Study of an information extraction method using a single-dot pattern method with clear ink

†Yuki Matsue †Kitahiro Kaneda †Keiichi Iwamura and ‡Isao Echizen

†Dept. of Electrical Engineering, Tokyo University of Science 6-3-1 niijuku, Katsushika-ku, Tokyo, 102-0073 Japan
‡National Institute of Informatics 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430 Japan
E-mail: †(matsue, kaneda, iwamura)@sec.ee.kagu.tus.ac.jp, ‡iechizen@nii.ac.jp

ABSTRACT

In this paper, we propose a single dot method [1][2]. This is a nearly invisible, new pattern information embedding method aimed at satisfying quality, capacity, accuracy, and processing speed. On the other hand, clear ink for high-quality professional printing is currently becoming popular. To further decrease the visual effects of a single dot method, we try to apply error-correcting code using convolutional code with the Viterbi algorithm for a single dot method and with clear ink. We were able to extract all of the single-dot effects by using convolutional code.

1. INTRODUCTION

A paper medium is popular as an information transmission means in the electronic society of modern times. On the other hand, illegal printing acts have become easier. Copy-protection technologies for printed materials have not been studied as much as such techniques for electronic media. Therefore, prevention technology of information leaks in the paper medium is more necessary. In this paper, we propose a new protection technology of printed materials using a single-dot pattern method with a special ink.

2. SINGLE-DOT PATTERN METHOD

2.1 Embedding

In the single-dot method, a pattern with a one-pixel dot in the center of an 18×18-pixel tile is defined as “1” and a pattern with no dot in the center is defined as “0”. This defines a single-dot pattern method and is shown in Fig.1.

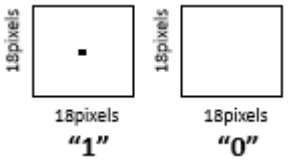


Figure1. “1” and “0” using single-dot

2. 2 Detection

Kitazawa [1] showed the information extraction method that takes into account the scan shift. Before performing information extraction, the pattern indicates any information as "1" in the vertical and horizontal directions of the printed paper, as shown in Fig.2. When performing information extraction, the pixel values around the intersecting point formed by the marker are read. We obtain pixel values around this intersection. If the calculated value in a certain block is larger than the threshold value, this block represents “1", if not this block represents "0". Conventional methods embed dots unobtrusively. Because each dot is embedded in black ink, we are aware of the existence of that dot. We propose a method for the purpose of further reduction of subjective discomfort.

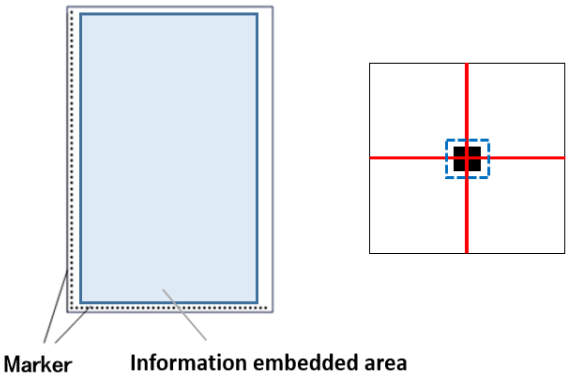


Figure.2 Detection method

3. CLEAR INK

Clear ink (transparent ink) is printed onto photographs or logos. Printed matter by using clear ink is generated, e.g. varnished in the reflection of the light condition. Clear ink can be expressed in terms of gloss, clarity, and solidity that could not be expressed in the color printing of conventional ink. Examples of the use of clear ink include wedding invitations, high-quality photos, and business card.

3.1 SINGLE-DOT PATTERN METHOD USING CLEAR INK

In this study, we propose a new method for embedding in clear ink as a new embedded method of the single-dot method. Although clear ink was used originally in high-quality printing, it is used the characteristic visual discomfort from transparent ink. This proposed method is aimed at overcoming the further reduction in document quality and visual

discomfort. In addition, we improve the information extraction rate by using the convolutional code with the Viterbi algorithm.

It is impossible to embed a dot by the conventional method using clear ink due to the printer hardware. For preliminary experiments, a pattern with a 25-pixel dot in the center of a 51 × 51-pixel tile is defined as "1" and a pattern with no dot in the center is defined as "0".

