Fast communication

Watermark detection from clustered halftone dots via learned dictionary

Chang-Hwan Son *, Hyunseung Choo

College of Information and Communication Engineering, Sungkyunkwan University, 300 Chunchundong, Jangangu, Suwon 440-746, South Korea

Article info

Article history:
Received 19 December 2013
Received in revised form 27 February 2014
Accepted 11 March 2014
Available online 20 March 2014

Keywords:
Halftoning
Clustered-dot dithering
Hardcopy watermarking
Dictionary learning

Abstract

Modulating the orientation of elliptically clustered dots in each halftone cell enables binary data to be embedded into the clustered halftone dots. In this paper, a new decoding method is proposed for recovering hidden binary data from clustered halftone dots by using learned dictionaries, which are optimized to represent clustered dots with different elliptical shapes. The basic idea is that the reconstruction errors of the clustered dots in a halftone cell are differentiable according to the dictionaries used. The experimental results showed that determining which of the learned dictionaries provides a minimum reconstruction error in a halftone cell can reveal the orientation of the clustered dots and thus indicate the embedded binary data.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

In digital printers, continuous-tone images with 255 discrete levels are converted into halftoned images [1], which are binary images with uniformly distributed black and white dots. The pixel locations of the black and white dots determine whether or not black toner or ink will be printed on a substrate. In other words, binary quantization occurs during printing. Conventional watermarking methods [2,3] try to embed data into the continuous-tone images; however, the embedded data suffers due to the binary quantization during halftoning. To solve this problem, hardcopy watermarking methods [4–6] that directly embed data into the halftoned images have been developed. Since the dot patterns and shapes in the halftoned images can be formed differently depending on which halftoning techniques are used, such as dithering [7], error diffusion [8], or direct binary search [9], the characteristics of these varied techniques have been considered for data hiding [10–12]. In dispersed-dot dithering [10], two types of dither matrices that can produce different halftone patterns are used to embed the binary data into each halftone cell. In direct binary search [11], halftone texture orientations are controlled by changing the shape of the contrast sensitivity function to represent different watermark values. In block-based error diffusion [12], a user-controlled shape and size are provided to embed the information into the hardcopy image. Since these halftoning methods vary the dot density to achieve the tonal rendition, higher spatial resolution can be obtained, but the isolated dot reproduction can be an issue with electrophotographic (laser) printers [13].

Thus, clustered-dot dithering [13] has been mainly used with electrophotographic printers because of its stability and reproducibility. In clustered halftoned images, where tonal rendition is achieved by varying the size of each dot, clustered dots can be generated in various shapes and orientations in order to effectively encode the watermark [14,15]. Recently, Bulan et al. [15] proposed a new data coding method, which uses clustered halftone dots to embed binary data. The basic idea of their method is to modulate the orientation of each clustered dot in a halftone cell to embed binary data. In this paper, we propose a new decoding method for recovering hidden binary data from clustered halftone dots using learned dictionaries, which are optimized to represent clustered dots with different elliptical shapes.

* Corresponding author. Tel.: +82 31 299 4642; fax: +82 31 299 4134.
E-mail addresses: jjdhompy@hotmail.com, changhwan76.son@gmail.com (C.-H. Son).

http://dx.doi.org/10.1016/j.sigpro.2014.03.016
0165-1684/© 2014 Elsevier B.V. All rights reserved.
hiding scheme that embeds binary data into the clustered halftone dots by generating elliptical clustered dots and modulating their orientation. In each halftone cell, the orientation of the elliptically clustered dots is determined by the binary data, '0' or '1', to be embedded. For instance, the elliptical clustered dot is oriented vertically for a bit value of 1 and horizontally for a bit value of 0. In other words, given a binary data sequence, it can be encoded as vertically or horizontally clustered dots forming a watermark. However, the moment-based decoding method for inferring the clustered-dot orientation within each halftone cell cannot provide satisfactory decoding accuracy, where accuracy is evaluated based on the bit error rate (BER) [15]. To improve the decoding accuracy, we propose a new decoding method for learning the dictionaries that optimally represent vertically or horizontally clustered dots. The proposed method reconstructs the clustered dots in each halftone cell via the learned dictionaries to infer the clustered-dot orientation that indicates the embedded '0' or '1' data.

