

リアルタイム検出可能な動画向けモバイル電子透かし

中村 高雄[†] 山本 奏[†] 北原 亮[†] 宮武 隆[†] 片山 淳[†]

[†] 日本電信電話株式会社
NTT サイバースペース研究所
〒 239-0847 横須賀市光の丘 1-1

あらまし カメラ付き携帯電話機を用いてアナログ画像とデジタル情報を連携させるモバイル電子透かし技術があるが、従来は一枚の写真から電子検出を行なう静止画方式であった。しかし端末の能力向上および高速手法の提案により、リアルタイム検出可能な動画向けモバイル電子透かし技術の可能性が考えられる。本稿では技術の位置付けを明らかにして要件定義を行なった上で具体的な方式検討を行ない、さらに実機上に実装し評価実験により有効性を確認する。

キーワード 電子透かし, 携帯電話, 動画, カメラ, リアルタイム

Video watermarking scheme for real-time detection on camera phone

Takao NAKAMURA[†], Susumu YAMAMOTO[†], Ryo KITAHARA[†], Takashi MIYATAKE[†],
and Atsushi KATAYAMA[†]

[†] NTT Cyber Space Laboratories,
Nippon Telegraph and Telephone Corporation
1-1 Hikarinooka, Yokosuka, Kanagawa, 239-0847 Japan

Abstract Digital watermarking technology can be applied to mobile phone; the application system provides information related to pictures taken by camera phone. The watermarking technique for still image has been applied to mobile phone because of low computation capacity. There exists, however, some possibility of real-time detection from captured video, according to the progress of computational performance and the fast algorithms. In this paper, we clarify the requirements for the technology as a whole, and propose a watermarking scheme designed to be robust and detectable on real-time. Finally, we summarize experimental results and confirm the efficiency of the proposed scheme.

Key words digital watermark, mobile phone, video, camera, real-time

1. はじめに

電子透かし技術の応用として、印刷やモニタ表示などでアナログ出力された透かし入り画像をカメラ付き携帯電話機で撮影して電子透かし検出を行ない、検出した透かし情報を用いてアナログ情報とデジタル情報を連携させるモバイル電子透かし技術が提案されている [1] [2]。従来のモバイル電子透かし技術は、PC などと比較して処理能力が低い携帯電話機上で比較的处理量の多い画像処理である電子透かし検出を行なう必要があるため、基本的に静止画像を一枚撮影して検出処理を行なう静止画向け電子透かし方式

であった。

しかし、携帯電話機の処理能力の向上、および高速処理可能な携帯電話機向け静止画電子透かしアルゴリズムの提案 [2] などにより、カメラで連続的にキャプチャされるフレーム画像からの電子透かし検出をリアルタイムに行なうことが現実味を帯びてきた。これにより、TV 映像などの動画からのリアルタイム検出といった新たな技術の可能性が考えられる。

本稿では、技術の位置付けを明らかにして要件定義を行なった上で、リアルタイム検出可能な動画向けモバイル電子透かし方式の検討を行ない、さらに実機上に実装し評価



図 1 動画向けモバイル電子透かしの利用シーン

実験により実現可能性を確認する。

2. 技術の位置付けと要件定義

リアルタイム検出可能な動画向けモバイル電子透かし技術の利用シーンを図 1 に示す。動画を視聴している利用者が視聴中のコンテンツに関連する情報を取得する際は、カメラ付き携帯電話機を表示画面に向けてビデオキャプチャし、リアルタイムに電子透かしを検出する。検出成功した場合は関連情報提供サーバに透かし情報である ID 情報を送信し、ID 情報に対応する関連情報を取得して様々なサービスを利用することができる。

2.1 関連技術

動画向けモバイル電子透かし技術と同様の機能を実現可能と思われるいくつかの関連技術について考察する。

2.1.1 タイムスタンプ利用

放送中の TV 番組に関する情報を取得したい場合に、現在時刻 (と放送チャンネル情報) をサーバに送信し、サーバ側で番組表に基づいて視聴シーンを特定して関連する情報を提供する方式である。しかし、この方法では PVR (Personal Video Recorder) などに録画したコンテンツをタイムシフトして再生している場合や、DVD などのパッケージ媒体に記録されたコンテンツの再生の場合に関連情報を得ることが不可能である。

