

JPEG のイデオロギーを利用したステガノグラフィー

Steganography Utilizing the Ideology of JPEG

概要

本稿は、JPEG 画像への情報埋め込み、いわゆるステガノグラフィーの新たな手法を提案する。しかし、ここに提案する手法は情報隠蔽と保護という、一般にステガノグラフィーに求められる要件の完全性を指すものではない。有用なステガノグラフィーというよりは、一つのアイデアの提出である。JPEG のイデオロギーを明らかにし、そのイデオロギーを流用することによって生まれるアイデアだ。ここで言う JPEG のイデオロギーとは、次のような表現に要約される。「見えないものは存在しない」。JPEG というシステムがその圧縮効率を高めるために存在しないものとした領域に、まったく別の情報を埋め込むというのがこのアイデアの骨子である。手法の提案と合わせ、筆者は概念実証として実装も提示する。この実装を用いて画像に情報を埋め込んだテスト結果を Figure 1 に示した。処理された画像とオリジナル画像を見比べても、視覚的には差異が認められないことがわかるだろう。

(ただし、テスト画像は諸事情により掲示することができない。諸事情の詳細については、本稿の内容と関係がないのでここでは説明しない。)

Abstract

We propose a novel method of hiding information, steganography, into JPEG images. However, the method proposed here does not aim for the completeness of the general requirements in steganography such as perfect concealment and protection. It is a submission of an idea rather than useful steganography. The idea is produced by revealing and appropriating the ideology of JPEG. The ideology of JPEG discussed here can be summarized as follows: "invisible things do not exist." The essence of this idea is to embed completely foreign information in areas that the JPEG system defines as non-existent due to its compression efficiency. Along with the proposal of the method in this paper, the author will also present an implementation as a proof of concept (PoC). The test result of hiding information in an image using the PoC implementation is shown in Figure 1. It is obvious that there is no visual difference when comparing the processed image and the original image.

(However, the test images cannot be posted for various reasons. The details of these various reasons are not relevant to the content of this paper, so they will not be explained here.)

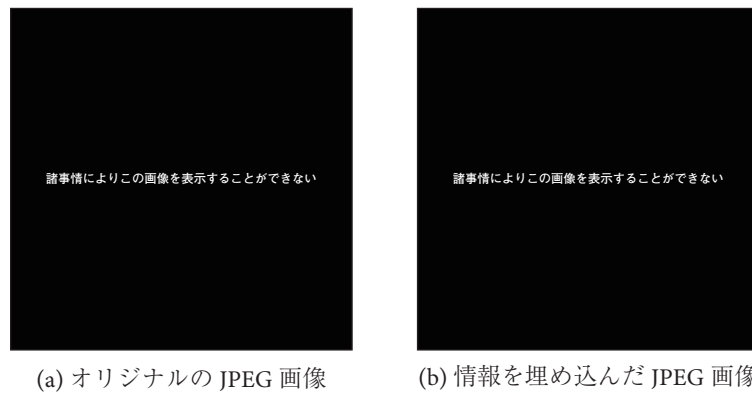


Figure 1. 256×256 ピクセル、クロマ・サブサンプリング 4:2:0、品質 50% の標準量子化テーブルを用いて JPEG エンコードされたオリジナル画像 (a) と、オリジナル画像の量子化された DCT 係数の一部を書き換えることで情報を埋め込み、さらにこの埋め込み部分を処理する量子化テーブルの値を 0 に書き換えるこ (ただし、ここで使用した画像は諸事情により掲示することができない。諸事情の詳細については、本稿の内容と関係がないのでここでは説明しない。)

1 イン트로ダクション

情報の隠匿とその通信は、闘争の歴史である。20 世紀の世界大戦における各国の軍事的な暗号解読への挑戦がなければコンピュータの開発と発展がなかったことは、よく知られた事実だ。情報を秘匿して通信する方法には、大きく分けて2つある。ひとつはクリプトグラフィーであり、もうひとつはステガノグラフィーだ。前者は送信者と受信者以外の誰かに見られることを前提とする隠蔽である。たとえ見つかっても読解のルールを知らない者は読むことができず、ルールを解明することもできない。すなわち暗号だ。一方、後者は秘匿情報を見えなくすることを目的とする。ステガノグラフィーはクリプトグラフィーと並んで、またはクリプトグラフィーとの組み合わせで使用されてきた。ステガノグラフィーはヘロドトスの歴史にすでに見られ、世界大戦の時代には戦局を左右する情報を多く運んだ。机のワックスを剥がしてメッセージを刻んでワックスを塗り直す、というものから、頭皮に入れ墨でメッセージを書き込み髪が伸びたその人をメッセンジャーとして送るという方法、温度や薬品で読めるようになる見えないインク、文章内の単語の n 文字目だけを読むことで現れるメッセージ、ピリオドなどの小さな点の中に印刷されたマイクロ文字、といったものが用いられてきた。インターネット以降のデジタル環境でも、データ内に情報を隠す手法が多く提案されている。[1]

