社団法人　情報処理学会　研究報告
IPSJ SIG Technical Report

2003－ＣＳＥＣ－22　(31)
2003／7／18

MPEG映像に対する改ざん検出用電子透かし方式

乗富 賢一† 井上 尚† 江島 將高†

†松下電器産業株式会社 マルチメディア開発センター
〒820-0067 福岡県飯塚市川津 693-47
E-mail: †{noridomi.ken-ichi, inoue.hisashi, ejima.masataka}@jp.panasonic.com

あらまし MPEG映像に対する改ざん検出が可能である電子透かし手法を提案する。提案手法は、フレーム単位の空間上の改ざんに加えて、フレーム編集等の時間軸上の改ざんも検出可能である。本稿では、提案手法の改ざん検出精度とセキュリティについて考察を行い、提案手法の有効性を示す。
キーワード 改ざん検出，電子透かし，MPEG，DCT係数

A Fragile Watermarking Method for MPEG Video

Ken-ichi NORIDOMI† Hisashi INOUE† and Masataka EJIMA†

† Multimedia Development Center, Matsushita Electric Industrial Co., LTD.
693-47 Kawazu, Iizuka City, Fukuoka, 820-0067 Japan
E-mail: †{noridomi.ken-ichi, inoue.hisashi, ejima.masataka}@jp.panasonic.com

Abstract We propose a fragile watermarking method for detecting alteration in MPEG video. This method makes it possible to detect alteration both on the temporal domain, e.g. frame-base editing, and on the spatial domain. In this paper, we consider this method from the point of view of detection accuracy and security. Furthermore we show the validity of this method.
Keyword Tamper Detecting, Watermarking, MPEG, DCT Coefficient

1. はじめに

近年、防犯などのセキュリティ上の観点から監視システムの需要が拡大している。特に、監視用記録装置では、ユーザーから長時間録画、高画質化が強く求められている。このため、デジタル化した画像データを圧縮し、デジタルのまま記録するデジタルディスクレコーダが急速に普及している。

その一方で、デジタルデータは、市販の画像処理ソフトウェアを利用して簡単に編集・加工などの改ざんを行うことできるため、記録したデジタル画像において原本性の保証が課題となっている。

この課題を解決するものとして、デジタル画像に改ざんが加えられているか否かを判定可能な改ざん検出技術がある。改ざん検出技術の一つとして知られているものに電子透かしがある。電子透かしとは、デジタル画像データに人間には知覚できないような形で、情報を画像と不可分に埋め込む技術である[1]。

現在、監視用記録装置おいては、記録される画像のフォーマットとして、Motion JPEG，Waveletが主に使用されている。しかしながら、今後はMPEGが広く使用されることが推測される。これは、(1)対符号量の画質が良い、(2)エンコーダ／デコーダ等のMPEGを扱う機器が数多くあり多様なシステム形態をとることができるからである。

しかし、これまでに提案されている改ざん検出手法[2]～[6]は、MPEGに対応していない。そこで、本稿では、MPEG映像に対する改ざん検出が可能である電子透かし手法を提案する。提案手法は、動画像に対する以下の改ざんを検出することが可能である。

(1) 空間上（フレーム内）の改ざん
　領域（顔等）のすり替え、領域消去等。
(2) 時間軸上の改ざん
　ある区間のフレーム抜き、差し替え等。

本稿では、まず、2章においてMPEG映像に対する改ざん検出が可能である電子透かし手法を提案する。そして、提案手法の改ざん検出精度とセキュリティについてアルゴリズムの観点で考察を行う。次に、3章において機能モデルを用いた改ざん検出機能と、電子透かしが画質に及ぼす影響について検証する。最後に、4章においてまとめを行う。

2. 提案手法

2.1. 概要

埋め込み側では、MPEG映像データから各フレームにユニークな固有情報を算出し、この固有情報を電子

透かし（WM）として埋め込む。

検出側では、WM 入り MPEG 映像データから各フレームの固有情報を算出する。また、WM 入り MPEG 映像データから埋め込み時の固有情報である WM データを検出する。算出した固有情報と検出した WM データとを比較照合して、一致する場合には改ざん無し、そうでない場合には改ざん有りと判定する。

