The 44th Symposium on Information Theory and its Applications (SITA2021)
Nishinomiya, Hyogo, Japan, Dec. 8–10, 2021

視覚暗号を用いた秘匿音声伝送
Privacy-preserved Speech Transmission by Using Visual Cryptography

太刀岡 勇気 *
Yuuki Tachioka

Abstract— Visual cryptography that share secrets into multiple share images and prevent from restoring secrets by using insufficient share images has been rapidly improved because of a numerical optimization. In the speech processing field, sound source separation is a difficult problem especially when the number of observation channels is smaller than the number of sources. We exploit the difficulty of this problem and propose a privacy-preserved speech communication method that uses visual cryptography to extract secret message from mixed signals that contains secret message and many interference signals.

Keywords— Visual cryptography, Sound source separation, Speech transmission, Spectrum masking

1 はじめに

公衆に聞き取れる音声に秘匿したい音声を埋め込む技術は有用である。暗号化した音声を埋め込む方法 (音響透かし) には、付加情報を位相シフトキーイングし情報をスペクトル拡散させた拡散符号系を直接利用する直接拡散方式がある [1]。またほかの信号にマスクされる場所に付加情報を埋め込む方法 [2] や、識別情報を可聴域外に変調することで埋め込む方法も提案されている [3]。

ただし信号を変調する方式では変調方式がわかれば簡単に復号されてしまう。一般に暗号化の安全性は普通復号にかかる計算複雑性により担保されているので、どんな方法も暗号化手段が漏れた場合に安全でなくなる。これに対して、暗号化手法が漏れても安全な技術として、秘密画像を複数枚のシェア画像に秘密分散し、一定枚数以上持ち寄らないと秘密が漏れない視覚暗号技術 (visual cryptography scheme; VCS)[4] がある。VCS は、近年数値的最適化法が確立されたことで急速に性能が向上している [5]。一方、音声処理において、複数の信号が混ざった信号から各音源に信号を分離する音源分離は、観測チャネル数が音源数よりも少ない劣決定条件では解くのが難しいことが知られている。著者はこの音源分離の難しさを利用し、秘匿したい信号に加えて妨害信号を多数付加した混合信号から視覚暗号技術により、秘匿信号を抽出できる秘匿音声通信を行うための検討を行った。

* 〒 150-0002 東京都渋谷区渋谷 2-15-1 渋谷クロスタワー 28F (株) デンソーアイティーラボラトリ 研究開発グループ, Research and Development Group, Denso IT Laboratory, Shibuya Cross Tower 28F, 2-15-1 Shibuya, Shibuya-ku, Tokyo, Japan. E-mail: tachioka.yuki@core.d-itlab.co.jp

2 視覚暗号

視覚暗号は、1 枚の秘密画像 (secret image; SI) から n 枚のシェア画像を生成し、そのうちの任意の t 枚以上の異なるシェアを持ち寄ることで元の秘密画像を復元でき、$t-1$ 枚のシェアからは秘密が漏れないというものである。このような方式を (t,n)-VCS と呼ぶ。復元はシェアを重ね合わせるだけで、複雑な復号処理は必要ない。

基本行列の生成が課題であり、従来は多項式を代数的に処理する方法 [6] が使われていたが、$t \leq 4$ の場合に、このようなシェアを閉じた形で生成する方法が長らく不明であった。このため、利用が限られていたが、近年その方法が確立され [7, 8]、以来応用事例が増えてきている。

2.1 整数計画法を用いた視覚暗号の構成法

任意の (t,n)(ただし $2 \leq t \leq n$) に対してシェアを生成するには、まずシェアに対する参加者のアクセス構造 Γ を定める必要がある。n 枚のシェアの集合 $P = \{1, 2, ..., n\}$ があったときに、どのシェアを保持しているかで場合分けしたべき集合は 2^P となる。このときシェアを重ね合わせることで、秘密画像 SI を復元できるシェアの集合を有資格集合 Γ_Q とする。t 枚のシェアが集まれば SI を復元できるので、Γ_Q は t 枚のシェアが集まっている場合が極小となり、この時を極小有資格集合 Γ_Q^* と呼ぶ。逆にシェアから SI に関する一切の情報を得られない集合を禁止集合 Γ_F とする。容易にわかるように、Γ_F は全体のべき集合 2^P に対する Γ_Q の補集合となり、$\Gamma = (\Gamma_Q, \Gamma_F)$ となる。この時に $t-1$ 枚のシェアが集まっている場合を極大禁止集合 Γ_F^* と呼ぶ。ここで SI が 2 値 $\mathbb{B} = \{0, 1\}$ であるとすると、画素拡大の倍率を m としてブール行列 ($\mathbb{B}^{n \times m}$) の組 (X_0, X_1) が以下の 2 条件を満たすとき、上記アクセス構造 Γ を実現する基本行列であるという。

