LOGATOM RECOGNIZABILITY OF FINGER ALPHABET IN VIDEO

Heribanová P[*], Polec J.[+]

[*] FMFI Comenius University, Mlynská dolina, 812 19 Bratislava, Slovakia petra.heribanova@gmail.com
[+] FEI Slovak University of Technology, Ilkovičova 3, 812 19 Bratislava, Slovakia jaroslav.polec@stuba.sk

Abstract

This paper deals with the problem of the cued speech (fingers alphabet) recognition methods in video. Cued speech is a specific gesture language used for communication between deaf people. The aim of this paper is to show new objective method of testing consonant sign recognizability in single-handed finger alphabet (dactyl) analogically to acoustics. We used the "sign logatoms" to testing intelligibility because they have no meaning and cannot be part of common words. From the results we construct the minimum coded bit-rate recommendations for every spatial resolution.

Introduction

Cued speech is visual and spatial language and it uses three-dimensional space for communication. The basic communication element is gesture (= word). Finger alphabet is a system of finger and movement configurations that represent alphabetic characters. The number of characters is related to the number of speech sounds (phonemes) of the language. It is commonly used for purposes of clarification, such as unfamiliar words, names of persons, geographical names, or with words, for whose the asking person doesn't know the appropriate gesture. There are two types of alphabets in the Slovak cued language: the single-handed alphabet (Fig. 1) and the double-handed alphabet. [2]

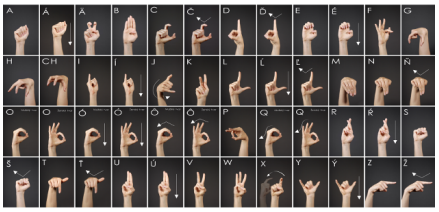

Fig. 1 Example of single-handed alphabet (dactyl) [3].

Intelligibility

The main difference between the terms quality and intelligibility is that the term "quality" describes the appearance of decoded video signal ("how" the viewer sees it) and the "intelligibility" is just one aspect of quality saying if the received information gives any sense ("what" the viewer sees in it).

Sound articulation

The inspiration of designing the methodologies for intelligibility is based on testing in acoustics.

$$Z = \frac{a}{b} * 100\% \quad (1)$$

The intelligibility is defined as a percentage of the correctly received elements or parts of speech "a" to the total number of transmitted "b" [4]. We can explore the intelligibility of video recordings: sentence and word intelligibility using gestures, while logatom and consonant recognizability using the finger alphabet.

Logatom recognizability

We shows a new objective method for examining the logatom recognizability (as used in telephonometry for speech sound articulation) with a use of subjective ACR method (full categorical evaluation). In logatom recognizability evaluation we use artificial monosyllabic words without meaning (logatoms) to mitigate people's tendency to correct the incorrectly understood consonants or words according to the meaning. We create so-called "sign logatoms" for the deaf – one sign in finger alphabet represents one speech sound in logatom. We produced video previews with seven different logatoms in Slovak single-handed finger alphabet (around 42 consonants). The length of the video previews is about one minute. For the whole experiment we used four different video formats of 1280x720, 640x360, 320x180 and 160x90 pixels per frame with 25 frames per second. Subsequently, these recordings were encoded by the H.264 codec in various bit rates (QP = 30, 40, 50 that corresponds to rates from 390 kbit/s to 4.5 kbit/s respectively).

The whole test consists of two parts:

1. Subjective, where the video was evaluated according to given voting options shown in Tab. 1.The respondent chooses the answer following his subjective feelings. General feelings of the respondent have an impact on the testing. The video sequence may cause different feelings.

1	Completely understandable
2	Partially understood, but understood the content
3	Partially understood, but misunderstood the content
4	Not understandable

Tab. 1 Proposed voting options for testing of logaom recognizability

2. Objective, where the respondent had to rewrite the consonants organized into logatomes to the letters of the Slovak alphabet. It is important to understand corect, because the meaning cannot be guessed from the context.

