

文字への加工を利用したハードコピーへの情報ハイディングの一提案

林 克明[†] 上田 芳弘[†]
岩田 雅士^{††} 木村 春彦^{†††}

近年はインターネットをはじめとしたコンピュータネットワークの発展、普及にともない、電子文書が流通する機会が増えてきた。それにともない、情報を伝達する媒体として用いられてきた紙の使用頻度は低下していると考えられる。しかし、情報を記録する媒体としての紙は、多くの利点が存在するため、早急になくなるとは考えられず、今後も使用されると予想される。そこで、本研究では紙、すなわちハードコピー状態のテキスト文書を対象とした、埋め込み情報の抽出を視認で行うことができる情報ハイディングの一手法を提案する。さらに、実験によりその有効性を示す。テキストを対象とした情報ハイディングに関する報告の多くは、文字間隔や行間隔の微小変化を利用するものが多いが、これらの方法ではハードコピーにした場合のハイディング情報の抽出には電子機器が必要である。本論文の提案手法では、テキストへの情報ハイディングを行うために、文字の一部分に切断加工を施した。その実用性についても実験により実証できた。これにより、ハードコピーのテキスト文書を対象とした、視認による抽出を可能とした情報ハイディングの有効性を示すことができた。

A Proposal for Information Hiding to Hardcopy by Using Cutting Processed Characters

KATSUAKI HAYASHI,[†] YOSHIHIRO UEDA,[†] MASASHI IWATA^{††}
and HARUHIKO KIMURA^{†††}

Recently, opportunities of exchange for electronic documents are increasing with development of computer networks such as the Internet. And the use frequency of paper as an information transfer medium is decreasing. However, it will be seemed that paper as an information record medium will continue to be used for its some merits. In this research, an information hiding method for text on hardcopy is proposed. The characteristic of the method is to extract embedded information in visual. Furthermore, three experiments show the validity of the method. Many of reported methods of information hiding into text are using line-shift coding or word-shift coding. These methods need to use image scanner to extract embedded information. We give a cutting processing to a character in order to hide information into a text. And the validity of our cutting method is shown by experiments. Therefore, the validity of proposed information hiding for text on hardcopy is shown.

1. はじめに

近年は、インターネットをはじめとしたコンピュータネットワークの発展、普及にともない、電子的な方法で文書を流通する機会が増えてきた。また、文書の流通量が増えるに従い、著作権保護や著作権侵害抑制ということが問題化し、その解決のために、電子透かしなどの技術が注目されている^{1),2)}。このことにより、

ますます文書の電子化が促進されるものと考えられ、情報の伝達媒体として用いられてきた紙の使用頻度は、低下するものと考えられる。しかし、情報を記録するための媒体としての紙は電子媒体に比して、以下の理由から今後も必要であり、使用され続けるものと予想される。

- (1) 保存媒体としての優秀性（水に濡れても読める、任意の折り曲げに耐えられる、など）
- (2) 読みやすさ（ページめくりがしやすい、など）
- (3) 証拠性の高さ（改ざんや複写の事実が分かりやすい、など）

そこで、本論文ではハードコピー状態でのテキスト文書を対象とした、情報ハイディングの一手法を提案する³⁾とともに、手法を検証した評価結果についても

[†] 石川県工業試験場

Industrial Research Institute of Ishikawa

^{††} 金沢工業大学

Kanazawa Institute of Technology

^{†††} 金沢大学工学部

Faculty of Engineering, Kanazawa University

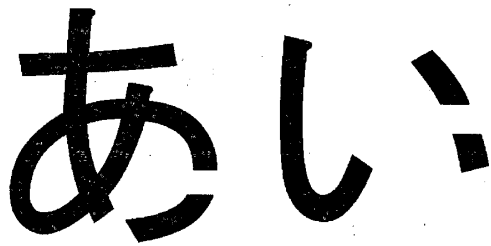


図 1 切断加工された文字
Fig. 1 Cutting processed character.

報告する。本研究での検証対象は和文テキストとしたが、提案手法は欧文テキストにもそのまま適用可能である。

現在のハイディング技術は情報をハイディングするためにコンテンツの冗長性を利用するものが多く¹⁾、冗長性が低いテキスト文書への適用は困難であり、その研究事例は、画像などへの適用事例に比して多いとはいえない。本研究では、図 1 のように文字フォントの一部切断という文字への加工を利用して、テキスト文書への情報のハイディングを行う。さらに、本論文では文字に切断加工を施しても文書品質が低下する問題のないことや、実験により検証した結果について述べる。提案手法は、以下に述べる特徴を有する。

- (1) ハードコピーの和文および欧文のいずれのテキスト文書にも適用ができる。すなわち、提案が比較的少ないテキストを対象とした情報ハイディングの新しい手法を提案できた。
- (2) ハイディングされた情報の抽出を視認によって行うことができる。すなわち、情報の処理に電子機器などを必要とせず、簡便に利用できる。

