Speech recognition: HEAD acoustics launches analysis software VoCAS

Evaluate voice control systems objectively and reproducibly

HEAD acoustics launches VoCAS (Voice Control Analysis System), the efficient software for evaluating speech recognition systems. From voice control in vehicles to the use of speech commands for smartphones, tablets or telephone hotlines: Automatic Speech Recognition (ASR) is already part of numerous applications today. VoCAS allows quick and objective quality evaluation of ASR systems under realistic and reproducible test conditions. The software takes crucial factors into account, such as background noise, language or accent which significantly influence the performance of voice control systems. VoCAS allows the use of use predefined test sequences for such ASR systems to determine their quality, to analyze the weaknesses of the systems and to optimize them based on the results.

Customized, reproducible and realistic test sequences

Depending on the device under test (DUT) and the requested test case, the appropriate test sequence can be defined in VoCAS. From speech commands for vehicle navigation (“Navigate to New York airport”) to speech commands for a call via mobile phone (“Call John Doe“): All possible commands for controlling an ASR system can be evaluated. Each test sequence consists of different elements and is processed sequentially. These elements are, for example, playing test sentences or background noises, inserting pauses for acoustical feedback of the voice control system or the evaluation of the DUT. All elements can be arranged flexibly, added as often as required and adjusted individually (volume, length etc.). Each test sequence is reproducible. For background noise, a wide range of realistic sound scenarios are available (e.g. cafeteria, vehicle, train station). The user can test the measurement object by choosing various parameter sets such as different speaker, languages, background noises, destination address or person to be called.

Database with existing voice recordings can be extended individually

For testing different voice control systems, audio source databases with appropriate speech commands are required. Audio databases in VoCAS can be individually expanded by importing own speech recordings. In addition, VoCAS provides an integrated recorder for recording individual speech commands easily and fast. Larger lists of imported or recorded audio files can be cut, filtered and adjusted to defined speech levels automatically. VoCAS also offers the possibility to manually tag keywords to each speech command. There are often recordings, which contain the same command, but which are nevertheless available in different acoustic variants, because different languages, speakers or user accents were used for the recordings. With the help of the tagging system, VoCAS systematically guides the user through the requested variants, creates the appropriate measurement sequence and helps to keep an overview.

Easy to interpret representation of test results

VoCAS convinces with clear representation of test results. Both, percentage values (e.g. 60 % of speech commands recognized, 40 % not recognized) as well as a colorful accentuation can be chosen for an optimal interpretation. Furthermore, a direct comparison of different voice control systems is possible. All available attributes (e.g. utterance, speaker, language, background noise) can be selected for result presentation. This enables the user to check which test sentence has passed or failed the test with certain attributes. The results can be exported to Microsoft® Excel® for further post processing.
VoCAS allows fast benchmarking of different ASR systems and software versions under realistic and reproducible test conditions. The analysis software is compatible with other HEAD acoustics products. The front end MFE VI.1 can be controlled via VoCAS for playback and monitoring of speech recordings via the artificial head measurement system and for mouth equalization. Furthermore, the background noise simulation systems 3PASS, HAE-BGN or HAE-car can be managed via the Voice Control Analysis Software.

About HEAD acoustics – Telecom Division

HEAD acoustics was founded in 1986 and has been involved in noise and vibration, electroacoustic and voice quality testing since its inception. HEAD acoustics is based in Herzogenrath, Germany, with affiliates in France, Great Britain, Japan, South Korea and USA as well as a world-wide network of representatives. The Telecom Division of HEAD acoustics manufactures telecom test equipment and provides consulting services in the field of speech and audio quality. Moreover, HEAD acoustics closely co-operates with DECT Forum, ETSI, ITU-T, 3GPP, TIA CTIA, GSMA and other standardization bodies with regard to the development of quality standards for voice transmission and speech communication. In many partnership projects, HEAD acoustics has proven its competence and capabilities in conducting tests and optimizing communication products with respect to speech and audio quality under end-to-end as well as mouth-to-ear scenarios.
Each sequence consists of different elements (e.g. play test sentences or background noises, pause, evaluation, level measurement) and is processed sequentially. The elements can be arranged flexibly, added as often as required and can be adjusted individually (volume, length etc.). Each test sequence is reproducible.

A typical test sequence for measuring in-vehicle navigation systems: Realistic background noises can be added. The speech command is played (Call HEAD acoustics in Brighton Michigan), followed by a pause for the acoustical feedback of the voice control system. At the end, the evaluation is made to check if the system executes the call.
Clear result representation: The results are displayed in percentage values and colorful for an easy and fast interpretation. Direct comparison of different voice control systems is possible.