SI の復元可能条件　すべての $S \in \Gamma_Q^*$ に対して定数 $\alpha > 0$ が存在して,

$$HW(OR(X_0[S])) + \alpha m \leq HW(OR(X_1[S])) \quad (1)$$

となる。ここで $X.[S]$ は $X.$ の内から S に対応する行のみを抜き出す操作、OR は列ごとの OR、HW はハミング重みである。

安全性条件　すべての $S \in \Gamma_F^*$ に対して、$X_0[S]$ と $X_1[S]$ は適当な列の並び替えで等しくできる。

このような基本行列を何らかの方法で求めれば、X_0,X_1 をそれぞれ SI の $\{0,1\}$ に対応する画素拡大した際のシェアの画素値に対応付けることで、シェアを生成できる。上述の 2 条件のうち、「SI の復元可能条件」は式 (1) を満たすようにすればよいので、線形制約の形で書けるが、「安全性条件」を満たすのが難しかった。Iwamoto, Shyu らは独立に、「X_0,X_1 は列の適当な並び替えで等しくできること」と、

$$HW(OR(X_0[S])) = HW(OR(X_1[S])), \forall S \subset \{1, ..., n\}$$

が同値であることを利用して、上記「安全性条件」が

安全性条件 (同値)

$$HW(OR(X_0[S])) = HW(OR(X_1[S])), \forall S \in \Gamma_F \quad (2)$$

と同値であること[1]を示した [7, 8]。これは線形制約の形で書けるので、以下の手順により画素拡大 m の最小化を目的とする整数計画問題に変換できる。これにより適用範囲が大幅に拡大され、場合によっては従来の構成法より小さい行列で基本行列を構成できるようになった。

ここで基本行列の m 本の列ベクトル ($\in \mathbb{B}^n$) は $E_0 = [0, ..., 0]^\top$,$E_1 = [1, ..., 0]^\top$,...,$E_{2^n-1} = [1, ..., 1]^\top$ の 2^n 通りしかない ($\top$ は転置) ことから、X_0 と X_1 の列数が等しいという制約下で、上記は E_j $(0 \le j \le 2^n - 1)$ から x_{0j} 個、x_{1j} 個ずつ選んで X_0 と X_1 を構成するという問題にできる。その際の目的関数は画素拡大 $m(= \sum_j x_{0j})$ の最小化である[2]。制約条件は上記 (1)(2) に対応する

$$\sum_j x_{0j} OR(E_j[S]) + 1 \le \sum_j x_{1j} OR(E_j[S]), \forall S \in \Gamma_Q^*,$$
$$\sum_j x_{0j} OR(E_j[S]) = \sum_j x_{1j} OR(E_j[S]), \forall S \in \Gamma_F$$

に加え、非負条件 ($x_{0j} \ge 0$,$x_{1j} \ge 0$) と基本行列の列数が等しいこと ($\sum_j x_{0j} = \sum_j x_{1j}$) である。これを線形計画法で解くことで m を最小化する基本行列を求められる。シェアを生成する際には X_0、X_1 を同じ形のまま使うとシェアに特定のパターンが現れてしまうため、画素ごとに列ベクトルをランダムに置換してシェアを生成する。

2.2　整数計画法を用いた拡張視覚暗号の構成法

VCS の拡張としてシェア単体からは SI と異なるカバー画像が見え、t 枚集めると SI が見えるという拡張視覚暗号 (Extended VCS; EVCS) が提案され [9]、その基本行列の最適化法 [10, 11] が提案されている。これを用いるとシェアにも情報を載せられるので、SI が暗号化されていることに気づかれにくいという利点が得られる。SI の画素値に加えて、カバー画像 $c_1, ..., c_n$ の画素値 (2 値) で場合分けすると、2^n 個のブール行列 ($\mathbb{B}^{n \times m}$) の組 $(X_0^{c_1,...,c_n}, X_1^{c_1,...,c_n})$ が以下の 3 条件を満たすとき、アクセス構造 Γ を実現する基本行列であるという。

