Acoust. Sci. & Tech. **27**, 3 (2006)

ACOUSTICAL LETTER

Audio secret sharing for 1-bit audio

Norihiro Fujita*, Ryouichi Nishimura and Yôiti Suzuki

Research Institute of Electrical Communication, Tohoku University

(*Received 19 August 2005, Accepted for publication 8 December 2005*)

Keywords: 1-bit audio, Secret sharing scheme, Hiding communication
PACS number: 43.60.−c, 43.60.Dh [DOI: 10.1250/ast.27.171]

1. Introduction

The Internet has become one of the most important infrastructures. However, there are security risks in communication on the Internet such as theft and tapping of information. In this paper, a new secret sharing scheme (SSS) [1] for audio signals, called "binary audio secret sharing (BASS)" is proposed. This can be used to protect audio signals on the Internet by making it robust against theft and tapping of information. SSS is an encryption method and produces n shared data from original data to realize hiding communication. In the original SSS proposed [1], with any $k-1 (k \leq n)$ shared data, nobody can access information of original data. Only with all the k shared data, the original data can be obtained. Thus, even if any $k-1$ shared data were stolen, no original data would be stolen. Consequently, applying SSS to audio communications on the Internet can help to make it more robust against theft than normal communications on the Internet. Moreover, we propose a method of making each shared data heard as its intended decoy sound.

The BASS algorithm is based on visual secret sharing (VSS). Of several proposals regarding VSS, "k out of k VSS [2]" is a well-known method targeting binary images and shares an original image into k random dot images. The original image can be seen by stacking all of the k shared images. Observers, however, cannot see any useful image with any $k-1$ shared data. Figure 1 shows a sample of 3 out of 3 VSS. The process of encryption of 3 out of 3 VSS is depicted schematically in Fig. 2, where each pixel of the original image is represented by four subpixels.

2. BASS algorithm

2.1. 1-bit audio

1-bit audio is a high-quality digital audio format employed in the superaudio CD (SACD). It has a sampling frequency of 2,822.4 kHz and a resolution of 1-bit.

2.2. Algorithm of encryption and decryption

BASS shares each sample of the original signal, which has a value of '0' or '1,' according to sharing tables $t_b (b = 0, 1)$. These sharing tables t_b are $k \times n$ boolean matrices, where k is the sharing number and $n = 2^{k-1}$. Defining $A(m, b)$ as the number '1' in the logical sum of rows of $m (1 \leq m \leq k) \times n$, submatrices of t_b, t_b should be constructed so that $A(m, b)$ satisfies Eq. (1).

$$A(m,b) = \begin{cases} \displaystyle\sum_{i=1}^{m} \frac{n}{2^i} & (m < k) \\ n-1 & (m = k, b = 0) \\ n & (m = k, b = 1) \end{cases} \tag{1}$$

Defining $\boldsymbol{T}_b$ as the collections of matrices obtained by exchanging any rows of $\boldsymbol{t}_b$ with other rows, shared signals $\boldsymbol{s}_i (1 \leq i \leq k)$ are obtained by Eq. (3) from the original signal $\boldsymbol{o}$ represented in Eq. (2).

$$\boldsymbol{o} = (o_1 o_2 \ldots o_j \ldots o_L), \quad o_j \in \{0, 1\} \tag{2}$$

$$\begin{pmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \\ \vdots \\ \boldsymbol{s}_k \end{pmatrix} = (\boldsymbol{t}_{o_1} \boldsymbol{t}_{o_2} \ldots \boldsymbol{t}_{o_L}), \quad \boldsymbol{t}_{o_j} \in \{\boldsymbol{T}_{o_j}\} \tag{3}$$

When the sharing number k is $3 (k = 3)$, for example, $\boldsymbol{t}_0$ and $\boldsymbol{t}_1$ may be 3×4 matrices as in Eq. (4).

$$\begin{pmatrix} 1001 \\ 0101 \\ 1100 \end{pmatrix} \in \boldsymbol{T}_0, \quad \begin{pmatrix} 1001 \\ 0101 \\ 0011 \end{pmatrix} \in \boldsymbol{T}_1 \tag{4}$$

Decryption of the original signal in BASS requires all the k shared signals and uses the characteristic represented in Eq. (1). The $m \times (n \cdot L)$ matrix represented in Eq. (3) is divided by every n columns resulting in $m \times n$ matrices of L. $A(m, t)$ for the $j (1 \leq j \leq L)$th $m \times n$ matrix is then obtained as

$$A(m,t,j) = \begin{cases} \displaystyle\sum_{i=1}^{m} \frac{n}{2^i} & (m < k) \\ n - 1 + o_j & (m = k). \end{cases} \tag{5}$$

Consequently, we can decrypt the original signal $o_j (1 \leq j \leq L)$ from all the k shared signals using Eq. (5) since the information of the original signal o_j can be obtained only if $m = k$.

3. Shared signals heard as intended sound

BASS based on the original k out of k VSS makes shared signals heard as noise. However, we can encrypt an original signal into $k-1$ shared signals heard as distinct and intended sounds and one random noiselike signal by properly constructing the sharing tables. Shared signals can be any audio signals. Note that having shared signals heard as intended sounds does not lessen the secrecy of BASS.

