IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications
8-10 September 2003, Lviv, Ukraine

Intelligent Control of the Lift Model

P.Cernys, V.Kubilius, V.Macerauskas, K.Ratkevicius
Department of Control Technology, Kaunas University of Technology,
Studentu 48-320, LT-3031 Kaunas, Lithuania, v.macerauskas@ktu.lt, http://ediaclit.vtk.ktu.lt

Abstract: *The paper provides the analysis of the lift model controlled by voice and sensor control panel. The model is constructed by using average powered controller, it has voice recognition, programmable terminal and logical lift program, which connect them all. The model can be easily reprogrammed in order to have various working regimes and thus is suitable for training students with the broad automation specialization. Presented system could be used in the smart house projects and, especially for disabled people. The modification of the well-known DTW (Dynamic Time Warping) algorithm was used. The set of voice commands consisted of eight Lithuanian words (floor number, go, stop) and two phrases (greeting and goodbye). The recognition accuracy of eight voice commands was equal 100%, two commands were recognized with some errors. The voice controlled lift model was demonstrated in the exhibition "Science-2002" in Lithuania.*

Keywords: - controller, driver, frequency converter, voice command, speech recognition

1. INTRODUCTION

Rational and accurate movement control of separate mechanisms, devices and other parts is very important in lots of technological equipment. Most frequently, executive devices are electric drives and collection, processing and forming of control algorithm is conducted by electronic-digital means in such systems. Automation specialists of the broad qualification are required for creating and exploiting such mechatronic systems. A lot of up-to-date control elements and software is used for these specialists' training. Therefore the model of the lift, which is suitable for training purposes because it contains all parts of a control system used in any industrial lift is presented in this article.

A lift is a characteristic electromechanical device controlled in electronic digital way containing controlled electric drives, differently realized positioning systems, quite a complex logic control system and having high requirements for reliability and safety. In order to have a model, which is as realistic as it's possible, it has an industrial design except the lift cage and its shaft model. It has no relay contact and up-to-date control elements are chosen for the model: cage drive is with the asynchronous motor and frequency converter, control logic is realized through easily programmed controller, sensors of the cage position are non-contact, control commands are given in a voice mode (reserve version is sensor programmed control panel).

Commanding/controlling of the computer and applications, using speech for handling the environment (smart house) are very promising fields, especially for disabled people. For those who have difficulties in entering data with other input tools like keyboard, mouse, etc., speech recognition is an effective alternative to alter or to combine input methods. People with motor disabilities can control various devices via speech input. One of smart house elements could be the control of a lift by voice commands. There are four main factors related with speech recognition problem:

- Speaker – each voice is unique; hence creating techniques that can accurately and reliably recognize anyone's voice and any dialect of a given language is a major challenge;
- Coarticulation – the spectral characteristics of a spoken word (or sounds within the word) vary depending on what words (or sounds) surround it;
- Speaking rate and style – people speak at different rates and with different pronunciations of the same sounds, thereby making it difficult to get stable patterns for sounds or words that can be used with all speakers and speaking rates and styles;
- Environmental conditions – speech can be difficult to recognize in home environments (background speech from radios or TV), when spoken in a car (road noise distortions), or in noisy backgrounds (airports, train stations).

Each of the above factors contributes some degree of variability to the speech signal. These sources of variability must be carefully considered when developing applications based on speech recognition technology.

2. THE PROJECTION BASED SPEECH RECOGNITION ALGORITHM

Model of a lift can be controlled by a small number of voice commands (floor number, stop and go), but the high recognition accuracy of voice commands is necessary. Speaker independent recognition would be more convenient, because it would not be necessary to train the recognition program for a new user. But speaker independent recognition is more complicated than speaker dependent one and ensures less recognition accuracy. Therefore speaker dependent projection based recognition algorithm was selected [1]. It is based on very small number of non-uniformly spaced feature vectors in opposition to 5-10 msec stated speech preprocessing.

The projection based speech recognition algorithm was developed as a highly phonetically motivated alternative to DTW (Dynamic Time Warping) or HMM (Hidden Markov Models) speech recognition models [2]. The projection method is based on phonetically segmented speech features. The segmentation function is smoothed using linear and nonlinear filtering. Each utterance then is divided to stationary and transient phoneme like segments. The average number of the stationary - transient segment pairs is approximately equal to number of phonemes per utterance.