SI の復元可能条件　すべての $S \in \Gamma_Q^*$ に対して $0 \le l_s \le h_s \le m$ を満たす整数 l_s と h_s が存在して,

$$HW(OR(X_0^{c_1,...,c_n}[S])) = l_s$$
$$HW(OR(X_1^{c_1,...,c_n}[S])) = h_s \quad (3)$$

がすべての $c_1, ..., c_n$ に対して成り立つ。

安全性条件　すべての $S \in \Gamma_F^*$ と $c_1, ..., c_n$ に対して、$X_0^{c_1,...,c_n}[S]$ と $X_1^{c_1,...,c_n}[S]$ は適当な列の並び替えで等しくできる。

カバー画像の視認条件　すべての $j = 1, ...n$ に対して、$0 \le l_j < h_j \le m$ を満たす整数 l_j と h_j が存在して、

$$HW(X_{\{0,1\}}^{c_1,...,c_n}[\{j\}]) = l_j \ (c_j = 0)$$
$$HW(X_{\{0,1\}}^{c_1,...,c_n}[\{j\}]) = h_j \ (c_j = 1) \quad (4)$$

が j を除くすべての $c_1, ..., c_n$ に対して成り立つ。

安全性条件に関しては VCS の場合と同様に、同値な以下の条件に直す。

安全性条件 (同値)　すべての $S \in \Gamma_F$ と $c_1, ..., c_n$ に対して、

$$HW(OR(X_0^{c_1,...,c_n}[S])) = HW(OR(X_1^{c_1,...,c_n}[S])) \quad (5)$$

が成り立つ。

これにより、制約条件 (3)(4)(5) が線形制約の形になるので、整数計画法が使える。

3　音声秘匿通信への応用

3.1　概要

現在異なる音声信号が混ざった信号から各音源に成分を分離する音源分離問題が盛んに検討され、観測チャネル数が音源数よりも少ない劣決定条件では解くのが難しいことが知られている。ここではこの事実を利用し、秘密音声に対して、妨害音声を I 個分混合した音声に対して、基本行列 X_0,X_1 を用いてマスキングを行うことで、秘密音声を秘匿するシェア音声を n 個生成する。これを公衆にスピーカーで放射し、VCS により t 個集めた場合にマスキングにより秘密音声を抽出できるようにする。

3.2　送信者によるシェアの作成

3.2.1　スペクトログラムマスキング

まずは秘匿したい秘密音声に対して、妨害音を任意個数 I だけ用意する。秘密音声と I 個の妨害音声を短時間

[1] ここで元の命題の極大禁止集合 Γ_F^* が禁止集合 Γ_F になっていることに注意。

[2] 画素拡大は基本行列の列数に等しい。

フーリエ変換し、それぞれのスペクトル $s(\tau,f)$,$\sigma_i(\tau,f)$ を得る。ここで τ は時間フレーム、f は周波数ビンの ID である $(1 \leq i \leq I,\, 1 \leq \tau \leq T,\, 1 \leq f \leq F)$。秘密音声の音声のレベル $|s|$ が妨害音声のレベル $|\sigma_i|$ よりも閾値 θ 以上大きい時間周波数ビンに対してマスク M を 1 とし、それ以外の点を 0 とする。

$$\begin{aligned} &M(\tau,f) = 1,\ \forall i, |s(\tau,f)|/|\sigma_i(\tau,f)| > \theta \\ &M(\tau,f) = 0,\ \text{otherwise} \end{aligned} \tag{6}$$

混合信号にマスク M をかける

$$s'(\tau,f) = M(\tau,f)\left[s(\tau,f) + \sum_i \sigma_i(\tau,f)\right] \tag{7}$$

と $s(\tau,f)$ の主要な部分 $s'(\tau,f)$ が取り出せるので、この M を SI とみなして、シェアを作成する。

3.2.2 画素拡大の取り扱い

シェアマスクはマスク M の m 倍に拡大されているため、シェア音声を作成する際には、画素拡大の必要がある。$m_1 m_2 = m$ を保ち、時間方向の拡大 $(T \to m_1 T)$ と周波数方向の拡大 $(F \to m_2 F)$ を行う。短時間フーリエ変換の際に、時間フレーム数を増やすには、フレームシフトを $1/m_1$ 倍に、周波数ビン数を増やすには、窓長を m_2 倍にすればよい。

