



Using AI to improve speech quality

INVISIO[®]

Introduction

Once the subject of science fiction, artificial intelligence is now a reality of the present day and is transforming the way we live as well as the way industries work all over the world. From manufacturing robots to self-driving cars, AI has already changed many sectors and functions beyond recognition compared to their predecessors or earlier iterations. Where traditional technology reaches its limits, AI offers the opportunity to redefine problem-solving and find solutions that would otherwise remain elusive.



Tactical communication within the military and public safety is no exception. While improvements have seen the provision of state-of-the-art communication technology that can operate within the most extreme environments, traditional technology has its limits. Until now.

INVISIO's latest generation of headsets and control units not only improve on previous models with traditional technology advancements, but also incorporate artificial intelligence to improve sound quality and speech intelligibility far beyond levels previously achieved.

INVISIO Using AI to improve speech quality

This paper will explain how AI has proved a game-changer in new technology from INVISIO, increasing speech intelligibility and removing unwanted noise to

achieve the aim - core to all audio communications specialists - to provide hearing protection while ensuring secure communication and a natural level of situational awareness. INVISIO has used AI to achieve a 30% increase in speech quality via its control units, as well as reducing the power consumption of the original algorithm by a factor of 80. The paper outlines the background

to the creation of the AI itself, the collection of data, and initial results from the products created using that AI, as well as plans for the future.

Artificial Intelligence

At its simplest form, artificial intelligence is a field, which combines computer science and robust datasets, to enable problem-solving inspired by the human mind.

Source: IBM Cloud Education

The story so far

For 20 years INVISIO has been designing military-grade communications systems used worldwide by military personnel, special forces and police. Its uses and applications mean ensuring high quality audio, speech intelligibility and hearing protection is paramount. INVISIO's personal communication and intercom systems not only provide situational awareness but hearing protection in some of the world's most extreme environments from loud noise to excessive heat or cold and even underwater.

Each new iteration of INVISIO's creations has addressed potential obstacles, difficulties and challenges in the operating environment, maximising technology to ensure it improves on previous versions. While the most recent versions of its headsets offer the latest in hearing protection and water protection as well as providing versatility, comfort and safety, there was still the potential for improvements.

INVISIO's X5 in-ear headsets use bone-conduction microphones to transmit speech from one user to another. The technology within the headset conducts the vibrations from the jawbone with a specially developed in-ear microphone. The signals then travel through a digital filter in the control unit and are converted into clear speech. The construction of the bone conduction microphone and its location inside the ear blocks surrounding sounds and only transmits the wearer's voice to the receiver.

However, the nature of bone-conducted speech can lead to sound degradation due to the lack of high-frequency speech, resulting in a muffled effect. That degradation, combined with an aggressive audio compression in digital radios, can lead to compromised quality when it comes to communication. While such

INVISIO Using AI to improve speech quality

a compromise may seem minor in some situations, in mission-critical environments it can have a significant impact. Put simply, if a message has to be repeated several times or cannot be understood by the receiver, the effects can be catastrophic. In addition, such sound degradation can have a more insidious impact, increasing cognitive burden on users and resulting in fatigue or a lack of situational awareness. Research has shown that a reduction in noise received through headsets or hearing aids can lower sustained listening effort¹, remedying some of these effects and therefore reducing cognitive burden on users.

Why AI?

Despite efforts to reconstruct the bone-conducted speech in various ways, the solution proved elusive. It became clear that the solution would require more than simply reconstructing something that wasn't there or trying to replace the missing high frequency. To solve the problem required first identifying the problem and that needed an entirely different approach. Enter AI.

The traditional way of solving problems in this field involves building a model or algorithm based on a series of assumptions. But that only works if you can clearly identify the problem and build a model accordingly. What happens when the problem is too complex for humans to model?

It is in complex situations like this that Artificial Intelligence can be key. AI's ability to absorb and analyse large amounts of data allow it not only to identify the obstacle but to solve it, without humans giving it a defined task or specific algorithm. Rather than telling the AI what the problem is, researchers feed it the audio data from a range of users and allow it to provide a solution.

This approach is a departure from traditional methods and has not only helped bring about a leap forward in the audio quality and communication capabilities provided by INVISIO's headsets and control units, but has opened the door to potential developments elsewhere in the world of mission-critical communication and also in other industries and sectors.

INVISIO's V-Series Gen II and AI: How we made it

INVISIO recruited several experts who were already working together with INVISIO on the very same problem at Danish Technical University (DTU) and had ascertained the need for AI to find a solution. They collected a range of data - recording people speaking into INVISIO's X5 headset with its bone conduction microphone as well as high quality microphones in order to produce degraded, low quality speech as well as good quality speech.

INVISIO Using AI to improve speech quality

Using the bone conducted speech as the input and the good quality speech as the output, the AI could then create its own solution to transform one into the other. It would essentially solve the issue itself and 'fix' the audio. The concept was proven - AI had provided the answer needed to improve sound quality within the headsets. But in order to deploy the algorithm in the relevant environments more work was needed to ensure as stable a solution as possible.

