Human-Oriented IoT Solutions Using Hearable Technology from NEC

FURUTANI Satoshi, KOSHINAKA Takafumi, OOSUGI Kouji

Abstract

The relationship between IoT, AI, and the network technology that connects it is often described using metaphorical comparisons with the brain, the nervous system, and the five senses. The Internet experience, for example, is shifting from one where user and media are isolated from one another, their interaction mediated by monitors and speakers, to one where technologies such as augmented reality (AR) and mixed reality (MR) are used to partially or totally immerse the user in a virtual physical space. A related concept - already in use - is called omnichannel, in which all potential communication channels - both physical and digital - are integrated to provide the customer with a seamless, flawless experience that combines real-world physicality with mobile convenience. Now, more and more people are talking about the coming wave of hearable technology, which goes beyond earbuds and hearing aids to advanced computational devices that you wear in your ear and which interface you with AI, robotics, and IoT, all without interfering with your activities in physical space. This paper introduces NEC-developed ear acoustic authentication technology and a geomagnetic indoor positioning system which will be core technologies of NEC's hearable systems.

Keywords

wearables, hearable technology, hearables, AR, MR

1. Introduction

IoT, AI and the network technologies that connect them are often compared to the human nervous system and brain. It goes like this: When IoT devices are connected to networks as real-world sensors, sensations - we could call them heat or pain, for example - cause the actuators to activate semiautomatically or automatically, instantly transmitting the relevant data to the AI - or artificial brain - on the server, which then provided appropriate real-world feedback. According to Dr. Takeshi Yoro, an anatomist who has written numerous books on the subject, while there are five (six) senses that provide inputs to the brain, the brain only outputs muscle contractions. In other words, when you receive input from your senses, you might respond (output) by vibrating your vocal chords, closing your eyes, or moving their limbs. To put it another way: To impel a human to take a particular action, it is necessary to transmit an appropriate input through one or more of the five senses.

So, for instance, if you mistake a shriveled tree for a ghost, you may be paralyzed with fear or you may turn and flee. Your responses are the same whether your sensory input is real or imagined - which means that in terms of obtaining a desired "output," it doesn't much matter whether the "input" is real or not.

We have now reached the point where new technologies such as AR and MR make it possible for us to access information in the real world without having that information mediated by a display. In other words, these technologies have the potential to induce people to take action in the real world based on data that appears real but exists only in virtual space.

Until now, the strength of a "real" business resided precisely in its ability to provide an experience that could not be known without "being there." Now even that advantage is being eroded by advancement in digital technology that promise not only to recreate reality, but to create an alternate world on top of the real world as was evident in the Pokémon Go phenomenon that took the world by storm in 2016. This kind of mixed or layered reality is where we are headed and it is critical that we have the tools to navigate it safely and effectively.

So far we have talked about sensory input and how it relates to new "reality" technologies that are fast coming on-stream, in which the visual, auditory, and

tactile senses in particular all play important roles. In this paper, we will focus on what is known generically as "hearable technology", that is, technology based on the auditory sense and implemented in or near the ears.

Up to now, the internet experience has been largely two-dimensional, mediated through a browser and relying on various tools such as a mouse, keyboard, and display. The design of online services, including electronic commerce, assumes the user will be interacting via this two-dimensional interface.

However, from now on, it is expected that we will have to cope with situations where a connection with cloud AI is up and running all the time, on the assumption that users are not in front of the display.

In the near future, humans will no longer need to learn how to use user interfaces ingeniously designed in a two-dimensional perspective. Instead, they will be able to operate computers using natural tools acquired as infants, such as talking, using their hands to pick up or manipulate objects, nodding in agreement, shaking their head, and so on.

When you think about how, since the advent of the smartphone, we now operate devices using just our index finger and thumb or just our voice alone, it is safe to say that such a future will almost certainly come soon.

However, if the user is no longer tied to a monitor, this introduces new difficulties for the service providers in the cloud. Nothing is more important to building and maintaining customer relationships than the ability to capture the activity and behavior of the user in the real world. Hearable technology is an ideal solution, providing a smart, wearable interface that can help businesses capture "who," "where," and "what condition."