3.2.3 シェアマスク・シェア音声の生成

上述の M からシェアのマスク $M'_{\{1,2,\ldots,n\}}$) を基本行列 X_0,X_1 に従い構成する。得られたマスクにより秘密音声と妨害音声を混合した混合信号をマスキングし、n 個のシェア音声を生成する。すなわち j 番目 $(j = 1, ..., n)$ のシェア音声は $M'_j(\tau',f') * (s(\tau',f') + \sum_i \sigma_i(\tau',f'))$ を逆短時間フーリエ変換して得られる。ここで τ' と f' は画素拡大したスペクトログラムの時間フレームと周波数ビン $(1 \leq \tau' \leq m_1 T,\, 1 \leq f' \leq m_2 F)$ を表す。

3.3 受信者による秘密音声の抽出と局所畳み込み

ここでは受信者はシェアマスクを受け取るとする[3]。シェアマスク $\sum_j M'_j(\tau',f')$ を t 個取得できれば、j が Γ_Q の要素からなる場合、各時間周波数ビンごとにマスクの値の OR を取ることで音源のマスクを推定できる。この場合には、画素拡大 m_1,m_2 がわかっていれば隣接 m_1,m_2 個のマスク成分を局所的に畳み込んで足し合わせ、それがある閾値[4]を超えていれば 1 超えていなければ 0 とすることで、よりマスクの推定精度を向上させられる。すなわち、元のマスク M での (τ,f) 成分は、シェアマスク $M'_{\{1,2,\cdots,n\}}$ では $(\tau',f') = (m_1(\tau-1)+k_1, m_2(f-1)+k_2)$ $(1 \leq k_1 \leq m_1,\, 1 \leq k_2 \leq m_2)$ に拡大されるので、シェアマスクから拡大されたマスク M' を加算や乗算により推定し、$M'_{est} = \chi_\eta(\sum_{k_1,k_2} M'(m_1(\tau-1)+k_1, m_2(f-1)+k_2))$ を生成する。$\chi(x)$ は x が閾値 η 以上であれば 1, 未満であれば 0 を返す指示関数である。このとき、M'_{est} の大きさは元のマスク M に等しい。このマスクを混合信号にかけると秘密信号が抽出できる。

3.4 EVCS の利用

秘密音声と n 個のカバー音声を用意することで EVCS が利用できる。それらを短時間フーリエ変換し、それぞれのスペクトル $s(\tau,f)$, $\kappa_j(\tau,f)$ を得る $(1 \leq j \leq n)$。秘密音声の音声のレベル $|s(\tau,f)|$ がカバー音声のレベル $|\kappa_j(\tau,f)|$ よりも閾値 θ 以上大きい時間周波数ビンに対してマスク M を 1 とし、それ以外の点を 0 する。

$$\begin{aligned} &M(\tau,f) = 1,\ \forall j, |s(\tau,f)|/|\kappa_j(\tau,f)| > \theta \\ &M(\tau,f) = 0,\ \text{otherwise} \end{aligned} \tag{8}$$

これに加えて、カバー音声のレベル $|\kappa_j(\tau,f)|$ がそのほかのカバー音声のレベルと秘密音声の音声のレベル $|s(\tau,f)|$ よりも閾値 θ 以上大きい時間周波数ビンに対してマスク M_j^c を 1 とし、それ以外の点を 0 とする[5]。

$$\begin{aligned} &M_j^c(\tau,f) = 1,\ \forall j' \neq j, |\kappa_j|/\max(|\kappa'_j|, |s|) > \theta \\ &M_j^c(\tau,f) = 0,\ \text{otherwise} \end{aligned} \tag{9}$$

max は引数のうちの最大値である。ここで $c_j = M_j^c(\tau,f)$ とすれば、EVCS の基本行列により、シェアマスク M' を VCS と同様に生成できる。シェア音声 j にはカバー音声 j が聞こえる。t シェア加算すると秘密音声を VCS と同様に得ることができる。

4 実験

4.1 条件

シミュレーション混合音声を、人の音声を混合して作成 (testA,testB) した。testA,B は人の音声である。発話長がほぼ等しい複数話者の音声から[6]、testA は 6 人を、testB は testA に加えてさらに 5 人を混合した計 11 人を混合して評価音声を作成した。音源のマスク M の作成時には SNR が 10 dB より大きい成分のみが 1 となるように、$\theta = \sqrt{10}$ とした。分離性能は音源ごとの信号対ひずみ比 (SDR)[12] により評価した。

線形計画法のソルバーである PuLP[7]により基本行列の生成を行った。最適化された m は VCS の場合が 3、EVCS の場合が 6 であった。復元された秘密音声の有用

[3] 受信者がシェア音声を受け取る場合は文献 [13] を参照されたい。

[4] 閾値は、その領域内の成分が全部 1 の場合すなわち $m_1 m_2$ の場合にのみ 1 とする、もしくは、全体のヒストグラムを作成し、ある一定成分が 1 となる値とする、中央値をとるなどの方法が考えられる。

