events happening in a football match and involves:

- production of a **detailed report** about the movements of each player (acting in its own role as defender, midfielder, etc.) during the match;
- collection of **team dynamics**, to analyze forwarding and defending tactic movements;
- labelling a single action under two different points of view: (i) single player actions, expressed as *tag combinations* (e.g., MIDFIELDER > KILLER PASS > POSITIVE), and (ii) team actions (actions which involve more players, e.g. COUNTERATTACK);
- in addition to active plays (i.e., those whose object is the ball), there are a plenty of **passive plays** which need to be labelled, to properly track the team dynamics (eg., man-to-man/zone marking).

The FootAPP enhanced GUI



Since a football match lasts ninety minutes (excluding extra times), the entire tagging activities can take up to ~ 8 hours for, on average, **about 2,000 events**. The aim of the integrated VUI interface is then that of reducing the time needed for event tagging, reached by applying a **voice interface**, so to fasten up the tagging process, and by adapting the original GUI with **combined mode** (voice plus touch).

The FootAPP GUI is organized as shown in the figure above:

- the **TOP LEFT BLOCK** (in red), that features a video player, to reproduce the match; directional arrows ease the forward and backward skips of the video;
- the **BOTTOM LEFT BLOCK** (in cyan), which shows a virtual field in which the team lineups and shape are shown, with clickable player buttons;
- the **TOP RIGHT BLOCK** (in orange), which contains a summary of the tag combination records;
- the **FIRST MIDDLE RIGHT BLOCK** (in pink), with the voice interface that automatically activates when the user casts a command;
- the **SECOND MIDDLE RIGHT BLOCK** (in green), that shows the keyboard which activates as the user selects a player from the virtual field or spells his number (first level tags);
- the **POPUP BLOCK** (in yellow), that appears when a main event is activated, and provides second/third level options.

VUI Integration



The voice user interface has been developed by using the **Web Speech API** [4], which defines a complex interface (*SpeechRecognition*), providing a set of methods for transforming speech into text. A high-level overview of the implemented VUI structure is shown in the above figure.

Empirical Tests

From preliminary empirical tests, we noticed that total migration to a VUI would have partially obtained the expected benefits (actually, an average reduction of the 13% of the tagging time).

We adapted the original interface to optimize the use of a combined tagging mode (voice and touch), resulting in an **average reduction of the 28% of the time** (~ 2 hours for each full tagged match). This happens since both:

1. the processing time of too long voice commands (which, despite the robustness of the exploited API, can last two or more seconds), could negatively impact on the global time saving;
2. some tagging patterns still result quicker in the manual mode, whereas other tagging practices can obtain real benefits from the voice interface.

References

- [1] Kurtenbach G. P. *Graphical user interface for defining and invoking user-customized tool shelf execution sequence*. U.S. Patent No. 5,867,163. 2 Feb. 1999.
- [2] Wagner A., and Gray J. *An empirical evaluation of a vocal user interface for programming by voice*. International Journal of Information Technologies and Systems Approach (IJITSA) 8.2 (2015): 47-63.
- [3] Bouguettaya A., Yu Q., Liu X., Zhou X., Song, A. *Efficient agglomerative hierarchical clustering*. Expert Systems with Applications, 42(5), 2785-2797.
- [4] Adorf J. *Web speech API*. KTH Royal Institute of Technol.