

誤りパターン埋込み型ステガノグラフィに関する一考察

索手 一平 †

松江工業高等専門学校 †

コードの列とした。

1 はじめに

本研究はグレースケール画像にテキスト情報を埋め込むようなステガノグラフィ技術を対象とする。埋め込み方式の一つである LSB 法は画像の LSB 平面をテキスト情報のバイナリ表現とそのまま置き換えることによって埋め込みを行う。これに対して、誤りパターン埋め込み法では、テキスト情報を冗長かつハミング重みが小さなビット列である誤りパターンに変換し、誤りパターンと画像のビット列との排他的論理和でビット列を置き換える。

埋め込みに使用したビット数とそれによって変化したビット数との比を誤り率、埋め込んだ情報量との比を埋め込み率と言い、誤りパターン埋め込み法は誤り率において LSB 法に比べ非常に優れている。その一方で情報を冗長なビット列に変換することから埋め込み率において劣っている。

本研究では誤りパターン埋め込み法における画質劣化と埋め込み率のトレードオフ関係を実験的に明らかにした。また、その実験結果をもとに画質劣化の起こりやすい画像の特性について調査するための実験を行った。

2 誤りパターンへの変換方法

誤りパターンへの変換方法の一つに埋め込みデータと誤りパターンを対応付けたテーブルを用意する方法が知られている。しかし、この方法の問題点として埋め込みデータの大きさに比例してテーブルが膨大となりメモリ制約の大きい環境での実装が困難になるという点があげられる [1]。そこで本研究では Shalkwijk の数え上げ符号 [2] を用いた埋め込みデータから誤りパターンを動的に生成する手法を提案する。

3 埋め込み率と画質劣化のトレードオフ関係

実験手順の概要を図 1 に示す。以下に示す手順で実験を行い、実験には 256×256px の 8bit グレースケール Bitmap である自然画像 100 枚を使用した。また、埋め込みに使用するメッセージはほぼ当確率で発生する 8bit

- (1) メッセージの各コードを Shalkwijk の数え上げ符号を用いて動的に誤りパターンへと変換し、画像のビット列との排他的論理和を埋め込む。
- (2) 埋め込み前後の画像を比較し、誤り率、PSNR、SSIM を算出する。ここで誤り率 r は埋め込み前後で変化したビットの数を d 、埋め込みに使用した画像のビット数を n として式 (1) で算出する。

$$r = \frac{d}{n} \times 100 \quad (1)$$

- (3) 誤りパターン長を 8~256bit、埋め込み範囲を LSB のみから全ビットまでの範囲で変化させながら (1)、(2) を繰り返す。
- (4) 画像を入れ替えて (3) を繰り返す。

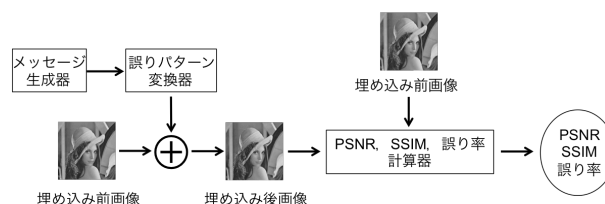


図 1 実験概要 1

画像ごとの SSIM の変化について図 2 に示す。図 2 から画像によって画質劣化の度合いに差があることがわかる。つまり画像によっては同様の画質劣化、より多くの文字を埋め込むことができるようになる。そこで、実験に使用した画像に対し主観的に観察を行ったところ、同じような色で構成される領域が大きい画像ほど画質劣化が大きくなるという特徴があるということがわかった。その例を図 3、図 4 に示す。

図 3 は善意的にさまざまな色のピクセルが混ざり合って構成された画像であり、一方で図 4 は白やグレーが連続した領域が大きい画像である。私はこの結果が主観によるカバーデータ選択の指標できるのではないかと考え、この結果を客観的に評価するための実験を行った。

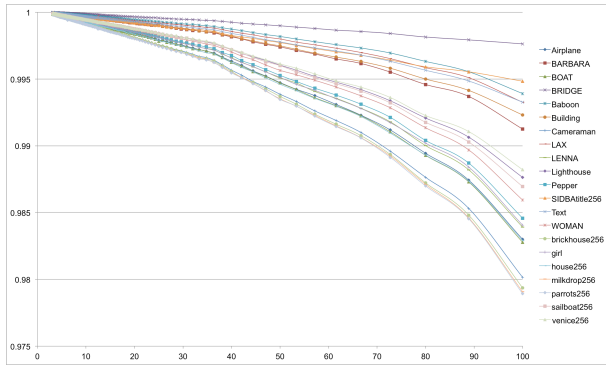


図2 画像ごとの SSIM の遷移



図3 劣化の少ない画像 (左:埋め込み前 右:埋め込み後)