[5] ここでは κ_j と s の (τ,f) は省略した。

[6] http://datashare.is.ed.ac.uk/bitstream/handle/10283/1942/clean_trainset_wav.zip よりクリーン音声を取得。

[7] https://coin-or.github.io/pulp/index.html

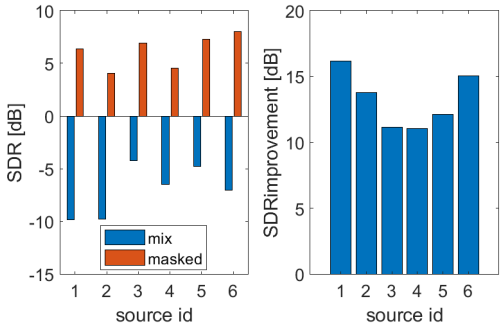

図 1: SDR [dB] of mixed signals and masked signals (testA).

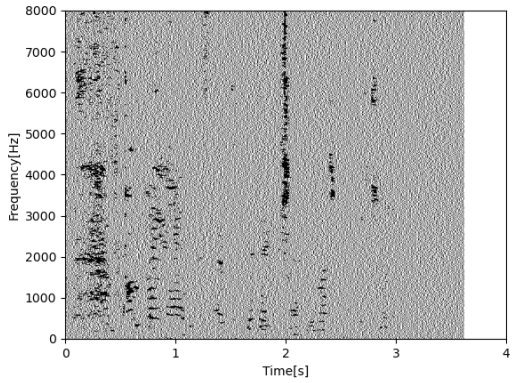

図 2: Spectrogram of a restored mask (testA).

性とシェア信号の安全性の観点からの評価を行った。純音を使ったより基礎的な検討 [13] により、シェア音声の同期加算では復元ができないため、シェアマスクの共有を行う必要があること、時間方向の拡大が周波数方向の拡大に比べて有効であること、EVCS によりカバー信号が受信できることがわかっている。

4.2 testA の場合

6 人の音声を混合した場合の、混合信号とマスク信号の SDR を図 1 に示す。混合信号の SDR は十分低く、またマスクによって著しく改善されることもわかる。

VCS 前節での検討の結果から時間方向拡大したものの、シェアマスクの OR を取ったものが有効であることがわかったので、以降その設定で評価する。VCS により得られた復元マスクのスペクトログラムを図 2 に示す。このように音声による調波成分の部分に 1 が集中していることがわかる。これに対してシェアマスクは 0,1 は一様に分布していた。

EVCS EVCS により得られた復元マスクのスペクトログラムを図 3 に示す。この場合も test1 と同様に対象話者 1 人に加えて、3 人の対象話者以外の話者をカバー信号としてカバーマスクを作成した。復元マスクに音声の調波成分は良く見えるが、対象音声以外の部分が濃くなっていることから VCS の場合に比べて SNR が低下し

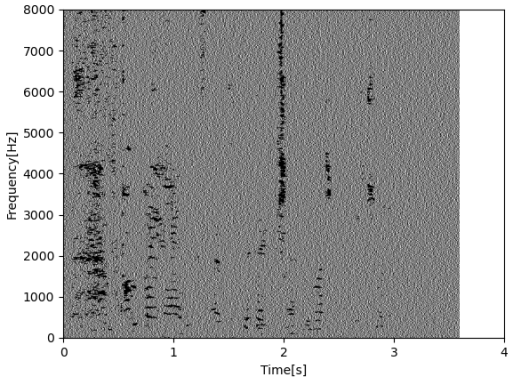

図 3: Spectrogram of restored mask (testA, EVCS).

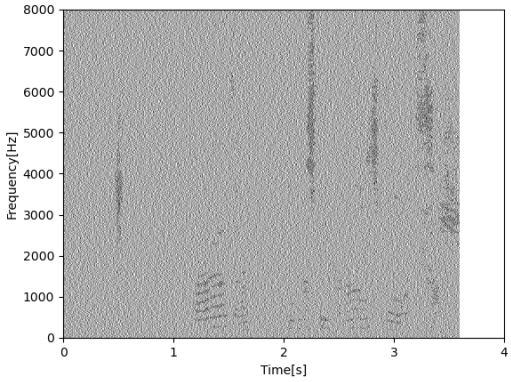

図 4: Spectrogram of a share mask (testA, EVCS).