2. Limitations of the conventional moment-based decoding method

Various clustered halftone patterns can be generated according to the comparison between the input image’s pixel values and the dithering matrix’s elements. The conventional method [15] uses moments to estimate the orientation of the clustered halftone dots. However, the moments used can provide an inaccurate estimation. For example, clustered halftone dots, as shown in Fig. 1, are obtained using a dithering mask, providing vertically clustered halftone dots. In this case, the moment value \( \sigma_x \) calculated along the horizontal axis is larger than the moment value \( \sigma_y \) calculated along the vertical axis. Thus, the moment-based decoding method will detect the horizontal orientation from the given clustered halftone dots in Fig. 1. This produces an inaccurate estimation. The primary limitation of the conventional moment-based decoding method is that it cannot decode various types of clustered halftone dots, such as circular, dual, or octa dots [15,16].

3. Proposed decoding method

Dictionaries known as textons or visual codewords are the redundant basis vectors used for sparse linear representation, which has been extensively applied to image denoising, face recognition, texture classification, and inverse halftoning [17–20]. The basis functions of the discrete cosine transform or discrete wavelet transform have been directly used for the dictionaries; however, it has recently been recommended more often that the many natural patches extracted from training image sets be learned for a better dictionary generation [17].

In this paper, dictionaries are used to detect the embedded binary data in clustered halftone dots. The basic idea behind this approach is that the reconstruction errors of the clustered dots in a halftone cell are differentiable according to the dictionaries used. The concept of the proposed method is illustrated in Fig. 2, wherein the learned dictionaries correspond to the basis vectors, which can be linearly combined with the coefficients \( \alpha \) to represent the input halftone patch. It is assumed based on the statistics of natural images [17] that most coefficients \( \alpha \) are zero, i.e., \( \ell_0 \)-norm \( \|\alpha\|_0 \) is less than the constant value, TH. Given the input halftone patch with vertically clustered dots on the leftmost side, dictionaries with vertically clustered dots can more accurately reconstruct the input halftone patch than can dictionaries with horizontally clustered dots. Therefore, two types of learned dictionaries that optimally represent vertically or horizontally clustered dots will be generated to infer the clustered-dot orientation that indicates the embedded '0' or '1' data.

3.1. Dictionary generation for clustered halftone dots

Assuming that there are two dictionaries, \( \{D^H, D^V\} \), that optimally represent the horizontally- and vertically-clustered dots, respectively, the shape of the clustered dots in a halftone cell can be more accurately represented by one of the dictionaries. The dictionary that provides better representation can determine the orientation of the clustered dots in a halftone cell, and thus reveal the embedded binary data.

\[
\min_{\alpha} \|x^j - D^j \alpha\|_2, \quad \|\alpha\|_0 \leq TH
\]

(1)

In the preceding equation, the column vector, \( x^j \), contains the values of the \( j \)th patch that was extracted from an input halftoned image. In this paper, it is assumed that the patch is identical in size to the dithering matrix that was used to generate the clustered halftone dots, \( \alpha \) is the predicted representation column vector of \( x^j \) resulting from the gradient pursuit algorithm [21], which is the fast

![Fig. 1. Drawback of the moment-based decoding method.](image)

![Fig. 2. Concept of the proposed decoding method.](image)
version of the orthogonal matching pursuit [22], and its sparsity is controlled by the constant value, $TH$. The preceding equation shows that either $D^H\mathbf{a}'$ or $D^V\mathbf{a}'$ more accurately represents $\mathbf{x}'$. If $D^H\mathbf{a}'$ is closer to $\mathbf{x}'$ than $D^V\mathbf{a}'$, $\mathbf{x}'$ will represent horizontally shaped clustered dots. Otherwise, $\mathbf{x}'$ will represent vertically shaped clustered dots. Therefore, the embedded binary data that is encoded by the vertically or horizontally clustered dots in a halftone cell can be estimated by evaluating the reconstruction errors, as shown in (1). If $j=V$, the hidden binary data will be ‘1’ and if $j=H$, ‘0’. The discrimination between the reconstruction errors with two types of dictionaries, $D^H$ and $D^V$, can be influenced by the sparsity, i.e., the value of $TH$, which will be checked in the experimental results.