ここで、本論文で使用する用語を次のように定める⁴⁾。情報をハイディングすることを、“埋め込み”という。このときに、埋め込まれる情報を“埋め込み情報”といい、埋め込み先のテキストメディアを“カバーテキスト”という。カバーテキストに埋め込み情報が埋め込まれた状態のテキストを“ステゴテキスト”という。ステゴテキストから埋め込み情報を取り出すことを“抽出”という。また、本論文で独自に用いる“読みとり”とは、ステゴテキストから、情報埋め込みのために切断加工を施された文字を読む行為とする。抽出は、さらにそこから埋め込み情報を得ること、と区別する。

以下、2章でテキストを対象とした情報ハイディングの在来研究と本研究の手法について述べ、3章では、本手法の検証実験とそのまとめについて述べ、最後に4章で本研究で得られた結果と今後の課題について述べる。

2. テキストを対象にした情報ハイディング

現在の電子透かしを含めたハイディング技術は、画像や音楽への情報埋め込みに関しての研究例が多く報告されている。これは、情報ハイディングのためにコンテンツの冗長性を利用することが多く、冗長性が低いテキスト文書への適用は困難なためであり、その結果としてテキストを対象とした研究事例が比較的少ない。

そのような中で、テキスト文書を対象にした情報ハイディングの手法は、次のように大別できる。

- (1) 文字間隔や行間隔を微小に変化させ、間隔の相違を利用して情報を表現したり、表示されない文字を挿入したりする方法^{5),6)}。この方法では、テキストを2値画像として扱うことが多く、広義には改行コードの挿入位置を調整する方法⁷⁾も含まれる。
- (2) テキスト文書中の言語表現の意味を変えずに別の言語表現へ置換する辞書変換法と呼ばれる方法⁸⁾。この方法では、テキストそのものがステゴテキストとなる。
- (3) 文字形状の一部を微細に変化させ、変化の有無で情報を表現する方法^{9),10)}。この方法もテキストを2値画像として扱うことが多い。

また、上記の方法の多くは、ハードコピー状態のステゴテキストから埋め込み情報を抽出するために、スキャナなどの電子機器による機械処理を必要とする。ここで、(1)に例示した文献7)は、改行コードの挿入位置の変更、すなわち1行に含まれる文字数を増減させ、たとえば1行の文字数が34字ならばその行はビット0を表し、文字数が35字ならばその行はビット1を表すという手法である。そのため、表現できる情報ビット数は、行数と一致する。よって、視認によっても埋め込み情報を抽出することが可能と思われる。すなわち、情報の抽出方法によって、(a)機械処理が必要か、(b)機械処理が不要か、という分類もできる。なお、我々の提案手法では、1行で8ビットを表現しているの、文献7)に比して、埋め込み情報量で8倍多い利点がある。

本研究では、(3)の文字形状の一部を変化させる方法を採用し、視認によって埋め込み情報を抽出することを可能としている。文字形状の変化方法は、図1のように、文字フォントの一部を切断するという手法である。文字フォントの切断方法については3章で述べる。

さらに、本論文で提案する手法は、埋め込み情報を

人間の視認で抽出させることを前提とし、文字フォントへ施した切断加工は熟視すれば分かる程度のものとなっている。また、切断幅は微小なため、情報ハイディングのために文書中の一部のフォントにのみ加工が施された場合でも、第三者に対しては、後述する理由により簡単には見破られることがないと考えている。このような、視認による抽出を前提とした方法についての報告はこれまでにあまり見られない。たとえば、上述した文献(9)、(10)では、微少な変化を文字フォントへ与える方法であるが、文字のフォントサイズが72ポイントや、48ポイントと通常の文書には使用されないサイズであり、情報の抽出にスキャナなどの機械処理を必要とする。しかし、これらの手法も目的によっては適切な場合があると考えられる。

ここで、一般に文字の一部が切断などにより欠損することは、文書のコンテンツとしての品質低下に直結すると考えられるが、コンテンツ品質低下を最低限におさえながら、かつ、情報読みとりを可能とできるサイズの切断幅を実験により求めた。ここで、コンテンツとは、「情報の内容」のことを指すものとされる¹¹⁾。このときの情報とは、放送、映画、文書、画像などである。本研究では、それらの内容の中で文書のことを指すものであり、カバーテキストのことである。また、本論文ではハードコピーを対象としているので、印刷された文書とも換言できる。さらに、コンテンツ品質の低下とは、文書が読者にとって読みにくい、または、読むにあたって違和感を覚えるなどの程度が大きくなること、とする。ただし、本研究の手法では文意の変更はまったく行われないため、元のテキストの意図を損なうという種類の品質の低下⁸⁾は発生しない。

