3-3 Multimedia Information Hiding

TAKIZAWA Osamu, SONODA Kotaro, YOSHIOKA Katsunari, and SUZUKI Jouji

We research the multimedia information hiding (MIH) which is a technology to overlay digital information on multimedia contents such as picture, sound, document, etc. As application to disaster management of MIH, we have developed a new technology called "PAIH" (Public Address Information Hiding) which is a technology to overlay information such as position or destination on siren sound of urgent vehicles. Siren sound is allowed to degrade of sound quality, but PAIH is required to decode correct information from siren sound with pitch shift (Doppler effect) or under noisy environment.

Keywords
Digital watermarking, Steganography, Information hiding, Public address sound, Siren

1 Introduction

Multimedia information hiding (MIH) is a technology that overlays digital information onto images, sounds, text and other multimedia contents. In this day and age when the amount of multimedia contents created and distributed is rapidly increasing, MIH is an important application security technology[1]. MIH can be broken down into subcategories depending on the objective of the information overlay. Subcategories include digital watermarking, which is used to claim rights by adding copyright information to contents; digital fingerprinting, which is used to identify the party distributing information by adding unique information for each recipient; and steganography, which safely communicates confidential information to specific recipients while hiding the fact that the information was overlaid.

We have continued our research on a new method of information hiding focusing on documents[2][3]. In striving to further apply MIH technology to disaster management, we have developed Public Address Information Hiding (PAIH), a technology that communicates detailed warning information such as position and details about events to surrounding areas by overlaying that information onto siren sounds emitted by ambulances and other emergency vehicles[4]-[7]. We will summarize and evaluate PAIH in this paper.

2 Summary and purpose of PAIH

2.1 PAIH summary

PAIH is a technology that allows siren sounds overlaid with information to be generated and transmitted through the air and to be extracted through terminals that receive sounds. The hardware consists of a transmitter that generates siren sounds embedded with information and a receiver that receives those siren sounds and extracts and displays the embedded information.

2.2 Differences between digital audio watermarking and PAIH

PAIH is different from common digital audio watermarking in that it is intended simply for warning sounds whose only purpose is to attract attention. The information overlay allows a tolerance of original sound distortion as large as is perceptible. While common digital audio watermarking is used mainly on digitized music files to protect musical copyrights and the like, PAIH is intended for sounds transmitted through the air, so the embedded information has a reliable robustness and does not disappear even if the sound quality suffers because of interference from noise in the surrounding area.

Audio Orthogonal Frequency Division Multiplexing (OFDM) aimed at overlaying information related to broadcast speech has been proposed as a digital audio watermarking technology for sound transmitted through the air[8]. Audio OFDM is a method that keeps only low-frequency components, in which most basic frequencies of broadcast audio signals are concentrated, in their original state while substituting diffusion signals through information for high-frequency components. Relaxing restrictions on sound quality deterioration has allowed stronger resistance for transmission by air and enabled greater amounts of information to be communicated by this method. However, Audio OFDM is essentially a medium for embedding long, complex sounds with a lot of space, such as music and speech signals. In contrast, PAIH is intended for very simple audio signals like sirens, whose acoustic properties are extremely simple and have little redundancy. Applying information hiding to these types of sounds has never been examined before because it was difficult for past digital audio watermarking methods to deliver sufficient quantities of transmittable information for such sounds.

Furthermore, since warning sounds are often emitted by emergency vehicles traveling at high speeds, and since signal decoders are often on the move as well, receivers receive sound signals that have been shifted by the Doppler effect. The Doppler effect is a pitch scaling effect that alters the pitch of a sound in relation to time with respect to the speed at which the source of the sound is moving toward the point of reception and vice versa.

Digital watermarking for music is also expected to withstand a pitch scaling attack of around 5% (in other words, the watermarked information does not disappear even if the pitch changes up or down 5%) as an assessment item on its resistance[9]. However, the pitch shift added to the sound received from a vehicle traveling at high speeds on a highway is greater than 5%, and the change in relation to time is pronounced. What's more, the sound changes discontinuously at the moment of intersection. For example, if the sound source passes in front of the point of reception at 80 km per hour, there is a ±7% pitch shift in relation to time. Music contents lose their value when sound quality deteriorates because of discontinuous pitch shifting, and it has been determined that the need to detect watermarking vanishes in those cases. This is why we did not decide to apply digital audio watermarking to cases where the sound source and point of reception are moving.

2.3 PAIH applications

PAIH could possibly function as a secret transmission of information to specific recipients or as a broadcast to the surrounding area of detailed warning information related to sirens. An example of the former would be a mobile telephone terminal equipped with PAIH. If a fire engine passed by with its sirens wailing as you walked through town, you could very quickly acquire information about where the engine was headed from the sound of the sirens. An example of the latter would be car navigation systems equipped with PAIH. If an emergency vehicle approached with its sirens wailing, the navigation systems could receive information, such as which direction the engine is headed, that could help drivers avoid getting in the way of the emergency vehicles.

It is easily possible to use wireless communication to transmit the kind of information transmitted by PAIH without overlaying information onto sound. However, with respect to universal information communication, the relationship between warning sounds and events to be communicated need to be connected in a format that people can perceive is important. In other words, since information is embedded in the warning sound itself, PAIH will enable people to intuitively understand the relationship between the warning sound that they can hear and information received and extracted from it by a machine. In addition, transmitting information via warning sounds is sufficient when the sounds can reach the area in which all intended recipients are located. Using wireless communication in those cases would cause the information to travel outside of that area without purpose and could cause confusion. Furthermore, it is important to prepare redundantly as many methods of communicating information as possible for disasters, emergencies and other unforeseen situations, and PAIH could be one of those methods.