In this paper, the two types of dictionaries, $(D^H, D^V)$, were obtained by minimizing the following cost function:

$$\min_{D^H, D^V} \sum_{k} \|A(k)\|_0 \text{ subject to } \|X - D^H X\|_2^2 \leq \varepsilon$$

where $A(k)$ is the column vector of the matrix $A$ which indicates the representation column vector corresponding to $X(k)$, which is filled with the values of the extracted halftoned patch from the training halftoned images. To learn $D^H$, a dithering matrix for horizontally clustered dots is first applied to the continuous-toned training images in order to generate the corresponding halftoned images. Next, the extracted patches from the halftoned images are reshaped with the column vectors and then inserted into $X(k)$. The K-SVD algorithm [22] is used to minimize (2). Another dithering matrix for vertically clustered dots can be used to learn $D^V$. In this paper, a modified Pellar threshold function [15] was used to generate the dithering matrices for the vertically and horizontally clustered dots. The size of dictionary $D^V$ was fixed at 36 × 1024 and the column size of the dictionary $X$ was 215,442 (representing the total number of halftoned patches in the dictionary learning).

### 3.2. Estimating embedded data via learned dictionaries

Given the learned dictionaries $(D^H, D^V)$, the procedure for estimating the hidden binary sequence from an input halftoned image is as follows.

- **Step 1:** Put a sliding window on the halftoned input image and extract the patch covered by the sliding window.
- **Step 2:** Reshape the extracted patch as the column vector $\mathbf{x}'$, and then solve (1). If $j=V$, the hidden binary data is ‘1’ and if $j=H$, ‘0’. Store the currently estimated binary data.
- **Step 3:** Move the sliding window to the next position without any overlapping along the raster scanning order. Unless the sliding window reaches the last position, extract the patch covered by the current sliding window and go back to Step 2. Otherwise, stop the algorithm.

In the procedure, the size of the sliding window should be the same as that of the dithering matrix. Also, the halftone input image should already contain the embedded binary sequence with the data encoding method [15].

### 4. Experimental results

#### 4.1. Data decoding from vertical and horizontal clustered dots

Fig. 3 shows examples of the dithering matrices for generating circular clustered dots, vertical clustered dots, and horizontal clustered dots. The round dot screen function [16] and the modified Pellar threshold function [15] were used with clockwise ordering to develop the dithering matrices. The three halftoned images that resulted from the use of the corresponding dithering matrices are given in Fig. 4. To check the clustered dots, the red rectangles at the top left are magnified and then placed at the bottom right. In the leftmost image, the circular clustered dots can be observed, whereas in the rightmost image, the horizontal clustered dots can be found. The visualized learned dictionaries that are responsible for the representation of horizontally clustered dots $(D^H)$ and vertically clustered dots $(D^V)$ are shown in Fig. 5. To visualize $D^H$ and $D^V$, each of their column vectors is reshaped as a block and then normalized to [0–255]. In Fig. 5, each 6 × 6 patch is the visualized dictionary. For example, the black patches at the top left represent the constant patch. Note that the left image tends to have horizontally clustered dots, whereas the right image tends to have vertically clustered dots. More details on how to visualize the dictionary can be found in [17,22].

Fig. 6(a) shows the original watermark with the binary data, i.e., 0 or 255, to be embedded into the halftoned image. For each halftone cell, one of the two dithering matrices, as shown in the middle and right images in Fig. 3, is selected to generate the vertical or horizontal clustered dots, using the binary data in the original watermark. From the result, the original watermark can be reconstructed as either vertical or horizontal clustered dots, as shown at the bottom right image in Fig. 7(a). If the original

![Fig. 3. 6 × 6 dithering matrices for generating circular clustered dots, vertical clustered dots, and horizontal clustered dots, respectively (left to right).](image-url)
watermark is directly embedded into the halftoned image, the watermark can be visually detected. To prevent this, the binary sequence of the original watermark was rearranged according to the generated random permutation of the integers. The same approach was adopted in [23], see Section 3 for further detail. The original watermark was estimated from Fig. 7(a) with both the conventional and proposed methods, by which Fig. 6 (b) and (c) was obtained, respectively. As these images indicate, with the proposed method, the estimated watermark is similar to the original watermark. In contrast, the conventional moment-based detection method [15] yields a less accurate watermark.

