

音声付き動画を対象とした改ざん検出手法

A Tamper Detection Method for Movie with Sound Data

荻原 昭夫†

田中 啓介†

岩田 基†

汐崎 陽†

Akio Ogihara

Keisuke Tanaka

Motoi Iwata

Akira Shiozaki

1. まえがき

音声・画像・動画・音声付き動画などのデジタルデータは、コンピュータを使用することで痕跡を残すことなく編集・加工などの改ざんを行なうことが出来てしまう。そこで、これらのデジタルデータを証拠として用いるためには、デジタルデータに一切の改ざんが加えられておらず、原本であることを検証する必要がある。音声・画像・音声を伴わない動画については電子透かしを利用した種々の改ざん検出手法[1][2]が既に提案されているが、音声付き動画を対象とした改ざん検出に関する研究はほとんど報告されていない。そこで本稿では、相関型の電子透かしを利用し、時系列の改ざんを検出することのできる音声付き動画を対象とした改ざん検出手法を提案する。

2. 音声付き動画への透かしの埋め込み

2.1 透かし情報の決定

提案手法では、音声付き動画の「音声データ」および「動画データ」を 1/30 秒ごとのフレームに分割し、各フレームに一連の異なる透かし情報を埋め込む。各フレームへ埋め込むべき透かし情報の決定過程を図 1 に示す。まず、秘密鍵 k を乱数のシードとして、1~ L の整数により構成される乱数系列 $\alpha_k(i)$ を生成する ($i = 1, 2, 3, \dots$)。次に、+1, -1 により構成される長さ P ビットの L 種類の擬似乱数

$$R_j = \{r_j(m) \mid r_j(m) = \{1, -1\}, 0 \leq m \leq P-1, 1 \leq j \leq L\} \quad (1)$$

からなる擬似乱数表を準備する。

例として、第 1 フレームに埋め込むべき透かし情報を求めてみる。図 1 では $\alpha_k(1) = 5$ であるので、事前に準備した擬似乱数表の 5 番目に格納されている擬似乱数 R_5 を第 1 フレームへ埋め込む透かし情報として用いる。第 2 フレームについては、 $\alpha_k(2) = 53$ であるので、擬似乱数 R_{53} を透かし情報として用いる。以下同様に、第 i フレームに対しては擬似乱数 $R_{\alpha_k(i)}$ を透かし情報として使用する。

2.2 動画データへの透かしの埋め込み

動画データをフレーム単位に分割することにより、静止画像が得られる。図 2 に示すように、第 i フレームの静止画像にウェーブレット変換を施して、 LL_3 成分のウェーブレット係数 (図中の赤線で囲まれた部分) を求める。これらの係数から P 個の係数 $W_i(m)$ を選択し、2.1 節で決定した透かし情報 $r_{\alpha_k(i)}(m)$ を次式に従って埋め込むことで、透かし埋め込み後のウェーブレット係数 $W'_i(m)$ を得る。

$$W'_i(m) = W_i(m) + T_V r_{\alpha_k(i)}(m), \quad 0 \leq m \leq P-1 \quad (2)$$

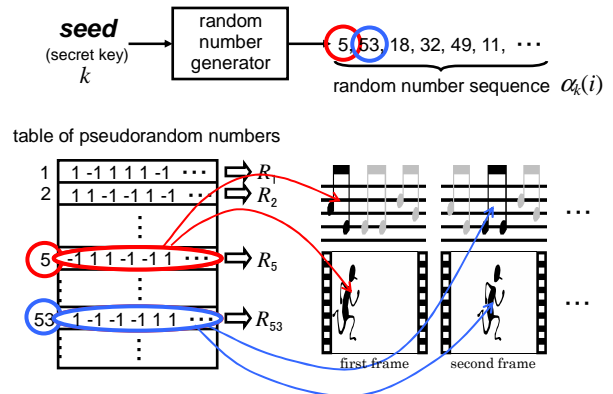
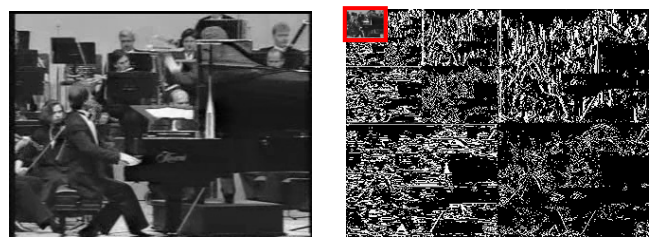