2. IoT Solution Example Using Hearables – Kinetic Management of Workers at Factories

Hearable technology is a new technology incorporated in wearable in-ear-devices. It provides businesses with the ability to acquire and act upon user data, while providing users with the ability to interface with online services without being conscious of the UI (**Fig. 1**).

When the workers plug the hearables into their ears, their hands and eyes are free to concentrate on their work and other real-world activities. At the same, they will be able to acquire auditory information as required and interact by speaking.

For example, in a human kinetic management solution, the hearable technology offers functions that include constant sensing of not only who the wearer of the hearable is and where they are - even when they are indoors such as inside a factory or underground
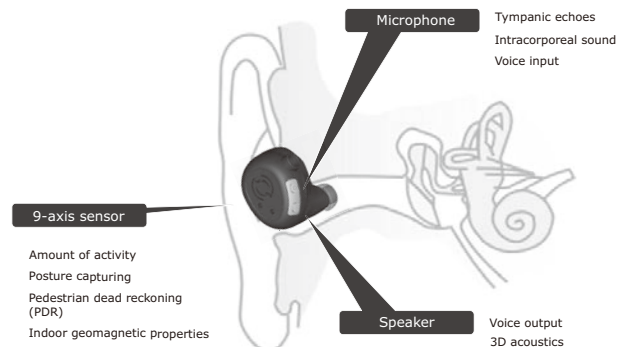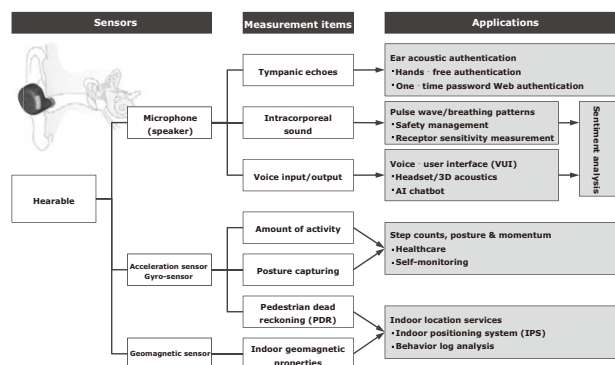

Fig. 1 Features of a hearable.

Fig. 2 Examples of information that can be obtained by a hearable.

where GPS cannot reach, but also how they are moving, whether or not they are standing still or falling, and what they are talking about (**Fig. 2**).

The acquired and measured data is transmitted to a cloud-based server via a smartphone using wireless networking such as Wi-Fi or Bluetooth. Analysis of the acquired and measured data in combination with other sensor data and system information will also make it possible to surmise various other user conditions.

What's more, the results of the analysis performed in the cloud can be transmitted back to the devices using a synthetic voice based on rules specified in advance. This enables the wearable system to be also used in situations where both hands are occupied with a task and where notification should be confidential.

In the next section, we will look at biometric technology which takes advantage of the characteristics of the ear and indoor positioning technology which utilizes the properties of geomagnetism - as representative examples of the hearable technology offered only by NEC.

3. Ear Acoustic Authentication Technology

User authentication is one of the most fundamental

functions of a mobile device. Some familiar examples range from unlocking the screen of a smartphone to settlement of online transactions. Undoubtedly, as Internet services grow ever more diverse and complex, we will see an even wider range of possible applications. Developing an easy, convenient, yet secure method of user authentication is crucial under these conditions. NEC is conducting R&D into unique biometric authentication system that will streamline the authentication procedure for users and at the same time will work with hearables.

3.1 Operation Principles

The human ear has a different shape from individual to individual. The ear canal, in particular, has various characteristics which vary widely from one person to the next. These characteristics include length, width, as well as the number, angles, and positions of the bends (**Fig. 3**). These characteristics are as unique as facial features and an individual can be accurately identified by acoustically measuring the characteristics of the ear canal — in other words, by using sound to measure the differences in the shapes inside the ear. This is the basic idea of ear acoustic authentication.