2.1.2 2次元コード利用

電子透かしの代わりに 2次元コードを用いて、動画中に 2次元コードを表示し、これをカメラ付き携帯電話機で読みとって関連情報を取得する方式である。しかし、この方法ではコードパターンが表示画面上で一定の面積を一定時間占有するため、コンテンツの視聴の妨げとなるという問題がある。

2.1.3 静止画電子透かし技術利用

従来の静止画向けモバイル電子透かし技術 [2] を用いて、電子透かしが埋め込まれたロゴ画像などの静止画像を表示し、表示画面を撮影し透かしを検出して関連情報を取得する方式である。この方法では透かし入り静止画像を一定時間表示しなければならず、動画コンテンツ制作に制約が生

表 1 関連技術との比較

	品質 /制作 の制約	既存 環境 対応	実現 容易 性	タイム シフト 対応	個体化 (メディ ア区別)
タイムスタンプ	○	○	○	×	○
2次元コード	×	○	○	○	○
静止画透かし	△	○	○	○	○
データ放送	○	△	○	○	○
特徴量マッチング	○	○	△	○	×
動画透かし	○?	○	○?	○	○

(△:困難, ×:原理的に不可能)

じるという問題がある。

2.1.4 データ放送利用

チューナ装置などによって再生中コンテンツの関連情報をデータ放送などのチャンネルを通じて取得し、必要ならチューナからの赤外線通信などによって携帯電話機で関連情報を取得する方式である。しかし、この方法では地上デジタル放送などデータ放送チャンネルを持つことを前提とし、またチューナ装置と携帯電話機双方にハードウェア的な機能追加が必要であるため、既存のメディアやハードウェアでは利用できない。実際、2005 年 3 月末での携帯電話総契約数 8,700 万契約のうち動画カメラ機能を有する携帯電話機は 63.8%利用されており、台数に換算すると 5,550 万台に上る。一方、2005 年 4 月末の地上デジタル放送受信機の出荷台数は累計 433.5 万台であり、現時点ではモバイル電子透かし技術よりも適用範囲が狭いという問題がある [3]。また、例えば屋外のオーロラビジョンから関連情報を取得するようなことがデータ放送利用方式では困難である。

2.1.5 特徴量マッチング利用

撮影したコンテンツの特徴量を用いて、特徴量 DB を持つサーバにおいてマッチングにより撮影されたコンテンツを特定して関連情報を取得する方式である [4]。まず、静止画特徴量マッチング利用方式について考えると、静止画電子透かし利用方式と同様、特徴量が登録されているロゴ画像などを一定時間表示しなければならず、動画コンテンツ制作に制約が生じるという問題がある。また、動画特徴量マッチング利用とした場合でも、動画のリアルタイム処理という観点では、特徴量算出処理量、サーバへの送信データ量、サーバ側で保持すべき特徴量 DB のレコード数などに問題があり実現が難しいと考えられる。また、端末側で得られた特徴量から関連情報が得られるか否かはサーバ側でマッチングしなければ分からないため、ピンボケや手ぶれなどによって正しく特徴量が得られていない場合でも必ずサーバに要求を送らなくてはならず、利便性の面で難がある。さらに、同一の絵柄のコンテンツから得られる特徴量は一つのみであるから、同一のコンテンツに対応付け可能な関連情報は一通りのみである。例えば同一のコン

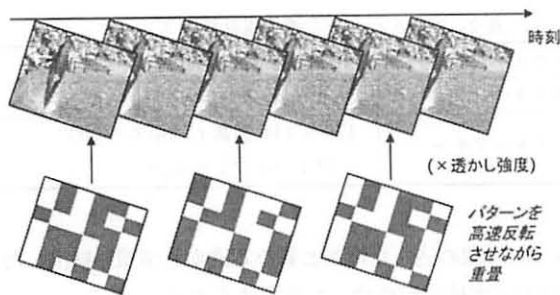


図2 時間方向高周波パターン埋め込み

テンツに対応付ける関連情報を、TV放送用とDVD用などメディアの違いを区別して異なるようにすることができないため、pay-per-click型の広告ビジネスへの展開が不可能など、適用範囲が狭くなるという問題がある。