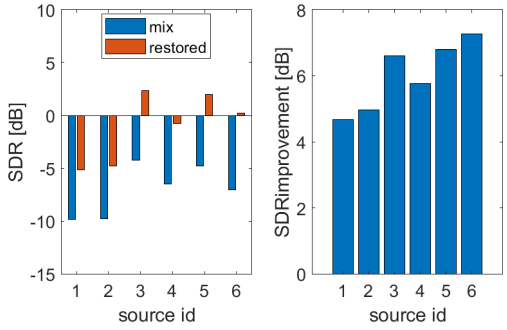

図 5: SDR [dB] of mixed signals and restored signals (testA).

ていることが示唆される。図 4 は対象信号と異なる話者の音声をカバー信号にした例である。カバー信号の調波構造が確認される。

SDR での評価 図 5 に、SDR での評価を示す。対象話者に依存して SDR の改善量が異なることがわかる。これにより秘匿しやすい話者としにくい話者がいることがわかる。この例では話者 6 は話者 1 に比べて秘匿しやすいということがわかる。シェアの SDR も確認したがシェアからは秘密情報は得られないことがわかった。図 6 の SDR を示す。VCS と比べて性能は劣るものの、秘密音声を抽出すること自体は可能であることがわかった。

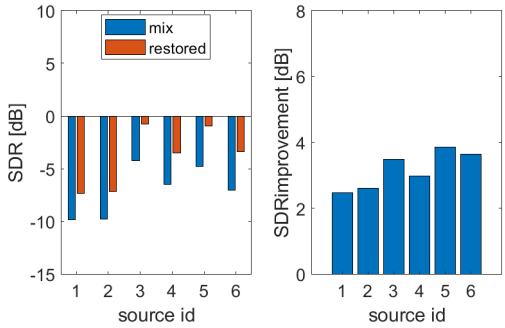

図 6: SDR [dB] of mixed signals and restored signals (testA,EVCS).

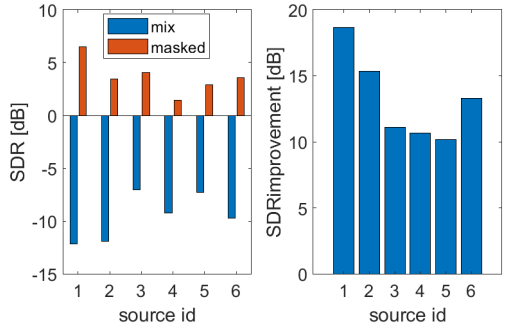

図 7: SDR [dB] of mixed signals and masked signals (testB).

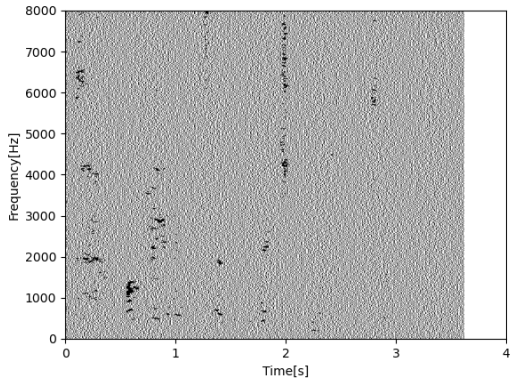

図 8: Spectrogram of a restored mask (testB).

4.3 testB の場合

testA に加えて 5 人の妨害話者を足した。対象話者 (EVCS の場合はカバー音声の話者も) は testA と同一である。図 7 には混合信号とマスク信号の SDR を示す。当然ながら混合信号の SDR は混合話者を増やすほど低下するが、図 1 と比べてマスク信号の SDR はあまり変わらない。これは音声にはスパース性があるため、混合話者を増やしても重なりはそれほど増えないためである。

VCS　図 8 には復元マスクのスペクトログラムを示す。このように図 2 に比べるとだいぶ情報は失われているものの、主要部分は残っていることがわかる。シェアマスクから情報を得られないのは同様である。

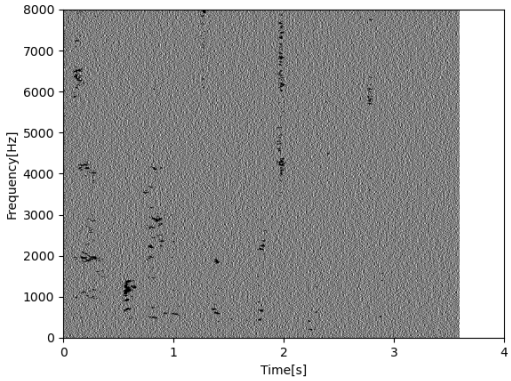

図 9: Spectrogram of restored mask (testB, EVCS).