次に、提案手法の適用例を述べる。本手法の利点は、視認での抽出が可能であり、電子機器を必要としないことである。そこで、ハードコピー媒体に、ある簡略な情報を埋め込み保存する、情報を埋め込んで他人へ伝え、簡単に抽出する、という利用例として次のようなものが考えられる。

(1) 文書を受領してすぐに、ハイディングされた情報を基に処理しなければならない。

組織においては、稟議文書や物品、出張決済伝票などが多く使用されている。具体的な内容は、書面に記載されているが、稟議の決裁を受けた後の保管先分類や、外部への開示に関することなど文書の性格についての情報を埋め込んでおくことにより迅速な事務処理が可能と考えられる。また、伝票の記入者名などを埋め込んでおくことで、伝票の改ざんを防止する効果も期待できる。

(2) 出張先など、電子機器を使用できない状況でハイディング情報を抽出しなければならない。

最近の電子機器は小型軽量化されているが、スキャナはまだ可搬性が良いといえるサイズではない。また、出張先ではスキャナを自由に使用できるとは考えにくい。このような場合に、大衆が存在する場所で他者に情報を見られないようにして、ハイディング情報が取り出せるならば、その利便性は高いと考えられる。

(3) 災害情報掲示板の中にパニック回避のため一般大衆には解読されず、消防士だけに解読させたい情報を埋め込んでおく¹²⁾、など。

これまでにスキャナやコンピュータなどの電子機器を使用した電子媒体向けの情報ハイディングはいくつも報告されている。また、ハードコピーを対象とした情報ハイディングも前述のように報告されている。しかし、これらの方法は大量の情報の埋め込みを可能にするが、視認が前提のハードコピーを対象とした、電子機器が不要の情報ハイディングには適用しにくい。そこで、本研究が目的とした視認による情報抽出を可能とする方法では、情報の埋め込み量は比較的少なくなるが上述の他方式の欠点を補うことが可能であると考える。

そこで、視認できる情報埋め込みについて考察する。簡単に視認できるように情報をカバーテキストに追記するならば、文書の余白などに平文で書き込んでおけばよい。たとえば、脚注などがこれに相当するであろう。しかし、このようにして書き加えられた情報を第三者に知られたくない、また、関係者のみに知らせたいという場合も考えられる。すなわち、情報をハイディングしたいときに、視認を前提とすると情報埋め込みの手段がかなり制限されてしまう。たとえば、追記情報を平文ではなく、暗号文で記載するという方法についても、この場合には内容が判明しなくても、何かが記載されているという事実が第三者に分かってしまう。そこで、第三者に対して、追記された情報の存在自身も気づかせず、かつ、視認でハイディング情報を読みとる手段として、筆者らは文字そのものに着目し、その一部分に図1のように切断加工を施すという手法を選択した。

このときに、文字に施された切断加工を情報の受け手が確実に視認でき、かつ、第三者が気づかないか、という問題点について検討する必要がある。このうち、より重要な点は、文字の一部に施された切断加工が、第三者に気づかれてしまわないか、また、ステゴテキストを読む場合に違和感を与えてしまわないか、であると考えられる。この点について、以下の理由から、

第三者には気づかれにくく、また、ステゴテキストを読むにあたって違和感を与えることはないと思われる。さらに、3章での実験によってそのことを実証できたと考える。すなわち、本研究が提案する手法は妥当であると考えられる。

- ・ 視覚的補完機能により文字の切断加工部分が気にならない。

視覚入力にあるものが推定されて知覚される場合、視覚的補完が生じていると報告されている¹³⁾。知覚とは、感覚の質（たとえば視覚なら色の違い、など）を識別することとされている¹⁴⁾。また、視覚的補完とは、視覚入力すなわち目で見た物体が網膜上で結像したものを、脳内において知覚し、過去の経験などから照らし合わせてその物体が何であるかを認知する場合、照らし合わせる対象がないときに不足した情報を補う機能のこととされている。上記を本論文の場合にあてはめると、ステゴテキストとしての文字および切断加工された文字が知覚された場合、文字に切断加工がなされていることの予備知識を与えられていない人間には、切断加工の存在を認知できず、過去の経験に基づき完全な文字、つまり切断加工を施されていない文字として認知される、と考えられる。換言すると、一部が切断されている文字についての予備知識がない人間は、切断部分を無意識レベルで補完して、完全な文字（切断加工がない文字）として認識しているのではないかと考えられる。ここで、文字を熟視すれば、多くの人間が文字に施された切断加工に気づくが、通常人間は、文書を読む場合1文字1文字を認識しながら知覚するのではなく、単語などの、ある程度の意味のあるまとまり単位で知覚するものと考えられる。さらに、人間の形状認知は、文脈や主体の目的に合わせて変化し、網膜上での形状の変化（ここでは、文字の切断加工の有無）が結像しても、同一のものとして（ここでは、切断加工に気づかず通常の文字として）認知できる¹³⁾。とも報告されている。以上から、文字の切断加工部分は、第三者に気づかれにくいと考えられる。