本稿が主題とするのはデジタル画像に情報を埋め込む技術だ。デジタルにおけるステガノグラフィーは多岐にわたるが、静止画像に対する手法でよく知られるのは LSB (Least Significant Bit) 法である。LSB 法は複数桁の二進数の下一桁にデータを埋め込む。このとき、ビット深度 8 のピクセル情報はビット深度 7 に減衰し、表現できる精細度は半分になる。LSB 法は、ビット深度が 8 の画像と 7 の画像の差異を人間の目が見分けられないことを期待している。この期待は、人間の知覚閾に合わせて符号の精度を調整するという考え方、いわゆる知覚符号化の可能性に触れていると言える。

静止画像フォーマットで知覚符号化を大きく採用しているフォーマットは JPEG だ。JPEG は人間の目に見えづらい情報を削減しながらエンコードし、そのことによって高い圧縮率を可能にしている。LSB 法は知覚符号化と近い考え方をしている。しかし JPEG のそれはさらに巧妙だ。この巧妙さをステガノグラフィーに利用できないだろうか。

JPEG は、一般的には元データを 10 分の 1 程度に圧縮すると言われる。120 分の 1 に圧縮した JPEG でも視認性は確保されるという報告もある [3]。画像のデジタル保存の課題は、単純に考えれば、精細度とデータサイズとのトレードオフであるが、JPEG が実現する圧縮はこの手の論理を超越していると言える。つまり精細度が高くかつデータサイズが小さいという論理矛盾を JPEG は実現している。同じようなトレードオフはステガノグラフィーにもある。秘匿データの量と発見される危険の関係だ。多くの情報を埋め込めば見つかりやすくなる。画像の場合、秘匿データが大きければ元画像が破損して不自然になる。LSB をはじめとするステガノグラフィー手法は、発見の危険を軽減するために大きなカバーデータに少量の秘匿データを埋め込むことを好む。しかし、デジタル上で論理矛盾を実現できると JPEG が証明したことは、ステガノグラフィーにおいても同種の論理矛盾が可能だと示唆しているようにも思われる。

論理矛盾が現実のものとなっているということは、単に既存の論理が別の論理に入れ替わっているという理解もできる。JPEG の場合、別の理論とは「見えないものは存在しない」というものだ。人間の視覚の限界を超える情報を極力削除しているがゆえに、120 分の 1 のデータ量まで情報を減らしても画像として認識可能なのだ。JPEG はこの理論の具体化をエンコード・アルゴリズムの中で徹底的に行う。この徹底は、アルゴリズムを思想に似たものに変える。もはやそれはイデオロギーであり、政治である。このイデオロギーの裏をかき、存在しないとされた領域にまったく別のデータを挿入することは可能か。本稿の試みはこの問いから発している。本稿が提案するステガノグラフィーは、この問いへの回答である。

JPEG に対するステガノグラフィーには事例がいくつかある。最初期の実践としてよく知られるのは、デレク・アッパムによる実装、Jsteg である [4]。Jsteg は JPEG の圧縮の過程で周波数成分の量を表現する「量子化された

DCT 係数」に対して、LSB 法を用いて情報を埋め込む。ピクセルデータのビット深度を下げるのではなく、処理過程の係数のビット深度を下げることで、ピクセルデータへの直接の LSB 法よりは秘匿性が高くなる。しかし、極端に小さな係数への LSB 法は影響が大きくなりすぎるため、Jsteg では値が 0 と 1 と -1 の係数は処理対象としない。だが、0 である係数が多いときに JPEG は最も圧縮効率を発揮する。言い換えれば、係数が 0 であるということは、JPEG が人間の目に見えないと判断し決定したということだ。本稿が目にするのは、この目に見えない情報だ。見えない情報に秘匿情報を埋め込むにはどうすればよいだろうか。簡潔に言えば、デコードの際に量子化された DCT 係数を逆量子化するとき用いる量子化テーブルの値を 0 にすることでそれが可能になる。

量子化テーブルの値を部分的に 0 にする、この手法で出来上がる JPEG 画像に対して、デコードも人間の視覚もステガノグラフィーに気づくことはない。ただし、量子化テーブルは通常 0 を含んではいけない。極めて異常な量子化テーブルである。であるから、一般的にはステガノグラフィーとして脆弱だろう。しかし、提案する手法が行う隠匿は、監視者や傍受者に対してなされているのではない。そうではなく、JPEG のイデオロギーそのものの裏をかこうとしているのである。

本稿は、提案する手法のことを WDITDNE と名付ける。これは「We Decided Invisible Things Do Not Exist（見えないものは存在しないと私たちは決めた）」というフレーズの略記である。