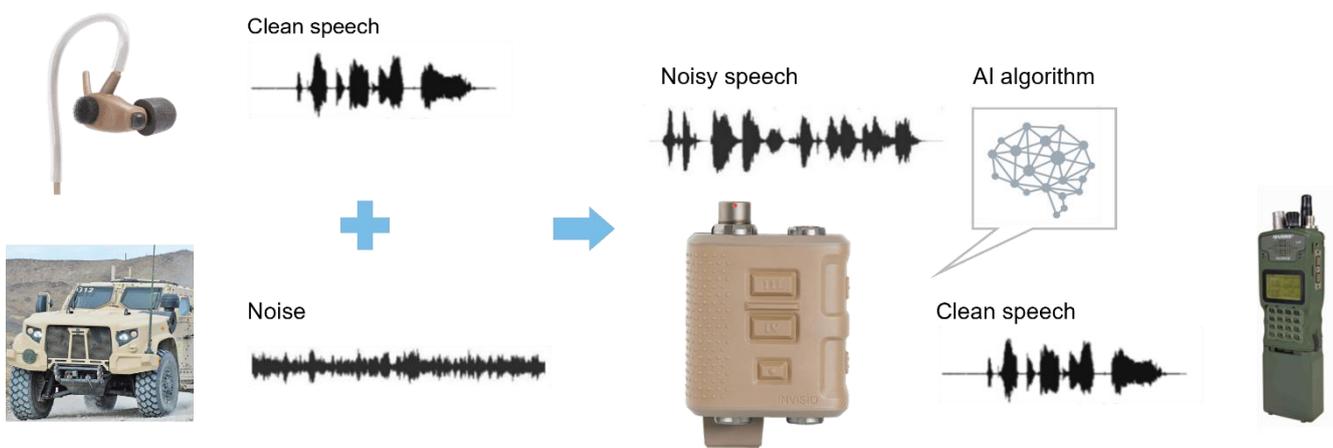


Figure 1: A mix of bone conducted speech and surrounding noise is picked up by the control unit and transformed into a high-quality speech signal that is transmitted over the radio network without degradation of speech intelligibility.

The team recognised that in order to make the AI more stable they needed more data. They recorded speech from multiple nationalities and different genders as well as including data involving a range of noise environments, mixing clean speech segments with a database of noises from machine noise to gunshots and helicopters, feeding that information to the algorithm so it would learn to operate in different environments.

Using more advanced equipment at INVISIO and a larger data set, they developed their initial proof of concept into an improved product that could be deployed into real-world situations.

Succeeding with Edge Computing

Whilst successful at solving the problem of sound quality, AI also brought its own challenges. Its power demands in contrast with the low-power nature of the headsets and units created by INVISIO meant the technology had to be reduced in size so it could practically be used within them in the field.

While the AI element had been tested on much larger networks and could run easily on a PC, it had to be 'squeezed' into a much smaller, lightweight device

INVISIO Using AI to improve speech quality

that piggybacks on radio power. That meant removing any unnecessary elements while testing to ensure it still provides the desired result. Such a complex task required more advanced techniques and investigation into various network structures to achieve the optimal balance between power and size.

Edge Computing

is computing that's done at or near the source of the data, instead of relying on the cloud at one of a dozen data centers to do all the work.

Source: The Verge

Work in this area achieved a reduction by 80 times - making the AI 80 times more efficient than the original algorithm that had been developed. That reduction in size and power consumption is what has made INVISIO's latest creation truly ground-breaking.

In addition, once downloaded to the control unit, the AI then exists locally on the soldier, ensuring secure, efficient communications. This development is part of a wider move towards edge computing within tactical networks that sees high performance computing and storage deployed to the tactical network's edge, supporting a range of applications including situational awareness, signal and image gathering, and other Internet of Things-based applications². The ability of the AI to operate offline provides added reassurance in terms of security, protecting communications between users from cyberattacks and any compromise of the data passed between users.

Running AI at such low power on the edge of the network is a vast step forward, not only in the world of military communication, but for potential applications across other industries and sectors where both practicality and security are key.

Success: Our results

INVISIO analysed the results of the AI installed within its Gen II headsets and found clear improvements to speech quality. Analysis using the Perceptual Evaluation of Speech Quality (PESQ) - an accepted family of standards comprising a methodology to assess speech quality - found a 30% increase in speech quality, with such improvements key to the ability of users to perceive what is being said in critical situations.