2.1.6 動画モバイル電子透かし技術利用

提案する動画モバイル電子透かし技術においては、関連情報を表すID情報をコンテンツデータ自体に埋め込むことにより、再生時刻によらず関連情報を取得可能である。また、メディア毎に異なるIDを埋め込むことでコンテンツを個別化し、同一コンテンツであってもメディア毎に異なる関連情報を提供可能である。そして、既存のTV放送やDVDなどに容易に適用することが可能である。また、電子透かし検出を端末上でスタンドアロンで行ない検出結果の信頼性を評価することにより、サーバとの不必要な通信処理を避けることができる。

電子透かし技術の必然である画質劣化が知覚できないレベルであるか否か、および本当に実現可能であるか否かについては現時点では不明である。これらの不明点に関して、本稿で示す提案方式の有効性が確認できれば、関連情報提供サービスに動画向けモバイル電子透かし技術を用いることについて、関連技術と比較して十分なメリットがあると考ええる。以上の議論をまとめると表1のようになる。

2.2 要件定義

以上の利用シーンや関連技術との関係を考慮し、本技術に求められる要件は以下のようになると考えられる：

- (a) 画質劣化が視聴の妨げにならないこと
- (b) 様々な撮影条件で検出可能なこと
- (c) 撮影を開始してから短時間で検出成功すること
- (d) 検出試行に対する応答の遅延が少ないこと
- (e) 検出した透かし情報の信頼性が保証されていること

電子透かしはコンテンツデータの変更により別の情報を埋め込む技術であるため、コンテンツの劣化が必然である。そこで要件(a)では埋め込みによる画質劣化が知覚されないことを要求する。

要件(b)～(d)は利便性に関する要件である。要件(b)について、例えば検出のためにカメラを三脚に固定することが必要では、関連技術と比較して利便性が低いと言わざる

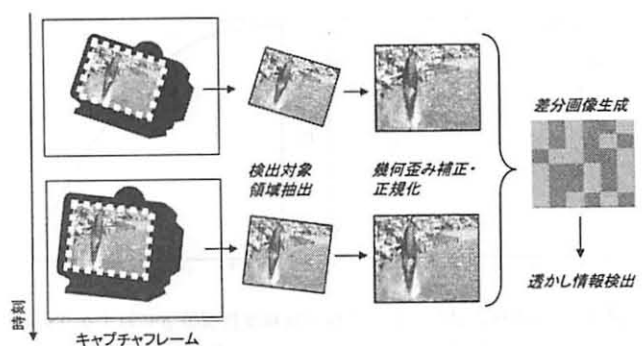


図3 検出対象領域抽出と差分処理

を得ない。また、要件(c)では、利便性の観点から携帯電話機スタンドアロン処理を前提とした上で、短時間で結果が得られるような高速な方式を要求する。許容可能な応答時間については、電話をダイヤルしてから応答音を聞くまでと同様に、検出を開始してから電子透かし検出結果を得るまでの時間が2秒以内となること、いわゆる「2秒ルール」がユーザインタフェース上の要件として適切であると考ええる[5]。さらに要件(d)については、検出状況をリアルタイムにフィードバックすることによる操作性向上を期待するものである[6]。

要件(e)は、実サービスへの展開には必須の要件である。これが満たされないと、例えば本来埋め込まれている透かし情報とは異なる情報を誤検出してしまい、全く無関係な関連情報を提示してしまうといった問題が生じてしまう。

3. 提案方式

前節で示した要件を満たすために、我々が以前に提案した携帯電話機向けの高速な静止画電子透かし方式[2]を元に以下のような方式を検討した。

3.1 時間方向高周波パターン埋め込み

まず、文献[2]の方法と同様に、透かし情報を誤り訂正/検出符号化し、符号語をスペクトラム拡散変調した後に2次元パターンに変換して透かしパターンを生成する。そして、劣化の程度/耐性のバランスを制御する透かし強度で透かしパターンを増幅し、動画像の各フレームに対して重畳して埋め込みを行なう。ただし後述する検出処理との兼ね合いから、埋め込みパターンの正負を時間方向で高速に反転させながら重畳を行なうようにする(図2)。