Each musical instrument has a different tone color depending on its configuration. Larger instrument have a lower pitch, while smaller instruments have a higher pitch. If we liken ears to musical instruments, then ear acoustic authentication is a technology that recognizes an instrument based on its color tone.

Now let's take a look at how user authentication is performed in actual cases. Unlike a musical instrument, the human ear does not make any sounds. So sound is sent to the ear canal from outside and the reflected sound is observed (**Fig. 4**).

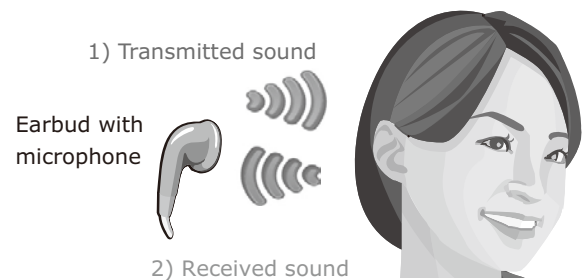To use ear acoustic authentication, an earbud with microphone is required. Unlike ordinary headphones that

Fig. 4 Usage scenario of ear acoustic authentication.

are provided only with speakers that reproduce sound, earbuds with microphone also incorporate a function to receive sound. There are many of these products currently on the market, so there is nothing particularly special about them. In recent years, noise cancelling earbuds have become popular because they allow users to listen to music or telephone conversation even in a noisy environment. These are simply another variation of earbuds with microphone, so they can also be used for ear acoustic authentication.

Once sound has been transmitted from the earbud with microphone and reflected sound has been observed, the procedure from then on is purely in the realm of software processing. In other words, the frequency of reflected sound for about a second is analyzed to compute the respective strengths contained in that sound, ranging from low-frequency (low-pitched sound) components to high-frequency (high-pitched sound) components. The computed results are then aggregated into about 20 feature values. These aggregated feature values are then compared to the user's previously obtained feature values to see if there is a match. In this way, it can be determined whether the person wearing the earbud with microphone is the registered user or not[1].

3.2 Main Features

Ear acoustic authentication offers many features ideal for user authentication.

1) High-speed
Transmission and reception of audio signals for only 1 second of acoustic signals is all that is required. Authentication processing is completed instantly.

2) High-precision
At a false acceptance rate of 0.01 to 0.1% and a false rejection rate of 2 to 3%, practical accuracy for user authentication has already been confirmed (based on the NEC testing conditions).

3) Low-load
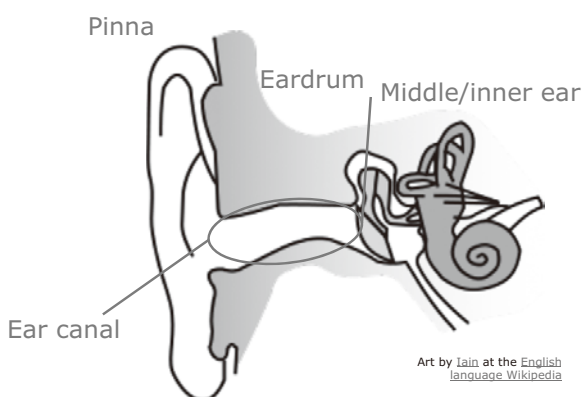No particular action - such as placing a finger over

Fig. 3 Anatomy of the human ear.

a sensor - is required, so the load on the user is very low. Authentication can be performed as many times as necessary. Continuous authentication is also possible.

4) Spoofing-resistant

The authentication process is completely invisible from outside, making it impossible to even know whether or not authentication itself is being performed. Stealing other people's biometric information is extremely difficult.

3.3 Future Prospects

Ear acoustic authentication is a new technology. As NEC's original AI technology continues to progress, accuracy and speed will be further increased. Currently, a dedicated sound called a sweep signal is used for authentication. In the future, however, it is expected that it will be possible to use almost any sound - for example, the music the user listens to every day or telephone conversations - for authentication. Also expected to take place in the future is authentication using inaudible (ultrasonic) sounds that the human ear cannot pick up.