0-7803-8138-6/03/$17.00 ©2003 IEEE

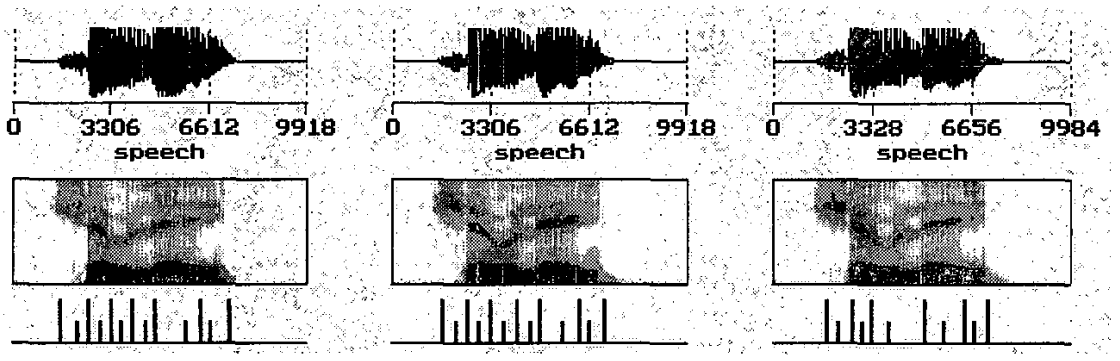

Fig. 1 – Three phonetically segmented speech examples. From top: a) speech, b) sonogram, c) stationary (lower) and transient (higher) segment markers.

Above are presented (Fig.1) the stationary and transient markers of the three different utterances of the same word, which has the following phonetic content: fricative-1st diphthong-sonant-2nd diphthong. The only single phonetic segmentation error in the right column between 1st diphthong and sonant is observed. The internal segmentation of diphthong was correct.

The comparison of test and reference patterns is performed using local similarity matrix where only segmented feature vectors are implemented. The projection is the max value of similarity, bound with the given row or column of similarity matrix and satisfying the restrictions of the global similarity measure. The essence of the restrictions: the global similarity measure is evaluated only from 2 or 3 diagonals, which show the greatest sum of projections.

The projection algorithm outperforms DTW (error decreases about twice) for the same phonetically segmented sequences [1]. The projection algorithm allowed to implement the "averaging" (the repetitions of a given phonetically segmented phrase). This feature can be applied for the averaging of many pronunciations of the same word and, in this way, it could help to create the more reliable word references or to detect and to collect the same phonetic units automatically [3]. The original method of detecting boundaries of words was used in projection based recognition algorithm [4].

3. THE MODEL OF A VOICE-CONTROLLED LIFT

A scheme of the model of a voice-controlled lift is shown in a Fig. 2.

All system is controlled by OMRON controller C200HX. Three connection units are used for connection with the object. These units enable to have three ways of cage position control:

- an exact stopping of the cage using two speed modes which can be switched on by using floor contacts;
- an exact stopping by switching on sensors of the cage position in the floor as the cage approaches the stopping place and forming the contour of the cage position control;
- the position of the cage is continuously controlled along all it's way, using a sensor of motor turning angle (Encoder).

Relay signals of the cage position are transmitted to the controller using the digital signal input unit ID212 (Input Control Unit). The impulsive signals of cage position sensors are controlled through the impulse meter unit CT021 (Input Control Unit). Continuous control of the cage position enables to conduct an exact stopping of the cage, to use various regulators of position contour and to tune their parameters. The asynchronous frequency drive of a lift "Frequency converter – Motor" is controlled by analogous signal through the output unit (Output Control Unit). That allows to obtain the smooth racing, stopping and the wanted speed of the cage. A way of the cage stopping is chosen with the help of a program. Its stopping is predestined in 6 positions or 6 floors.

The cage moves in a rectangular construction, which is divided into 6 parts, or "floors". There are relays and continuous action sensors situated in each "floor". The movement is transmitted from the motor (Motor) via the reducer (Reducer) and steel rope between pulleys to the cage. The motor is controlled by the frequency converter 393EV (Frequency converter).