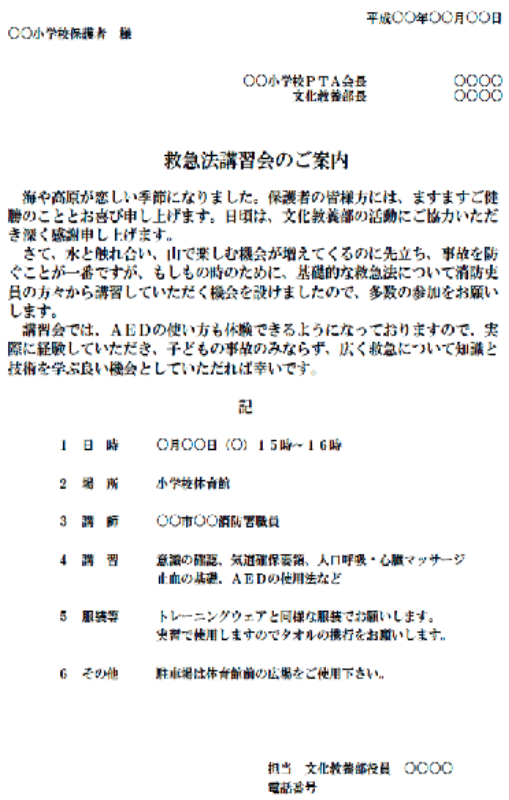
4. EQUIPMENT

(1) Equipment for printing
- •Canon PIXUS PRO-1 inkjet printer
  (Print resolution: 4800×2400dpi)

(2) Equipment for scanning
- •Canon MX923 inkjet printer/scanner
  (Scan resolution: 600dpi)

(3) Paper media
- •Canon photo paper, glossy gold GL-101
  (Canon Corporation)

(4) Printed material used for verification print with characters
- ・http://www.template-sozai.com/template/2333



平成○○年○○月○○日

○○小学校保護者　様

○○小学校ＰＴＡ会長　○○○○
文化教養部長　○○○○

救急法講習会のご案内

海や高原が恋しい季節になりました。保護者の皆様方には、ますますご健勝のこととお喜び申し上げます。日頃は、文化教養部の活動にご協力いただき深く感謝申し上げます。

さて、水と触れ合い、山で楽しむ機会が増えてくるのに先立ち、事故を防ぐことが一番ですが、もしもの時のために、基礎的な救急法について消防吏員の方々から講習していただく機会を設けましたので、多数の参加をお願いします。

講習会では、ＡＥＤの使い方も体験できるようになっておりますので、実際に経験していただき、子どもの事故のみならず、広く救急について知識と技術を学ぶ良い機会としていただければ幸いです。

記

1	日　時	○月○○日（○）１５時～１６時
2	場　所	小学校体育館
3	講　師	○○市○○消防署職員
4	講　習	意識の確認、気道確保要領、人口呼吸・心臓マッサージ 止血の基礎、ＡＥＤの使用法など
5	服装等	トレーニングウェアと同様な服装でお願いします。 実習で使用しますのでタオルの携行をお願いします。
6	その他	駐車場は体育館前の広場をご使用下さい。

担当　文化教養部役員　○○○○
電話番号

Figure3. Template image

(5) Computer and software
- •Intel(R)Core(TM)i7-2600CPU@3.40GHz
- •AdobePhotoshopCC（14.0）
- •Matlab2016

5. EXPERIMENTAL METHOD

This experimental procedure is carried out in the following steps.

(STEP1) Create a random dot pattern image
(STEP2) Print and scan
(STEP3) Image processing
(STEP4) Extraction of marker
(STEP5) Extraction of a random dot
(STEP6) Calculation of error rate
(STEP7) Error-correction simulation

6. 1. EXPERIMENTAL METHOD

(STEP1) Create a random dot pattern image

The image of random dots and marker dots is formed in accordance with an A4 size.

(STEP2) Print and scan

The dot pattern image generated (STEP1) is printed with a clear ink on an A4 document as shown in Fig.3 as a background. Next, this document is scanned at a resolution of 600dpi.

Printing a single-dot pattern using clear ink does not interfere with the visibility of the document content. However, the dots are difficult to extract from the scanned image.

(STEP3) Image processing

In order to recognize printed dots of clear ink easier, we apply a combination of several image processing techniques: minimum brightness, maximum contrast, and gradation reversal. Fig.4and Fig.5 shows that dots become recognizable due to image processing.