図 1 各フレームへ埋め込む透かし情報の決定



(a) 原画像

(b) ウェーブレット変換後

図 2 ウェーブレット変換処理と透かしの埋め込み位置

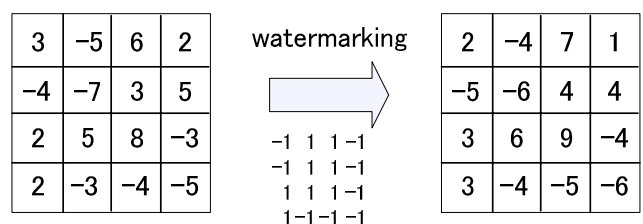


図 3 画像の LL_3 成分への透かしの埋め込み処理

ここで、式中の $T_V (>0)$ は透かしを埋め込む強度である。

$T_V = 1, P = 16$ の場合の埋め込み処理例を図 3 に示す。左側の図に示した透かしを埋め込む前のウェーブレット係数 $W_i(m)$ に対して、矢印下側に記した透かし情報を式(2)に従って埋め込む ($T_V = 1$ なので、対応する透かし情報の各ビット値を加算する) ことにより、右側の図に示した透かしを埋め込んだ後のウェーブレット係数 $W'_i(m)$ を算出する。

P 個のウェーブレット係数 $W_i(m)$ を $W'_i(m)$ に変更した後に、逆ウェーブレット変換を施すことで、透かし入りの静止画像が得られる。

† 大阪府立大学大学院工学研究科, Graduate School of Engineering, Osaka Prefecture University

2.3 音声データへの透かしの埋め込み

動画データのフレーム時間長 (1/30 秒) と同期をとりながら、音声データもフレーム単位に分割する。図 4 に示すように、分割された音声データの各フレームに変形離散コサイン変換 (MDCT) を施して、MDCT 係数を求める。MP3 圧縮などにより透かしが消失してしまうことの無いように中域部から P 個の係数 $Z_i(m)$ を選択し (中段の図における青線部分)、2.1 節で決定した透かし情報 $r_{\alpha_k(i)}(m)$ を次式に従って埋め込むことで、透かし埋め込み後の MDCT 係数 $Z'_i(m)$ を得る。

$$Z'_i(m) = Z_i(m) + T_A r_{\alpha_k(i)}(m), \quad 0 \leq m \leq P-1 \quad (3)$$

ここで、式中の $T_A (>0)$ は透かしの埋め込み強度である。

$T_A = 1, P = 16$ の場合の埋め込み処理例を図 4 に示す。中段の図において青線で記された透かしの埋め込み前の MDCT 係数 $Z_i(m)$ に対して、矢印右側に記した透かし情報を式(3)に従って埋め込む ($T_A = 1$ なので、対応する透かし情報の各ビット値を加算する) ことにより、下段の図において赤線で記された透かしの埋め込んだ後の MDCT 係数 $Z'_i(m)$ を算出する。

P 個の MDCT 係数 $Z_i(m)$ を $Z'_i(m)$ に変更した後に、逆 MDCT 変換を施すことで、透かし入りの音声データが得られる。

3. 改ざん検出

もし改ざんが行なわれていないのなら、同一フレームの静止画像と音声とは同一の透かし情報が埋め込まれているので、両者間から高い相関が検出できる。画像のウェーブレット係数 $W'_i(m)$ と音声の MDCT 係数 $Z'_i(m)$ の内積値を求め、 i 番目のフレームの静止画像と音声の改ざんの有無を検証する手順を以下に示す。