Resolution	QP	30	40	50
160x90	Objective evaluation [%]	90,24	71,95	0
	Subjective evaluation	2	3	4
	Bitrate [kbit/s]	18,3	7,2	4,5
320x180	Objective evaluation [%]	93,90	76,19	45,23
	Subjective evaluation	2	3	4
	Bitrate [kbit/s]	51,6	17	7,5
640x360	Objective evaluation [%]	96,49	83,30	50,00
	Subjective evaluation	1	2	3
	Bitrate [kbit/s]	149,2	42,1	18,4
1280x720	Objective evaluation [%]	95,23	95,12	93,90
	Subjective evaluation	1	1,5	2
	Bitrate [kbit/s]	389,4	126,9	56,2

Tab.2 Results from previous experiment

See Tab.2 and Fig. 2, the result in single-handed finger alphabet. Minimal video transfer speed depends on video format settings. With decreasing recognizability there was an increasing number of consonant interchanges, mostly between 'a' and 's', 'o' and 'f', and there was also higher frequency of missed or extra added consonants.

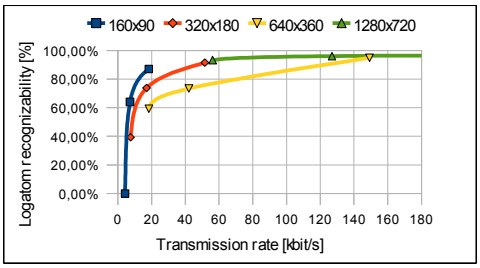

Fig. 2 Dependency logatom recognizability on the transmission rate.

This method is designed according to previous evaluation methodology clarity transfer Slovak sign language through video-conferencing and a subjective method of ACR, which is adapted for assessing the quality of video conferencing for the deaf. To interpret the results we apply the standard scale for subjective and objective methods for the quality and intellibility [1] and the percentual evaluation of intelligibility as in acoustics.

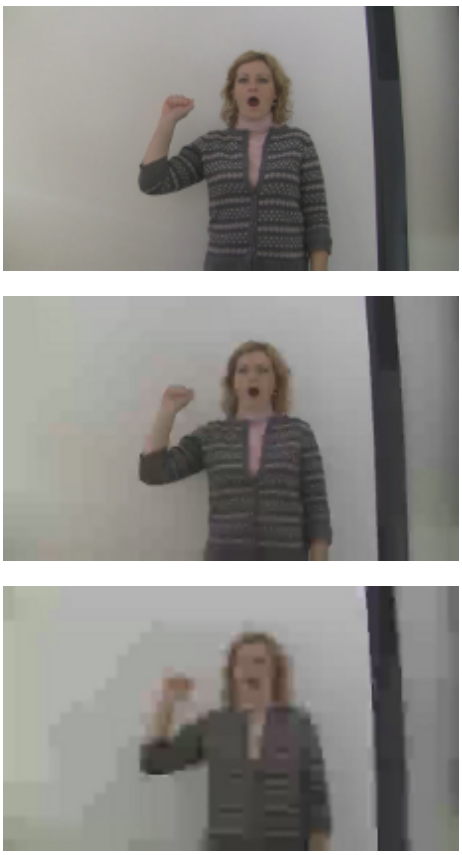

Fig. 3 Picture taken from experiment: first is original next decoded frames by H264 codec with parameter QP40, 50.

Conclusion

This paper describes the methodology of evaluating the quality of video signals based on logatom recognizability using so-called „sign logatomes" and also shows our obtained results in single-handed finger alphabet used by children and young people.

Acknowledgements

Research described in the paper was financially supported by the Slovak Research Grant Agencies: KEGA under grant No. 119-005TVU-4/2010 and VEGA under grant No. 1/0602/11.

References

[1] GRANÁT, M. *Objective methods for evaluation of audio signal quality* (in Slovak), Brno: Brno University of Technology, 2009.
[2] TARCSIOVÁ, D. *Pedagogics of hearing-impaired* (in Slovak), Bratislava: MABAG spol. s r. o., 2008.
[3] HEFTY, M. *Finger alphabet* (in Slovak), In Organization Myslím – development of thinking not only for hearing-impaired (in Slovak), 2009, www.zzz.sk
[4] MAKÁŇ, F. Elektroacoustics (in Slovak), Bratislava: STU, 1995.