さらに、EPSON社製スキャナ、GT-7000 WINSおよびOCR (Optical Character Reader) ソフト (読ん de!! ココ, Ver.6) を用いての機械的耐性についても実験を行った。実験に使用している12ポイントサイズのフォントにおいて、切断加工を有する全ひらがなについて、切断加工がない正常な文字として認識できた。対象とした文字がひらがなのみの理由は、本章末で述べる。OCRソフトは、比較的入手が容易で、一般的に使用されているソフトウェアである。すなわち、汎用ソフトウェアによる処理においては、切断加工を

有する文字についての機械的耐性は十分であり、切断加工文字の存在を知られたいくない第三者に対して、機械処理によってその存在が知られてしまうことはないと考えられる。

また今回の実験では、1文字につき、1カ所のみ切断加工を施した。しかし、情報埋め込みのために複数の切断加工を施すことや、数を1個に限定した場合でも、その位置を変更することにより、別の情報を持たせることも考えられる。しかし、本論文では、切断箇所数を1個とし、その位置も固定的とする。これは、本研究はまだ基礎段階であり、手法の有効性を検証することが主目的のため、情報埋め込みのためのパラメータは必要最小とした。

さらに、今回の実験では、切断加工を施したフォントはひらがなフォントとした。同様に、漢字や英数字などのすべての文字に切断加工を施すべきであるかもしれないが、今回は切断加工文字そのものについての検証も行うため、簡易的にひらがなのみを対象とした。しかし、ひらがなは、約50種類であるが日本語文書の51~62%、平均では57.3%を占めており¹⁵⁾、情報を埋め込むカバーテキストの実験条件を十分満たしていると考えられる。

3. 実験による検証

本研究においては、提案した手法を検証するために、以下の3実験を順次実施した。

- (1) 文字の切断加工幅決定のための実験
- (2) 埋め込み情報の抽出実験
- (3) コンテンツ品質の確認実験

3.1 切断加工幅決定実験

3.1.1 実験の概要

提案する手法は、図1のように文字の一部に切断加工を施した文字フォントを用いるものである。そのためまず、文字に施す切断加工の幅を決定する必要がある。このための実験を決定実験と呼ぶ。なお、図1の切断幅は、例示用に実際より広がっている。また、図2は、切断加工を有する文字を実際に使用したテキストの一部である。ここで、フォント加工のために、シェアウェア（一定期間の無料試用後に正式版を購入するという形態の商用ソフトウェア）である外字・TrueTypeフォントエディタ TTEdit¹⁶⁾ のバージョン3.20を用いた。TTEditを動作させたパソコンの環境を表1に示す。本アプリケーションは、ドロー系グラフィックソフトと同様な操作でフォントを作成できるものである。また、フォント作成のために参照したのはMSゴシック体フォントである。この参照し

本システムの利用によって、どの程度のため、全処理件数のうち、どの程度がした帳票を用いる帳票承認方式で処理された電子決裁と帳票承認の件数及び電子決裁システム利用開始後の業務の全件数裁処理された全件数は 264 件であり、残来と同様の方式で処理された。後者の 274 件、システム評価の対象ではないが、べを試みた。この 274 件の起案は、電子化ししかし、電子化可能な案件については、

図 2 実験に使用したテキスト例
Fig.2 Sample text of experiment.

表 1 フォント作成ソフトウェアの動作環境
Table 1 Computer environment of software.

OS	Windows Me
CPU	Celeron 500 MHz
メモリ	128 MB
ハードディスク容量	20 GB

表 2 実験の共通条件
Table 2 Common conditions of experiments.

文字サイズ	12 ポイント
文字の種類	ひらがな
フォント	ゴシック体

たフォントに切断コマンドを適用し、その幅を調節して、必要な幅の切断加工を有する文字フォントを作成した。

切断幅のサイズは、絶対値ではなく、文字を構成する線幅に対する割合 (%) で表示する。これは、文字のサイズが変更されても、扱いを統一するためである。実験に用いた文字の基本条件などを、表 2 にまとめる。この条件は、本実験だけではなく、以後のすべての実験に共通するものである。ここで文字のフォントとして、ゴシック体フォントを使用したのは、ほぼ均一の線幅で文字が構成されており、その結果として切断幅 (%) を表現するのに最も適していると考えたためである。切断幅は、広ければ広いほど視認しやすいが、文字として不自然な感じを読者に与え、コンテンツ品質の低下を招く。逆の場合は、切断加工の視認がしにくい。そこで、決定実験の目的は、文字に施された切断加工を視認することが容易かつ、コンテンツの品質低下を最小限に抑えられる切断幅を決定することである。

実験は、切断加工された文字を何文字視認できるか、という方法で行った。まず、1) 420 文字の文書を用意し、2) そのうち 20 文字を切断加工された文字に置換

表 3 決定実験の条件

Table 3 Conditions of determining experiment.