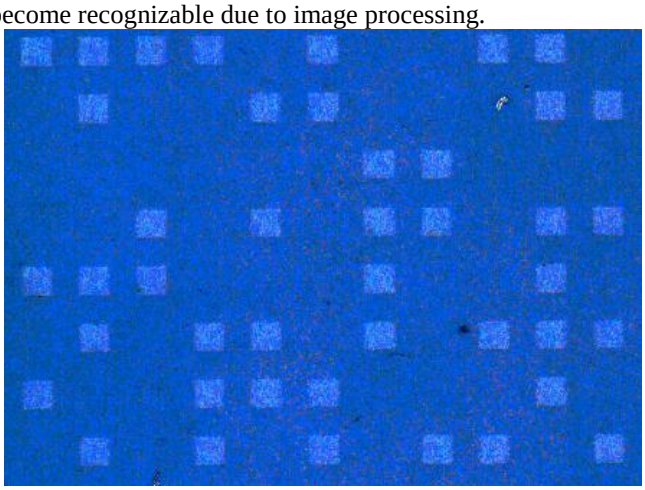


Figure 4. Image processing (gradation reversal)

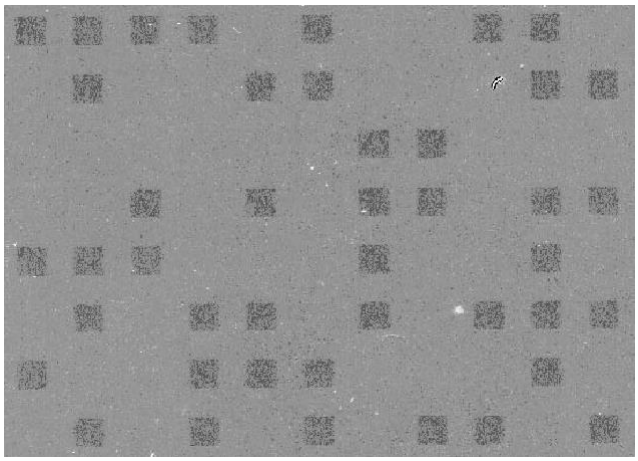


Figure 5. Image processing (gray scaled)

(STEP4) Extraction of marker

A coordinate of a marker of the height and width of scanned image is checked. Pixel values near the intersecting point are calculated. If the calculated value in certain block is larger than the threshold value, this block represents "1", if not this block represents “0".

(STEP5) Extraction of a random dot

When we edit a scanned image to recognize the dots, all the dots on the letters turn white. All the dots on the letters are extracted by mistake. In order to recognize the dots on the letters, we apply a combination of several image processing techniques: maximum brightness and minimum contrast. Fig.6 shows that dots on the letters reach recognizable levels through image processing. Thus, the extraction method is a case analysis of the dot existence and dot non-existence on the letters.

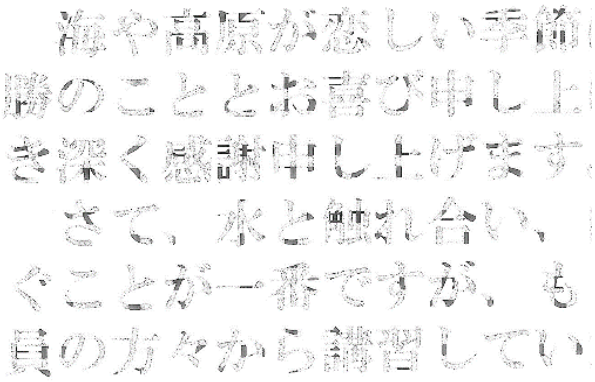


Figure 6. Part of dot on the letters

(STEP6) Calculation of error rate

Extracted dot information in (STEP 5) and an original random dot information defined in (STEP 1) are compared and an error rate is calculated.

(STEP7) Error-correction simulation

The conventional method [1] applies error-correcting code using LDPC (low-density parity-checking code). Because the conventional method has a large capacity for embedded information, LDPC code can store a longer code and therefore perform better. Proposed methods with a small amount of information are not suitable for LDPC code. Therefore, convolutional code, which does not have a restriction code length was used for the proposed method.

Fig.7 shows a process of error-correction by convolutional code. Detailed steps are explained below; the steps will be described in this order.

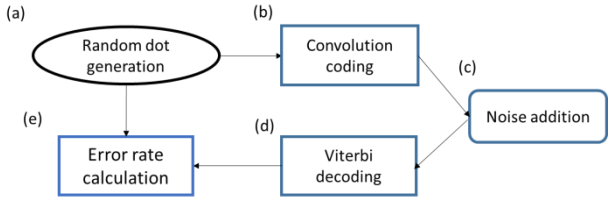


Figure 7. Outline of the simulation.

(a)Random dot generation

A random dot is a dot indicating data "1" or data "0". Total bit length of embedded information is set to 13056 [bit]. Convolution coding encodes the random dot as the input signal. A parameter of a convolutional encoder is indicated below. In addition, this time using the convolution encoder shown in Fig.8.