NEC's commitment to biometric authentication technology has been recognized in prestigious international evaluations where we have been repeatedly ranked at the top in testing performed by the U.S. National Institute of Standards and Technology (NIST). This is a testimonial to NEC's internationally superior technology. Although ear acoustic authentication is a technology still under development, we will continue to provide our customers with reliability and safety as we move toward the forthcoming age of the hearable technology.

4. High-Precision Indoor Positioning Technology Using Geomagnetism

NEC is currently conducting R&D into indoor positioning technology using geomagnetism with a view to applying it to hearables. While it is well-known that the geomagnetic field can be disturbed indoors by ambient factors such as steel frames, it is less well-known that this disturbance is staticized. By taking advantage of this property, we can measure and record the distribution of these geomagnetic disturbances beforehand and subsequently compare that data with the measured data of a positioning location.

4.1 Capturing Geomagnetic Data by Using Fluctuations

Because geomagnetic data is expressed using XYZ orthogonal coordinates obtained from a magnetometer, single-point data alone results in the generation of massive amounts of data, making positioning difficult. This problem can be solved by treating geomagnetism as something subject to change. Due to the properties of the magnetometer, a phenomenon called offset deviation - in which a reference point deviates - takes place. The degree of changes, however, remains constant regardless of the occurrence of offset deviation. In other words, once these changes have been recorded and the magnetometer settings have been adjusted appropriately, there is no need to perform calibration. This is of critical importance when implementing this technology in actual social situations.

4.2 Utilization of RAPID Machine Learning

To convert the geomagnetic-based indoor positioning technology into a practical tool, NEC is utilizing the image analysis version of our AI technology called RAPID Machine Learning. There are three reasons to use this.

The first reason is that it is based on convolutional neural network (CNN) technology - one of the deep learning technologies. Although disturbances of the geomagnetic field are static, they occur irregularly. If you try to apply any type of machine learning other than deep learning or estimation theory to geomagnetic disturbances, you will not be able to design a hypothetical model that serves a basis for inference because there is no fixed law that applies to these disturbances. This is why it is considered difficult to create a positioning system using geomagnetism. With deep learning, however, not only is the irregularity of these disturbance a problem, it is considered a distinguishing characteristic. Moreover, what is derived from the results of deep learning is a logic to specify positions - which are mere network models. As an added benefit is that this system does not increase storage requirement by generating excessive data - a common problem in sensor information analysis.

The second reason is that RAPID Machine Learning uses image analysis. In other words, location identification is performed by looking for similarities between the learning data and positioning data. Conversion of the data into images facilitates analysis in case the positioning does not work well. When conducting technological development, it is more convenient and less time-consuming for us to compare how similar the images are in the learning and positioning data than to compare lists of data. This is because comparison of similarities is intuitive. Additionally, when the data is converted into images, we can take full advantage of a wealth of traditional cartographic techniques that emphasize various features. It is also easy to compress the data in order to

reduce learning time.

The third and most important reason for the adoption of the image analysis version of RAPID Machine Learning is that the users do not need to be aware of or understand the configurations of neural networks. With conventional deep learning libraries such as TensorFlow, users have to configure the neural networks themselves - something they will be unable to do without expert knowledge. The image analysis version of RAPID Machine Learning, on the other hand, allows anyone to perform expert-level operations. This is something that can only be achieved by RAPID Machine Learning because it is extremely difficult to decide on an optimal neural network that is best suited to each location to be positioned. Also the high-speed learning capability of RAPID Machine Learning significantly contributes to the speed of technological development.