被験者数	延べ 80 人
被験者の年齢	19~21 歳
切断加工文字数	20 個 (1 文書あたり)
読みとり試行回数	3 回

し、3) 計 80 人の被験者に切断加工された文字を視認で探し出してもらう、という流れで行った。そして、この文書を異なる切断幅ごとに用意した。1 枚の用紙には、1 種類の切断加工幅の文字だけが含まれる。用意した切断幅は、40%、50%、60%そして 70%の 4 種類である。これは、事前に筆者らでいくつかの切断幅で検討したところ、以下の理由から、上述の 4 種類を選定した。(1) 40%以下の切断幅では、肉眼での認識が大変困難となること。(2) 70%を超えると、文書中において目立ちやすく、違和感を覚えやすいこと。(3) 5%のサイズ変更では、違いが感じられにくいこと。なお、4 種類の文書中における切断加工文字の位置は、それぞれ異なるように配置した。

3.1.2 決定実験の結果

本実験の条件を表 3 に示す。本実験では、まず被験者 80 人を 20 人ずつ、A、B、C、D の 4 グループに分けた。そして、1 つのグループの被験者は、1 種類の切断加工幅を有する文字を埋め込んだ文書を読んでもらった。たとえば、A グループは、切断加工幅が 40%のものを、B グループは、同 50%のものを読んでもらうという要領で行った。そして、A グループの被験者は、他のサイズの加工幅 (50%、60%、70%) の文章が埋め込まれた文書を読むことはない。これは、たとえば被験者が 40%の切断加工幅の文字を読んだ後での 50%以上の文字での認識率が、50%の文字だけを読んでいる被験者よりも向上すると想定されたため、それを避けるためである。

実験は、埋め込まれた 20 個の切断加工を有する文字を何個視認できるか、という指標で評価した。結果を図 3 に示す。なお、図 3 における誤差棒は 95%信頼区間を表す。実験結果から、

- (1) 切断加工幅が 40%および 50%の文字では、視認性が十分ではなく、情報ハイディングに使用することは困難である。
- (2) 切断加工幅が 60%および 70%の文字では、2 回以上読みとることにより認識率が 95%以上であった。その結果、我々は今後の実験を行っていくための基準として切断加工幅が 60%の文字を選定した。これは、1) 1 回目の試行では 83.4%の視認率であり、また、70%の文字の視認との差に大きな差が認められず、

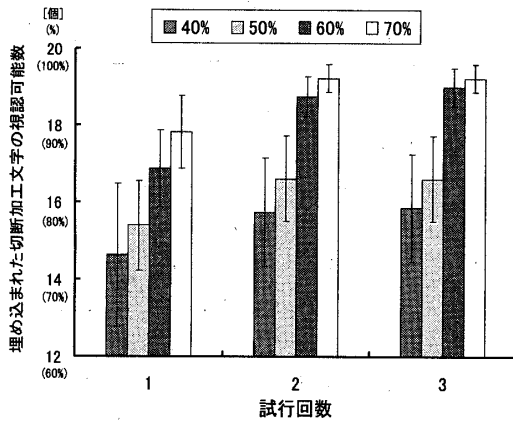


図3 決定実験の結果

Fig. 3 Result of determining experiment.

2) 2回目以降の試行において、70%の文字とほぼ同程度の視認性がある。

以上の2点により、切断加工幅が60%と70%の結果は視認可能数で60%がやや劣るが、有意差がないために両者は同程度と見なせる。よって、第三者により気づかれにくいと推定される、切断加工幅が60%の文字を使用するのがよいと考えた。

以後の実験に使用するすべての文字の切断加工幅は線幅の60%である。なお、文字の大きさが12ポイントサイズの場合、20文字を無作為に抽出してキーエンス社製のデジタルマイクロスコープで実測したところ、文字の線幅の平均値は387 μ m(標準偏差15.1)であった。また、切断加工幅の平均値は無作為抽出した12標本の平均で223 μ mであった。実測した文字の線幅387 μ mの60%理論値は232 μ mとなり、実測値とは3.9%の相違が生じているが、有意水準5%でt検定したところ、両者間に有意差は認められなかったため、理論値どおりと考えてよい。

