Press release

Institute for Information Systems successfully researches speech synthesis:
Human voices from the computer - barely distinguishable from the original

Hof, 29.03.2021 - Especially for blind or visually impaired people, computer applications that read texts aloud are already a great help in everyday life. Even when driving, people have long since become accustomed to the friendly voices from the navigation system, which save drivers from dangerous distractions. Naturally, the new technology also harbors dangers. The Institute for Information Systems at Hof University of Applied Sciences is conducting a study to determine the acceptance of artificially generated voices and is developing its own models for the German market.

The quality of so-called speech synthesis has improved considerably in recent years. Whereas for a long time voices sounded rather tinny or choppy, the sound is gradually giving way to an increasing naturalness and unobtrusive speech dynamics. This also makes listening to longer texts more pleasant.

Rapid improvement in speech quality
"This has been achieved in international research through the use of deep neural networks. In the English-speaking world in particular, it is already almost impossible to distinguish between a real person and a program," says Prof. Dr. Rene Peinl, Head of the Institute for Information Systems at Hof University of Applied Sciences. Accordingly, a number of freely available models are available that speak English very naturally if sufficient training data is used. Speech generation usually takes place in two stages. First, a so-called Mel spectrogram is generated, which is a representation of the speech frequencies. From this, a vocoder then generates the actual audio signal. Both stages are neural networks that must be trained separately.

Acceptance on the test bench
The DAMMIT program at Hof University of Applied Sciences, which focuses on the technology transfer between universities and small and medium-sized enterprises for digital transformation, is analyzing how high user acceptance is for computer-generated voices. Test subjects are read text content of medium length - for example, messages half a screen page long. The steady improvement in the quality of speech synthesis that has taken place in recent years increases the convenience and possible uses of the technology on the one hand, but also harbors dangers on the other, since machine voices that sound human can of course also be used for fraud or criminal acts.

Many possible applications

Automated text reading aloud is currently being found in more and more areas of application. Being able to take in information even though the eyes have to focus on another target is an invaluable advantage: "Speech synthesis is of course an essential part of accessibility for people with visual impairments. In very practical terms, however, orders can be verbalized for forklift drivers, among others, which can be very helpful and time-saving in their workflow. Or one can have the daily news read aloud in one's personal favorite voice. In general, speech synthesis is also an important part of voice-controlled applications such as smart speakers, e.g. Amazon's Alexa," says Prof. Dr. Peinl, explaining some of the possible applications.

Market demand is growing

Yet the demand for automatically generated, but human-sounding voices, is likely only just at the beginning. One example can be found on the campus of Hof University of Applied Sciences at the Einstein 1 start-up center: The start-up company ahearo offers a service that allows people to listen to audio podcasts of content that is otherwise only available as text. Until now, these texts have been read in by human speakers. "Such a production is of course cost-intensive and also reaches its limits due to the limited availability of professional speakers. The collaboration with Hof University of Applied Sciences therefore opens up completely new possibilities for us," says Johannes Garbarek, founder and CEO of ahearo.

High speed and low cost

"For ahearo and other companies looking for a cost-effective and fast way to incorporate high-quality speech synthesis into their products, we are developing a solution for generating German speech from text," said Prof. Dr. Peinl. Freely available, self-generated audio data provided by ahearo is used to train the speech synthesis models in the best possible way. The evaluation is based on objectively measurable parameters as well as on subjective assessments of the test persons.

Encouraging interim results

The results obtained so far are encouraging and give reason to hope that the software will soon be used in practice: "Short sentences are already read out very well in our model. The challenges are still pauses and stresses in more complex sentences, as well as abbreviations, compound words and proper names," explains researcher Peinl. A small anecdote shows that the computer program sometimes has the same problems as humans: "For example, we have the word "early summer meningoencephalitis (FSME)" in our test texts. It is no wonder that not only we, but also the computers, have difficulties with such word monstrosities," says Professor Dr. Peinl.

Support

The results of the study, as well as the software developed in the course of the research, will be freely published and made accessible. The project is funded by the ERDF Bavaria 2014-2020 program, by the European Union through the Regional Development Fund, and by the Bavarian State Ministry of Science and the Arts. Another project partner is smartlytic GmbH, a software development and data analysis company based on Hof University campus.

Translated with www.DeepL.com/Translator (free version)

Press contact:
Rainer Krauß, Communicationsn / PR
Alfons-Goppel-Platz 1, 95028 Hof
Telefon: 09281/409-3006
E-Mail: **pressestelle@hof-university.de**

About Hof University:
Practice orientation, internationalization and intelligent use of resources are the focus of teaching and research at Hof University of Applied Sciences. In the area of internationalization, the university places a further focus on India, and with regard to the topic of intelligent use of resources, the focus is on water and energy efficiency. Our wide-ranging and interdisciplinary study programs range from economics and business law to computer science and engineering. The Münchberg campus offers an education that is unique in Germany thanks to textile and design study programs that are closely interlinked with business. Regional companies also benefit from the establishment of competence centers and institutes at the university. The four research institutes focus on information systems, materials science, water and energy management, and biopolymers. At Hof University Graduate School, working professionals and managers can find German and English-taught continuing education programs at university level; their program includes part-time bachelor's and master's programs, certificate courses, academic continuing education courses and seminars. The Bavarian-Indian Center for Business and Higher Education BayIND, which is affiliated with Hof University of Applied Sciences, coordinates and promotes cooperation between Bavaria and India. Students with an interest in startups or entrepreneurship are advised and supported by the Digital Startup Center Einstein1 at the university campus.