なお、本稿においては、オリジナルと 1 画素でも異なれば改ざんと定義する。例えば、故意的な改ざんのない再符号化だけでも、それによってオリジナルとデータが異なるため、改ざんと判定する。

2.2. WM 埋め込みアルゴリズム

2.2.1. 処理の流れ

図 1 に WM 埋め込み処理の流れを示す。

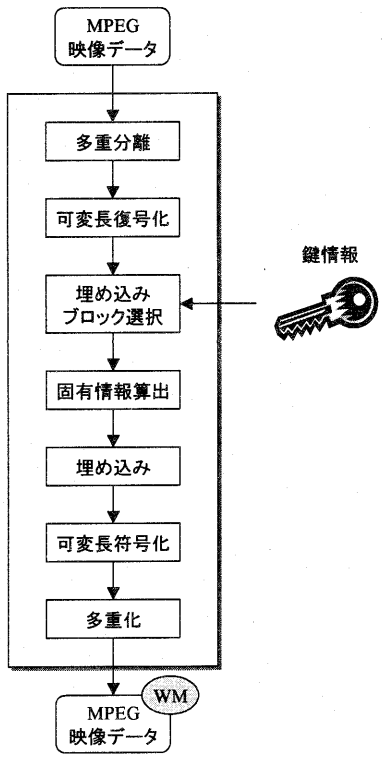

図 1 ：WM 埋め込み処理の流れ

以下の手順に従って、埋め込みを行う。

(1) まず、対象となる MPEG 映像データに対して多重分離を行う。次に、それぞれのフレームについて可変長復号を行い、一つのブロックが 8×8 個の DCT 係数からなる複数のブロックを得る。

(2) 鍵情報に基づいて、埋め込みを行うブロックを選択する。

(3) (1)で得た複数の DCT 係数から固有情報を算出する。

(4) (2)で選択したブロックに対して WM 埋め込みを行う。

(5) 可変長符号化、多重化を行い、WM 入りの MPEG 映像データを出力する。

なお、(3)固有情報算出と(4)WM 埋め込みで用いる DCT 係数は輝度成分とする。前記の手順によって WM が埋め込まれた MPEG 映像データを原本とする。

2.2.2. 埋め込みブロック選択

鍵情報 I_{key} は外部から与えられる情報である。鍵情報 I_{key} に基づいて、フレーム内の全てのブロックから埋め込み対象ブロックを選択する。埋め込み対象ブロックの個数は、埋め込む情報の総ビット数と等しくする。埋め込む情報の総ビット数については次項で説明する。

2.2.3. 固有情報の算出

以下の三つの情報を算出する。

(a) I ピクチャの固有情報 H_I

フレーム内の、全てのブロックの係数を入力としてハッシュ演算を行い、H_I を算出する。ただし、ハッシュ演算の入力値から WM を埋め込む対象となる DCT 係数は除く。これは、固有情報算出に用いた係数が WM 埋め込みによって変化すると、改ざん検出が意味をなさなくなることに起因する。

(b) P, B ピクチャの固有情報 H_{PB}

1 GOP 内の全ての P, B ピクチャをまとめて H_I と同様に、H_{PB} を算出する。ただし、P, B ピクチャには WM を埋め込まないため、固有情報 H_{PB} には、P, B ピクチャにおける全ての係数を用いる。

(c) 連続判定情報 I_{count}

連続性を示す情報 I_{count} を作成する。I_{count} は時間軸方向に連続する値である。

以上の三つの情報 H_I, H_{PB}, I_{count} を WM として埋め込む。H_I, H_{PB}, I_{count} のデータ長をそれぞれ N_I, N_{PB}, N_{count} ビットとすると、埋め込む情報の総ビット数は $(N_I+N_{PB}+N_{count})$ ビットである。

2.2.4. 埋め込み

埋め込む情報をそれぞれ 1 ビットずつ、選択したブロックに対応付けて、各 GOP 内の先頭 I ピクチャに対して埋め込みを行う。P, B ピクチャは符号化されているブロック数が埋め込む情報のビット数に満たない場合があり得るため、埋め込み対象とはしない。埋め込

