Engineering Applications of Artificial Intelligence 50 (2016) 245–255

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Deep neural network for halftone image classification based on sparse auto-encoder

Yan Zhang [a,b], Erhu Zhang [a,c,*], Wanjun Chen [a,c]

[a] Faculty of Mechanical and Precision Instrumental Engineering, Xi'an University of Technology, Xi'an 710048, China
[b] Department of Border-control and Immigration, The Armed Police Academy, Langfang 065000, China
[c] Faculty of Printing, Packing Engineering and Digital Media Technology, Xi'an University of Technology, Xi'an 710048, China

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form
14 January 2016
Accepted 20 January 2016

Keywords:
Halftone image classification
Sparse auto-encoder
Effective patch extraction
Majority voting

ABSTRACT

To restore high quality continuous tone images from each class of halftone images, halftone image fine classification is the key problem. In this paper, a novel feature learning method is proposed for classifying 14 kinds of halftone images produced by the most well-known halftoning algorithms. This study employs the stacked sparse auto-encoders (SAE) trained with unsupervised learning for extracting features of halftone images, and then uses softmax regression with supervised learning for fine-tuning the deep neural network and classifying halftone images. In order to reduce the run-time of deep neural network and improve the image correct classification rate, we propose an effective patch extraction method for testing halftone images by measuring the mean and variance of local entropy in a patch. Halftone image classification is determined by the classification results of all effective patches inside an image via majority voting (MV). The experimental results demonstrate that our proposed method achieves an average correct classification rate (ACCR) of over 99.44% for 14 kinds of halftone images on two public image sets. Compared with state-of-the-art LMS–Bayes and M_{10}–ML methods, the proposed SAE-MV method can distinguish the most categories of halftone images and achieve competitive ACCR, meanwhile, demonstrate better generalization performance.

1. Introduction

Digital halftoning is the rendition technique of continuous-tone images or graphics on a bilevel output device, which can approximate the original continuous-tone image when viewed from a distance based on the low-pass characteristic of the human visual system. This technique is widely used in the printing industry, desktop publishing systems, light emitting diode (LED) display and image compression. There are many well-known halftoning methods, such as ordered dither (Ulichney, 1986, 1987; Bayer, 1973), error diffusion (Floyd and Steinberg, 1976; Jarvis and Roberts, 1976; Image Dithering, 2012), dot diffusion (Ulichney, 1988; Knuth, 1987; Mitsa and Parker, 1991), iterative processing (Analoui and Allebach, 1992) and LUT (Mege and Vaidyanathan, 2000).

With the widespread use of the digital halftoning, a large number of digital halftone images need to be restored to continuous tone images for further processing such as image zooming, rotation, feature extraction, segmentation, and compression. Inverse halftoning is developed for this purpose as a technique of reconstructing a high quality continuous tone image from its halftoned version. Much research has been focused on inverse halftoning for digital image restoration since 1990s. There exist several inverse halftoning methods, including filtering (Kim et al., 1995; Chang et al., 2001; Xiong et al., 1999), optimization estimation (Stevenson, 1997; Liu et al., 2011), vector quantization (Lai and Yen, 1998) and machine learning (Mege and Vaidyanathan, 2001; Son and Choo, 2014) methods. Kim et al. (1995) developed a binary permutation nonlinear filter to reconstruct the order dithering halftone images. When using this filter to inverse error diffusion halftone images, the boundary conditions were not valid due to inherently different halftoning methods between the order dithering and the error diffusion. Chang et al. (2001) used the least-mean-square (LMS) adaptive filter to obtain the optimal mask shapes only on the three categories of halftone images, and then reconstructed the gray-level value by the minimum mean square error. Xiong et al. (1999) proposed a blind inverse halftoning using wavelets, moreover, the experimental results showed the better inverse halftoning performance could be achieved by using any priori knowledge of error diffusion kernel. The optimization estimation inverse halftoning methods proposed by both

* Corresponding author at: Faculty of Mechanical and Precision Instrumental Engineering, Xi'an University of Technology, Xi'an 710048, China.
E-mail addresses: zhangyanxq@163.com (Y. Zhang), eh-zhang@xaut.edu.cn (E. Zhang), wjchen@xaut.edu.cn (W. Chen).

http://dx.doi.org/10.1016/j.engappai.2016.01.032

Stevenson (1997) and Liu et al. (2011) could restore high quality continuous tone images, but their methods need to obtain the halftone image parameters. The machine learning inverse halftoning methods were proposed by Mege and Vaidyanathan (2001), Son and Choo (2014), which recreated good quality image with high efficiency, but these methods were only suitable for a certain type of images. As can be seen from the above mentioned methods, we can conclude: (1) A general purpose inverse halftoning scheme is difficult to restore the high quality visual images, since different halftone patterns have many discrepancies regarding their features such as dot directivity, dot distribution and pattern periodicity. (2) Given the type of halftone pattern and prior parameters of halftone image, an optimal reconstruction can be achieved. Therefore, halftone image classification is the crucial process in inverse halftoning for restoring good quality continuous tone images.