- Input bit = 1
- Output bit = 2
- Coding rate = 1/2
- Code length =6528 [bit]
- Constraint length = 3

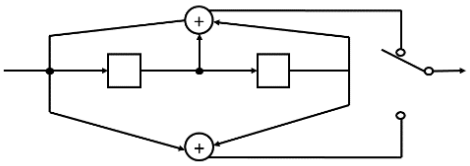


Figure 8. Convolutional encoder

(b) Noise addition

Noise (letters, dust, and dirt) is added to the coded bit sequence obtained in (b). When a single dot is printed and scanned, noise is added to the dots.

(c)Viterbi decoding

The extracted dots in (STEP5) are decoded.

(d)Error rate calculation

By comparing with the decoded bit stream and the coded bit stream, we can obtain the total number of different bits.

6. 2. Results

Table 1 shows the error rate and number of each 1 to 0 and 0 to 1 errors. The conventional method is also shown [1]. Total embedded information content of proposed method and conventional method is 13056[bit] and 98625[bit]. Error 0 to 1 indicates that original information "0" was erroneously recognized as information "1". Error 1 to 0 indicates that original information "1" was erroneously recognized as

information "0". The number of error bits divided by total embedded information content is the error rate.

TABLE 1. Error rate and number of errors

	Total rate	0to1	1to0
Proposed method	988[bit]	79[bit]	909[bit]
	7.567%	0.605%	6.962%
Conventional method	5929[bit]	5478[bit]	451[bit]
	6.011%	5.554%	0.467%

Table 2 shows the calculation of decoded error rate simulation convolutional code. The proposed method is applying error-correcting code using convolutional code. The conventional method applies error-correcting code using LDPC code.

TABLE 2. Simulation results

	Extracted amount	Error rate
Proposed method	12068[bit]	0%
	92.433%	
Conventional method	92696	0%
	93.989%	

6. 3. Discussion

From Table 1, it can be seen that the proposed method could achieve a high extraction rate, which is similar to the conventional method.

In the case of 0 to 1, when a single dot is scanned, noise is generated by dust and dirt adhered to the reading table. In addition, the conventional method has a higher error rate than the proposed method because we erroneously recognized the existence of a point overlapping characters in a block of information “0”.

In the case of 1 to 0, there are two causes. First, it deviated from the center of the intersection of the marker, also information "1", by the shift arising in printing. Second, almost no great difference as compared with the value of the average value of the RGB information has clear ink "1" and there is no clear ink information "0". Therefore, pixel value of the intersection point of the marker is determined to the threshold value "0" by the reflectance of light. In addition, the proposed method has a lot more errors than the conventional method because the extraction method is different from the conventional method. To extract dots, other than on the letters, was a classified case. Therefore, if the information amount of the dot is small, extracted information is recognized erroneously. An example of correctly extracted examples and error extraction shown in Fig. 9.

Figure 9. Extracted example
(left)：Failed example (right)：Successful example

From Table 2, dots decoded by the Viterbi algorithm contain no error. It showed that there was resistance to noise (letters, dust, and dirt).

7. Conclusion

In this paper, we aimed for further reduction of subjective discomfort of conventional method and we used clear ink. The proposed method could obtain a high extraction rate of a similar quality as the conventional method. It was also extracted without error by using a convolutional code with the Viterbi algorithm.

Although this experiment was fixed at 1/2 coding rate, we want to clarify the relationship between the coding rate and the error correction in future experiment. We need to evaluate the robustness for various noises (impact character density, stains, graffiti, cut and paste, copy) and validate resistant.

REFERENCES

[1] H. Kitazawa, K. Kaneda, K. Iwamura , “An improvement of a single-dot method for an information-hiding method by applying an error-correcting code,” The International Conference on Electrical and Electronics Engineering, Clean Energy and Green Computing (EEECEGC2013), United Arab Emirates, December 2013.

[2] K Kaneda, et al. “A study of information hiding performance using simple dot pattern with different tile sizes.” Intelligent Information Hiding and Multimedia Signal Processing, 2008. IIHMSP'08 International Conference on. IEEE, 2008.

[3] A. Varna, S. Rane, and A. Vetro, “Data hiding in hard-copy text documents robust to print, scan and photocopy operations,” in IEEE Intl. Conf. Acoustics, Speech, Sig. Proc. (ICASSP), April 2009.

[4] Hrishikesh C, Shefali S “Printed document watermarking using phase modulation.” In: 2nd International Conference on Emerging Trends in Engineering and Technology (ICETET’09), pp 222–227,2009

[5] K.Kawauchi,”Convolutional code with viterbi algorithm” Triceps Corporation,2008

[6] K.Imura,”PhotoshopCC2014SuperReferenceforWindows&Mac OS” Sotec Corporation,2014