

印刷文書を対象とした視覚認識可能な情報ハイディングの提案

清水賢[†] 林克明[‡] 阿部武彦[†] 石井和克[†] 上田芳弘[‡] 木村春彦^{††}
 金沢工業大学[†] 石川県工業試験場[‡] 金沢大学^{††}

1. はじめに

近年の急速な情報通信ネットワークの発展、普及により、コンピュータネットワークを介しての情報伝達が容易になり、コンピュータネットワークを介して電子文書を流通させる機会が増加している。その一方で、著作権保護や著作権侵害の抑制といった問題が生じている。その解決方法のひとつとして、電子透かしなど情報ハイディングの技術が提案されている^{1), 2)}。

しかし、情報ハイディング技法に関する研究の多くは電子メディアを対象としたものであり、ハードコピーを対象とした研究報告は少ない。

電子文書の流通量増加にともない、元来情報の記録、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予想されるが、以下に示す情報記録媒体としての紙の優位性により、紙の需要が今後急速に低下することはないと考えられる。

- 保存媒体としての優秀性
 - 耐水性、折り曲げ耐性
- 読みやすさ
 - ページめくりがしやすい
- 証拠性の高さ
 - 改ざん、複写事実がわかりやすい

上記により、今後も紙は使用され続けるものと予想されるため、ハードコピーを対象とした情報ハイディング技法もまた必要であると考えられる。

ハードコピーを対象とした情報ハイディング技術のひとつとして文字の加工を利用した情報付加手法³⁾が提案されている。この手法は文字に切断加工を施し、切断の有無によって情報を表す。しかし、埋め込みの対象はひらがな、カタカナに限定され、埋め込める情報量も少ない。そこで、本研究ではフォントの拡張を行い、

埋め込み可能な情報の増量を目的とした新たな情報埋め込み手法を提案する。

ここで、情報ハイディングにおいて用いられる用語を次のように定める。文字情報を埋め込み、隠匿することを“埋め込み”という。このときに埋め込まれる文字情報を“埋め込み情報”という。埋め込み情報が埋め込まれるテキストメディアを“カバーテキスト”といい、情報埋め込みがなされた状態のカバーテキストを“ステゴテキスト”という。そして、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、本論文においてステゴテキストから切断加工を施された文字を読み取る行為を“読み取り”という。抽出は、読み取り工程を経て、そこから埋め込み情報を得る行為とし、読み取りと抽出の2つを区別する。

2. 本研究の概要

情報を埋め込むことのできる文字フォントの対象として、既存フォント³⁾のひらがなおよびカタカナに加えて、漢字および英数字を追加したフォントを作成した。既存フォントを拡張したフォントのことを本研究では拡張フォントと呼ぶ。拡張フォントは既存フォントと同じく、10.5ptのゴシック体を基本とし構成されている。原則として、切断加工位置はフォントの上部、切断加工幅を切断対象となる画の太さに対して60%とした。これは、切断加工フォントにおける最適な切断加工位置、切断加工幅を再検討した研究⁴⁾の知見によるものである。また、切断加工は原則的に切断を施す画の筆道に対して垂直に施すが、一部アルファベット大文字の“I”や小文字の“l”のように、従来通り画に対して垂直に切断を施した場合の形状が小文字の“i”のようになり、小文字の“i”との識別が困難となるため、このような文字では斜めに切断を施すこととした。

本研究で作成した拡張フォントを図1に示す。拡張フォントにより、ひらがな、カタカナに加えて漢字および英数字への文書のほぼ全域を情報ハイディングに利用することができるようになる。

A Proposal for Visual Detectable Information Hiding for Printed Documents

[†]Masaru Shimizu, Takehiko Abe, Kazuyoshi Ishii, Kanazawa Institute of Technology

[‡]Katsuaki Hayashi, Yoshihiro Ueda, Industrial Research Institute of Ishikawa

^{††}Haruhiko Kimura, Kanazawa University

Information Hiding 情報ハイディング

図1 拡張フォント

3. 拡張フォントの検証実験

漢字および英数字の追加を行った拡張フォントの有効性を検証するための実験を行った。その方法は、ランダムで切断加工文字が配置されたハードコピー状態の文書を、被験者に文書の頭から目を通してもらい、切断加工が施された文字を“1”、切断加工が施されていない文字を“0”として書き出させた。さらに、ひらがな、カタカナ、漢字ごとに切断加工文字を正しく読み取れた数を集計し、切断加工文字の認識率を求めた。被験者は8人とし、実験に使用した文書中の切断加工文字は、全568文字中45文字とした。

4. 実験結果

ひらがな、カタカナ、漢字、それぞれの認識率を図2に示す。

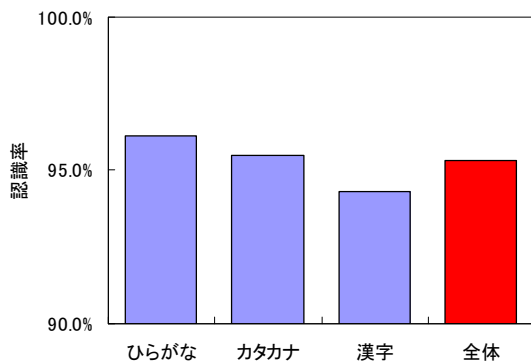


図2 拡張フォントの認識率

漢字の認識率が94.3%と最も低くなった。ひらがなは96.1%、カタカナは95.5%となっており、両者の平均値は95.8%である。これは、既存研究での切断加工フォントの認識率95.4%³⁾と比してみても統計的に有意な差はみられなかった。ここで、漢字の認識率が低くなった理由として、一部の文字の切断加工位置に不具合があったためと考えられる。例えば“隠”という文字では、フォントの左上部に切断加工を施すという原則により、文字左部の“こざとへん”の上部に切断加工を施しているが、切断加工が意図的なものなのかどうかの判断が難しくなっていると考えられる。実験に用いた“隠”のフォントを図3に示す。

“こざとへん”は筆の癖によっては1画目と3画目が接触せず、図3のように、2辺の間に空

白ができる場合がある。そのため、一部の被験者に対してこれが意図的な切断であるという確信を与えることができず、切断加工が施されていないと判断させてしまい、誤認識につながったものと考えられる。このため、こざとへんのように切断加工を施す位置によっては、意図的な切断か自然な切断かが紛らわしいフォントは切断加工位置をずらして、明らかに意図して切断されていると感じられる位置に切断加工を施すように修正した。例として、修正された“隠”のフォントを図4に示す。また、“こざとへん”のほかにも“くにがまえ”や“もんがまえ”などでも同様の修正を施し、切断加工文字の認識率向上を図った。



図3 修正前フォント 図4 修正後フォント

5. おわりに

本研究では、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文書への情報ハイディングを行う既存手法において、従来は埋め込みの対象がひらがなのみであったものを、漢字とカタカナ、英数字にまで拡張した。そして拡張フォントの有効性について検証を行い、その有効性を示した。

今後はより多くの文字情報を効率的に埋め込むことができる手法として、既存のJISコードを用いない方法を提案する。新埋め込み手法により埋め込み情報量の増加と、抽出率の向上を図り、その有効性を検証していく。

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ジェクト技術，森北出版
- 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
- 3) 林克明，上田芳弘，岩田雅士，木村春彦：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，情報処理学会論文誌，Vol. 44, No. 11, pp. 2817-2825 (2003)
- 4) 山本渉，梅原丈宏，林克明，上田芳弘，阿部武彦：ハードコピーを対象とした視覚認識可能な情報付加手法の拡張，平成16年度電気関係学会北陸支部連合大会，F-88 (2004)