We have developed PAIH to be used on siren sounds emitted by Japanese ambulances. The sounds we referenced are defined by an announcement from Commissioner of the Japanese Fire and Disaster Management Agency[10] as two repeating sounds, one high and one low, with the basic frequency of the high tone at 960 Hz and that of the low tone at 770 Hz, with the length of each tone set to 0.65 seconds for a complete cycle of 1.3 seconds. These are rough values; there is some wiggle room in information overlay to account for some measure of modification to the composition and length of the warning sounds.

In the following chapters, we will discuss the results of comparative examinations on multiple methods of applying PAIH intended for use on this type of siren.

3 Method 1: embedding for a continuous length of time

3.1 Summary

As we explained in Section **2.2**, PAIH requires a method which is not affected by the pitch shift. What sticks out about changes undergone during transmission through the air are audio signal component distortions. Reverberations and the Doppler effect cause the composition to change. However, a large percentage of those changes are signal expansions and contractions with respect to time, and the distortion is minor. Thus, it should be possible to extract information by aligning the continuous length of time of the simple tone emitted (0.65 seconds for each tone) with the embedded information, expanding and contracting that length of time, and measuring that expansion and contraction on the receiving end.

Figure 1 shows a summary of Method 1. Here, when we contract the length of time from 0.65 seconds to 0.55 seconds, the value is zero, and when we expand it to 0.75 seconds, the value is one. Since we embedded one bit for each 0.65 seconds or so, the transmission rate is around 1.8 bps.

3.2 Evaluation

We did not use sounds transmitted through the air in our evaluation; we performed a simulation with SN range +10 to −20 dB due to white noise inside the computer to evaluate the robustness. In addition, we performed experiments simulating the Doppler effect (sound source moving at 80 km/h; receiving end at 5 km/h; in random directions). The vertical axis of Fig. 2 represents the rate of information extracted properly (detection rate), and squares denote cases in which there was no Doppler effect while circles indicate that there was a Doppler effect.

The results of the experiments are as follows:

- It was possible to extract all information when there was no Doppler effect, even when the acoustic pressure of background noise was greater than the warning sound (negative SN ratio).

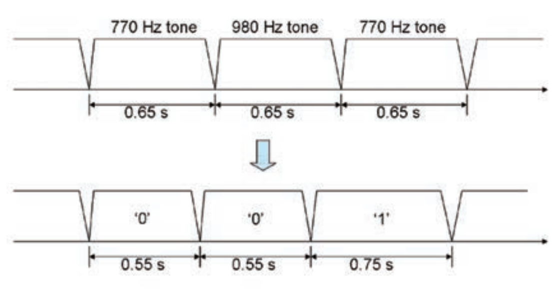


Fig.1 Summary of Method 1 (horizontal axis represents time)

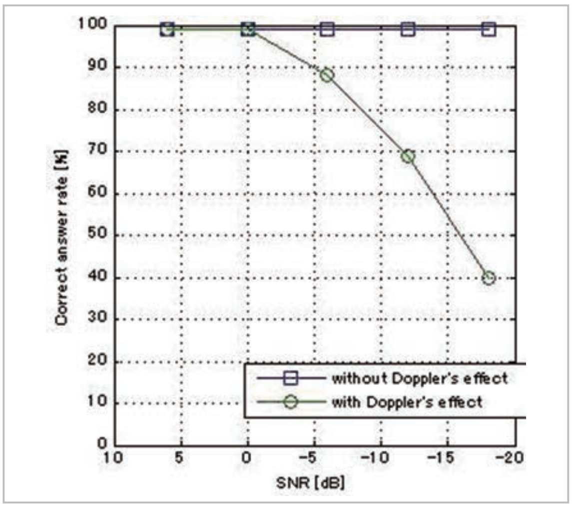


Fig.2 Results of Method 1 evaluation

- It was possible to extract all information when there was a Doppler effect when the acoustic pressure of background noise was less than the warning sound (positive SN ratio), but the detection rate plummeted when background noise was greater.

We performed our evaluations without knowing the extent to which the Doppler effect was present in simulations. Thus, we calculated the cross-correlation between each of the simple tones of 770 Hz and 980 Hz and the sounds received to detect the continuous length of time, which is the amount of time the relationship of the pitch of that value continued. We calculated some acute correlation values for cases where there was no Doppler effect, but the correlations shrank when there was a Doppler effect and, as a result, it became easier to be buried within background noise. Thus, robustness to background noise was strong in cases where there was no Doppler effect, but resistance could be considered weak in cases where there was a Doppler effect. As a way to improve on this, we could incorporate things like estimating the maximum velocity of emergency vehicles and considering the range of pitch shifts into basic frequency detection.

Problems other than the detection rate could be due to the fact that the amount of information that can be embedded is small, and that siren sounds falling out of rhythm will sound very strange. Thus, we have determined that it would be difficult to put Method 1 into effect.