むビットが0の場合には、対象係数の値を偶数に変更し、埋め込むビットが1の場合には、対象係数の値を奇数に変更する。この時、対象係数の増減値は最小になるようにする。

2.3. 改ざん検出アルゴリズム

2.3.1. 処理の流れ

図2に改ざん検出処理の流れを示す。

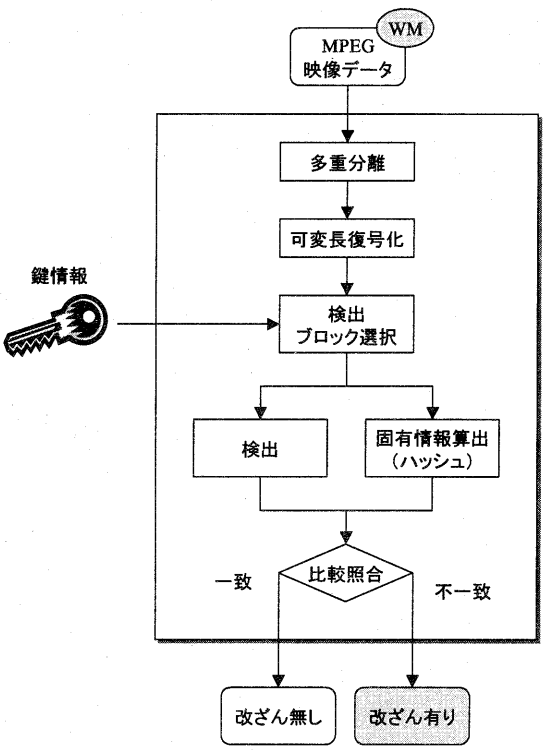

図2：改ざん検出処理の流れ

以下の手順に従って、改ざん検出を行う。

(1) まず、対象となるMPEG映像データに対して多重分離を行う。次に、それぞれのフレームについて可変長復号を行い、一つのブロックが8×8個のDCT係数からなる複数のブロックを得る。

(2) 鍵情報に基づいて、検出を行うブロックを選択する。

(3) (1)で得た複数のDCT係数から固有情報H_{cur}を算出する。

(4) (3)で選択したブロックからWM検出を行い、埋め込まれたWMデータH_{WM}を得る。

(5) H_{WM}とH_{cur}とを比較照合して、一致する場合には改ざん無し、そうでない場合には改ざん有りと判定する。

前記の手順において、固有情報H_{cur}の算出は埋め込み時と同じである。ただし、(2)において鍵情報は埋め込み時に使用した鍵情報と同じものを用いる。また、(3)固有情報算出と(4)改ざん検出で用いるDCT係数は、埋め込み時と同じく輝度成分とする。

2.3.2. 改ざん検出

固有情報算出によって得られた固有情報H_{cur}のうち、Iピクチャの固有情報、P，Bピクチャの固有情報に対応する情報をそれぞれH_{cur_I}，H_{cur_PB}とする。また、WM検出によって得られたWMデータH_{WM}は埋め込み時の固有情報に該当し、Iピクチャの固有情報、P，Bピクチャの固有情報、連続判定情報に対応するビット列をそれぞれH_I，H_{PB}，I_{count}とする。

空間上の改ざん検出

埋め込み時の固有情報H_IとH_{cur_I}を比較照合し、一致する場合には、Iピクチャに改ざん無しと判定し、そうでない場合には、Iピクチャに改ざん有りと判定する。

また、埋め込み時の固有情報H_{PB}とH_{cur_PB}を比較照合し、一致する場合には、GOP内に含まれるP，Bピクチャに改ざん無しと判定し、そうでない場合には、P，Bピクチャに改ざん有りと判定する。

このようにして、I，P，Bピクチャの空間上の改ざんをフレーム単位に検出する。

時間軸上の改ざん検出

GOP単位の差し替え、抜き取り等の時間軸上の改ざんを次のようにして検出する。連続判定情報I_{count}を繰り返し検出し、I_{count}の変化規則が埋め込み時と等しい場合に改ざん無しと判定する。そうでない場合に改ざん有りと判定する。

また、フレーム単位の差し替え、抜き取り等の時間軸上の改ざんも検出できる。この場合には、H_IとH_{cur_I}、またはH_{PB}とH_{cur_PB}が一致しなくなる。

