Signal Processing: *Image Communication* 29 (2014) 1181–1196

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Robust watermarking against print and scan attack through efficient modeling algorithm

S. Hamid Amiri, Mansour Jamzad*

Department of Computer Engineering, Sharif University of Technology, Tehran, Iran

ARTICLE INFO

Article history:
Received 8 January 2013
Received in revised form
21 July 2014
Accepted 21 July 2014
Available online 7 August 2014

Keywords:
Image watermarking
Print and scan attack
Modeling
Robustness
Image Quality

ABSTRACT

This article proposes a blind discrete wavelet transform-discrete cosine transform (DWT-DCT) composite watermarking scheme that is robust against print and scan distortions. First, two-dimensional DWT is applied to the original image to obtain the mid-frequency subbands. Then, a one-dimensional DCT is applied to the selected mid-frequency subbands to extract the final coefficients for embedding the watermark. To specify watermarking parameters, we utilize a Genetic Algorithm to achieve a predefined image quality after watermark insertion. Suitable locations for watermarking are determined by analyzing the effect of a modeling algorithm. This model simulates noise and nonlinear attacks in printers and scanners through noise estimation and system identification methods. The experimental results demonstrate that the proposed algorithm has a high robustness against print and scan attack such that its robustness is higher than related watermarking algorithms.

1. Introduction

Digital watermarking is a branch of information hiding that embeds a watermark into a multimedia object for several reasons like copyright protection, content authentication and tamper detection. A watermark can be considered as a kind of signature that reveals the ownership of a multimedia object [1]. The information of this signature is generated using a secret key and must be imperceptibly inserted into the host such that it remains robust against a variety of intentional or unintentional attacks.

There are three main requirements for watermarking algorithms: fidelity, robustness and capacity [1,2]. Fidelity refers to image quality after watermarking while robustness means that the embedded watermark must be detected after applying some attacks to the watermarked image. And, capacity indicates the maximum number of watermark bits that can be embedded such that the visual quality of watermarked image remains acceptable. These requirements compete with each other such that if one is improved, the other two might be reduced [3].

Digital images can be easily converted from the printed version to digital form and vice versa. During the process of printing and scanning some unintentional noise (that is due to the hardware of scanners or printers) are added to the output image. This process, known as print and scan attack, could disturb watermark extraction. Therefore it is necessary to design a robust watermarking method against print and scan operations.

There are several methods in the literature for robust image watermarking against print and scan attack. Dajun and Qibin [4] proposed a watermarking scheme that divides an image into smooth or texture blocks. Based on the block type, two different embedding methods in DFT and spatial domains are applied. Keskinarkaus et al. [5] used a multilevel watermarking method, in which they invert the geometrical attacks of operations using a grid

* Corresponding author. Tel.: +98 21 66166618; fax: +98 21 66019246.
E-mail addresses: s_amiri@ce.sharif.edu (S.H. Amiri), jamzad@sharif.edu (M. Jamzad).

http://dx.doi.org/10.1016/j.image.2014.07.004
0923-5965/© 2014 Elsevier B.V. All rights reserved.

structure proposed in [6]. Also, the watermark information is inserted into the wavelet domain using the additive spread spectrum techniques.

Chen et al. [7] proposed a technique that considers copy attack and embed two watermarks into the spatial domain to achieve the robustness against print and generation copy. Solanki et al. [8] have proposed two methods for print and scan-resilient data hiding. The first method called SELF hides data in the magnitude of selected low-frequency DFT domain whereas the second uses phase spectrum of the image for data hiding. These methods provide high capacity for watermarking but they are semi-blind algorithms. Pramila et al. [9] suggested a robust watermarking algorithm in multiple domains. In this algorithm, the geometric transformations such as rotation, scaling and translation are detected by inserting templates into DFT and spatial domains. After inserting these templates, the watermark is embedded into wavelet domain.