4 Method 2: embedding into harmonic overtones

4.1 Summary

Method 1 performed well against background noise, but when the Doppler effect was present, it was weak. When a pitch expands or contracts, it changes all frequency components of the original signal to m times that frequency. Here, we examined the method of adding harmonic overtones that fit with embedded information to the siren sounds. Harmonic overtones become chords when they overlap and, acoustically speaking, they are pleasant to hear. We assigned m_0 for situations where the harmonic overtone ratio showed “0” for information; and m_1 when it showed “1” for information. The sound source of an ambulance is actually emitting a single frequency component (f_0), but when harmonic overtone m_0 is added, the frequency component splits into two, f_0 and m_0f_0. Meanwhile, when the Doppler effect causes the system to undergo d times the pitch expansion and contraction, the frequency components at the point of reception are df_0 and dm_0f_0. In other words, since the basic frequency and harmonic overtones are both undergoing the same pitch expansion and contraction, the relationship between components that display information is maintained at m_0. If we extract information at that time, we first analyze the frequency of signals received and then identify basic frequency df_0. Next, we investigate the strength of each component on each harmonic overtone of df_0 and select the strongest elements to be embedded by estimating dmf_0 added to information overlay and calculating m, which is the ratio between dmf_0 and df_0 already identified. Figure 3 presents a summary of this method.

4.2 Evaluation

As we did for Method 1, we did not use sounds transmitted through the air to evaluate

Method 2. We performed a simulation with SN range +20 to −10 dB adding white noise within the calculator to evaluate the robustness. In addition, we performed experiments simulating the Doppler effect (sound source moving at 80 km/hr; receiving end at 5 km/h; in random directions) the same as in the Method 1. Figure 4 shows the simulation results. In the experiment, seven types of information bits were represented by seven different harmonic overtones. The bit rate was $\log_2 7/0.65$ because we embedded information once every 0.65 seconds, and we were able to transmit more information than we were when using Method 1.

The results of the experiments are as follows:

- It was possible to extract all information bits when the warning sound had acoustic pressure at least 10 dB greater than background noise, regardless of the presence of the Doppler effect. The SN ratio must be 5 dB or greater in order to achieve a detection rate of 70%.
- There were no significant differences when we compared results respective to whether or not the Doppler effect was present. Thus, this method can be said to be robust against the Doppler effect.

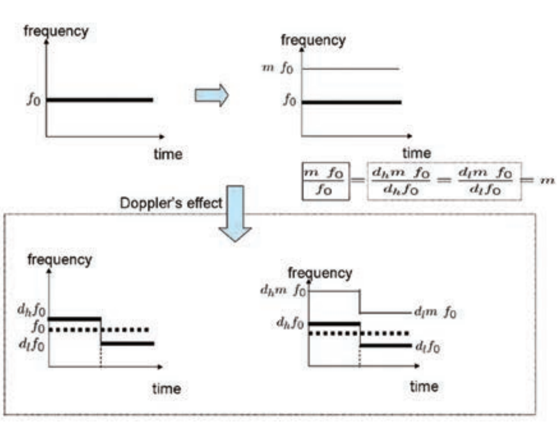


Fig.3 *Summary of Method 2*

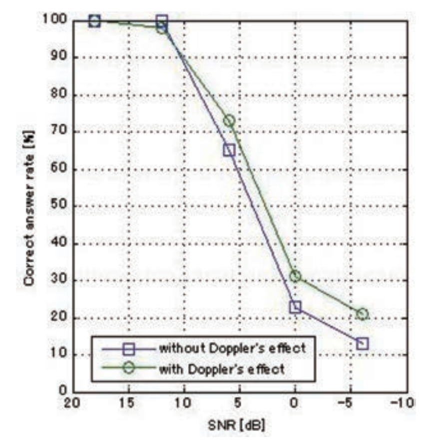


Fig.4 *Results of Method 2 evaluation*

This method is inferior to Method 1 in terms of resistance to background noise; however, it resists the Doppler effect better, is more acoustically pleasing, and can transmit a greater amount of embedded information than can Method 1. Thus, we are going to continue to add improvements to this method.

5 Method 2 improvements

5.1 Summary

Method 2 resists the Doppler effect well but is weak against background noise. We tried to raise robustness to background noise by incorporating spread-spectrum encoding and error correcting code into the transmit code strings for Method 2. The code strings transmitted showed values of 0 and 1. Corresponding harmonic overtone ratios were k_0 and k_1, respectively, which had values of 2 and 1/2, respectively. As Fig. 5 demonstrates, each code transmitted was diffused in the direction of the time axis by spread code strings (PN grouping) shared between the signal generating and signal decoding ends. We estimated code strings transmitted by taking the scalar product of the signal decoding end and the PN grouping. In addition, we added parity bit strings for every five bits of transmit code

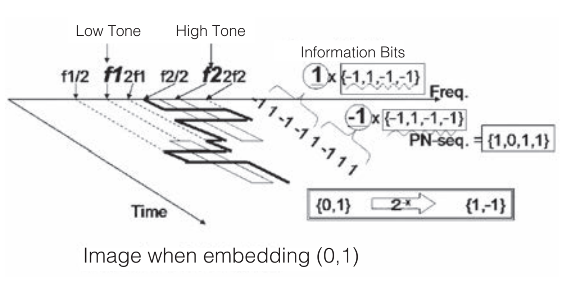


Fig.5 *Adding Spread-Spectrum to Method 2*

strings through (15,5) -BCH error correcting code. It was possible for the error correcting code to correct up to three errors for every 15 bits.