2.4. 検出精度

本節では空間上の改ざんと時間軸上の改ざんに対する提案手法の検出精度について考察を行う。

2.4.1. 空間上の改ざん

本手法では、MPEG映像データにおいて、フレーム内の周波数成分値の一つが1でも変化すれば、ハッシュ値が変化するため改ざん有りと検出される。

ただし、固有情報の算出において、WMを埋め込む係数は用いていないため、埋め込み対象係数のみを改ざんした場合に、誤検出してしまう可能性が考えられる。しかしながら、フレーム内の全係数に比べ、埋め込み対象係数の個数は微少である。例えば、画像サイズが720×480pixelsのMPEG2映像において、128個の

ブロックを選択することを想定した場合、埋め込み対象係数は、フレーム内の全係数の 0.04％に満たない。このようなわずかな個数の係数のみを操作して意味のある改ざんは実質上困難であり、検出精度は実用上問題無いレベルと考える。

更に検出精度を向上させる手法もある。WM 埋め込み後の、フレーム内の全ての係数から固有情報を算出し、算出した固有情報を別の画像へ埋め込む手法である。この手法では、上述した誤検出の可能性は無くなるが、改ざん検出に複数のフレームを必要とする。

本手法において、I ピクチャの固有情報を、同じ I ピクチャに埋め込んでいるのは、改ざん検出時に I ピクチャ 1 枚からでも検出可能であるというメリットを重要視しているためである。

2.4.2. 時間軸上の改ざん

まず、GOP 単位に編集された場合について考える。この場合、検出した連続判定情報が、改変した区間で不連続となり改ざんが検出できる。連続判定情報のビット数を I_{count} とすると、改ざん後に検出した連続判定情報が、偶然連続と判定される確率（誤検出率）は $2^{-Icount}$ である。I_{count} のビット数が十分とれれば実用上問題ないレベルとなる。

次に、フレーム単位に編集された場合について考える。I ピクチャが抜かれた場合には、それに続く P，B ピクチャは再生できなくなり、MPEG 映像データとして成り立たなくなる。一方、I ピクチャが加えられた場合には、そのI ピクチャには WM が埋め込まれていないため、改ざんが検出できる。また、P，B ピクチャが編集された場合には、参照元である I ピクチャに予め埋め込まれている WM と、P，B ピクチャの固有情報が一致しなくなり、改ざんが検出できる。この場合の誤検出率は、上述した空間上の改ざん検出の場合と同じであり、問題ないレベルと考える。

2.5. セキュリティ

電子透かしの場合、悪意のある第三者による埋め込み情報の改ざんが困難であることも重要な要求条件である。ここで、画像の改ざんと区別するために、埋め込み情報の改ざんを WM 改ざんと呼ぶ。WM 改ざんが可能な場合、改ざんを加えた画像に対して、新たに電子透かしを埋め込み、それを原本と主張することが可能となる。そこで、本節では本手法のセキュリティ強度について考察する。第三者に本手法のアルゴリズムを公開したと仮定する。それでも、使用しているハッシュ関数を解読しないと、埋め込み情報を改ざんすることはできない。つまり、アルゴリズムを公開してもハッシュ関数で守られる。

更に、使用しているハッシュ関数をも解読されても、外部から与えられる鍵情報で守られる。本手法では、鍵情報に基づいて埋め込みブロックの選択を行う。例えば、画像サイズが 720×480pixels の MPEG2 映像において、128 個のブロックを選択することを想定した場合、この組み合わせは ${}_{5400}C_{128}$ 通りとなる。このことから、たとえアルゴリズムとハッシュ関数が解読されても、悪意のある第三者が埋め込み情報を改ざんすることは困難である。

