

# 近傍ピクセルの性質を用いたデータハイディング — 抽出誤りの統計的考察 —

1 V-4

清水 周一      小林 誠士      上條 浩一  
日本アイ・ビー・エム株式会社 東京基礎研究所

## 1 はじめに

データハイディングは、動画やオーディオなどのデジタル・コンテンツに対して、知覚的に感知されないような方法で付加（ビット）情報を埋め込んで隠すための技術である。ビット情報は、コンテンツのデータ部を微小に変更することにより埋め込まれてコンテンツと一体化されるので、一般に元データとの分離は困難である。この一体性を利用して著作権情報管理や不正使用の抑止などの用途で利用され始めている。

以下では、まずビット情報の抽出に起こり得る誤りを分類し、統計的特性を利用したデータハイディングなどを例に、そのビット情報の信頼性を議論する。

## 2 抽出の誤り

コンテンツ・データ  $I$  に対する付加ビット情報  $B$  の埋め込み操作  $E_M$  を以下のように定義する。

$$E_M(I, B) \rightarrow I', \quad (1)$$

ここで、 $I'$  は埋め込み後のコンテンツ・データである。このコンテンツ  $I'$  からビット情報を抽出するには、以下に示すように、元データ  $I$  は不要とする [1]。

$$D_M(I') \rightarrow B, \quad (2)$$

ここで、 $D_M$  は抽出操作を表す。一般に、埋め込み操作直後に抽出の誤りを起こすことはない。

さて、埋め込み操作の施されていないデータ  $I$  に対して、抽出処理が行われる場合には、以下に示すように、十分な信頼性をもって埋め込み情報がなかったとの判定を下すことが望まれる。

$$D_M(I) \rightarrow \emptyset, \quad (3)$$

この判定の失敗、すなわち、埋め込み処理が行われていないにも関わらず、ビット情報を抽出してしまう、以下のような誤りを“false positive”と呼ぶ。

$$D_M(I) \rightarrow B_{garbage}, \quad (4)$$

データハイディングでは統計的確率モデルに基づき、この“false positive”誤り率に関して一定の基準を満たすように、埋め込みの有無を判定している [2, 3]。

埋め込み後のデータ  $I'$  が編集や圧縮、フィルタなどの処理  $F$  を施された場合、データとの一体性のため、埋め込んだビット情報も同時に影響を受ける。しかし、このような場合にも、以下に示すように、埋め込み情報が正しく抽出できることが望ましい。

$$D_M(F(I')) \rightarrow B, \quad (5)$$

ここでの誤り率が低いことを埋め込み情報の“耐性”と呼び、逆に、以下に示すように、埋め込んだ情報が正しく抽出できなくなる誤りを“false negative”と呼ぶ。

$$D_M(F(I')) \rightarrow \emptyset, \quad (6)$$

$$D_M(F(I')) \rightarrow B_{wrong}. \quad (7)$$

“false negative”誤りは、(1) 圧縮やフィルタなどの影響により埋め込みビット情報が損なわれて起こる場合、(2) 埋め込み情報は損なわれないが、切り取りや回転などの編集操作などにより抽出処理とのミスマッチのため起こる場合、の2通りに分類できるが、以下では、前者について考察する。

次の3章では、ビットに冗長性を持たせた埋め込み方法における“false positive”および“false negative”について考察する。4章では、特徴量の線形和を用いた統計的手法 [2, 3] の“false negative”について、フィルタ処理を絞ってその影響を考察する。

## 3 ビットの冗長性による信頼度

$n$  個のビットが偶然に互いに一致する確率は  $(1/2)^{n-1}$  であり、例えば  $n = 19$  の場合、その確率は  $3.8 \times 10^{-6}$  と非常に小さい。したがって、1 ビットの情報を表現するために、それを  $m$  回冗長に繰り返して埋め込むことにすれば、コンテンツ・データから得られた特徴量をビット情報に解釈するとき、その偏りから埋め込み情報の有無を判断することができる。 $n$  ビットのうち少なくとも  $m$  ビットが互いに一致しているとき埋め込みありと判定するルールでは、埋め込み情報がないのにも関わらず誤って埋め込みありと誤った判断をする確率 (“false

positive”)は、以下のように与えられる。

$$P_P = (1/2)^{n-1} \sum_{i=m}^n {}_n C_i \quad (8)$$

したがって、1ビットあたり  $n$  回繰り返しで合計  $k$  ビット情報を埋め込む方法では、 $k$  ビット全体での “false positive” 誤り率  $\epsilon_P$  は以下の通りである。

$$\epsilon_P = P_P^k, \quad (9)$$

例えば  $k = 32$  ビット情報の埋め込みにおいて、 $n = 19$  の繰り返しの中で  $m = 11$  以上が一致しているものを埋め込みありと判定するとき、“false positive” 誤り率  $\epsilon_P \approx 10^{-6}$  が達成できる。

