Russia,
St. Petersburg

Voice of the Future

# Speech Technology Center

# TECHNOLOGIES

## Recently Developed

### Voice recognition based on automatic voice search in a pre-recorded database of voice samples

This technology provides automatic voice searching in a previously recorded voice database. Both the voice of the target speaker and the voice samples in the database are recorded in the telephone channel. All types of telephone sets, lines and Automatic Telephone Exchanges currently used are accepted. The speech is of arbitrary content, and spontaneously produced. The minimum duration of the sample recording is 100 seconds; the speech of the target speaker should be at least 15 seconds in duration. The system has a reliability of over 90%.

### Voice command recognition technology for speaker- and language-independent command recognition

This technology is intended for speaker-independent recognition of a fixed set (several dozens) of isolated voice commands. To adapt the system to any lexicon in any language it is enough to collect a database as pronounced by 30-50 speakers of the desired language. The reliability of the system is over 97% provided the commands are more than 1 second in duration. The recognition process is based on STC proprietary patented algorithm. The technology makes no use of the floating point and thus can be used with even the cheapest microprocessors.

### Echo cancellation technology improving speech quality and facilitating user speech comprehension when using a mobile telephone in a car "hands-free" environment

When used in an automobile environment this technology allows the loudness level of the far-end speaker to be reduced by 15dB and more. As a result of combining echo suppression proper with noise cancellation, the magnitude of echo suppression grows as the noise level increases. The technology makes no use of the floating point and thus can be used with even the cheapest microprocessors.

### Automatic speaker segmentation of multi-speaker telephone conversations

This technology allows a recording of two-speaker telephone conversations to be automatically divided into two separate segments. Two files are created, each containing the speech of one of the speakers. Fragments of simultaneous speech are discarded or, if one of the speakers sounds louder, it can be written to a louder speaker's file. The reliability of speaker segmentation depends on channel- and speaker-related spectral differences. It averages at 95% reliability..

### Noise cancellation technology for speech recordings allowing removal or reduction of negative effects of most known interference types on speech intelligibility

This technology allows simultaneous application of several noise cancellation methods highly effective against complex noises - undoubtedly the most frequent kind of real life noises. It is possible to select and apply any set of the 11 available methods, as well as to use any desired method repeatedly for both simple and complex processing. Each of the methods allows flexible adaptation to specific noises through modifying the parameter set. To process typical noises one can apply any of over 10 available standard procedures. Using such a procedure demands no more than a standard computer's capacity.

### Automated segmentation of speech recordings (into phrases, words, phones)

The technology provides speech/non-speech and speech/pause detection and semi-automatic segmentation of speech files into isolated speech chunks (words and/or phrases). A HMM-based segmentation module has been developed for preliminary semi-automatic speech phones segments. This module can manually correct both boundary markers and transcription labels. This was initially developed to facilitate segmenting phonetic and labeling of Russian digits, but it can be adapted to be used for language-independent broad phonetic segmentation and labeling of noise-free speech data.

### Speaker-independent speech recognition from a distance of 50-70 cm in highly noisy automobile environments

This technology is intended for speaker-independent recognition of a fixed set (several dozens) of isolated voice commands inside a car moving at speeds of up to 200 km/h. The technology includes noise cancellation and speech recognition. The reliability of the recognition is over 95% provided the commands are more than 1 second in duration. The microphone is mounted above the windscreen about 50-70 cm from the driver. The technology does not use the floating point and thus can be used even with the cheapest microprocessors.

**Contact us to know more about our current achievements and extensive development resources**


Speech Technology Center

Office: 4, Krasutskogo str. 196084, St.Petersburg, Russia ◆ Tel.: +7 812 325 8848 Fax: +7 812 327 9297 ◆ Email: sales@speechpro.com http://www.speechpro.com ◆ Postal address: P.O. Box 114, Lappeenranta SF-53101, Finland