3. 検証

3.1. 目的

本手法の機能モデルを用いて、改ざん検出機能と、WM が画質に及ぼす影響について検証する。図 3 に本手法の機能モデルの構成を示す。

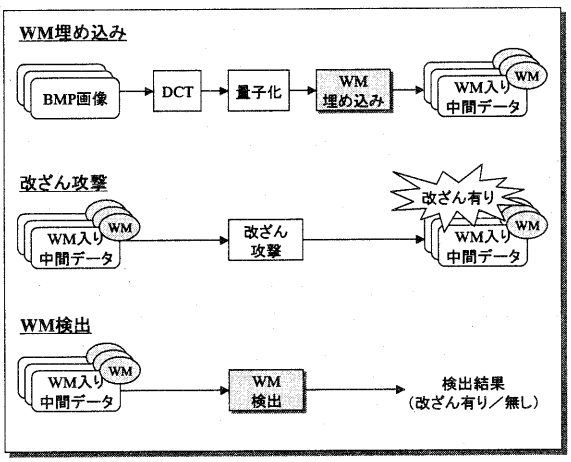

図 3 ：機能モデルの構成

図 3 において、WM 埋め込みの機能モデルは動画像を構成する連続フレームの BMP 画像を入力とし、DCT、量子化処理後の DCT 係数に対して WM 埋め込みを行い、WM 入り中間データを出力する。量子化の処理においては、MPEG2 規格のイントラピクチャ用のデフォルトマトリックスを用いている。機能モデルは、入出力を BMP 画像とした点で図 1 と異なる。しかし、量子化された DCT 係数に対して埋め込みを行うコアアルゴリズムは一致する。ただし、機能モデルにおいては、全てのフレームを I ピクチャと想定する。また、埋め込みに使用する固有情報は I ピクチャの固有情報 H_I のみであり、N_I＝128 ビットとする。

一方、WM 検出の機能モデルは WM 入り中間データから WM 検出を行い、検出結果を出力する。入力を WM 入り中間データとした点で図 2 と異なる。しかし、量子化された DCT 係数から検出を行うコアアルゴリ

ズムは一致する。

また、WM 入り中間データに各種改ざんを加える関数を準備し、各種改ざんを加えたデータを出力する機能も備える。

3.2. 検証方法

3.2.1. 改ざん検出機能検証

原本からの検出

図 3 における WM 埋め込み直後の WM 入り中間データ（原本）から、改ざん無しと正確に検出されることを確認する。

攻撃を受けたデータからの検出

原本に対して、各種改ざんを加える関数を用いて作成した改ざん有りデータから、WM の検出を行い、改ざん有りと検出されることを確認する。ここでは、以下の二つの攻撃を加えた。

(1) 再符号化：まず、原本に対して逆量子化、逆 DCT を行い再度 BMP 画像へ戻す。次に、再度 DCT、量子化を行い、原本と同じ形式の中間データに戻す。この時、逆量子化、量子化に用いる量子化スケールコードは、埋め込み時と同一とする。

(2) ランダム攻撃：各フレームにおいて、全ての DCT 係数から、ランダムに一つの係数を選択し、その値を“1”増加させる。

3.2.2. 画質検証

WM 有り画像と WM 無し画像を以下のように定義する。

WM 有り画像：図 3 において、WM 埋め込み直後の WM 入り中間データに対して逆量子化、逆 DCT を行った画像。

WM 無し画像：図 3 の WM 埋め込みにおいて、WM 埋め込み無しで作成した中間データに対して逆量子化、逆 DCT を行った画像。

WM 有り画像と WM 無し画像との S/N 比を求め、WM の画質への影響を確認する。なお、S/N 比の算出にはそれぞれの輝度値を用いる。

3.3. 検証結果

入力画像として、六つの標準画像（Flower Garden, Susie, Table Tennis, Mobile and Calendar, Football, Cheer-leaders）を使用した。入力画像の画像サイズはいずれも 720×480pixels で、それぞれ 600 フレームある。また、図 3 における量子化処理に用いる量子化スケールコード（QSC）は 1, 8, 16, 24, 31 の五つのパターンについて検証を行った。

3.3.1. 改ざん検出機能検証

原本からの検出

原本から改ざん検出を行った結果、全てのパターンにおいて、改ざん無しと検出された。

攻撃を受けたデータからの検出

再符号化を行われたデータからは殆どの場合改ざん有りと検出された。これは、再符号化処理が非可逆であり、この処理を行われたデータは DCT 係数が微小に変化しているためである。しかしながら、一部改ざん無しと判定されたフレームが存在した。その場合に、再符号化データにおけるフレーム画像内全てのブロックの DCT 係数と、埋め込み時の DCT 係数の比較を行ったところ、完全に一致した。つまり、改ざん無しと判定されたフレームについては、再度行った DCT、量子化処理が埋め込み時の DCT 係数に影響を与えなかった場合であり、実質改ざんされていないことが確認できた。

ランダム攻撃を受けたデータからは、全ての場合において改ざん有りと検出できた。したがって、フレーム内のブロックのうち、その一つの係数のみが 1 だけ変化した場合においても、改ざん有りと検出できることが確認できた。