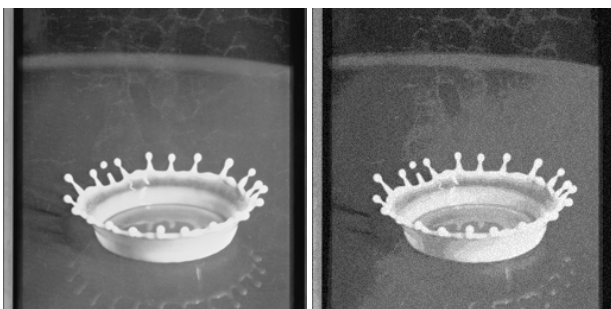


図4 劣化の大きい画像 (左:埋め込み前 右:埋め込み後)

4 画質劣化の起きやすい画像の特徴

実験手順の概要を図5に示す。以下に示す手順で実験を行い、実験には第3節で使用した画像と同じものを使用した。

- (1) グレースケールの画像に対し、0～7の画素値を持つピクセルを抽出し、抽出したピクセルとそれ以外のピクセルとで分け、二値化を行う。
- (2) 抽出する画素を8ずつずらし、(1)を繰り返す。
- (3) 二値化した画像に対し4近傍でラベリング処理を行い、各ラベルのついた領域の面積を計算する。
- (4) 全ての2値画像の中で最大の領域面積を求める。

(5) (1)～(4)をすべての画像に対し繰り返す。

(6) (5)で算出した最大面積と第3節で算出した SSIM とのピアソンの相関係数を算出する。

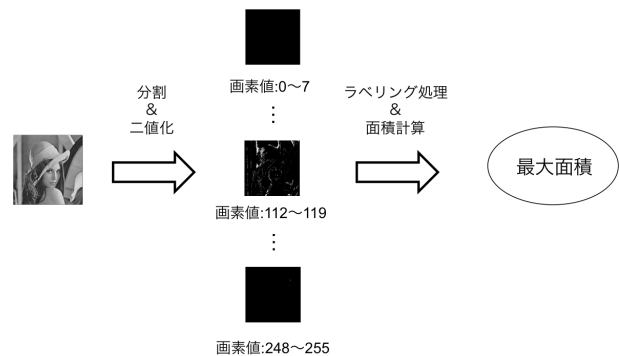


図5 実験概要2

私は画素値を小さな範囲で区切りその面積を計算することで同じような色で構成された領域の大きさを算出できると考えた。シンプルな方法ではあるものの各パターンにおける SSIM と 0.6 前後の相関を得ることができた。これは細かいコントラストの中で埋め込みによって生じるノイズが目立たなくなっているのだと考えられる。

5 考察

以上の研究で同じ条件での埋め込みにおいて画像ごとに画質劣化の度合いが異なることがわかり、そしてシンプルな方法ではあるが同じような色で構成された領域の大きさが画質劣化の度合いに対して影響しているということが明らかになった。

これらは人間の視覚によるカバーデータの選択において、この特徴が視覚的認識の回避のための効果的な指標となり得ることを示している。ただ一方で単純な SSIM は FBS や WBS などの代表的なステガノグラフィの解析アルゴリズムに対し有効であるとは言えない。今後は上記の特徴に関してより正確かつ信頼できる計算方法を用いて実験を行うと同時に、実際に解析アルゴリズムに対してこれらの特徴がどのように作用するのかを研究していきたいと考えている。

6 まとめ

参考文献

- [1] 合田翔, 渡辺峻, 松本和幸, 吉田稔, 北研二. コスト付き符号化を用いたステガノグラフィ. 信学技法 IT, Vol. 113, No. 153, pp. 5–9, 7 2013.
- [2] J.PIETER M.SHALKWIJK. An algorithm for source coding. *IEEE TRANSACTIONS ON INFORMATION THEORY*, Vol. IT-18, No. 3, pp. 395–399, May 1972.