3.2 検出対象領域抽出、幾何歪み補正

カメラによる動画キャプチャによって連続的に得られるキャプチャフレーム画像に対し、高速処理可能な矩形領域抽出アルゴリズム[7]を適用して、ディスプレイ装置の表示領域をキャプチャフレーム画像中から検出対象領域として抽出する。さらに撮影角度によって生じる平面射影歪みを補正しサイズを正規化して検出対象領域画像を生成する(図3)。

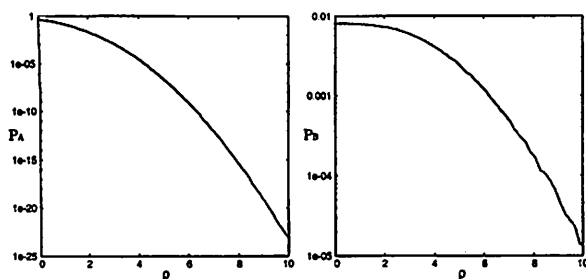


図4 1回の検出試行における検出信頼度指標値 ρ に対する Type A. 誤り率 P_A (左) と Type B. 誤り率 P_B (右)

3.3 フレーム間相関を利用した高精度検出

文献[2]の方法を含むスペクトラム拡散変調利用型などの加法的なアルゴリズムでは、透かし信号に対して原画信号が雑音として振舞う。しかし、動画画は一般的に時間的に近接したフレーム間の相関が高いため、近接した2枚のフレーム画像の差分を求めれば原画信号を大きく抑圧することが可能となる。さらに透かしパターンの正負を反転させながら埋め込みを行なっているため、2枚のフレームに埋め込まれている透かしパターンが正負逆転しているとすれば、差分画像中の原画信号が抑圧されるのみならず透かしパターンの振幅が倍となり、S/N比の向上により電子透かしの耐性向上が期待できる。また、各キャプチャフレーム毎に幾何歪み補正を行なって得られる検出対象領域画像の差分を取ることで、撮影角度およびカメラの動きにも対応することができる(図3)。

次に、差分画像をノイズの付加された透かしパターンとみなして、文献[2]と同様にスペクトラム拡散復調、誤り訂正/検出復号処理を行ない、後述する信頼度判定を行なって有意な透かし情報を検出したと判定した場合に、検出成功とする。

なお、差分のタイミングによっては差分画像中の透かしパターンが本来のパターンの正負反転となることがある。文献[2]のように、スペクトラム拡散復調時に相関値の正負で検出ビットを決定し、さらに誤り訂正/検出符号としてBCH符号のようなビット反転した符号語からはビット反転した情報が正しく検出されてしまうような符号を用いた場合、差分のタイミングによってビット反転した透かし情報を誤検出してしまうおそれがある。この問題に対して、今回は簡単のため透かし情報中の1ビットを反転フラグとして利用して検出透かし情報のビット反転を防ぐようにする。

3.4 電子透かし検出の信頼性評価

2.2節で挙げた要件(e)のために、電子透かし検出処理によって得られた透かし情報の信頼性を定量的に保証することを考える。検出時に起こり得る検出の誤りについては、Type A. 「電子透かしが埋め込まれていない画像にもかかわらず検出が成功したと出力する誤り」、あるいは、

表2 実装を行なった携帯電話機のスペック

プロセッサ	104MHz 32bit RISC
動作プラットフォーム	Symbian OS 7.0s Series 60 (C++)
ビデオキャプチャ	176 × 144 画素 (4:2:0 format) 平均 14.92 フレーム/秒

Type B. 「埋め込んだ情報と異なる透かし情報を検出したにもかかわらず検出が成功した出力する誤り」の2通りが考えられる。

文献[2]ではスペクトラム拡散復調の際の相関値に基づいて検出信頼度指標を導入し、Type A. の誤りに対しては統計的検定の議論によって、Type B. についてはBER分布と誤り訂正/検出符号の特性分析によって、どちらのケースについても誤り率を定量的に抑えるための閾値を導出するための設計基準を示している。この議論に基づいて、評価実験に用いたパラメータにおける1回の検出試行に対するType A. 誤り率 P_A 、Type B. 誤り率 P_B と、検出試行時に得られる検出信頼度指標値 ρ の関係を図4に示す。例えば検出信頼度指標値 ρ に対して、閾値8.5以上ならば検出成功という基準を設定した場合、1回の検出試行について $P_A < 10^{-17}$ 、 $P_B < 10^{-4}$ を保証することができる。