2 JPEG

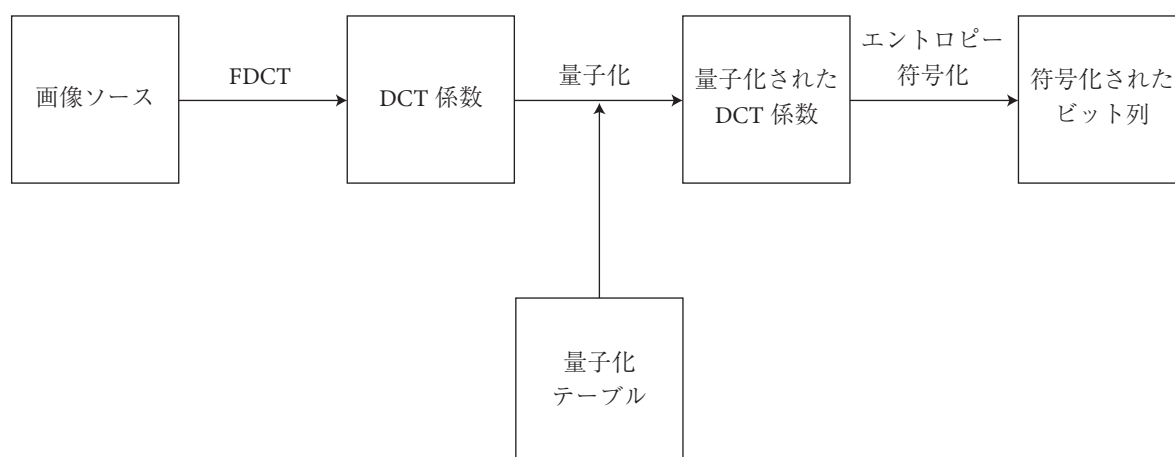


Figure 2. JPEG エンコード・プロセス

2.1 JPEG のエンコード・プロセス

まず、JPEG のエンコード・プロセスを要約しておく。(Figure 2)

JPEG は 8×8 の 64 ピクセルを基本単位としてエンコードを行う。この基本単位を MCU (Minimum Coded Unit) と呼ぶ。色空間が 3 次元で表現される場合、 8×8 のピクセル情報は 3 つの MCU に分割される。この MCU の重なりが画像平面に重複なく並んでいるとみなし、左上から文章を読むのと同じ順番で右下に向かって処理していく。

それぞれの MCU に対してまず行われるのは、DCT と呼ばれる変換処理 (FDCT) である。DCT は離散コサイン変換の略であり、フーリエ変換に類似の計算を用いて 8×8 の 2 次元平面に含まれる周波数成分の割合を算出する。8 種のコサイン波を縦横に組み合わせた 2 次元周波数成分 64 種類が定義されており、それらが対象の MCU にどの程度含まれているかを計算で導く。結果、DCT 係数と呼ばれる 64 の数値が得られる。技術文書では慣例的に DCT 係数を 8×8 のグリッドで示す。低周波成分が左上に示され、右と下に行くに従って周波数が高くなる配置だ。最も左上の DCT 係数は DC 成分と呼ばれ、特別な値とされる。この値は 8×8 ピクセル全体

| | | | | | | | | | | | | | | | |
|----|----|----|----|-----|-----|-----|-----|----|----|----|----|----|----|----|----|
| 16 | 11 | 10 | 16 | 24 | 40 | 51 | 61 | 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 |
| 12 | 12 | 14 | 19 | 26 | 58 | 60 | 55 | 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 |
| 14 | 13 | 16 | 24 | 40 | 57 | 69 | 56 | 24 | 26 | 56 | 99 | 99 | 99 | 99 | 99 |
| 14 | 17 | 22 | 29 | 51 | 87 | 80 | 62 | 47 | 66 | 99 | 99 | 99 | 99 | 99 | 99 |
| 18 | 22 | 37 | 56 | 68 | 109 | 103 | 77 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 24 | 35 | 55 | 64 | 81 | 104 | 113 | 92 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 49 | 64 | 78 | 87 | 103 | 121 | 120 | 101 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |
| 72 | 92 | 95 | 98 | 112 | 100 | 103 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 99 |

(a)
(b)

Figure 3. 標準量子化テーブル．輝度用のもの (a) と色差用のもの (b)

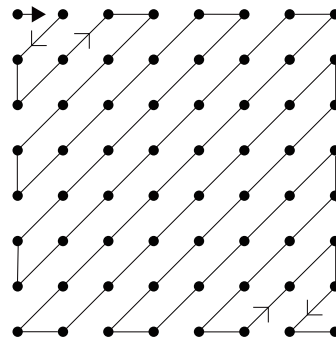


Figure 4. ジグザグ・シーケンス

の基盤を示しており、これが壊れることは画像の破綻につながる。一方、DC 成分以外の 63 値は AC 成分と呼ばれ、右下にいくに従って重要度を減じる。高い周波数の成分は人間の目に認識されづらいからだ。重要度が低い数値は無視できるはずだと JPEG は考える。そして事実無視する。