3.2 埋め込み情報の抽出実験

3.2.1 実験の概要

決定実験により求められた60%の切断加工幅を有する文字を用いて、文書に情報ハイディングを施し、これを被験者に抽出してもらう実験を行った。この実験を抽出実験と呼ぶ。本実験で行ったハイディング情報の構成方法を図4に示す。基本的に、8文字のカバーテキストであるひらがなにより、1文字の埋め込み文字を表現する。表現できる文字は、8bitのJISコードで表現できる英数字、記号とカタカナのおよそ157文字である。また、今回は、1文字の埋め込み文字を構成するためのカバーテキスト文字の8文字は、文章中の1行に収まるように配置した。すなわち、1行で埋め込み文字の1文字を表現できるようにした。本実験の条件を表4に示す。実験対象文書は3種類を用

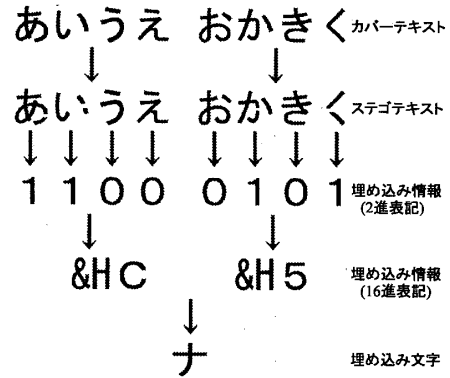


図4 情報ハイディングの概要

Fig. 4 Outline of information hiding.

表4 抽出実験の条件

Table 4 Conditions of extracting experiment.

被験者数	18人
被験者の年齢	22~24歳
文書の文字数	平均420文字
埋め込み対象文字数	80文字
埋め込み文字数	10文字
埋め込み文字列(文書1)	CP/M86 オーエス
埋め込み文字列(文書2)	TDFd ホウカイセキ
埋め込み文字列(文書3)	$S' = xY / (1 + t)$

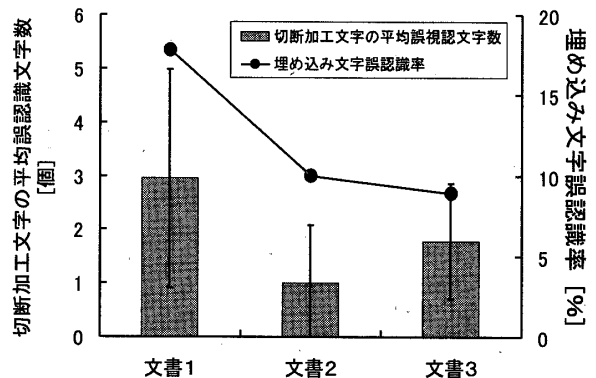


図5 抽出実験の結果

Fig. 5 Result of extracting experiment.

意し、埋め込み文字列の文言は、意味がある単語または数式とし、それぞれ異なる。そして、3種類の文書を被験者に順次読んでもらった。さらに、埋め込み文字列は被験者にとって、できるだけ前後の文字からの推定が困難な単語を選択した。これは、正確な情報抽出のために、前後の文字からの類推や、修正を防止するためである。

3.2.2 抽出実験の結果

実験結果を図5に示す。なお、図5における誤差棒は95%信頼区間を表す。平均誤視認文字数とは、切断加工を施された文字を正しく視認できなかった個数である。埋め込み文字列として、3文書とも10文字を埋め込んだ。折れ線グラフでは、この埋め込み文字

を正しく読みとれなかった割合を示す。

ここで、図5における文書2と文書3での切断加工文字の誤認識数と埋め込み文字の誤認識率が逆転しているのは、誤視認した切断加工文字の配置が原因であった。埋め込み文字は、同一行中の8文字で構成されているため、切断加工文字の誤認識数が多くても、同一行中であれば、埋め込み文字への影響は1文字である。しかし、誤認識した文字が複数行にわたれば、埋め込み文字への影響も複数文字になるために図5のようになってしまう。しかし、文書2と文書3においては、有意差があるとはいえないため、文書3における埋め込み文字の誤認識率が向上しているとはいえない。抽出実験では、各被験者には、文章を1度だけ読んでもらうようにした。このとき、決定実験において読みとり回数を2回、3回と繰り返すことにより、視認可能な切断加工文字数が向上していた。すなわち、埋め込み文字を正しく抽出するには、以下の必要性が考えられる。

(1) 図4のレベル1のエラーの低減

切断加工文字の認識率を向上させる。このために、同じ文書を2回以上読むことが有効である。決定実験の結果を援用するならば、10%程度の向上が期待できる。

(2) 図4のレベル2のエラーの低減

視認した切断加工文字から、埋め込み文字への変換ミスをなくす。これには、ユーザである被験者の慣れに期待する必要があるかもしれない。

しかし、上記の対策で埋め込み文字の平均誤認識率が低下しても完全に0になる可能性は低いと考えられる。けれども、視認による抽出が可能な情報ハイディングの手法を提案できた意義は大きいと考える。また、平均誤認識率を低減するためには、本手法を利用するユーザが正しく埋め込み文字を読みとれるようにする必要があり、そのためのチェックサムの導入などを検討する必要があると考える。