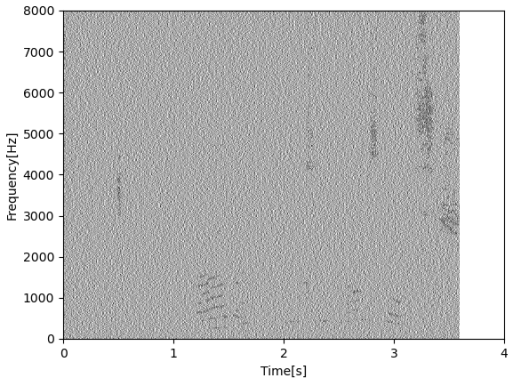

図 10: Spectrogram of a share mask (testB, EVCS).

EVCS　図 9 には EVCS での復元マスクのスペクトログラムを示す。testA と同傾向である。図 10 からわかる通り、カバー音声の調波構造も保持できている。

SDR での評価　図 11 には復元信号の SDR を示す。図 5 の testA の場合には 5~7 dB の SDR 改善が見られたが、ここでは 3~5.5 dB となっており、混合話者を増やすと SDR 改善量も低下した。また今回は話者 1 と話者 6 の間の SDR 改善量の差は少ない。秘匿しやすい話者かどうかは混合話者との関係によっても影響を受けることがわかる。図 6 の場合には EVCS の場合の復元信号の SDR を示す。testA の場合に 2~4 dB の SDR 改善が見られたがここでは 2 dB 程度となっている。

4.4 局所畳み込みを用いた場合

局所畳み込み (3.3) による場合の SDR を示す。図 13、図 14 がそれぞれ testA,B の場合で、図 5、図 11 に対応している。局所畳み込みの効果は大きい。

図 15、図 16 には EVCS の場合を示す。EVCS であっても局所畳み込みを用いれば高い復元性能となる。

5 まとめと今後の課題

本報では視覚暗号を音声秘匿通信に使うことを目的とした検討を行った。音源分離の問題が劣決定条件の場合

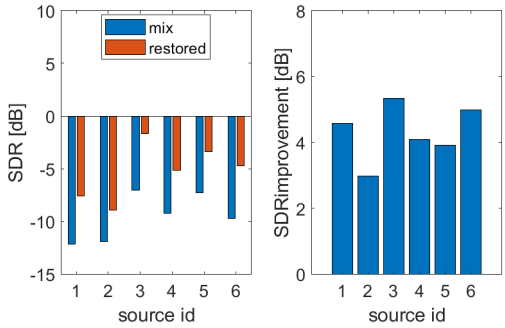

図 11: SDR [dB] of mixed signals and restored signals (testB).

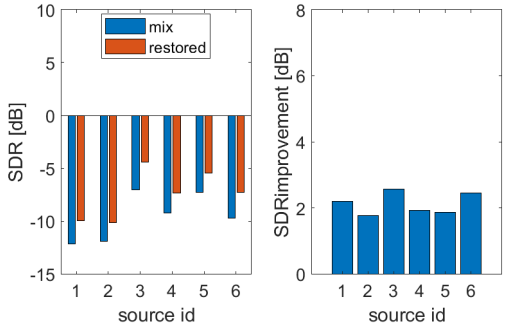

図 12: SDR [dB] of mixed signals and restored signals (testB,EVCS).

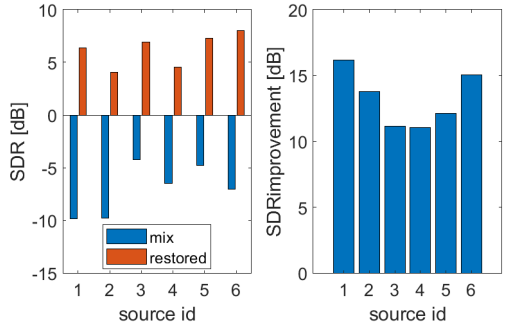

図 13: SDR [dB] of mixed signals and restored signals after local convolution (testA).

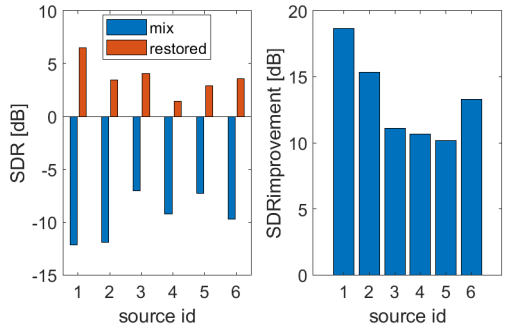

図 14: SDR [dB] of mixed signals and restored signals after local convolution (testB).