次に行うのは量子化と呼ばれる処理である。得られた DCT 係数のそれぞれの値を量子化テーブルの値で除算し、少数以下を切り捨てることで、量子化された DCT 係数へと変換する。それぞれの DCT 係数に対応したテーブル値がある。つまり、量子化テーブルもまた 64 の数値である。このとき、量子化テーブルの値が大きければ、DCT 係数は小さな値へと除算され、逆にテーブル値が小さければ DCT 係数は大きい数値へと除算される。量子化の処理は、除算し、さらに切り捨てを行うことだ。除算して小数値となった値は整数として保存される。Figure 3 は、JPEG 標準が提示している量子化テーブルの例である。左上のテーブル値は比較的小さく、右下は大きい。周波数が高い DCT 係数に対しては、大きな数で除算することになる。高い周波数成分はそもそも量が小さい上に、それを大きな数で除算すれば、容易に 1 以下の値になってしまい、量子化された結果 0 になる。事実、エンコード中の JPEG を覗き込み、量子化された DCT 係数を観察すれば、いかに 0 が連続しているかが分かるだろう。0 となった値は完全に無となり復元されることはない。JPEG の量子化の戦略は、いかに 0 を連続で出現させるかである。64 個の DCT 係数は、こうして 0 を大量に含む 64 個の量子化された DCT 係数に変化する。使用した量子化テーブルはファイルヘッダに記入され、デコードで用いられることになる。

この後行われるのは、ファイルに保存するためのビット列への変換である。エントロピー符号化と呼ばれる、ハフマン符号化とランレングス符号化を組み合わせた JPEG 独自の符号化だ。ハフマン符号化とは、頻出する記号をより小さなビット列で表現することによりデータ量を減らす符号化手法である。ランレングス符号化とは、同じ記号が大量に連続する場合、そのまま書き込むよりもある記号が何個連続したという表現にすることで、データ量を減らす古典的な符号化手法である。ランレングス符号化は記号が連続する箇所が少なければデータ量が逆に大きくなる可能性があるため、汎用的な符号化手法とはいえない。だが、JPEG は 0 が必ず連続する

ことを知っている。このゼロ・ランレングスが JPEG の驚異的な圧縮率を支えている。64 の量子化された DCT 係数は、ジグザグ・シーケンスと呼ばれる特殊な順序で処理されることになっている (Figure 4)。この順序で処理すると、0 が連続しやすくなる。左上から斜めに走査しながら右下へ。低周波から高周波へ至る順序であり、量子化によって高周波が消え去っているため、後半に至るほど 0 の出現率が高くなる。0 の連続に関する符号化は、もう 1 段階工夫されている。連続をあらわす最大長である 16 個、0 が連続する場合には ZRL という特殊な符号に変換し、以降 0 しか出現しない場合には EOB という符号に変換する。ZRL や EOB が頻出する場合には、それらを小さなビット長のハフマン符号で表現することでデータ量を削減できる。

エントロピー符号化という呼び名は、クロード・E・シャノンの「情報源符号化定理」によって示された復号可能な符号化による圧縮の限界 [6] が含意されている。量子化された DCT 係数は、それ以上のデータロスを生むことなく、復号可能な形でビット列に変換される。JPEG はここで、復号可能な圧縮を限界まで行うことを明言したいのだ。シャノンが示したこの限界は、可逆圧縮の限界であり、JPEG が行ってきた非可逆圧縮の限界ではない。JPEG がここで単に符号化と言わずエントロピー符号化と名付けたことに、なんらかの政治性を垣間見ることが可能かもしれない。あるいは、傲慢さを。エントロピー符号化という命名は、これ以上削減しなくても十分なほどすでにデータは消え去ったのだ、と宣言しているようでもある。

2.2 イデオロギー

見てきたように、JPEG はピクセルを周波数に変換することで値にヒエラルキーを与える。ピクセル情報として等価の 64 値は、DCT 係数という異なる価値を持つ 64 値になる。その価値は量子化テーブルによって決定される。一般に低い周波数の係数は重要であり量子化しても値が大きく残っている。一方、高い周波数の係数は 0 になっても構わないと考えられている。人間の目が認識しづらい要素は無にしてもかまわないという考え方、極端な知覚符号化である。JPEG とは、これを合目的的に徹底するアルゴリズムのことだ。エンコードの結果、JPEG によって人間には見えないであろうと判断された要素は抹消される。ゆえに JPEG のイデオロギーを次のように言い表すことができる。「見えないものは存在しない」そういった意思であり、規定であり、決定である。