To test the proposed algorithm on a real printed image, the digital halftone image in Fig. 7(a) was printed with an HP laser P2055dn printer at the 600 dpi mode and then scanned at the same mode. However, the print-scan path could generate unwanted geometric distortion, so the image was aligned via perspective transformation using the four corner points of the scanned image. In addition,
the scanned image was not a perfect binary image, so simple thresholding [24] was conducted based on the halftone cell unit. Fig. 7(b) shows the scanned image, in which it can be observed that the clustered dots are similar to those in the digital halftoned image in Fig. 7(a). Due to the print-scan path, including the geometric distortion and the physical dot gain [16], the accuracy of the estimated watermark decreased, as shown in Fig. 6(d).

For the quantitative evaluation, BER [15] was tested on six 1002 × 1002 testing images, as shown in Fig. 8. The original 167 × 167 watermark that is shown in Fig. 6(a) was embedded into the six testing images. The BERs were calculated by dividing the number of the inaccurately estimated binary data with the total binary data number. Table 1 shows the BERs of the six testing images, according to the change in the $TH$, as shown in (1). In Table 1, $TH = 3$ provides the smallest averaged BER. If the value of $TH$ increases, the reconstruction accuracy of $D^H\alpha^i$ or $D^V\alpha^i$ can be improved due to the increased number of the column vectors of $D^i$ with the nonzero representation coefficients in $\alpha^i$. This can reduce the differentiation between $D^H\alpha^i$ and $D^V\alpha^i$, thus degrading the BER. If $TH = 2$, the BER can be decreased because $D^H\alpha^i$ or $D^V\alpha^i$ might be inaccurately estimated with the increased number of column vectors.
with zero representation coefficients. The BERs for the real printed images are given in the brackets. Even if the BERs of the proposed method are increased due to the print-scan path, the proposed method would still be better than the conventional method that uses the moment. To ensure that the proposed method is robust to the print-scan path, more accurate geometric correction, dot gain modeling, or an advanced thresholding technique will be considered in future studies.

Table 2 shows the decoding times for the proposed and conventional methods, which were coded in C++ on the Windows system and then simulated on a desktop PC with i3-2012 3.3 GHz. As shown in Table 2, the computation time of the proposed method is 16.416 s for an image size of 1002 x 1002, which corresponds to a printed size of 23 cm x 23 cm. This means that the proposed method can classify various types of clustered halftone textures, and different clustered halftone dots can be combined to generate image barcodes [12]. In contrast, the conventional moment-based method [15] cannot decode these types of clustered halftone dots. As a reference, in Fig. 9, each subfigure was cropped from the halftoned version of the ‘Sunflower’ testing image. The learned dictionaries that are responsible for the sparse representation of the dual, octa, and circular dots are given in Fig. 10.

5. Discussion

5.1. Local geometric correction problem of the scanned image

The perspective transformation cannot correct local geometric distortion, especially when the printed image is large. Many algorithms [25] have been developed to be robust to geometric distortion based on the use of the mesh model, invariant subspace, registration pattern, or feature points; however, those methods target continuous-tone images, not halftone images. Thus, these methods cannot be directly applied to clustered halftone images. Global geometric distortion, e.g., rotation, can be estimated based on the periodicity of the clustered halftone dots [26]; however, local geometric distortion is a more challenging problem. One solution is to use inverse halftoning [27] to generate a continuous-tone image from a scanned halftoned image, and then the mesh model correction approach [28] can be applied to the reconstructed continuous-tone image with the assumption that the original image is known. Another solution is to insert a registration pattern into the clustered halftone image. The proposed decoding method can classify various types of clustered halftone dots, as shown in Fig. 9, and thus dual, octa, or circular dots can be used as registration patterns, which can be inserted into the clustered halftone image with vertical and horizontal clustered dots. Estimating the positions of the inserted registration patterns via learned dictionaries enables correction of local geometric distortion. The method based on decision directed synchronization briefly introduced in [15] can also be considered for correction of local geometric distortion.