1.1. Related work

So far only several halftone image classification methods are proposed. In 1998, Chang and Yu (1998) proposed a classification algorithm using enhanced 1-D correlation to extract the features of halftone image and classifying these features by back-propagation (BP) neural network. Their method classified only four types of halftone images produced by clustered-dot ordered dithering, dispersed-dot ordered dithering, constrained average, and error diffusion. Kong et al. (2011) used enhanced 1-D correlation, gray level co-occurrence matrix and gray run-length matrix to extract the periodic and texture features of halftone image and then employed multistage decision to classify halftone images into nine categories. Liu et al. (2011) separated halftone images into nine classes according to extracted features in Fourier spectrum by using naive Bayes classifier. Wen et al. (2014) sought the optimum class feature matrices by minimizing the total square error according to the characteristic of error diffusion filters and then used maximum likelihood to classify six types of ED halftone images. However, the features of halftone images extracted by different methods are only adapted to a certain class of halftone images, such that the method proposed by Wen et al. (2014) is suitable for error diffusion. To date there is no a general method to extract halftone image features from different classes, as well as the feature extraction from the same class but with different kernels or templates. The difficulty lies in non-unified distribution of dot features presented among various classes of halftone images and almost inconspicuous differences shown in halftone images produced the same class halftoning with different halftoning patterns or diffusion filters. Thus, for real application in inverse halftoning techniques, it is necessary to develop a general classification mechanism for various halftone images.

1.2. Main contributions

To address the issue of intrinsic feature extraction, we study the classification mechanism for 14 halftone image categories. The main contributions of this paper are as follows: (1) We propose a novel halftone image classification scheme adopting deep neural network based on sparse auto-encoder. By unsupervised learning, the stable and universal features of different halftone images are extracted to efficiently represent 14 classes of halftone images for accurate classification. (2) We propose image partition and efficient image patch extraction methods to train the deep neural network and classify images for reducing the execution time and improving the classification accuracy. (3) The experimental results demonstrate that the proposed method achieves excellent accuracy for 14 classes of halftone images produced by two types of cluster-dot order dither matrices, two types of dispersed-dot order dither matrices, three types of dot diffusion matrices, direct binary search and six types of error diffusion filters. To the best of our knowledge this is first work to adopt the deep learning network mechanism to address the problem of halftone image classification. Moreover, the proposed algorithm is suitable for the most categories of images and provides much better generalization performance than the former classification method.

The rest of the paper is organized as follows: In Section 2, we introduce four categories of the most well-known halftoning methods: the ordered dither, the error diffusion, the dot-diffusion and direct binary search, and corresponding halftone image features. Section 3 gives an overview of the proposed classification mechanism. Section 4 details the deep neural network of classification mechanism and presents the effective image patch extraction method and majority voting for halftone image classification. Section 5 presents the experiment results and Section 6 concludes this paper.

2. Digital halftoning and halftone image features

2.1. Digital halftoning methods

Digital halftoning is a representation technique to transform a continuous tone image into a binary image. Among all advanced halftoning algorithms, four categories: ordered dithering (OD), error diffusion (ED), dot diffusion (DD) and direct binary search (DBS) are the most popular in printing industry and desk publishing systems at present.

Ordered dithering is a method that provides a fixed pattern to indicate the order of turning pixels within a selected screen. In this method the threshold matrix is tiled over the image using periodic replication. According to the arrangement of the threshold values, the ordered dithering methods are mainly divided into two groups, clustered dot dithering (Ulichney, 1986, 1987) and dispersed dot dithering (Bayer, 1973). In clustered dot dithering the consecutive thresholds are located in spatial proximity. For a constant threshold halftoning patch, this method turns pixels on that are adjacent to one another, forming a cluster, as listed Clu4 and Clu8 in Table 1. In the dispersed dot dithering method, the threshold matrices are arranged in a way that the values of threshold grow separately. The popular methods, Disp8 and Bayer, are listed in Table 1.

Another class of halftoning method is error diffusion. In this method, the pixels are quantized in a specific order, such as raster ordering and peano scan, and the quantization error of the current input pixel is transferred to the latter input pixels by using the linearly filter in order to keep the local average intensity of the binary image closing to the original continuous tone image. A schematic diagram of error diffusion method is given in Fig. 1. The input pixel $f(i,j)$ is modified by adding certain past quantization errors, producing a modified input $\tilde{f}(i,j)$. This pixel is then converted to a binary value by the quantizer Q, using some threshold T. The error $e(i,j)$ of quantizing of the current pixel is diffused to latter pixels by means of a two-dimensional weighting filter H, known as the diffusion filter. The six classical error diffusion algorithms are listed in Table 1 adopting different diffusion filters (Floyd and Steinberg, 1976; Jarvis and Roberts, 1976; Image Dithering, 2012; Ulichney, 1988).

Dot diffusion technique is a compromise between ordered dithering and error diffusion, which not only enhances the restored image visual quality but offers parallelism processing. Knuth (1987) introduced firstly the dot diffusion halftoning by using a class matrix to retain the similar quality as for error diffusion. Based on this, Mese and Vaidyanathan (2000) optimized