*e-mail fuji@ais.riec.tohoku.ac.jp

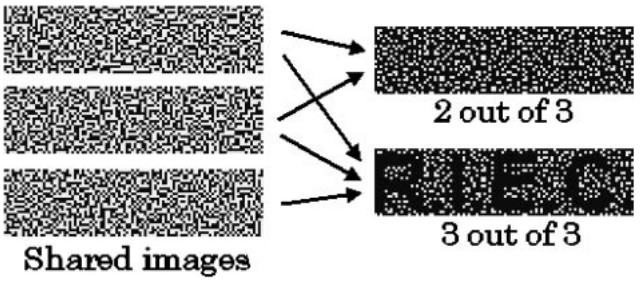

Fig. 1 Sample of 3 out of 3 VSS.

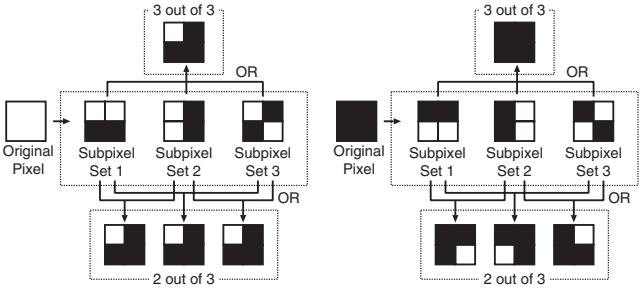

Fig. 2 Sample pattern of subpixels.

3.1. Algorithm

Defining $1 \times n$ matrix $\boldsymbol{p}_i (1 \le i \le k)$ as Eq. (6) for any sharing table $\boldsymbol{t}_0$,

$$\begin{pmatrix} \boldsymbol{p}_1 \\ \boldsymbol{p}_2 \\ \vdots \\ \boldsymbol{p}_k \end{pmatrix} = \boldsymbol{t}_0 \in \boldsymbol{T}_0. \tag{6}$$

We also define $n \times L$ samples of intended sounds as $\boldsymbol{d}_i$, and $\boldsymbol{d}'_i (1 \le i \le k-1)$ as the decimation of $\boldsymbol{d}_i$ by $1/n$, resulting in L samples. Using these notations, shared signals $\boldsymbol{s}_i (1 \le i \le k-1)$, which are heard as the intended sounds, and a shared signal $\boldsymbol{s}_k$ heard like a noise are generated as

$$\begin{pmatrix} \boldsymbol{s}_1 \\ \boldsymbol{s}_2 \\ \vdots \\ \boldsymbol{s}_i \\ \vdots \\ \boldsymbol{s}_k \end{pmatrix} = \begin{pmatrix} \boldsymbol{p}_1^{(d'_{11})} & \cdots & \boldsymbol{p}_1^{(d'_{1j})} & \cdots & \boldsymbol{p}_1^{(d'_{1L})} \\ \boldsymbol{p}_2^{(d'_{21})} & \cdots & \boldsymbol{p}_2^{(d'_{2j})} & \cdots & \boldsymbol{p}_2^{(d'_{2L})} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{p}_i^{(d'_{i1})} & \cdots & \boldsymbol{p}_i^{(d'_{ij})} & \cdots & \boldsymbol{p}_i^{(d'_{iL})} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \boldsymbol{p}_k^{(d'_{k1})} & \cdots & \boldsymbol{p}_k^{(d'_{kj})} & \cdots & \boldsymbol{p}_k^{(d'_{kL})} \end{pmatrix}, \tag{7}$$

$$\boldsymbol{d}'_i = (d'_{i1} d'_{i2} \ldots d'_{iL}), \tag{8}$$

$$d'_{kj} = o_j + \sum_{i=1}^{k-1} d'_{ij}, \tag{9}$$

$$\boldsymbol{p}_i^{(0)} \equiv \boldsymbol{p}_i, \tag{10}$$

$$\boldsymbol{p}_i^{(1)} \equiv \overline{\boldsymbol{p}_i}, \tag{11}$$

where $\overline{\boldsymbol{p}_i}$ (11) denotes the logical negation of $\boldsymbol{p}_i$.

The original signal can be decrypted from $\boldsymbol{s}_i (1 \le i \le k)$ due to the following reason. Define any column of Eq. (7) as $\boldsymbol{t}'_j$ as represented in Eq. (12).

$$\boldsymbol{t}'_j = \begin{pmatrix} \boldsymbol{p}_1^{(d'_{1j})} \\ \boldsymbol{p}_2^{(d'_{2j})} \\ \vdots \\ \boldsymbol{p}_i^{(d'_{kj})} \end{pmatrix} \tag{12}$$

Here, $\boldsymbol{t}'_j$ is a matrix obtained by exchanging any columns of an arbitrary matrix in $\boldsymbol{T}_0$ and taking logical negation $\sum_{i=1}^{k} d'_{ij}$ times for each component. A sharing table has the following characteristic.