5.2. Learning issues

To classify the halftone patterns generated with the orientation modulation integrated into direct binary search, the Bayes classifier was used with feature vectors obtained via LMS filtering [11]. This method sought to establish a feature descriptor able to discriminate the orientations of halftone patches, whereas the proposed method sought to learn the dictionaries, i.e., the basis

Table 1
BER evaluation.

<table>
<thead>
<tr>
<th>Images</th>
<th>Methods</th>
<th>Proposed method</th>
<th>Conventional method</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Text</td>
<td>0.0105 (0.0893) 0.0135 (0.1042) 0.0236 (0.1029)</td>
<td>0.1934 (0.3394)</td>
<td></td>
</tr>
<tr>
<td>Butterfly</td>
<td>0.1619 (0.0911) 0.0260 (0.0806) 0.0167 (0.0109)</td>
<td>0.1619 (0.2317)</td>
<td></td>
</tr>
<tr>
<td>House</td>
<td>0.0398 (0.0136) 0.0250 (0.0120) 0.0277 (0.0120)</td>
<td>0.0109 (0.1703)</td>
<td></td>
</tr>
<tr>
<td>Sunflower</td>
<td>0.0213 (0.0568) 0.0266 (0.0806) 0.0180 (0.0849)</td>
<td>0.1109 (0.3230)</td>
<td></td>
</tr>
<tr>
<td>Logo</td>
<td>0.0046 (0.0023) 0.0050 (0.1220) 0.0058 (0.1703)</td>
<td>0.0849 (0.3230)</td>
<td></td>
</tr>
<tr>
<td>Golf field</td>
<td>0.0486 (0.0386) 0.0581 (0.1330) 0.0604 (0.1703)</td>
<td>0.2197 (0.3230)</td>
<td></td>
</tr>
<tr>
<td>TH</td>
<td>2 3 6 9</td>
<td>16.416 (s)</td>
<td>0.01 (s)</td>
</tr>
</tbody>
</table>

4.2. Application of the proposed method to other types of clustered halftone dots

The proposed decoding method is not limited to vertically or horizontally clustered halftone dots; it can also be applied to various types of clustered halftone dots, such as dual, octa, or circular dots [16], as shown in Fig. 9. This means that the proposed method can classify various types of clustered halftone textures, and different clustered halftone dots can be combined to generate image barcodes [12].
vectors that optimally represent vertically or horizontally clustered halftone patches. The sparse constraint of the learned dictionaries can classify the clustered halftone dots. Thus, the learning technique used in [11] differs from the dictionary learning method. The goal of this paper is not to establish the difference between these learning techniques [11,22]. The contribution of this paper is the determination that clustered halftone dots can be sparsely represented by a linear combination of the learned dictionary and the coefficients, and that reconstruction errors with learned dictionaries can be used for data decoding. However, since the feature descriptor used in [11] can classify the halftone patches, the intensity values, i.e., 0 or 255, of the halftoned patch in (1) may be replaced by the feature vectors obtained via LMS filtering [11] for the dictionary learning. The performance of the proposed method depends on how sparsely the dictionaries can be learned with higher discriminating power; thus, a feature descriptor for upgrading dictionary learning can be considered.

6. Conclusions

A new decoding method for learning the dictionaries that optimally represent vertically or horizontally clustered dots was proposed to estimate the embedded binary data into clustered halftone dots. The experimental results showed that determining which of the learned dictionaries provides a minimum reconstruction error in a halftone cell can reveal the orientation of the clustered dots and thus indicate the embedded binary data. Moreover, the proposed decoding method can also be used to classify clustered halftone textures.

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2013R1A1A061165 and NRF-2010-0020210)

References