に解くことが難しいことを利用して、信号のマスキングに視覚暗号を使う方式を提案した。人の音声を用いた実

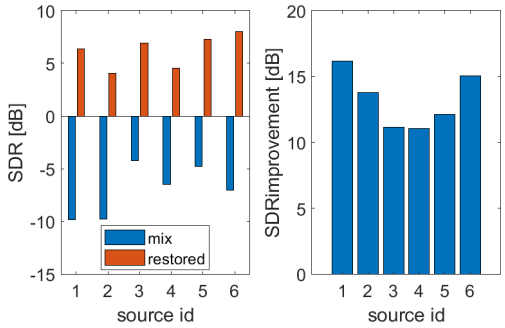

図 15: SDR [dB] of mixed signals and restored signal-safter local convolution (testA,EVCS).

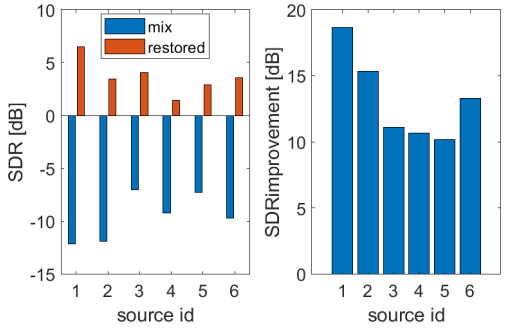

図 16: SDR [dB] of mixed signals and restored signals after local convolution (testB,EVCS).

験により、提案法の有効性を明らかにした。今回は最小構成の (t,n)=(2,3) で検討したが (t,n) を大きくした場合の検討、最適な混合音の作成法や、相対差を最大化する VCS を用いた場合の検討 [5] も今後の課題である。

参考文献

[1] 鵜木祐史，西村竜一，伊藤彰則，西村　明，近藤和弘，薗田光太郎，音響情報ハイディング技術，コロナ社，2018.

[2] 中山　彰，陸　金林，中村　哲，鹿野清宏，“心理音響モデルに基づいたオーディオ信号の電子透かし，” 電子情報通信学会論文誌. D-2，vol.83，no.11，pp.2255–2263，nov 2000.

[3] 仙場祐二，椎原浩雅，“情報送信装置、音響通信システムおよび音響透かし重畳方法，” 特許公開 2019-023762.

[4] M. Naor and A. Shamir, “Visual cryptography,” Advances in Cryptology - EUROCRYPT, vol.950, pp.1–12, 1994.

[5] 古賀弘樹，“視覚暗号の最近の進歩，” 電子情報通信学会技術研究報告，IT2020-25，pp.7–12，Dec. 2020.

[6] H. Koga, “A general formula of the (t,n)-threshold visual secret sharing scheme,” Advances in Cryptology - ASIACRYPT, pp.328–345, Dec. 2002.

[7] S.J. Shyu and M.C. Chen, “Optimum pixel expansions for threshold visual secret sharing schemes,” IEEE Trans. on Inf. Forensics and Security, vol.6, no.3, pp.960–969, 2011.

[8] M. Iwamoto, “A weak security notion for visual secret sharing schemes,” IEEE Trans. on Inf. Forensics and Security, vol.7, no.2, pp.372–382, 2012.

[9] G. Ateniese, C. Blundo, A.D. Santis, and D.R. Stinson, “Extended capabilities for visual cryptography,” Theoretical Computer Science, vol.250, no.1, pp.143–161, 2001.

[10] S.J. Shyu, “Threshold visual cryptographic scheme with meaningful shares,” IEEE Signal Processing Letters, vol.21, no.12, pp.1521–1525, 2014.

[11] K. Sekine and H. Koga, “Optimal basis matrices of a visual cryptography scheme with meaningful shares and analysis of its security,” ISITA, pp.422–426, 2020.

[12] E. Vincent, R. Gribonval, and C. Févotte, “Performance measurement in blind audio source separation,” IEEE Trans. on Audio, Speech, and Language Processing, vol.14, pp.1462–1469, July 2006.

[13] 太刀岡勇気，“視覚暗号を用いたスペクトログラムマスキング，” 電子情報通信学会技術研究報告 121，EA2021，Oct. 2021.