電子透かしのアタックに対する耐性評価尺度の構築

5 G − 6

大網 亮磨　　宮本 義弘　　太田 睦
NEC C&C メディア研究所

1. はじめに

電子透かしに対する関心は，近年急速に高まりつつある．しかし，そのアタックについて体系的に検討している例[1]はまれであり，異なる方式間での耐性比較が困難である．本稿では，通常の画像処理等で生じ得るアタックを分類し，カテゴリ毎にアタックに対する耐性評価値を算出する方式の枠組みを提案する．

2. アタックの分類

電子透かしに対するアタックは，

(1) 画像処理等で生じる劣化

(2) 埋め込み情報の意図的な改竄

の２つに大別できる．(1)は容易に実行できるだけでなく，通常の画像処理で不可避的に生じるため，一般に強い耐性が要求される．(2)は，(1)よりも高レベルのアタックであり，電子透かし挿入，検出方式に強く依存する．本検討では，(1)について，信号処理的観点から分類を行った(表 1)．

表1 アタックの分類

カテゴリ	アタックの種類
非可逆圧縮	JPEG 符号化
雑音付加	一様雑音，ガウス雑音など
幾何変換	拡大縮小，回転，反転など
画素値変換	白黒化，色相/彩度/明度の変化，限定色変換，2値化など
画像処理	ぼかし，鮮鋭化，メディアンフィルタ処理，ディザ化，ハーフトーン化など
画像編集	切り取り

3. 耐性評価値算出手法

3.1 評価値算出の概要

前節で述べたカテゴリ毎に耐性評価値を算出する．まず，画像に電子透かしを挿入し，透かし入り画像を作成する．次に透かし入り画像にアタックを行った後，透かしの検出を行う．これを複数の画像に対して行い，検出率を求める．この検出率の算出を，アタックの強度を調節するパラメータを変化させて繰り返す．そして，得られた検出率データに後述する統計処理を行い，アタックに対する耐性評価値を算出する．最後に，アタックに対する耐性評価値をカテゴリ内で平均し，カテゴリ毎の耐性評価値を算出する．

検出率データからアタックに対する耐性評価値を求める方法には様々なものが考えられる．本検討では，

・加重平均に基づく評価方式 (3.2 節)

・閾値処理に基づく評価方式 (3.3 節)

の2方式により，アタック毎の耐性評価値を算出した．

3.2 加重平均に基づく評価方式

アタックの強度を表すパラメータを x，x に対する検出率を $f(x)$ とし，アタックに対する評価値 v を

$$v = \int_{-\infty}^{\infty} w(x) f(x) dx \qquad (1)$$

により算出する．ここに $w(x)$ は重み付け関数であり，アタックの生じる頻度，アタックによる画質劣化などを考慮して決定する．

実際に得られる $f(x)$ の値は，限られた強度 x に対する離散値のみである．そこで式(1)の $f(x)$ として，検出率の実測値から折れ線近似した関数を用いる．

3.3 閾値処理に基づく評価方式

3.2 節と同様に，まず，アタック強度パラメータと検出率の関係を求める．次に，検出率が閾値以上となる強度パラメータ x の区間を求め，この区間長を基準となる区間長で正規化した値を評価値とする．すなわち，評価値 v を

$$v = \frac{1}{L} \int_{-\infty}^{\infty} Th(f(x), \alpha) dx \qquad (2)$$

Study on establishing a robustness measure for digital watermarking algorithms
Ryoma Oami, Yoshihiro Miyamoto, and Mutsumi Ohta
C&C Media Research Labs., NEC Corporation
1-1, Miyazaki 4 Chome, Miyamae-ku, Kawasaki, Kanagwa, 216-8555, Japan

により求める．ここに L は基準となる長さで，$Th(t,\alpha)$ は t を閾値 α で2値化する関数である．式(2)の計算では，3.2節と同様，折れ線近似した $f(x)$ を用いる．

3.4. 本方式の特長

本評価値算出方式は，以下の特長を有する．

- 重みづけ関数などを適切に設定することで．アタック特性の異なる用途毎に評価値を算出できる．
- 検出率を実測する強度 x の値の違いによらず，方式間の比較が可能である．
- 段階的な評価値の精度向上が可能である．