提案方式では動画キャプチャからの検出試行を繰返し行なう。1試行あたりの検出誤り率を p_0 とするとき、 N 回の繰返し検出試行における検出誤り率 p は、

$$p = 1 - (1 - p_0)^N \simeq Np_0 \quad (1)$$

となる。後述する評価実験により、2秒間の検出試行回数は $N = 28$ である。これにより、例えば1試行に対する閾値を8.5としたときの2秒間の検出試行時については $P_A < 2.8 \times 10^{-16}$ 、 $P_B < 2.8 \times 10^{-3}$ を保証することが可能である。

また、同一の閾値がType B. よりもType A. により厳しい基準となること、およびType A. の誤り評価が透かしの存在有無の評価に相当することに注目すれば、 P_A に対する閾値 α と P_B に対する閾値 β を $\alpha < \beta$ なる異なる値に設定することによって、

- (i) 「電子透かしが存在しない」($\rho < \alpha$)
- (ii) 「電子透かしは存在するが検出結果に誤りがある可能性がある」($\alpha \leq \rho < \beta$)
- (iii) 「電子透かしを正しく検出」($\rho \geq \beta$)

の3つ検出信頼度ステータスを得ることができる。(ii)のケースを用いて後述するように電子透かしの検出可能性を利用者にフィードバックすることが可能である。

4. 実装と評価

提案方式の有効性を確認するため、携帯電話実機への実装を行ない、評価実験を行なった。

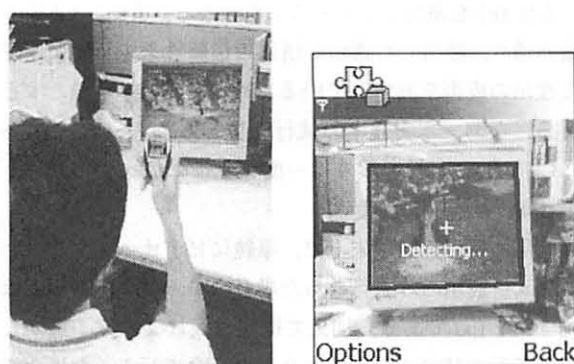


図5 実験風景(左)と実験プログラムの実行中画面(右)

4.1 実装

実験プログラムの実装を行なった携帯電話機のスペックを表2に示す。この機種は一代前のものであり既に動作周波数が倍以上の機種も存在するが、提案方式の高速性を確認する意味で敢えて用いている。

検出対象領域抽出には輝度成分画像を用いたが、透かしパターンの埋め込みは視覚特性を考慮してCb成分を用いた。キャプチャフレームが4:2:0 formatであるため、電子透かし検出は実質88×72画素内の検出対象領域から行なわれることになる。

誤り訂正/検出符号には(31+1,16)拡大BCH符号(2重ビット誤り訂正)を用い、各フレームに16ビットの情報を埋め込んだ。このうち1ビットをビット反転フラグに用いるため有効な透かし情報長は15ビットとなる。また、予備実験から透かしパターンの反転頻度を15反転/秒と定めた。

図5に実験風景と実験プログラムの実行中画面を示す。次節の評価のため、検出対象領域探索処理および電子透かしの有無判定を常時行ない、ボタン押下後に電子透かし検出を2秒間連続して行なうようにしている。

検出信頼度指標に対する閾値については3.4節の議論に従い、検出信頼度指標値 ρ が $\alpha = 4.8$ 以上の時($P_A < 10^{-6}$ に相当)に透かしが存在すると判定し、 $\beta = 8.5$ 以上のときかつBCH復号によって誤り訂正可能なとき($P_B < 2.8 \times 10^{-3}$ に相当)に電子透かし検出成功と判定する。

実験プログラムでは、ボタン押下時でなくてもリアルタイムに検出対象領域抽出結果および電子透かしの存在有無をフィードバックし(電子透かし有と判定した場合はレビュー画面上の検出対象領域を示す線の色を変化させる)、利用者に電子透かし検出可能性を提示したり、利用者側で電子透かしが検出しやすくなるように撮影角度などを調整することが可能である。