加えて、JPEG にはクロマ・サブサンプリングの慣例がある。デジタル画像とは一般に RGB で表現される。汎用的な画像フォーマットのほとんどは内部でも RGB をそのまま扱うし、JPEG も仕様の RGB でピクセルデータをエンコード可能である。だが、現在私たちが日常的に目にしている JPEG 画像のほぼすべては YCbCr という色空間を用いている。RGB の色空間の 3 値を YCbCr という別の色空間の 3 値に変換した上で処理している。なぜ 3 値を別の 3 値に変換するのかといえば、YCbCr では色空間の特性上、データの間引きが可能になるからだ。YCbCr は、輝度信号 Y と 2 つの色差信号 Cb と Cr で構成される。輝度信号は人間が画像を認識するときに最も重要な要素であり、これだけで表現された画像はモノクロ画像のように輪郭や陰影を示すものになる。一方、色差信号は色を示す情報だが、これを減衰させても人間の目は違和感を感じないことが知られている。YCbCr のように輝度と色差を組み合わせた色空間は、テレビ伝送の時代から使われる古いもので、色差を間引いて転送量を小さくする目的で用いられてきた。間引きはクロマ・サブサンプリングと呼ばれる。クロマ・サブサンプリングにはいくつかの段階があるが、JPEG が一般に採用するのは、最も大きい間引きの段階である 4:2:0 である。これは色差信号を 4 分の 1 に削減するという意味だ。こうすることで、輝度信号はそのまま残るが、色差の 2 値はそれぞれ 25% の量となり、総体として RGB に比べ、半分のデータ量にできる。ここで削除された情報は復元できない。非可逆圧縮は「情報源符号化定理」が示した限界を嘲笑うかのようにデータを欠落させる。クロマ・サブサンプリングは、非可逆圧縮の歴史の中でも最も過激で貪欲なもののひとつだろう。JPEG は、この貪欲さを取り込む。

YCbCr の採用は JPEG のエンコードにおいて必須とされているものではない。だが、仕様としてその利用を可能にした。慣例的に JPEG はデフォルトでこれを採用し、4:2:0 のクロマ・サブサンプリングを行う。仕様、データ構造、処理単位と順序、アルゴリズム、すべてがこのことをアフォードしている。同じことは量子化テーブルについても言える。仕様だけを見れば、高い周波数成分は必ずしも切り捨てを求められていない。JPEG エンコーダは、JPEG ファイルひとつひとつに、ファイル固有の量子化テーブルを自由に定義し書き込むことが可能

だ。高い周波数成分を保持し、低い周波数成分を切り捨てるという普通とは逆の量子化テーブルも定義できる。だが、仕様上可能であっても、そうなることはまずない。なにより、エンドユーザーがこれを変更したり操作したりする余地はない。「見えないものは存在しない。」これは決定事項である。JPEG はそう決めたのだ。

一方、ステガノグラフィーは「存在するものが見えない」ことを目指している。見えない細部への埋め込み、見えないインク。これは JPEG のイデオロギーと相容れない。なぜなら、JPEG は存在するものを消してしまう。それを見えなくすることはもはやできない。もう無いのだから。ゆえに、JPEG のイデオロギーに基づいてステガノグラフィーを構想するなら、「存在しないものが存在する」という不可能な状態を目指すことになる。

2.3 JPEG へのステガノグラフィー

JPEG に対して情報を埋め込むステガノグラフィー手法はこれまでいくつか提案されてきた。ここでは、その最初期の実装である Jsteg [4] と、Jsteg のバリエーションであり量子化テーブル改変を行う、C.C. チャンらの手法 [5] を参照する。

Jsteg は、量子化された DCT 係数に対して秘匿情報を埋め込む。対象となっている量子化された DCT 係数とは、これ以上欠損せず、あとは符号化によって整理されビットとして収められるだけの存在であり、ゆえに欠落を避けたい情報隠蔽の埋め込みステージとして利用される。Jsteg は、MCU それぞれの量子化された DCT 係数に対して、LSB 法でデータを埋め込む。ただし、値が 0 と 1 と -1 の係数は避ける。というのも、0、1、-1 はいずれも 2 ビットで表現される値であり、JPEG が扱う係数の中で最小ビット桁の値だからだ。十分なビット桁がない値に対する LSB 法は、データを大きく破損させるため、これらの値を避けているのである。Jsteg には改善の余地があり、改良する提案がいくつも出されてきた。ただし後続の提案も、基本的なコンセプトは踏襲して

| | | | | | | | |
|----|----|----|----|-----|-----|-----|-----|
| 16 | 11 | 10 | 16 | 1 | 1 | 1 | 1 |
| 12 | 12 | 14 | 1 | 1 | 1 | 1 | 55 |
| 14 | 13 | 1 | 1 | 1 | 1 | 69 | 56 |
| 14 | 1 | 1 | 1 | 1 | 87 | 80 | 62 |
| 1 | 1 | 1 | 1 | 68 | 109 | 103 | 77 |
| 1 | 1 | 1 | 64 | 81 | 104 | 113 | 92 |
| 1 | 1 | 78 | 87 | 103 | 121 | 120 | 101 |
| 1 | 92 | 95 | 98 | 112 | 100 | 103 | 99 |

Figure 5. C.C. チャンら [5] の提案する量子化テーブル

おり、0 や 1 や -1 を避ける点も同じである。量子化された DCT 係数に対して LSB 法でデータを埋め込む方法は、JPEG のステガノグラフィーの一般的なテクニックだ。Jsteg はその嚆矢であり、すべての JPEG ステガノグラフィーはなんらかの形でその影響下にあると言える。