The programmable terminal NT20S (Programmable Terminal) is used for indication of the cage position, the lift working regime and imitation of "call" and "go" buttons. It works together with voice recognition program, which is installed in the computer. Recognized voice command (command number) is transmitted from personal computer to controller through ASCII unit OMRON C200H-ASC02 (ASCII Unit).

The programmable terminal is programmed with the package NTWIN. Connection with the controller is supported through control panel of devices. Programmable terminal is supposed to have control panel variants, so called "windows". Certain "windows" in the panel display can appear when the cage stops or starts moving or by pressing buttons, which imitate states of the cage: "busy", "call", "send" or "emergency stop". Configuration of the cage buttons in separate windows imitates the cage control from landings or the cage itself. While the cage is moving it is not allowed to "open" "windows" of control buttons.

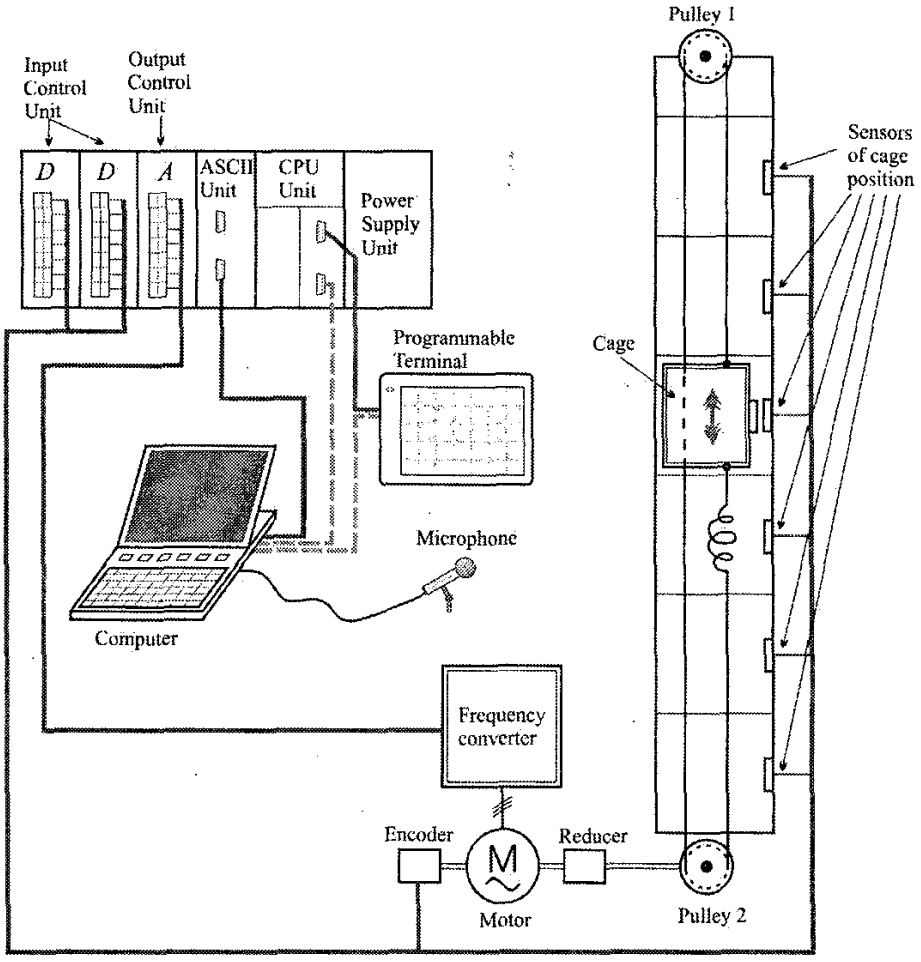

Fig. 2 – A scheme of the model of a voice-controlled lift.

Commands of the voice recognition program and control panel get into logical program of the lift control which carries out full control of the cage movement. It is created by using "Ladder diagram" method of programming on the basis of SYSWIN software package. Functional destination of the program:

- to set the position of the cage before starting to work;
- to call the cage to a certain floor;
- to conduct the accurate stopping of the cage in floors, according to chosen stopping regime as well as to conduct emergency stopping;
- to control all blocking signals.