4.2 評価実験

図6に示す映像情報メディア学会の標準動画像3種(Horse Race, Whale Show, Baseball, シーケンス長:15秒, 29.97 fps, 640×512画素)を用いて、以下のような評

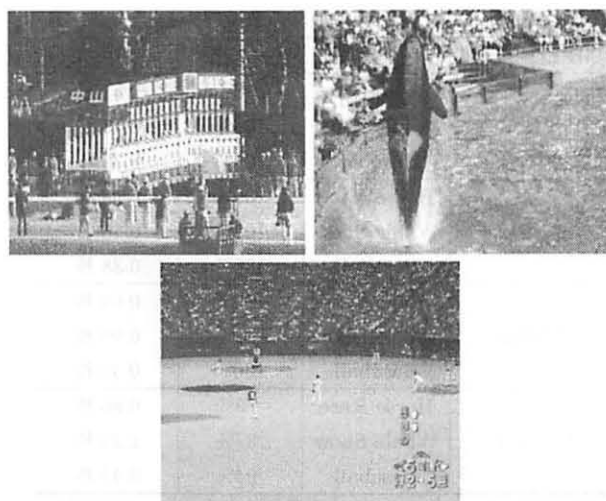


図6 評価用動画像: Horse Race(左上), Whale Show(右上), Baseball(下)

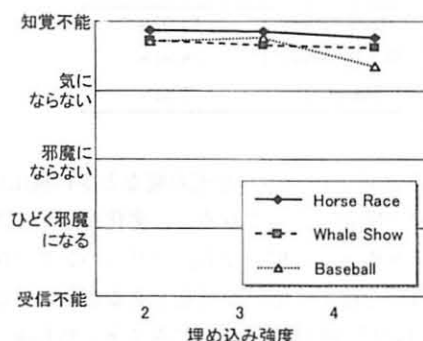


図7 主観評価結果

価実験を行なった。

4.2.1 画質評価

各評価用動画像に対して、3段階の埋め込み強度(256階調のCb成分に対する透かしパターンの振幅値)で電子透かしを埋め込んだものを18インチの液晶モニタに出力して、原画と比較して画質劣化について主観評価(被験者12名)を行なった結果を図7に示す(評価方法: 原画 → 透かし入り画像 → 原画の順に提示し、2番目の画像が透かし入りであることを事前に伝えた上で、原画と比較してどの程度劣化しているかを回答してもらった)。今回の評価環境においては、設定した透かし強度やコンテンツの種類によらず、ほぼ劣化が知覚できないレベルであることが分かる。これにより2.2節の要件(a)を満たすことが分かる。

4.2.2 耐性評価

埋め込み強度=3で透かしを埋め込み、さらに4段階のMPEG-2圧縮(非圧縮, 6, 3, 1.5Mbps, FFMPEG [8])を使用を行なって得られる動画像を、18インチ液晶モニタに出力し実機を手で保持して実験プログラムで検出試行した場合の検出成功率および成功時の平均所要時間を表3に示す(ボタン押下後2秒以内(約28回試行)に検出信頼度指

表 3 検出成功率と成功時の平均所要時間 (透かし強度=3)

MPEG-2 圧縮	画像	検出成功率	平均所要時間
非圧縮	Horse Race	100%	0.57 秒
	Whale Show	100%	0.43 秒
	Baseball	100%	0.39 秒
6 Mbps	Horse Race	100%	0.59 秒
	Whale Show	100%	0.46 秒
	Baseball	100%	0.38 秒
3 Mbps	Horse Race	100%	0.69 秒
	Whale Show	60%	0.99 秒
	Baseball	100%	0.39 秒
1.5 Mbps	Horse Race	73%	0.86 秒
	Whale Show	37%	1.25 秒
	Baseball	94%	0.43 秒

(各画像についてランダムな検出開始タイミングで 30 回試行)