C.C. チャンらの手法は、Jsteg の改良版であり、埋め込みの対象となる量子化された DCT 係数をより選択的に絞り込む。この手法では修正された量子化テーブル (Figure 5) が用いられる。テーブル内で値が 1 になっている箇所に該当する中間周波数帯が、埋め込み対象となる。量子化テーブルの値が 1 であるとは、DCT 係数が除算されずにそのまま整数となり登録されることを意味する。言い換えれば大きな係数は大きなまま残る。量子化された DCT 係数ができるだけ大きくなるように中間周波数を選択しているのだ。高周波域はそもそも値が大きくなりやすく、低周波域は人間の目にとって重要度が高いため画像の破損につながるため、対象としない。さらに中間周波数帯では大きな数残りやすいため、LSB 法のビット数を 2 ビットに拡大している。埋め込み対象を絞り込んだ分、LSB の量を増やし情報量を確保しているのだ。Jsteg と同様に、0、1、-1 の係数は避ける。この手法は Jsteg をより戦略的に修正していると言える。著者らの主張によれば、Jsteg よりも画像劣化を 1 割ほど軽減した上、隠匿できる情報の量を約 3 倍にまで増やすことができるという。ただし、C.C. チャンらの手法に

は、問題点として量子化テーブルのあからさまな不自然さがある。量子化テーブルの1の連続は、何らかの秘匿を示唆するように見えるのではないか。量子化テーブルを目視することのない一般ユーザーの目から隠すことができるのは確かだろう。だが、監視者や傍受者を想定した場合はどうか。著者らは、埋め込む前にメッセージを暗号化することを提案している。

Jsteg とその派生は、0 を避ける。無には何も埋め込むことができない。しかし、JPEG 圧縮の特徴は、量子化された DCT 係数に 0 が大量に出現することだった。JPEG の内部では 0 の連続があまりにも多い。JPEG が掲げたイデオロギーがそれを望み、推進するからだ。0 を避けるということはこの JPEG の荒野のほとんどすべてを利用できないことを意味する。本稿が提案する手法は、これを可能にする。

3 提案する手法

3.1 0 を含む量子化テーブルを用いたステガノグラフィー

提案する手法、WDITDNE もまた、量子化された DCT 係数にデータを埋め込む。埋め込んだ値を隠匿するた

| | | | | | | | | | | | | | | | |
|-----|----|----|----|----|-----|----|----|-----|----|----|----|----|----|----|----|
| 16 | 11 | 10 | 16 | 24 | 40 | 51 | 61 | 17 | 18 | 24 | 47 | 99 | 99 | 99 | 99 |
| 12 | 12 | 14 | 19 | 26 | 58 | 60 | 55 | 18 | 21 | 26 | 66 | 99 | 99 | 99 | 99 |
| 14 | 13 | 16 | 24 | 40 | 57 | 69 | 56 | 24 | 26 | 56 | 99 | 99 | 99 | 99 | 99 |
| 14 | 17 | 22 | 29 | 51 | 87 | 80 | 62 | 47 | 66 | 99 | 99 | 99 | 99 | 99 | 99 |
| 18 | 22 | 37 | 56 | 68 | 0 | 0 | 77 | 99 | 99 | 99 | 99 | 99 | 99 | 99 | 0 |
| 24 | 35 | 55 | 64 | 81 | 0 | 0 | 92 | 99 | 99 | 99 | 99 | 99 | 99 | 0 | 0 |
| 49 | 64 | 78 | 87 | 0 | 0 | 0 | 0 | 99 | 99 | 99 | 99 | 99 | 0 | 0 | 0 |
| 72 | 92 | 95 | 98 | 0 | 100 | 0 | 99 | 99 | 99 | 99 | 0 | 0 | 0 | 0 | 0 |
| (a) | | | | | | | | (b) | | | | | | | |

Figure 6. WDITDNE が用いる量子化テーブル . 輝度用のもの (a) と色差用のもの (b)

めに、量子化テーブルを改変する。埋め込み箇所に対応するテーブル値を 0 に変更して保存するのだ。JPEG デコーダは係数を逆量子化するために、量子化された DCT 係数と量子化テーブルの乗算を行う。量子化テーブルの値が 0 の箇所は、どのような異常値が潜んでいても 0 と解釈される。

Figure 6 に、この手法が改変した量子化テーブルの例を示す。元となっているのは、Figure 3 で示した標準テーブルで、それぞれ 10 個の値が 0 に書き換わっている。ジグザグ・シーケンスの後方寄りの、全体の中でより大きなテーブル値を書き換え対象にする。言い換えれば、JPEG がより積極的に 0 にしたがつている、0 になる可能性が高い周波数が対象となる。この箇所にあたる量子化された DCT 係数は、デコード時に強制的に 0 になるので、自由な数値に変更して構わない。LSB 法などのビット操作は必要ない。たとえば ASCII の 1 バイトをそのまま挿入してかまわない。量子化テーブルに 0 をいくつ挿入するかは自由に決められる。