5.2 Evaluation

We evaluated how information detection for this improved method resisted white noise and pitch shifts (Doppler effect on the signal encoding side due to motion of sound source) by performing a computer simulation. Figure 6 shows detection error rates when messages were sent at 4 bps. The p_w in the figure is the maximum width of the coefficient for pitch shifts undergone by the siren signals at 0.5-second intervals, and it represents what arrived on the signal decoding end after all frequency components of signals transmitted were multiplied by $1-p_w$ to $1+p_w$ over 0.5-second intervals.

Through the results of this experiment, we verified that error rates for detection of transmitted bits were held below 10% when pitch shifting was $p_w=0.1$ or less, even in environments with white noise at a −5 dB level. As we discussed in Section **2.2**, p_w is around 0.07 when the sound source is moving at a maximum of 80 km/h, so this method could be precise enough for practical use.

We evaluated the proposed method at first not by transmitting sounds through the air but by adding white noise and pitch shifts to the inside of the computer. The results of this evaluation showed us that the improved version of Method 2 had promise, so we continued by building a prototype system based on the improved version Method 2 and by evaluating the detection rate on sounds that were actually transmitted through the air. Figure 7 shows what the evaluation experiment looked like. The speaker connected to the computer on the right side emitted siren sounds, and the microphone connected to the computer on the left collected those sounds and detected information.

The results of this experiment showed us that, though it was possible to detect information in the open air to some extent, it was nearly impossible to do so in closed spaces such as meeting rooms. We suspected that the effects of reverberating waves were the cause. In environments where sound reverberates, such as closed spaces, direct waves from the transmitted sound reach the signal decoding end as well as reverberations from the original sound some time afterward (essentially, duplications of the direct waves). This significantly disturbs the process of decoding the code diffused throughout the time axis.

5.3 Further improvements and implementation of Method 2

We decided to make an additional improvement to Method 2 to lessen the effects of reverberations described in Section **5.2** by diffusing code over the frequency axis instead of the time axis. The harmonic overtone ratio k(n) with respect to diffused code grouping m(n) is defined by the following formula:

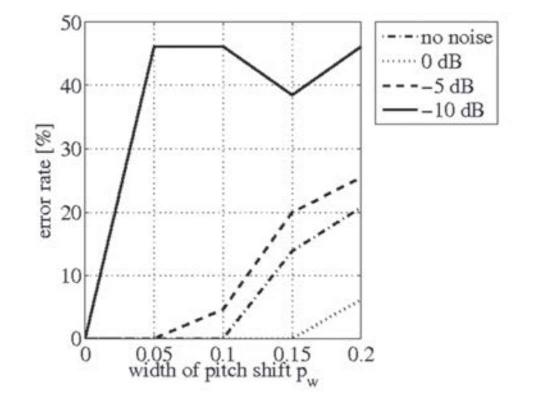


Fig.6 *Improved Method 2 detection error rate on white noise and pitch shifts*

Fig.7 *Experiment on detection rate from transmission through Air – Evaluation Results*

$$k(n) = n + 1 + \alpha \times (-1)^{m(n)}$$
$$(n=1 \ldots N,\ 0 < \alpha < 0.5)$$

If k(1) makes a harmonic overtone of 2, $k(1)=2+\alpha$ if $m(1)=0$; and $k(1)=2-\alpha$ if $m(1)=1$.

In that case, message m=1. Thus, if the harmonic overtone is skewed a bit high, it will send the "0" message, and if it is skewed a bit low, it will send the "1" message. Siren sounds overlaid with information will simultaneously emit the basic frequency of the siren itself as well as n simple tones from harmonic overtone components determined by the harmonic overtone ratio. Figure 8 describes siren sounds embedded with information by using this method. The vertical axis represents frequency, and the horizontal axis represents time. We can see that each harmonic overtone vibrates up and down in response to the "0" and "1" messages.

When we first started developing this method, we imagined being able to free text to overlay as information since we were thinking of space to transmit arbitrary messages. However, the slow transmission rate and detection errors showed us that allowing free text would not be practical. Thus, we decided that the signal generating and signal decoding sides would share a message table and send only strings of symbols defined in relation to the messages to be sent. As a result, the freedom of options for messages would be limited, but both the apparent message transmission volume and the rate of information detected correctly would likely improve drastically since information to be overlaid would be limited to short strings of symbols defined on the message table.

We developed a prototype system equipped with the above improvements in 2009. Figure 9 shows the structure of this system.

This system is composed of a signal generator that generates sirens overlaid with information and a signal decoder that extracts information from those siren sounds that come out of speakers. We will discuss the details of each component below.