Kang et al. [10] utilized uniform log-polar mapping (ULPM) to reach robustness against both geometric distortion and print and scan process. In the embedding stage of this approach, an informative watermark is combined with a tracking sequence to form the hidden data. The tracking sequence is used to resynchronize the informative watermark at the extraction process. The hidden data is then inserted into DFT coefficients of the host image according to its log-polar mapping. Keskinarkaus et al. [11] proposed another approach that forms a periodic pattern from the watermark bits and embeds the pattern into image blocks. The authors also suggested a different approach for watermarking in [12] in which watermark bits are split into different segments and then each segment is embedded in a coordinate system defined by two salient feature points. These points provide robustness against geometric transformations. In addition to the above approaches, there are several methods for watermarking in halftone images [13–16] aiming to prevent forgery in security documents such as ID cards and expensive tickets.

Designing a robust watermarking algorithm needs to study the influence of print and scan attack on the images. However, the complex nature of the attack makes it difficult to understand all aspects of changes imposed on printed and scanned images. Until now, different models have been proposed for print and scan attack. The model in [17] separates distortions into systematic and operational types. The source of systematic distortions is the structure of printers and scanners while the operational type is from the operator itself [17]. Solanki et al. [18] divide operations into three broad categories: geometric transformations, nonlinear effects and colored noise. Then, the effect of each part on the DFT coefficients has been studied.

In this paper, to develop a robust watermarking algorithm, we propose a modeling algorithm that simulates noise addition and nonlinear effects of print and scan attack. Indeed, this attack includes various types of distortions and, therefore, in spite of other attacks such as noise addition and JPEG, it is not straightforward to analyze the influence of print and scan attack on the coefficients of transform domain used for watermarking. Using an accurate model, we can efficiently specify the coefficients that slightly change by print and scan distortions. By embedding the watermark into these coefficients, we achieve a robust watermarking algorithm.

The proposed modeling algorithm statistically estimates the noise of printers and scanners imposed on the digital document by determining the probability distortion of noise and its parameters based on the extracted samples from real printed and scanned images. Furthermore, to analyze nonlinear effects of the attack, we consider multiple nonlinear systems in the model such that each system corresponds to a class of images. To determine the structure of each system, we propose an identification method based on some samples extracted from printed and scanned images in the corresponding class of the system.

Besides the modeling algorithm, we propose a robust watermarking algorithm using the combination of DWT and 1-D DCT. After applying wavelet transform to the input image, we use the proposed model to select the subbands whose coefficients have taken minor effects from print and scan distortions. Using the model, we can also find suitable coefficients in the selected subbands for robust watermarking. To determine these coefficients, we propose a method that applies 1-D DCT to the selected subbands and extracts two sequences for embedding the watermark. Indeed, these sequences provide indirect access to suitable coefficients in the DWT domain for watermarking. An important property of the embedding algorithm is that it achieves robustness with small variations in the extracted sequences. In this way, the quality of the watermarked image is maintained.

The outline of this paper is as follows. In Section 2, we focus on the modeling algorithm for print and scan operations by studying the distortion of attack. The watermarking algorithm is proposed in Section 3 which illustrates the implementation of watermark embedding and extraction stages. Because one part of print and scan distortions are geometric transformations, we present an approach in Section 4 to undo the rotation and scaling attacks. Experimental results are presented in Section 5 and the conclusion remarks are given in Section 6.

2. Modeling the print and scan operations

We present a model for print and scan operations by identifying a system that simulates distortions of print and scan operations on an input image. In this degradation model, we divide print and scan distortions into three categories [19]:

- *Noise addition*: Different devices in printers and scanners such as laser source, lens and CCD sensors are the main sources of some noises that degrade the quality of the output digital image.
- *Nonlinear effects*: Some operations such as gamma tweaking, dot gain and digital half-toning in the printers and gamma correction during the scanning process are considered as a nonlinear transformation that change the values of pixels in the scanned image.
- *Geometric transformations*: Printed image could be rotated, cropped or scaled during scanning. These transformations belong to the operational distortions that are caused by the