4.3 Combination with Pedestrian Dead Reckoning

Yet another important technology in our indoor positioning system is pedestrian dead reckoning (PDR). This system uses an acceleration sensor and gyroscope to track movement in relative positions. Sole reliance upon either the acceleration sensor or PDR inevitably causes situations where you cannot perform positioning correctly. When they are combined, however, one technology can compensate for the disadvantages of the other. Due to its properties, the PDR technology does not function correctly unless the axis or the sensor is fixed or the movement of the sensor is constant. This requirement can be satisfied when the sensor is incorporated in the hearable because it is fixed on the user's body.

5. Conclusion

NEC will continue to expand the functionality of spatial information-oriented business tools and platforms including hearables. Currently, we are working to develop functionality that will support health management and mental healthcare of employees based on the vital data collected by the hearable. Similarly, we are developing functions to support communication with the elderly in which the hearable will be used as a "smart" hearing aid.

Active utilization of hearables not only lets users stay connected to what is happening around them, but also makes it possible to predict what will happen in the future more quickly and more correctly. Based on those predictions, we will create feedback systems responsive to user behavior. Through these activities, we are confident that we will be able to continue to offer solutions that will increase the safety and security of people's lives.

* Pokémon GO is a trademark of Nintendo, Creatures Inc., and Game Freak Inc.
* Wi-Fi is a registered trademark of Wi-Fi Alliance.
* Bluetooth is a trademark owned by Bluetooth SIG, Inc.
* FIDO is a trademark of FIDO Alliance.
* All other company names and product names and logos that appear in this paper are trademarks or registered trademarks of their respective companies.

Reference

1) T. Arakawa, T. Koshinaka, S. Yano, H. Irisawa, R. Miyahara, H. Imaoka, "Fast and accurate personal authentication using ear acoustics," Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Dec. 2016.

Authors' Profiles

FURUTANI Satoshi
Senior Manager
NTT DOCOMO Sales Division

KOSHINAKA Takafumi
Senior Principal Researcher
Data Science Research Laboratories

OOSUGI Kouji
Assistant Manager
Business Development Division

Information about the NEC Technical Journal

Thank you for reading the paper.
If you are interested in the NEC Technical Journal, you can also read other papers on our website.

Link to NEC Technical Journal website

Vol.12 No.1   IoT That Supports Digital Businesses

Remarks for Special Issue on IoT That Supports Digital Businesses
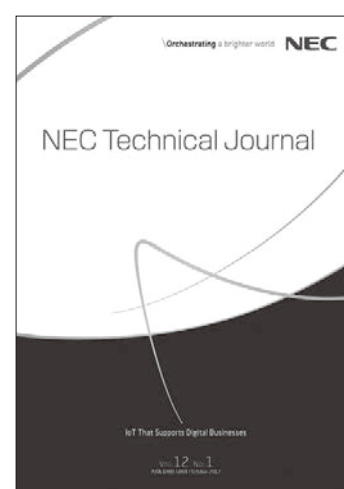NEC's IoT Operations That Support Digital Businesses

Papers for Special Issue

Platforms built to support IoT
An IoT Platform to Support Business Transformation - "NEC the WISE IoT Platform"
Edge Computing Supporting Customer Values in the IoT Era
Edge Computing Technologies to Connect the Missing Link of IoT
Case Studies of Edge Computing Solutions

IoT solutions that offer value to customers
NEC Industrial IoT - For Manufacturing in the Age of IoT
Warehouse Product Inspection System Achieves Work Efficiency and Quality Improvements
Warehouse Staffing Optimization Solution Using Autonomous and Adaptive Control - NEC's latest AI technology
Human-Oriented IoT Solutions Using Hearable Technology from NEC
Video Streaming Technology That Supports Public Safety
IoT and AI Innovations for the Retail Industry
Wireless Networking Technology for Real-time Remote Control of Factory Equipment: Wireless ExpEther
Lightweight Cryptography Applicable to Various IoT Devices
PoC of AI Demand Forecast Deployment in the NEC Group's Manufacturing Facilities from an Ethnographical Perspective

General Paper

"My Number" Collection Service Utilizes Several Key Image Recognition Technologies

NEC Technical Journal

Vol.12 No.1
October 2017

Special Issue TOP