(A) Signal Generator

1. Electric siren signal generating function
   Generates siren signals based on an announcement from Commissioner of the Japanese Fire and Disaster Management Agency[10].
2. Message selection function
   Presents defined messages that can be selected from the message table.
3. Error correcting code (ECC) function
   Creates error correcting code.
4. Diffusion treatment (PN) function
   Performs diffusion treatment on bit strings in the direction of the frequency axis.
5. Information embedding function
   Converts the message selected through the message selection function into strings of symbols based on the message table, and then overlays them onto the

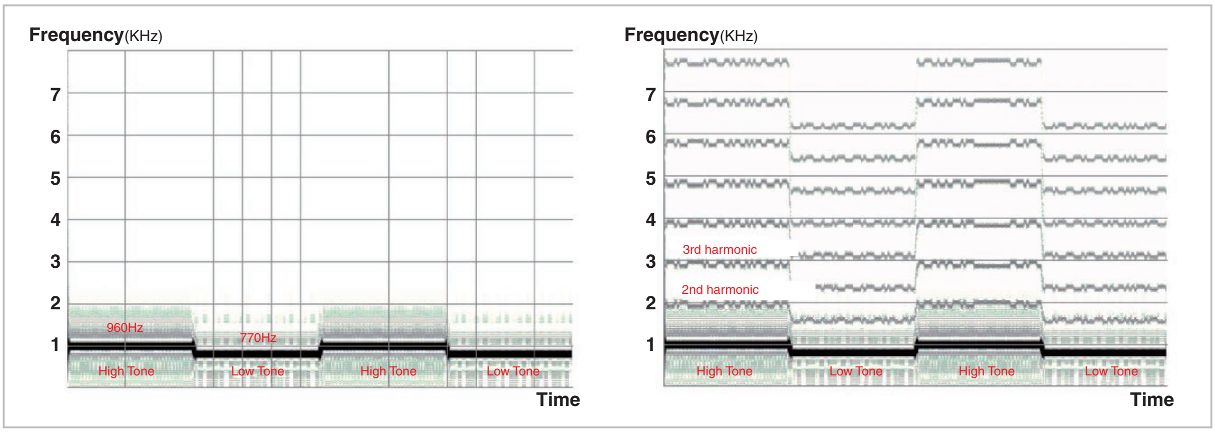


Fig.8 (L) Siren sound not be overlaid with information; (R) Siren sound overlaid with information by using Method 2 with additional improvements

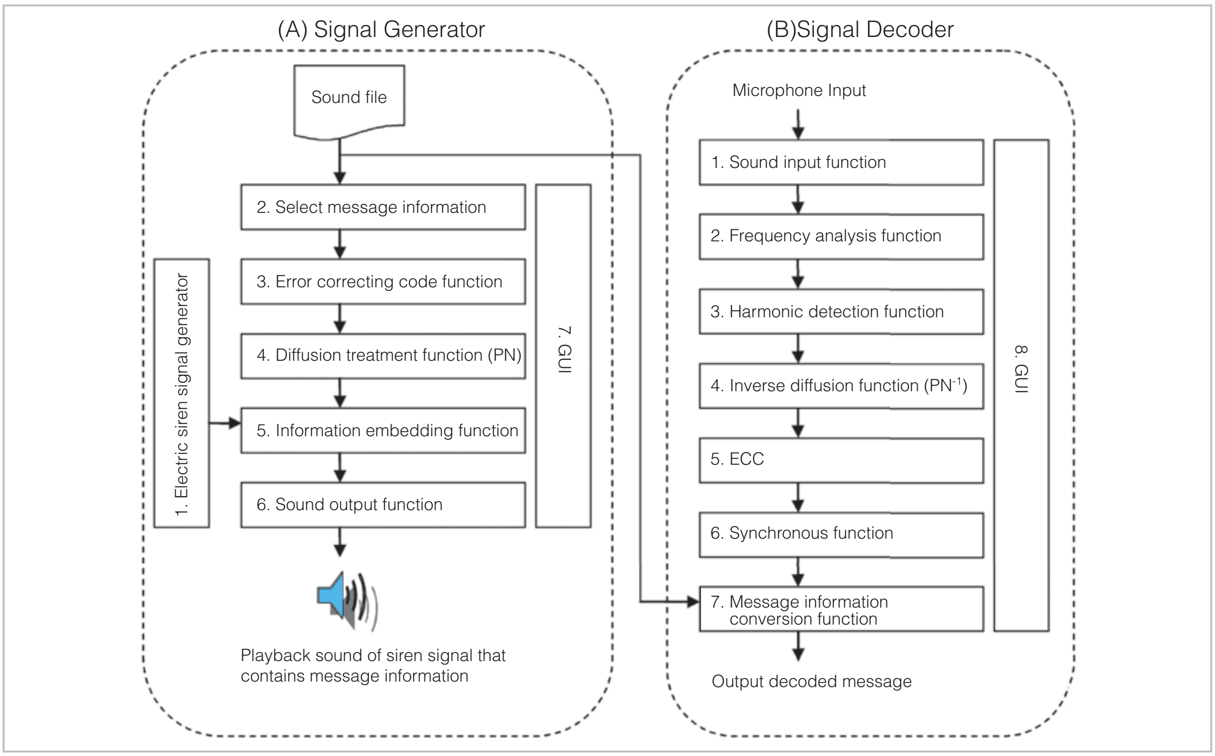


Fig.9 *Structure of prototype system developed in 2009*

siren signal.

6. Sound output function
   Generates and outputs siren signal sounds overlaid with information. As a simulation function for us, it has the ability to overlay outside distractions (background noise, Doppler effect). Output of signal sounds generated is done through speech files and speakers.

(B) Signal Decoder

1. Sound input function
   Inputs sound via microphone or speech files.
2. Frequency analysis function
   Analyzes the frequency of input sounds.
3. Harmonic detection function
   Detects harmonic waves and has the ability to determine siren signals.
4. Inverse diffusion function (1/PN)
   Performs inverse diffusion on bit strings corresponding to the diffusion treatment performed by the signal generator.
5. Error correction function (ECC)
   Correct errors on bit strings.
6. Synchronous function
   Performs synchronous processing that detects the beginnings of overlaid strings of symbols.
7. Message conversion function
   Converts decoded strings of symbols into messages from the message table.

