Improvement of Spotting Words Speech Recognition Agent for Tele-Operating a Manipulator Arm

Mohamed Fezari[1], Ibrahiem El-Emary[2],
Mohamed Serghir Boumaza[3], and Hamza Attoui[1]

[1] Laboratory of Automatic and Signals, Annaba, BP.12, Annaba, 23000, Algeria
[2] King Abdulaziz University, Kingdom of Saudi Arabia
[3] Department of Electronics University of guelma
Mohamed.fezari@uwe.ac.uk, omary57@hotmail.com,
hmza.atoui@gmail.com

Abstract. This paper details the simulation and design of a voice tele-operating system for a robot arm (manipulator). The methodology adopted is based on hybrid technique that combines classical patterns used in automatic speech recognition system such as Crossing Zero, energy, Mel frequency cepstral derivative MFCC coefficients as features and Dynamic time warping as classifier. To implement the approach on a real-time application, a Personal Computer interface was designed to control the movements of a six degree of freedom (DOF) robot arm by transmitting the orders via radio frequency circuits. A simulation GUI was developed in order to guide the user in his commands. The speech recognition agent is tested and results of hybrid and classical techniques were compared. The hardware of the slave part (the robot arm) is composed from: a microcontroller from Microchip PIC16F628, a radio frequency wireless module and the power circuits.

Keywords: Speech recognition, Robot Arm, tele-operation.

1 Introduction

Human-robot voice interface has a key role in many application fields. Robotics has achieved its greatest success to date in the world of industrial manufacturing. Robot arms, or *manipulators*, comprise a 2 billion dollar industry. Bolted at its shoulder to a specific position in the assembly line, the robot arm can move with great speed and accuracy to perform repetitive tasks such as spot welding and painting. In the electronics industry, manipulators place surface-mounted components with superhuman precision, making the portable telephone and laptop computer possible [1-2].

Yet, for all of their successes, these commercial robots suffer from a fundamental disadvantage: lack of human voice control. A fixed manipulator has a limited range of commands provided by a manipulator. Mainly using a keyboard joystick or a mouse.

This paper proposes a new approach to the problem of the robot arm command. Based on the recognition of isolated words, using a set of traditional pattern recognition approaches and a discrimination approach based on test results of classical methods [3][6] and [7] in order to increase the rate of recognition.

The increase in complexity as compared to the use of only traditional approach is negligible, but the system achieves considerable improvement in the matching phase, thus facilitating the final decision and reducing the number of errors in decision taken by the voice command guided system.

Moreover, speech recognition constitutes the focus of a large research effort in Artificial Intelligence (AI), which has led to a large number of new theories and new techniques. However, it is only recently that the fields of robot arm control and AGV navigation have started to import some of the existing techniques developed in AI for dealing with uncertain information.

Hybrid method is a simple, robust technique developed to allow the grouping of some basic techniques advantages. It therefore increases the rate of recognition. The selected methods are: Zero Crossing and Extremes (CZEXM), linear Dynamic Time Warping (DTW), Linear Predictive Coefficient (LPC) parameters, Energy Segments (ES), and cepstral coefficients (MFCC: Mel Frequency Cepstral coefficients). This study is part of a specific application concerning robot control by simple voice commands. The application uses a set of commands in Arabic words in order to control the directions of four DOF robot arm. It has to be implemented on a DSP [8] and has to be robust to any background noise confronted by the system.

The aim of this paper is therefore the recognition of isolated words from a limited vocabulary in the presence of background noise. The application is speaker-dependent. Therefore, it needs a training phase. It should, however, be pointed out that this limit does not depend on the overall approach but only on the method with which the reference patterns were chosen. So by leaving the approach unaltered and choosing the reference patterns appropriately, this application can be made speaker-independent [9].

As application, a vocal command for a four DOF robot arm is chosen. The robot arm is "TERGANE -TR45". There have been many research projects dealing with robot control, among these projects, there are some projects that build intelligent systems [10-12]. Since we have seen human-like robots in science fiction movies such as in "I ROBOT", making intelligent robots or intelligent systems became an obsession within the research group. Voice command needs the recognition of isolated words from a limited vocabulary used in Automatic Robot Arm Control system (RACS)[13] . The paper is presented as follow: in section 2, we present the application, in section 3 the speech recognition agent is detailed, the hardware part is presented in section 4 and 5, and finally in section 6 and 7 we present the results of some tests and a conclusion.

2   Description of Application

The application is based on the voice command for a set of degree of freedom for a robot arm T45. It therefore involves the recognition of isolated words from a limited vocabulary used to control the movement of selected parts of the arm.

The vocabulary is limited to nine words to select the arm part (upper limb, limb, hand and forceps or grip) and to command the selected part (up, down, left right and stop). These commands are necessary to control the movement of the T45, Up movement, Down movement, stop, turn left and turn right. The number of words in the vocabulary was kept to a minimum both to make the application simpler and easier for the user.