3.3 コンテンツ品質の確認実験

3.3.1 実験の概要

本実験の目的は、文字に施された切断加工が第三者に気づかれるか否かを確認し、切断加工を有する文字がどの程度コンテンツ品質を低下させるかを確認することである。切断加工が第三者に気づかれても、すぐに埋め込み情報が知られてしまうことはないが、埋め込み情報の漏洩につながることになる。本実験を確認実験と呼ぶ。すでに、切断加工文字による埋め込み情報の抽出については可能なことを検証した。そこで、埋め込み情報を必要としない、つまり埋め込み情報を

認識させたくない人間に対して、どの程度の確率で認識されてしまうかを検証した。

実験は、まず8種類の文章を用意し、1つの文章を1枚の用紙に印刷した。用意した文章の文字数は、決定実験および抽出実験とほぼ同等の440文字(8文章平均値、句読点含む)とし、文章の種類は技術的な内容で、さらに同一の文体で似通ったものを使用した。そのうちの半分の4枚の文章に切断加工を有する文字を用いて情報埋め込みを施した。埋め込まれた切断加工文字数は、42.25文字(4文章平均値)である。残りの4枚の文章は、通常のフォントのみで作成された文書である。これら、8種類、すなわち8枚の文章を、文字の切断加工による情報埋め込みについてまったく予備知識を与えていない被験者に渡して、分類してもらうという実験を行った。分類方法や分類数などの被験者からの実験に対する質問に対しては回答せず、被験者にすべてを任せた。なお、本実験の被験者には、切断加工文字の予備知識がないことが条件となるため、決定実験および抽出実験のいずれも経験していない人間を対象とした。

3.3.2 確認実験の結果

本実験の結果を表5に示す。結果として、15人中2人が切断加工を有する文字に気づき、切断加工文字の有無によってグループ分けを行うことができた正答者であり、他の13人の非正答者は切断加工文字の有無によるグループ分けができなかった。非正答者らに実験後個別に聞き取り調査を行ったが、切断加工文字の存在には気づいていなかった。

本実験から、切断加工文字についての予備知識がない人間の約87%である13人はその存在に気づかなかった。すなわち、文字に施された切断加工が、文章全体のコンテンツとしての品質を低下させることはあまりないと我々は考える。

3.4 実験のまとめ

切断加工を有する文字を用いたテキストへの情報ハイディングの有効性を示すために、3つの実験を行った。

- (1) まず、文字に施す最適な切断加工幅として60%が適していることを求めた。
- (2) 切断加工を有する文字を用いた埋め込み情報の抽

表5 確認実験の結果
Table 5 Result of checking experiment.

被験者数	15人 (うち男性9人、女性6人)
被験者の平均年齢	24.8歳
非正答者数	13人
非正答者割合	86.7%

出が可能なことを実証した。

- (3) コンテンツ品質の低下について、文字に施された切断加工によるハイディング情報を必要としない第三者への影響が少ないということを検証した。ただし、切断加工文字についての予備知識を有した人間に対するコンテンツ品質の低下についての検証は、今後の課題としたい。

4. おわりに

本論文では、文字の一部に切断加工を施すことによって、ハードコピーへ情報ハイディングを行う一手法を提案した。この方法によって埋め込まれた情報は、視認によって読みとることが可能なことを実験によって検証できた。また、文字に施された切断加工は、熟視した場合に視認できるサイズの加工である。このようなサイズでの切断加工は、第三者に気づかれにくく、かつ、コンテンツ品質の低下にあまり影響しないことを人間工学的に視覚の機構から仮定し、そのことを実験により検証した。さらに、スキャナとOCRソフトウェアを使用した機械的読みとりを行っても、切断加工された文字の存在に気づかれぬ耐性があることを確認した。これらから、本手法がハードコピーを対象とした情報ハイディングとして有効なことを示すことができた。

本論文で実施した実験による埋め込み情報の抽出では、1回の読みとりで約90%の情報認識が可能であった。この認識率は、2回以上の読みとりを行うことも有効であるが、チェックサム情報の利用などを検討することにより、さらに向上することが期待される。課題として、このチェックサム情報の利用があげられる。また、今回の報告では、悪意ある第三者の存在およびその行為についての考察までは行うことができなかった。これらについても十分に検討していくことが必要であり、今後の課題である。さらに、今回の文字への切断加工は、1文字につき1カ所であり切断位置も固定であった。今後は、1文字へ複数カ所の切断加工を施したり、切断位置を可変したりした場合の効果や情報抽出の検証についても行いたいと考えている。