この量子化テーブルは明らかに異常である。量子化テーブルの値に 0 が入ることを JPEG は想定していない。なぜならそのようなテーブル値はゼロ除算をひき起こし、エンコードを破綻させるからだ。しかし、エンコーダとデコーダは別々のエージェンシーである。デコーダは除算はせず乗算するだけであり、通常 0 をチェックするように実装されていない。そして、エンコーダは通常、量子化テーブルをその都度自ら生成しており、既存のファイルに入っているテーブルを参照することはない。WDITDNE はデコーダに気づかれることなく、ただ 0 の連続に傷をつけていく。量子化の際に挿入されるゼロランレングスの特殊符号、0 が 16 個連続することを示す ZRL、および最後まですべて 0 であることを示す EOB は、JPEG のイデオロギーの徴でもあるだろう。WDITDNE はそれらを消去し、かわりに JPEG が知り得ないメッセージを埋め込む。

3.2 PoC 実装

概念実証としての簡易実装を次の URL で公開している。 <https://github.com/ucnv/wditdne/>

3.3 効果検証

Jsteg と WDITDNE との比較を Table 2 に示す。オリジナル画像は Figure 1 (a) に示した JPEG で、標準量子化テーブルを用いたものだ。それぞれ、限界まで情報を埋め込み、その埋め込み許容量と埋め込み結果のファイルサイズ、および画像品質を示す。画像品質を示す PSNR は、両者のオリジナルとなった JPEG との比較値である。

WDITDNE は、許容量と品質において Jsteg を圧倒する。ただしファイルサイズにおいて不自然な肥大を見せる。WDITDNE は量子化テーブルの中の値を 10 個、0 に書き換えたものを用いた。許容量およびファイルサイズは書き換えた 0 の数で調整可能である。

Table 2. Jsteg と WDITDNE との比較

| | 許容量 (bits) | ファイルサイズ (KBytes) | 画像クオリティ (PSNR) |
|-------------|------------|------------------|----------------|
| オリジナルの JPEG | - | 8.5 | - |
| Jsteg | 12664 | 8.5 | 32.74 |
| WDITDNE | 122888 | 47 | 51.01 |

4 見えないものは存在しないと私たちは決めた

肉眼では見えなかった事象がカメラによって見えるようになることを、ヴァルター・ベンヤミンは「視覚的無意識」と呼んだ [8]。技術による視覚の拡張のことだ。彼の脳裏にはマイブリッジやマレーの連続写真があったに違いないが、彼はこれらの写真家たちに全く言及していない。そのエッセイの中で他の多くの写真家を名指しているにも関わらず。

＊

マイブリッジの馬の連続写真は賭けの勝負を決するために撮られた。見たいという純粋な欲求によるのではなく、誰かの賭け金のために技術を生み出し、マイブリッジは歴史に名を残した。一方、JPEG の発明者として個人名が名指されることはない。

＊

「視覚的無意識」によって得られた解像度を、JPEG は過剰な解像度と判断する。見えるようになったものを、見えないものとして存在の外へ返却する。技術による技術の否定だ。

＊

シャノンは役に立たない機械を作り所有していた。スイッチが付いた小さな箱で、スイッチをオンにすると中から指が出てきてスイッチをオフにする、という機械だ。[7] 自己目的であることが、祈りのように見えることもある。

＊

JPEG は、見えないという判断を行うが、この判断は肉眼に完全に寄り添っているわけではない。JPEG によって馬の宙に浮いた四足が判別できなくなるわけではない。JPEG には独自の閾値があり、独自の価値観がある。JPEG を使う者の目はその価値観に従わざるをえない。

＊

JPEG のエンコードには、量子化のプロセスがある。だがそもそも、デジタルデータとは量子化された情報ではなかったのか。JPEG は、量子化を経てきたデジタルデータに対して、さらに過酷な量子化を行い剰余をさらに削り取る。多重の量子化を行う画像フォーマットだ。もはや還元されえない剰余は存在しない。

＊

だから JPEG は速い。十分速いと感じているところへ、さらなる加速を加えてくる。到達できないだろうと思われた領野にたやすく踏み込んで走り去る。こんなに残酷な処置を行うアルゴリズムを私は他に知らない。

＊

ヒト・シュタイエルは、過度の圧縮によって劣化しながら速度を獲得したイメージファイルを称揚して「貧しいイメージ」と呼ぶ。資本主義構造によって失われそうになりながらネットワーク上に残存し続けるイメージに、解放の希望を託しながらマニフェストを記した [10]。だが、JPEG のイデオロギーに従うなら、失われそうになっているイメージというものは、まずもって消去され、存在しないことになる。だから貧しいイメージは、圧縮による劣化と速度によって生まれるのではない。その可能性はないに等しい。であるから、劣化と速度を戦略的に使いこなすことによって意図的に生み出すしかない。すなわち、闘争の本質は変化していない。