$$\boldsymbol{t}_b^{\langle I \rangle} \in \boldsymbol{T}_{((I+b) \bmod 2)} \ (b = 0, 1) \tag{13}$$

Here, $\boldsymbol{t}_b^{\langle I \rangle}$ denotes that I rows of $\boldsymbol{t}_b$ are taken as logical negation. That is, one in $\boldsymbol{T}_0$ becomes one in $\boldsymbol{T}_1$ when taking logical negation of any rows, and vice versa. Because the total number of rows taking logical negation in $\boldsymbol{t}'_j$ is represented by $\sum_{i=1}^{k} d'_{ij}$ due to $d'_{ij} \in \{0, 1\}$, Eq. (14) is satisfied.

$$\begin{aligned} \boldsymbol{t}'_j \in T_{(\sum_{i=1}^{k} d'_{ij}) \bmod 2} &= T_{(o_j + 2 \cdot \sum_{i=1}^{k-1} d'_{ij}) \bmod 2} \\ &= T_{o_j} \end{aligned} \tag{14}$$

3.2. Sound of sharing signals

A way to compose sharing signals heard as an intended decoy sound was discussed in the previous section. In this section, those sharing signals $\boldsymbol{s}_i$ are confirmed to be heard as the intended sound $\boldsymbol{d}_i$.

Exchanging $\{0, 1\}$ of each sample of a signal with $\{-1, 1\}$, $\boldsymbol{s}_i$ is generated from $\boldsymbol{d}_i$ using a block diagram as shown in Fig. 3, where $\boldsymbol{d}'_i$ is derived by decimating $\boldsymbol{d}_i$ by $1/n$. On the other hand, $\boldsymbol{d}''_i$ is the signal obtained by up-sampling $\boldsymbol{d}'_i$ by n. Therefore, $\boldsymbol{d}''_i$ is equal to $\boldsymbol{d}_i$ in the frequency domain under $f_s/2n$, where f_s is the sampling frequency of $\boldsymbol{s}_i$. Since the convolution of $\boldsymbol{d}''_i$ with $\boldsymbol{p}_i$ is equivalent to the multiplication of these signals in the frequency domain, $\boldsymbol{d}''_i$ becomes the signal $\boldsymbol{s}_i$ filtered by $\boldsymbol{p}_i$. If $f_s/2n$ is larger than the highest audible frequency, $\boldsymbol{d}_i$ is heard as $\boldsymbol{s}_i$ filtered by $\boldsymbol{p}_i$. Although f_s is very high for 1-bit audio, the generated sound signals are heard contaminated by quantization noise because of decreasing nominal sampling frequency and filtering by the linear filter $\boldsymbol{p}_i$.

Figure 4 shows some samples of amplitude characteristics of $\boldsymbol{p}_i$. No phase distortion is induced because of the linear phase characteristics of the filtering by $\boldsymbol{p}_i$. If it is necessary to make $\boldsymbol{d}''_i$ identical to $\boldsymbol{s}_i$, the decoy sound must be filtered with the inverse filter of $\boldsymbol{p}_i$ in advance.

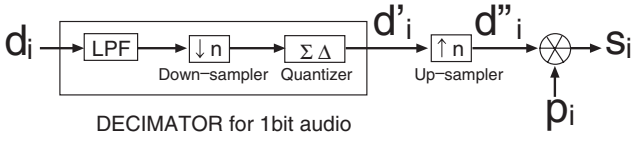

Fig. 3 Block diagram expressing the relationship between $\boldsymbol{d}_i$ and $\boldsymbol{s}_i$.

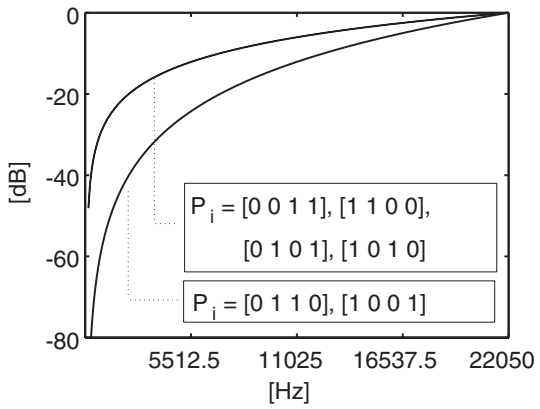

Fig. 4 Amplitude characteristics of $\boldsymbol{p}_i$ within audible range with $n = 4$.

4. Conclusions

We propose binary audio secret sharing, BASS, which can share an original audio signal into k shared audio signals. A method to compose the $k - 1$ shared signals heard as intended decoy sound is also proposed. Using these techniques, we believe that highly secure communication can be realized on the Internet.

Acknowledgments

We thank Prof. Shizuya and Prof. Manbo for their introduction of the secret sharing scheme. This study is partly supported by SCOPE(051302004).

References

[1] A. Shamir, "How to share a secret," *Commun. ACM*, **22**, 612–613 (1979).
[2] M. Naor and A. Shamir, "Visual cryptography," *Advances in Cryptology - EUROCRYPT '94*, LNCS 950, pp. 1–12 (1995).