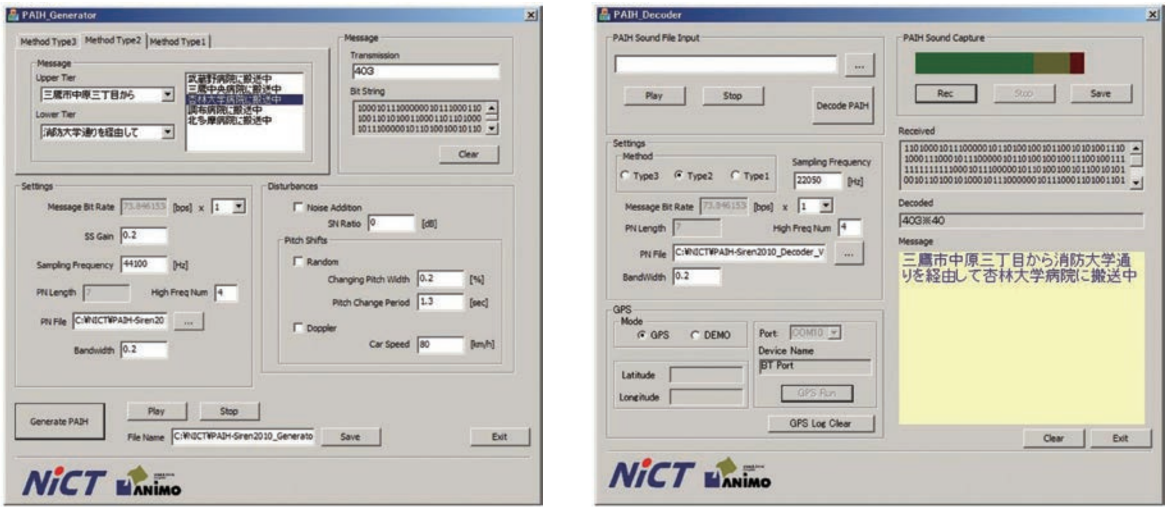
The 2009 version of this method was set up such that the signal generator selected three messages and the signal decoder extracted those messages. Specifically, the signal generator selected ambulance location, route, and destination and transmitted such messages as, “The ambulance departed from Location A and is en route to Hospital B via Road C.” This version of the method converted overlaid information transmitted through siren sounds into character strings that displayed those three pieces of information. A pound symbol (#) identified the end of the string and was followed by three numerals to create a string with a total of four two-byte characters. The method would post that single set of character strings in the time it took the siren to repeat two high-low tone cycles. Figure 10 shows the 2009 version application screen. The three pieces of information are selected in the upper left of the signal generator application to the left in Fig. 10, and the three-numeral codes that correspond to those pieces of information (defined in the message table) are displayed in the window on the upper right. The signal decoder detected those numerical strings in its application to the right in Fig. 10 and referred to the message table before displaying messages in the window to the lower right.

5.4 Developing applications associated with GPS and GIS

In 2010, we developed applications to be used with GPS receivers and Google Maps and installed them onto Windows PCs in order to realize our vision of having them installed in car navigation systems as described in Section **2.3**.

The signal generator application acquires information about longitude and latitude from GPS receivers, embeds it into siren sounds, and sends it. The signal decoder application takes information about its own longitude and latitude from the GPS receiver and combines it with the information generated by the signal generator, which it has extracted from the siren. It then shows this information on Google Maps.

The longitudes and latitudes must send off coordinates to five significant digits in order to achieve precision of position down to one meter, so the required amount of information is 00.00000 to 180.00000 for longitude (0˚ to 180˚; just short of 25 bits) and 00.00000 to 90.99999 for latitude (0˚ to 90˚; just short of 24 bits). Meanwhile, the Shift JIS character set includes spaces but not DEL, and the total of one- and two-byte characters is 11,438 (11,391 for Shift_JIS-2004). If we avoid adjoining characters and use 4,096 characters for bit strings in order to reduce the detection of errors, we can display 12 bits for each character. Thus, the number of characters required to send longitudes and latitudes is five (24+25)/12, and when the pound symbol (#) identifying the end is included, the method transmits strings with a total of six two-byte characters. The method posts that single set of character strings in the time it takes the siren to repeat three high-low tone cycles. Figure 11 shows the 2010 version application screen. In the signal generator application, character strings converted from the unit's own positional coordinates acquired through GPS (displayed on the upper left of the screen) are displayed in the window to the upper right of the screen. The signal decoder application displays both its own position acquired through GPS (displayed in the lower left of the screen) and signal generator position converted from the decoded character strings (displayed in the


PAIH_Generator

Method Type2

Message

Upper Tier: 三鷹市中原三丁目から

Lower Tier: 消防大学通りを経由して

武蔵野病院に搬送中
三鷹中央病院に搬送中
杏林大学病院に搬送中
調布病院に搬送中
北多摩病院に搬送中

Transmission: 403

Settings

Message Bit Rate: 73.846153 [bps] x 1

SS Gain: 0.2

Sampling Frequency: 44100 [Hz]