表 4 原画像のフレーム間差分 (Cb 成分) の分散

評価用画像	差分の分散の平均
Horse Race	14.938
Whale Show	28.818
Baseball	3.315

標値 p が閾値 β 以上かつ誤り訂正可能ときに検出成功)。

まず非圧縮の場合についてみると、劣化が知覚できない程度の強度で埋め込んだ場合でも、フリーハンドで保持した携帯電話機から確実に検出が可能であることが分かる。

次に MPEG-2 圧縮耐性についてみると、Whale Show が特に圧縮に弱いという結果となっている。この理由として、Whale Show はパン・チルトなどのカメラワークの他、水面のゆらぎなど複雑な動きを多く含むため、差分画像のエネルギーが大きく、これが透かしに対する大きなノイズとして影響を与えて検出が困難となるためではないかと推測した。この確認のため、各原画像のフレーム間差分 (実際のキャプチャレートと合わせるため 2 フレーム遅延したフレームとの差分を取っている) の Cb 成分の分散の平均値を求めた結果を表 4 に示す。予想通り、Whale Show は差分の分散、すなわちエネルギーが大きいたことが分かり、他の画像と比較して検出しづらい画像であることが分かった。非圧縮や 6Mbps 程度の MPEG-2 圧縮の状態では、電子透かしのロバスト性のマージンによって検出が可能であるが、他の画像よりもマージンが少ないため符号化ノイズなどに対する耐性が低くなり、検出性能が悪くなっていると考えられる。

しかしながら、DVD や放送品質レベル (6Mbps に相当) では全ての画像で実用上十分な圧縮耐性を有していると言える。特に Baseball のような動きの少ない画像では、1.5Mbps まで圧縮してもかなり高い検出成功率を実現している。以上から 2.2 節の要件 (b) をほぼ満たすと言える。

さらに、本実験の検出成功基準は 2 秒以内であることが

ら、要件 (c) を満たしていることが分かる。また、3.4 節の議論の通り、検出した透かし情報の信頼性を定量的に保証して検出の成否を判定しているため、要件 (e) も満たすと言える。なお、本実験中の試行回数 360 回のうち、Type A. および Type B. の誤りは一度も生じなかった。

4.2.3 検出処理速度

評価用の携帯電話端末上で、単純にビデオキャプチャとプレビュー表示のみを行なった場合のキャプチャレートは平均 14.92 fps であり、加えて提案方式による電子透かし検出処理およびフィードバック出力を全て行なった場合では平均 14.05 fps であった。若干の処理落ちが見られるがほぼリアルタイム処理を実現し、遅延の少ないフィードバック出力が可能であり、要件 (d) をほぼ満たすと言える。

5. ま と め

リアルタイム検出可能な動画向けモバイル電子透かし技術の位置付けを明らかにし要件定義を行ない、要件充足のために高速な矩形領域認識アルゴリズムと動画画像のフレーム間相関を利用した高速電子透かしアルゴリズムを組み合わせた方式を提案した。さらに提案方式は検出信頼性の定量的保証が可能であることも示した。実機への実装および評価実験により、全ての要件をほぼ満たすことを検証し、提案方式の有効性を確認した。今後は、より詳細な評価実験と分析を行ない、さらなる改善により実用化を目指す。

文 献

- [1] 中村他, カメラ付き携帯電話機を用いたサービス仲介システムののための電子透かし検出方式, 情報科学技術フォーラム FIT2003, N-020, 2003.
- [2] 中村他, カメラ付き携帯電話機を用いたアナログ画像からの高速電子透かし検出方式, 俗学論 (D-II), Vol. J87-D-II, No.12, pp.2145-2155, 2004.
- [3] 平成 17 年度情報通信白書, 総務省.
- [4] ケータイカメラでインターネットに誘導する技術, インターネットマガジン 2005 年 4 月号, (株) インプレス, pp.114-117, 2005.
- [5] Ben Shneiderman, Designing the user interface 3rd edition, Addison-Wesley Longman Publishing Co. Inc., 1998.
- [6] 中村他, カメラ入力画像からのリアルタイム電子透かし検出方式, 情報科学技術フォーラム FIT2004, J-036, 2004.
- [7] 北原他, カメラ付き携帯電話機による撮影画像中の矩形領域抽出手法の高性能化, 情報科学技術フォーラム FIT2005, J-027, 2005.
- [8] FFMPEG, <http://ffmpeg.sourceforge.net/index.php>