4. 評価結果

本耐性評価値算出方式を，市販あるいは公開されている既存の電子透かし 3 方式(以後 A, B, C 方式と呼ぶ)に適用し，耐性評価を試みた．5 枚の画像を使用し，市販の画像編集ツールにより表 1のアタックを実行(3 から 10 通りの強度)し，カテゴリ毎に評価値を算出した．重み付け関数は，アタックによる画質劣化とアタックの頻度を考慮してヒューリスティックに決定した．例えば，雑音付加のアタックの場合，雑音電力に対して単調減少する関数とした．また，検出率 0.8 を透かしの検出可否判定の目安と考え，式(2)の閾値 α を 0.8 とした．

加重平均による評価方式の結果を図 1に示す．これより，B方式は，幾何変換，画像編集に対して評価値が極端に低いことがわかる．これは B 方式が幾何変換，画像編集のアタックすべてに対して耐性が無いためである．一方他の2方式は，全体的に良い評価結果となっている．A 方式はぼかしに，C 方式は回転にやや弱いという実験結果であったが，それ以外のアタックに対する耐性は強く，得られた評価値は妥当であると考える．

閾値処理による評価方式の結果を図 2に示す．図 1と比較すると，全体の傾向は一致するが，図 1よりも厳しい評価になっている．これは，検出率の高い範囲のみが評価値算出に寄与するためである．また，方式間の優劣がより強く表れている．特に，A 方式がぼかしや切り取りにやや弱い点，B 方式が画素値変換に強い点，C 方式が回転に弱い点が評価値に反映されている．

このように，加重平均による評価方式は，アタックに対する耐性を総合的に把握するのに適し，閾値処理による評価方式は，高い検出率を保証する範囲や方式の特徴を把握するのに適していると考える．

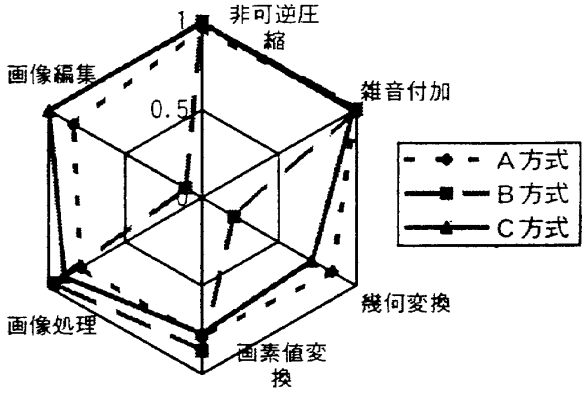

図1 加重平均による耐性評価値

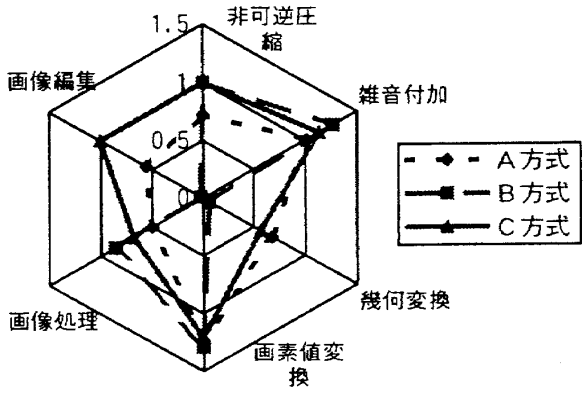

図2 閾値処理による耐性評価値

5. まとめ

本稿では，電子透かしのアタックに対する耐性を評価する方式の枠組みを提案し，既存の電子透かし方式に適用して，評価値算出を試みた．今後，さらに多くの方式に適用し，評価に必要なデータ量（画像枚数，アタック強度のサンプル間隔など）の検討，適切な重み付け関数の検討，およびアタック分類の妥当性の検証を行う予定である．

謝辞

本研究は，郵政省のプロジェクト「電子透かし技術に関する研究開発」の一環として，通信・放送機構（TAO）から委託されて実施した．

参考文献

[1] I. Cox, J. Linnartz: "Some General Methods for Tampering with Watermarks," IEEE J. of Selected Areas in Comm., 16, 4 (1998).