INVISIO Using AI to improve speech quality

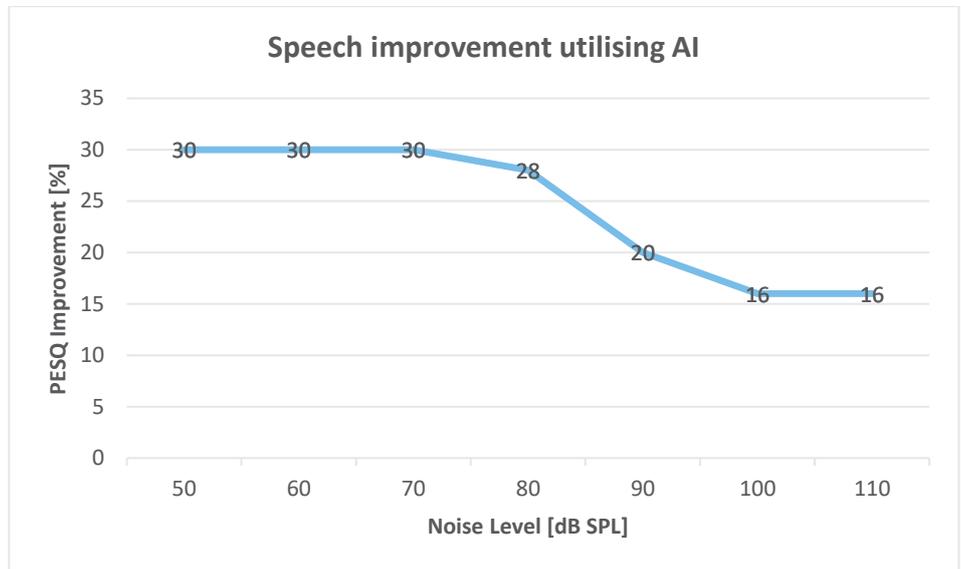


Figure 2: Up to 30% improvement was achieved when measuring speech quality in relation to noise and comparing the new AI-powered V-series Gen II and the X5 headset with the existing V-Series.

Additionally, a large field test with the Swedish Police produced anecdotal evidence to suggest that the use of the AI within the headsets contributed to a reduction in cognitive burden in those users.

“My colleagues have never heard my transmissions this clear with previous in-ear headsets ever!”

“I have not needed to repeat my transmission a single time after I started to use AI.”

Source: Feedback from large field test

After the field test, which involved 400 users, INVISIO collected feedback from 50 of them via questionnaires in which they were asked about their experience.

The feedback was very positive, yielding comments reflecting, that colleagues had heard transmissions more clearly than with previous in-ear headsets, and users broadcasting transmissions did not have to repeat themselves thanks to the AI.

What next?

The inclusion of AI within INVISIO’s technology has proved a game-changer in increasing speech intelligibility and removing unwanted noise. By undertaking its own research and developing its own in-house competencies in this area, the potential to develop and expand the technology is clear.

Results of the research have already exceeded expectations and brought unexpected benefits, with further research and developments already underway. While the algorithm itself used in the Gen II control units is fixed and operates on the soldier, there is the potential for work to build on the leap forward that it has achieved. This includes improvements to the algorithm

INVISIO Using AI to improve speech quality

itself to make it even better and the addition of further noise reduction capabilities. The use of more speech data can also be used to develop new firmware for the control units in the near future. There are also possibilities to expand the use of AI to other potential uses within INVISIO's area of expertise.

Conclusion

As in so many other sectors, artificial intelligence has proved a game changer in providing a solution to a hitherto unsolved problem within communication using bone conduction in-ear headsets. The introduction of AI in this scenario has allowed a huge leap forward, achieving a 30% increase in speech quality compared to previous technology. INVISIO's research has also provided a solution for the obstacle of power consumption, reducing it by a factor of 80 which in turn has allowed the practical deployment of the technology in the field. The use of AI and the ability to reduce its power consumption by such a huge factor has created a new generation of communication technology and paves the way for further advances in the future.

From INVISIO's perspective, the incorporation of AI into its latest generation of technology is simultaneously a giant step forward whilst also only scratching the surface of the capability of AI to problem-solve within the sector and to provide solutions to so-far-unsolved problems and obstacles when it comes to mission-critical communication.

It remains to be seen how AI will transform hearing protection and communication for the users of the future. But INVISIO is excited to be at the forefront of such an exciting frontier for the industry.

Related Links and Resources

1. https://journals.lww.com/ear-hearing/Abstract/9000/Hearing_Aid_Noise_Reduction_Lowers_the_Sustained.98497.aspx
2. https://www.army.mil/article/216392/four_future_trends_in_tactical_network_modernization

About INVISIO

INVISIO offers cutting-edge personal communication and hearing protection systems. The systems enable users to operate and communicate safely and clearly in all environments, even under extreme conditions, such as loud noise, heat, and underwater. INVISIO systems consist of headsets and advanced control units that interface to a wide range of communication devices. The systems provide hearing protection while maintaining the natural level of situational awareness. Sales are made via the headquarter in Copenhagen and sales offices in the USA, France, Italy as well as a global network of partners. INVISIO is listed on Nasdaq Stockholm (IVSO).

www.invisio.com

INVISIO®

© 2020 INVISIO A/S. All rights reserved. Specifications are subject to change without notice. The Soft Spring™ system and Bone Conduction Microphone Technology are patents of INVISIO A/S. INVISIO® and IntelliCable® are registered trademarks of INVISIO A/S.

INVISIO - info@invisio.com