3.3.2. 画質検証

WM 無し画像と WM 有り画像との S/N 比は 80～55dB の範囲であった。したがって、画質劣化は殆どないことが分かった。図 4 に Flower Garden の場合の一例を示す。

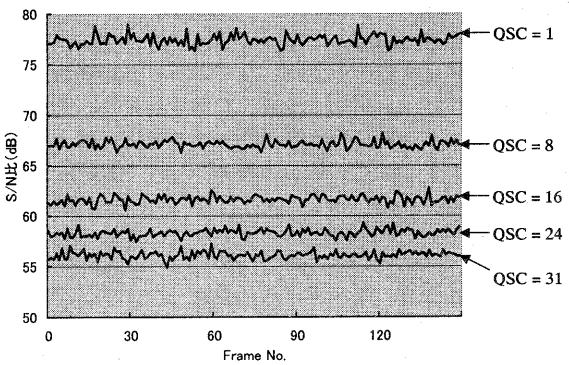

図 4 ：画質検証結果（Flower Garden）

画質を保持できる理由は、WM は検証に用いた画像（サイズ 720×480pixels）の 5400 個のブロックのうち、128 個のブロックのみに埋め込まれるため、埋め込み前後の差異が殆どないからである。また、WM は量子化された DCT 係数に対して埋め込まれるため、量子化時の QSC の値が大きくなるほど、WM 埋め込み画像の輝度値の変化量が大きくなる。しかしながら、図 4 において、QSC＝31（高圧縮）の場合でも、S/N 比は約 56dB であった。このように、高圧縮の場合でも数値的

には問題なく、WMの画質への影響は問題ないレベルであることが確認できた。また、主観評価においても、WM無し画像とWM有り画像の違いは知覚できなかった。

補足検証として、MSSG[7]のMPEG2エンコーダを用いて、本検証におけるQSCの値が対応するビットレートを検証したところ、本検証におけるQSC＝8でビットレートが約6 Mbps、QSC＝16程度でビットレートが約3 Mbpsであることを確認している。

4. まとめ

本稿では、MPEG映像に対する空間上の改ざんと時間軸上の改ざんの両方を検出可能な電子透かし手法を提案した。検出精度とセキュリティの観点においても、実用上は問題ないレベルであると考える。

本手法の主な特徴を以下にまとめる。

・フレーム単位の改ざんを検出可能

空間上の微小な改ざんと時間軸上の改ざんの両方を検出可能。

・画質劣化防止

一部のDCT係数にのみ最小の変化量で埋め込む。

・WM改ざんが困難なセキュリティ強度

本稿において、検証に用いた機能モデルは本手法のコアアルゴリズムの検証を行う機能限定版であった。今後は、実際の適用形態に基づいた検証を行う。MPEG映像データを入出力とするフルスペック版で改ざん検出機能の検証を行う。また、リアルタイム化を行い、実用化していく。更に、改ざん位置の特定や、悪意のある改ざんと、再符号化等のフォーマット変換との判別を可能とする改ざん検出技術についても検討を進める。

文　献

[1] 松井甲子雄，電子透かしの基礎－マルチメディアのニュープロテクト技術－，森北出版，1998.

[2] 森藤元，安細康介，渡辺大，吉浦裕，瀬戸洋一，“証拠写真を撮影可能なデジタルカメラ，”ITE Technical Report，vol.25，no.61，pp.43-48，Sep. 2001.

[3] 貴家仁志，松井勝之，“JPEG2000を用いた画像の改ざん防止法，”信学技報，vol.102，no.152，pp.25-30，Jun. 2002.

[4] 岩村恵市，林淳一，櫻井幸一，今井秀樹，“改ざん位置検出用電子透かしに関する考察と提案，”画像電子学会誌，vol.32，no.1，pp.22-28，2003.

[5] 汐崎陽，“改ざん前の状態が分かるデジタル画像の改ざん検出用電子透かし法－JPEG画像への埋め込み－，”SCIS2003講演論文集，Jan. 2003.

[6] 井藤浩，馬養浩一，鈴木光義，“JPEG画像の真正性を証明する電子透かしの方法，”電子情報通信学会2003年総合大会講演論文集，D-11-33，pp.33，Mar. 2003.

[7] MPEG Software Simulation Group，http://www.mpeg.org/MSSG/