Step1 隣接する静止画像フレームのウェーブレット係数の差分値 $\Delta W'_i(m) = W'_i(m) - W'_{i+1}(m)$ を求める。

Step2 差分値 $\Delta W'_i(m)$ と同一フレームの MDCT 係数 $Z'_i(m)$ との内積値 IP を次式で計算する。 ($j = i$)

$$IP = \sum_{m=0}^{P-1} \Delta W'_i(m) Z'_j(m) \quad (4)$$

Step3 異なるフレーム ($j \neq i$) の MDCT 係数 $Z'_j(m)$ に対しても式(4)を計算し、各音声フレームに対する内積値を求める。

Step4 各音声フレームに対する内積値の平均 μ と標準偏差 σ を求め、画像フレームと同位置の音声フレームの内積値のみが閾値 $\mu + 3\sigma$ を超えていれば、透かしの正しく抽出できたので「改ざんされておらず、原本である」と判断する。

4. 実験

提案手法の有効性を確認するために、MPEG 形式で 1.5Mbps に圧縮された音声付き動画を実験対象として改ざん検出実験を行なった。画像サイズ 320 x 240、動画の埋め込み強度 $T_V = 2$ 、音声の埋め込み強度 $T_A = 1$ 、埋め込みビット数 $P = 300$ とし、音声付き動画の第 100 フレームに改ざんが施されているか否かを判定した結果を図 5 に示す。図 5 に示すように第 100 フレームのみが閾値を上回る内積値であることから、第 100 フレームは「改ざんされておらず、原本である」と判定できる。

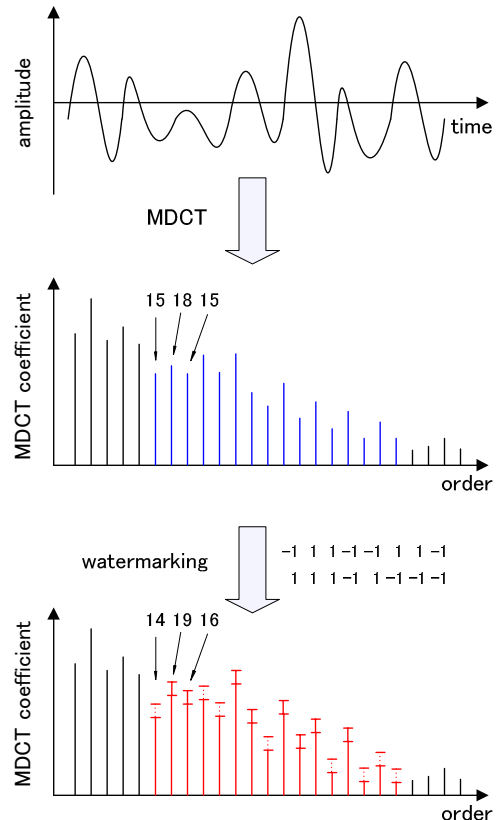


図 4 音声データへの透かしの埋め込み処理 (上段: 原音声信号, 中段: MDCT 変換後, 下段: 透かし埋め込み後の MDCT 係数)

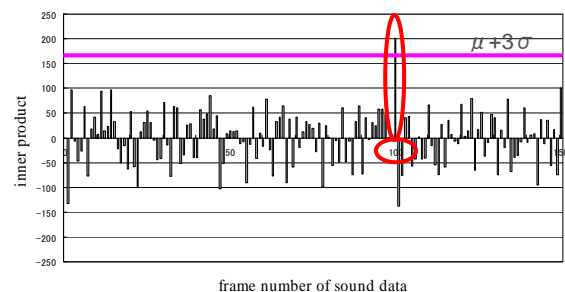


図 5 第 100 フレームの改ざん検出結果

5. むすび

音声付き動画に対する改ざんの有無を検出できる手法を提案した。本稿では秘密鍵の提示を必要としない改ざん検出処理について述べた。もし秘密鍵の提示が許されるのなら、各フレームに埋め込んだ透かしの正当性を検証することが可能となり、さらに厳密な改ざん検出を行なうことが出来る。

参考文献

- [1] 乗富賢一, 井上尚, 江島将高, “ MPEG映像に対する改ざん検出用電子透かし方式,” 情処学研報, 2003-CSCE -22 (31), 2003 .
- [2] 松井甲子雄, “ 電子透かしの基礎,” 森北出版, 1998 .