PN Length: 7

High Freq Num: 4

PN File: C:¥NICT¥PAIH-Siren20

Bandwidth: 0.2

Disturbances

☐ Noise Addition

SN Ratio: 0 [dB]

Pitch Shifts

☐ Random

Changing Pitch Width: 0.2 [%]

Pitch Change Period: 1.3 [sec]

☐ Doppler

Car Speed: 80 [km/h]

File Name: C:¥NICT¥PAIH-Siren2010_Generato

PAIH_Decoder

Settings

Method: ○ Type3 ● Type2 ○ Type1

Sampling Frequency: 22050 [Hz]

Message Bit Rate: 73.846153 [bps] x 1

PN Length: 7

High Freq Num: 4

PN File: C:¥NICT¥PAIH-Siren2010_Decoder_V

BandWidth: 0.2

GPS

Mode: ● GPS ○ DEMO

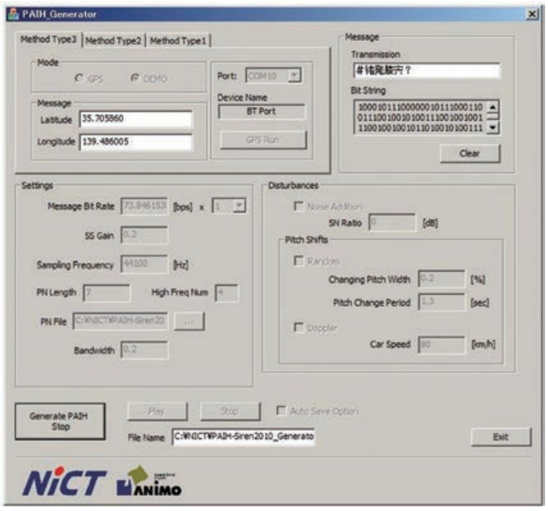
Port: COM10

Device Name: BT Port

Decoded: 403※40

Message: 三鷹市中原三丁目から消防大学通りを経由して杏林大学病院に搬送中

Fig.10 (L) 2009 version signal generator application screen; (R) 2009 version signal decoder application window



Mode: ○ GPS ● DEMO

Message

Latitude 35.705860

Longitude 139.486005

Port: COM10

Device Name BT Port

Settings

Message Bit Rate 73.846153 [bps] x 1

SS Gain 0.2

Sampling Frequency 44100 [Hz]

PN Length 7

High Freq Num 4

PN File C:¥NICT¥PAIH-Siren20

Bandwidth 0.2

Disturbances

☐ Noise Addition

SN Ratio 0 [dB]

Pitch Shifts

☐ Random

Changing Pitch Width 0.2 [%]

Pitch Change Period 1.3 [sec]

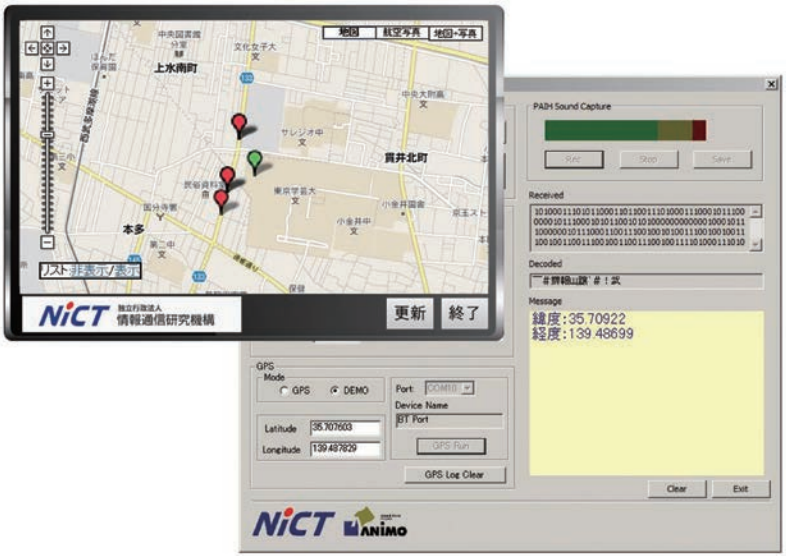
☐ Doppler

Car Speed 80 [km/h]

☐ Auto Save Option

File Name C:¥NICT¥PAIH-Siren2010_Generato

Signal Generator Application Screen



Message

緯度:35.70922
経度:139.48699

GPS

Mode: ○ GPS ● DEMO

Latitude 35.707603

Longitude 139.487829

Port: COM10

Device Name BT Port

Signal Decoder Application Screen

Fig.11 2010 version application screens

lower right of the screen) on Google Maps. A green pin denotes the signal decoder position, and a red pin shows the signal generator position. A line of pins denotes the paths taken by both the signal generator and the signal decoder when both of them are in motion.

6 Discussion

We have proactively demonstrated and explained PAIH at various exhibitions and public venues since 2009, when we were first able to achieve practical detection capabilities. For example, we presented PAIH at the National Research Institute of Fire and Disaster Open House on April 16, 2010; at the G-Spatial EXPO on September 19–21, 2010; and the Risk Control Convention in Tokyo on October 6–8, 2010. PAIH frequently appears in the media since it is a completely unique technology that is easy to visualize intuitively[11]-[13].