一方、1回の繰り返しあたりの誤り率を  $q$  とおくと、 $n$  回の繰り返しで表現するビット情報は、少なくとも  $(n - m + 1)$  回の誤りが生じたとき “false negative” となるので、その誤り率は、

$$P_N = \sum_{i=n-m+1}^n {}_n C_i q^i (1-q)^{n-i}, \quad (10)$$

となる。ここで、誤り率  $q$  は圧縮やフィルタなどの処理によって異なる値を取り予測できないので、実験により推定する。 $k$  ビット情報全体では、少なくとも1ビットに誤りが生じる確率は以下の通りである。

$$\epsilon_N = 1 - (1 - P_N)^k \approx k P_N, \quad (11)$$

ただし、近似は、 $P_N$  が十分に小さいとする。例えば、 $q = 0.05$  のとき、先と同じ条件では “false negative” 誤り率  $\epsilon_N \approx 3.6 \times 10^{-6}$  が達成できる。

## 4 統計的手法の信頼度

特徴量  $X_i$  の線形和を用いてビット情報を表現する統計的手法 [2, 3] では、“false positive” 誤りは確率モデルにより予測することができる。

$$\bar{X} = \sum_{i=0}^{n-1} X_i, \quad (12)$$

ここでは、 $n$  が十分に大きければ、 $X_i$  の分布に関わらず線形和  $\bar{X}$  は正規分布に近似できること (中心極限定理) を利用している。このとき、特徴量  $X_i$  の分散を  $\sigma^2$  とおけば、線形和  $\bar{X}$  の分散は  $n\sigma^2$  となる。この確率モデルに基づき、ビット情報の埋め込み有無の判定について “false positive” を以下に示すように予測し、目的の誤り率を達成するための有無判定の閾値を決定している。

$$P_P(|\bar{X}| > T_P) = 2 \int_{T_P/(\sqrt{n}\sigma)}^{\infty} \phi(t) dt \quad (13)$$

ここで、 $\phi(t)$  は平均0分散  $1^2$  の正規分布を表す分布関数であり、 $T_P$  は埋め込みの強さで、抽出時の閾値となる。

この確率モデルは、フィルタ処理などの影響により特徴量  $X_i$  が変化するとき起こる “false negative” 誤りについても予測に利用できる。本来、観測値  $\bar{X}$  は分散を持たないが、例えば、画像データを構成する色の数の減らす (e.g., GIF) とし、輝度に基づく特徴量はランダムに変化し、その線形和  $\bar{X}$  も同様にランダムに揺らぐ。各特徴量の揺らぎの分散を  $\sigma'^2$  とおいたとき、観測値  $\bar{X}$  の分散は  $n\sigma'^2$  で、 $n$  が十分に大きければ同様に正規分布に近似できる。この揺らぎが起きて、観測された線形和  $\bar{X}$  が閾値を下回らないようにするには、埋め込みの強さを大きくすれば良い。その増分  $T_N$  と、 $\bar{X}$  が閾値を下回る確率、すなわち “false negative” 誤り率との関係は、以下に示すように予測できる。

$$P_N = \int_{T_N/(\sqrt{n}\sigma')}^{\infty} \phi(t) dt \quad (14)$$

したがって、埋め込みの強さを  $T_P + T_N$  とすることにより、“false positive” 誤り率  $P_P$  および “false negative” 誤り率  $P_N$  が同時に達成できる。ただし、本来の分散  $\sigma$  と処理の影響による分散  $\sigma'$  とが互いに独立であるとき、線形和  $\bar{X}$  の分散は  $\sigma^2$  から  $\sigma^2 + \sigma'^2$  へと増えるので、 $T_P$  も  $\sqrt{1 + (\sigma'/\sigma)^2}$  倍に増やす必要がある。しかし、コンテンツ・データの質が保たれている限り、一般に  $\sigma' \ll \sigma$  であることが期待できるので、埋め込みの強さはあまり変わらない。

なお、 $\sigma'$  の大きさは各処理に依存するので、実験により推定する必要がある。 $\sigma'$  を大きく取れば、“耐性” の高い埋め込み処理となる。

## 5 おわりに

埋め込み処理の施されていないデータから誤ってビット情報を抽出する誤り (“false positive”)、および、埋め込まれているビット情報を正しく抽出できない誤り (“false negative”) について議論し、その確率的予測に基づく閾値や埋め込みの強さの決定方法を示した。

## 謝辞

本研究は、情報処理振興事業協会で実施された創造的ソフトウェア育成事業の一環として行っており、この機会を与えていただいたことに感謝の意を表する。

## 参考文献

- [1] 小出: “DataHiding 技術とその応用”, 精密工学会画像応用技術専門委員会, 1/23, 1998
- [2] 上條他: “近傍ピクセルの性質を用いたデータハイディング — 付加情報埋め込みと抽出 —”, In Proc. of IPSJ 56th annual conference, 1998
- [3] 小林他: “近傍ピクセルの性質を用いたデータハイディング — 近傍ピクセルの統計的性質 —”, In Proc. of IPSJ 56th annual conference, 1998