＊

ステガノグラフィーとは元来、闘争の技法のひとつである。ただし技法というものは、単体で対象を固定的

に示すわけではない。本稿による提案の特異な点は、闘争の対象を JPEG のイデオロギーと特定していることだ。

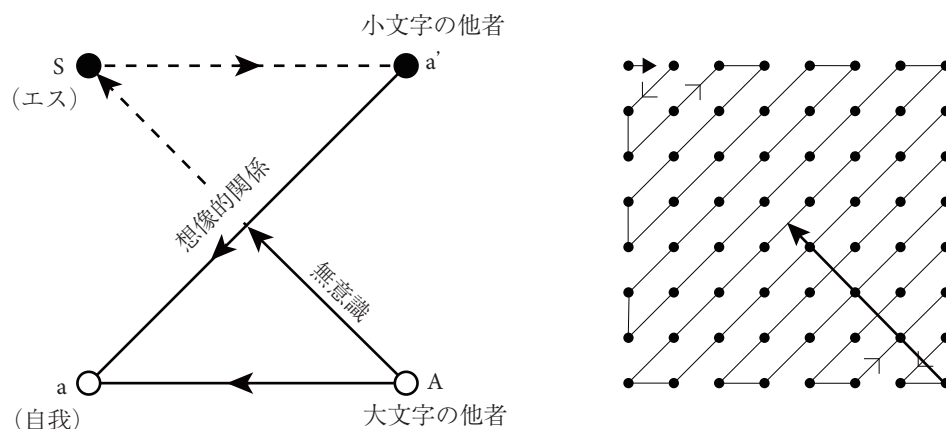


Figure 7. ラカンのシェーマ L (左) と矢印を追加したジグザグ・シーケンス (右)

この態度は自己目的的なニュアンスを孕む。自己目的的であることは遊戯的ということでもある。しかし遊戯的であることが政治的になることもある。

5 まとめ

ロザリンド・E・クラウスは、ベンヤミンの「視覚的無意識」を援用しながら、次のように問う。「視覚的な場——雲、海、空、森などの視覚的現象の世界——は無意識を持っているのか。」[9] この問いは、ベンヤミンがカメラを無意識として定位したことを批判的に検討した上で、視覚自体の中に無意識を措定しようとする試みと理解できる。カメラ、視覚、では、JPEG どうか。本稿は次のように問わなければならない。「JPEG は無意識を持っているのか。」クラウスが執着するラカンのシェーマ L では、無意識は右下から左上に向かって伸びようとする矢印だ。JPEG という主体にとっての大文字の他者の場所に JPEG のイデオロギーを措定するならば、「見えないものは存在しない」というメッセージから発し、JPEG 自身に向かってくる矢印に、無意識と記入することができる。この図と JPEG のジグザグ・シーケンスのマトリクスを併置してみれば、JPEG のシーケンスに重ねて、高い周波数成分から低い周波数成分への矢印を配置してみることができる (Figure 7)。WDITDNE はジグザグ・シーケンスの左下、つまり高周波成分を示す箇所、エンコードされて 0 が密集する箇所を重点的に利用し、これらを破壊する。大文字の他者による「見えないものは存在しない」というメッセージを否定し、そのことによって予期される「見えないものが存在する」という状況をもただちに置き去りにして、「存在しないものが存在する」という論理矛盾を実体化しはじめる。このことによって、JPEG の無意識は宙吊りになるだろう。

参考文献

- [1] Johnson, Neil F., Zoran Duric, and Sushil Jajodia. Information Hiding, Steganography and Watermarking-Attacks and Countermeasures, Kluwer Academic Publishers, 2001.
- [2] "Information Technology - Digital Compression and Coding of Continuous-Tone Still Images - Requirements and Guidelines" in , ISO as 10918-1 and ITU as T.81, 1992.
- [3] Haines, Richard F., and Sherry L. Chuang. "The effects of video compression on acceptability of images for monitoring life sciences experiments." IEEE Computer Society Data Compression Conference. No. NASA-TP-3239. 1992.

- [4] Upham, Derek. Jsteg. [Online]. Available: <http://zooid.org/~paul/crypto/jsteg/>, 1993.
- [5] Chang, Chin-Chen, Tung-Shou Chen, and Lou-Zo Chung. "A steganographic method based upon JPEG and quantization table modification." *Information Sciences* 141.1-2 (2002): 123-138.
- [6] Shannon, Claude Elwood. *The Mathematical Theory of Communication*, by CE Shannon (and recent contributions to the mathematical theory of communication), W. Weaver. Champaign, IL, USA: University of illinois Press, 1949. [クロード・E・シャノン、ワレン・ウィーバー『通信の数学的理論』植松友彦訳、ちくま学芸文庫、2009年]
- [7] Soni, Jimmy, and Rob Goodman. *A Mind at Play: How Claude Shannon Invented the Information Age*. Simon and Schuster, 2017.
- [8] Benjamin, Walter "Kleine Geschichte der Photographie"(1931), in *Gesammelte Schriften*, II (I), Frankfurt am Main: Suhrkamp, 1977; Zweite Auflage, 1989. [ヴァルター・ベンヤミン「写真小史」『ベンヤミン・コレクション (1)』浅井健二郎編訳、久保哲司訳、ちくま学芸文庫、1995年]
- [9] Krauss, Rosalind E. *The Optical Unconscious*. MIT press, 1993. [ロザリンド・E・クラウス『視覚的無意識』谷川雁／小西信之訳、月曜社、2019年]
- [10] Steyerl, Hito. "In Defense of the Poor Image." *e-flux journal* 10.11, <https://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image>, 2009.