Vehicle Information Communication Systems (VICS) and other such technologies that understand the positions of vehicles exist to contribute to moving traffic along, and VICS may even be able to make our vision of transmitting information about the positions of emergency vehicles a reality. However, PAIH does not rely on communicating information to a central location to be managed in a concentrated manner; all it needs is one sender vehicle and one receiver vehicle to work. Thus, one plus for PAIH is that the hurdle of introducing it into use is low. For example, suppose that it is being used by a fire department in a sparsely populated area with only one ambulance. From that point, it only has to be installed in the surrounding residents' car navigation systems to work successfully. This is a feature that cannot be duplicated by systems that manage information in a concentrated fashion.

7 Conclusion

In this paper, we discussed PAIH, a technology for applying digital watermarking to siren sounds emitted by ambulances and other emergency vehicles as a way to use MIH technology to manage disasters. To expand the prevalence of this technology, we must first focus on the sending end, working together with siren manufacturers to verify and evaluate whether the quality of these siren sounds is enough to work its way into common knowledge when emitted at actual volumes. We also have work to do on the receiving end, first involving manufacturers and getting this system installed into car navigation systems, and then considering installation into smartphones, which are rapidly gaining popularity in recent times.

Acknowledgement

We are grateful to ANIMO Limited for their cooperation with the development of PAIH.

References

1 H. Nakagawa, O. Takizawa, and S. Inoue, "Information Hiding on Documents," Journal of Information Processing Society of Japan, Vol. 44, No. 3, pp. 248–253, March, 2003.

2 O.Takizawa, T.Matsumoto, H. Nakagawa, I. Murase, and K. Makino, "Final Information Hiding in Documents Through Continuous Position Adjustment," Journal of the National Institute of Information and Communications Technology, Vol. 52, Nos. 1/2, pp. 153–171, 2005.

3 O. Takizawa and K. Makino, "Implementing Information Hiding in Plain Text (TextIH) Through Continuous Position Control," IEICE Trans., J93-A, No. 2, pp. 100–103, February, 2010.

4 K. Yoshioka, K. Sonoda, O. Takizawa, and K. Nakao, "Hiding information into emergency public address sound," The 9th Western Pacific Acoustics Conference (WESPAC IX 2006), p. 86, 2006.

5 K. Sonoda, K. Yoshioka, and O. Takizawa, "Information Hiding for Public Address Audio Signal using FH/FSK Spread-spectrum Scheme," The 3rd International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP 2007), pp. 279–282, 2007.

6 K. Sonoda, J. Suzuki, and O. Takizawa, "Overlaying Information on Digital Siren Sounds in Motion," IEICE, Multimedia Information Hiding, 5th Research Seminar, November 6, 2008.

7 K. Sonoda, K. Yoshioka, and O. Takizawa, "Information Transmission Methods and Devices that Use Identical Warning Sounds," Japanese Unexamined Patent Application Publication No. 2009-128590.

8 Y. Nakajima, H. Matsuoka, and K. Yoshimura, "Audio OFDM Characteristics During Airwave Transmission," IEICE, Tech. Report RCS 106 (168), pp. 139–144, July, 2006.

9 JASRAC, "Final Selections Being Made on Technology for Worldwide Popularization of Music's 'Digital Watermarking'," http://www.jasrac.or.jp/release/01/06_1. html, June 29, 2001.

10 Announcement #337 from Commissioner of the Japanese Fire and Disaster Management Agency, June 10, 1970.
11 “Communicating Information During Disasters: A Look at a New System Currently Being Tested,” News Terrace, NHK TV (Osaka), March 6, 2009, 18:00.
12 “Information Sneaks Into Siren Sound: Information Transmission Research Institute Develops Navigation Display for Cars in Vicinity of Ambulances,” Sankei Shimbun (Osaka HQ Vol.), December 4, 2010 Evening Edition, Page 1.
13 “Digital Audio Watermarking Technology,” DIME, Vol. 26, No. 3, p. 88, Shogakukan, February, 2011.

(Accepted March 30, 2011)

TAKIZAWA Osamu, *Ph.D.*[†1]

Group Leader, Disaster Management and Mitigation Group (April 2006 – March 2011) / Group Leader, Security Fundamentals Group (May 2008 – March 2010), Information Security Research Center

Telecommunication for Disaster Management, Application Security

SONODA Kotaro, *Ph.D.*

Assistant Professor, Computer and Information Science Program, School of Engineering, Faculty of Engineering, Nagasaki University
Former: Expert Researcher, Disaster Management and Mitigation Group, Information Security Research Center (April 2006 – Sept. 2009)

Acoustic Information Sciences, Multimedia Information Enrichment

YOSHIOKA Katsunari, *Ph.D.*[†2]

Assistant Professor, Interdisciplinary Research Center, Yokohama National University

Former: Researcher, Network Security Incident Response Group, Information Security Research Center (April 2006 – Nov. 2007)

Information System Security,
Network Security,
Malware Response

SUZUKI Jouji, *Dr. Eng.*

Former: Guest Researcher/Invited Advisor, Disaster Management and Mitigation Group,
Information Security Research Center (April 2006 – March 2010)

Former President, Communications Research Laboratory, M.P.T.

Speech Communication

[†1] Currently, Manager, Technology Transfer Promotion Office, Outcome Promotion Department

[†2] Currently, Associate Professor, Graduate School and Research Institute of Environment and Information Sciences, Yokohama National University