Process of drive racing is formed in the frequency converter. Process of drive stopping can be formed both in the frequency converter and in the controller by using the positioning contour with the cage position sensor.

The described structure of the model covers the spectrum of technical and program means widely used in electromechanical and mechatronic systems. Usage of the average powered controller enables to explore various regimes of the lift work, to change them as well as voice control and sensor control panel commands easily. It is impossible to carry it out by using only logical controllers, non-contact or contact control schemes.

4. EXPERIMENTS

Testing of voice commands recognition accuracy was carried out for one speaker. The set of voice commands consisted of eight Lithuanian words and two phrases: pirmas (first), antras (second), tretchias (third), ketvirtas (fourth), penktas (fifth), sheshtas (sixth), sustok (stop), vazhiuok (go), kiek dabar laiko? (what's the time?), atchiu, viso gero (thank you, good bye) (Lithuanian words presented without Lithuanian transcription marks). Two pronunciations of each voice command and one averaged reference of both pronunciations were used during the training procedure. Each voice command was pronounced 100 times and the recognition accuracy was calculated.

Recognition accuracy of voice command "penktas (fifth)" was equal 98%, recognition accuracy of voice command "ketvirtas (fourth)" - 99%, other commands were recognized without errors. The recognition accuracy of ten voice commands is shown in the Fig. 3.

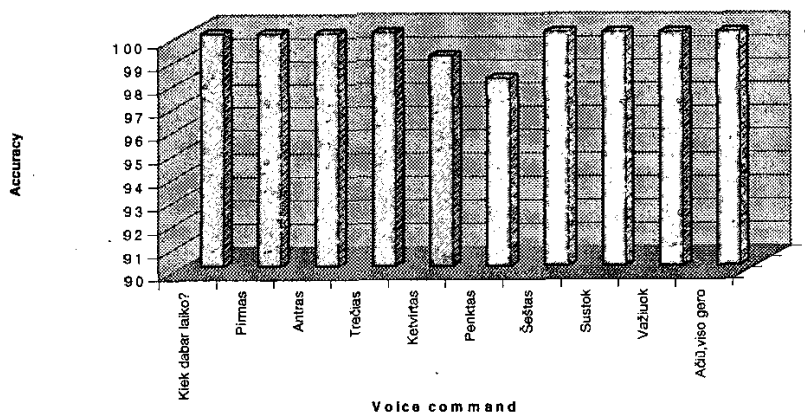

Fig.3 – Recognition accuracy dependence on the type of voice command.

5. CONCLUSIONS

1. The model of a voice-controlled lift is universal and it enables to realize regimes of real working lift by using up-to-date means.
2. A voice recognition program and its connection with the controller can supply a sufficient amount of commands necessary for the lift control.
3. The model of a lift is a useful tool for training students in specialization of automation, voice signal recognition and control technologies as well as for specialists' qualification improvement in similar specialization.
4. Voice controlled systems are especially useful for disabled people. Speaker dependent projection based recognition algorithm ensures a sufficiently good recognition accuracy of voice commands. It can be improved by increasing the amount of references and by selecting acoustically different voice commands. References can be collected from many speakers and averaged. The presented recognition algorithm in such way can be transformed into the "multi-speaker independent" one.

6. REFERENCES

[1] A. Rudzionis. *Isolated word recognition by fully phonetical word template. Contribution to the COST232 final report.* 1994. pp. 11-13.

[2] R. Cox, C. Kamm, L. Rabiner, J. Schroeter, J.Wilpon. Speech and Language Processing for Next-Millennium Communications Services. *Proceedings of the IEEEE*, Vol.88, No.8, August 2000, pp. 1314-1337.

[3] A. Rudzionis. Recognition by averaged templates. COST249: "Continuous Speech Recognition Over the Telephone". *Draft minutes of the 1st Management Committee Meeting*, Brussel, Belgium 1994, pp. 41-47.

[4] A. Rudzionis, V. Rudzionis. Noisy speech detection and endpointing. Voice operated telecom services. Do they have a bright future? *Workshop Proceedings*, Ghent, Belgium 11-12 May 2000, pp. 79-82.