参 考 文 献

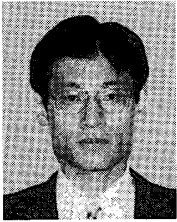
- 1) 松井甲子雄：電子透かしの基礎—マルチメディアのニュープロテクト技術，森北出版(1998)。
- 2) 佐々木良一，吉浦 裕，手塚 悟，三島久典：インターネット時代の情報セキュリティ 暗号と電子透かし，共立出版(2000)。
- 3) 岩田雅士，阿部武彦，林 克明，上田芳弘，木村春彦：ハードコピーを対象とした視覚認識可能な

情報付加方法，平成14年度電気関係学会北陸支部連合大会講演論文集，p.290 (Sep. 2002)。

- 4) 情報処理振興事業協会：「インフォメーションハイディングの技術調査」報告書(1998)。
<http://www.ipa.go.jp/security/fy10/contents/crypto/report/Information-Hiding.htm>
- 5) Brassil, J.T., Low, S., Maxemchuk, N.F. and Gorman, L.O': Electronic Marking and Identification Techniques to Discourage Document Copying, *IEEE Journal on Selected Areas in Communications*, Vol.13, No.8, pp.1495-1504 (1995)。
- 6) 中村康弘，松井甲子雄：著作権保護のための電子文書のハードコピーへの署名の埋め込み，情報処理学会論文誌，Vol.36, No.8, pp.2057-2062 (1995)。
- 7) 滝澤 修，牧野京子，村瀬一郎，松本 勉，中川裕志：改行位置の調整による自然言語テキストへの情報ハイディング，電子情報通信学会 暗号と情報セキュリティシンポジウム (SCIS2002) (2002)。
- 8) 中川裕志，三瓶光司，松本 勉，柏木健志，川口修司，牧野京子，村瀬一郎：意味保存型の情報ハイディング—日本語文書への適用，情報処理学会論文誌，Vol.42, No.9, pp.2339-2350 (2001)。
- 9) 辻合秀一，上辻雅義：文字形状を利用したレタリングへの電子透かし法，電子情報通信学会論文誌 (D-II) 研究速報，Vol.J82-D-II, No.11, pp.2175-2177 (1999)。
- 10) 岡本晃宏，宮崎明雄：モルフォロジカル信号処理を利用した電子透かし方式，電子情報通信学会論文誌 (A)，Vol.J84-A, No.8, pp.1037-1044 (2001)。
- 11) 相磯秀夫 (監修)：情報技術用語大事典，オーム社 (2001)。
- 12) 中川裕志，滝澤 修，井上信吾：インフォメーションハイディング 4.ドキュメントへのインフォメーションハイディング，情報処理，Vol.44, No.3, pp.248-253 (2003)。
- 13) 日本視覚学会 (編)：視覚情報処理ハンドブック，朝倉書店 (2000)。
- 14) 長町三生ほか：現代の人間工学，朝倉書店 (1986)。
- 15) 佐竹秀雄：各種文章の字種比率，国立国語研究所報告，Vol.71, pp.327-346 (1982)。
- 16) 武蔵オンラインソフト。
<http://www.interq.or.jp/www1/anzawa/ttedit2.htm>

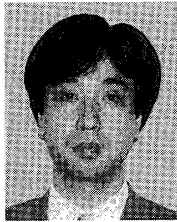
(平成15年4月28日受付)

(平成15年9月5日採録)



林 克明

昭和 39 年生。平成元年金沢大学大学院工学研究科電気工学修士課程修了。同年石川県工業試験場入所。現在、金沢大学大学院自然科学研究科博士後期課程在学中。ワークフローシステム、情報セキュリティに関する研究に従事。電子情報通信学会会員。



上田 芳弘 (正会員)

昭和 35 年生。昭和 60 年慶應義塾大学大学院工学研究科管理工学修士課程修了。同年松下電器産業(株)入社。平成元年石川県工業試験場入所。平成 13 年金沢大学大学院自然科学研究科博士後期課程修了。グループウェア、知識共有、感性工学の研究開発に従事。博士(工学)。電子情報通信学会、人工知能学会各会員。



岩田 雅士 (学生会員)

昭和 54 年生。平成 14 年金沢工業大学経営工学科卒業。現在、同大学院修士課程在学中。情報セキュリティに関する研究に従事。



木村 春彦 (正会員)

昭和 27 年生。昭和 54 年東北大学大学院工学研究科情報工学博士課程修了。同年富士通(株)入社。昭和 55 年金沢女子短期大学講師。昭和 59 年金沢大学経済学部助教授を経て、現在、同大学工学部情報システム工学科教授。その間、最適コード変換、プロダクションシステムの高速度の研究に従事。工学博士。電子情報通信学会、人工知能学会各会員。