Signal Processing: *Image Communication* 39 (2015) 1–16

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Low-complexity and robust comic fingerprint method for comic identification

Taegeun Oh, Nakyeon Choi, Doyoung Kim, Sanghoon Lee *

Department of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Republic of Korea

ARTICLE INFO

Article history:
Received 10 October 2014
Received in revised form
16 July 2015
Accepted 19 July 2015
Available online 8 August 2015

Keywords:
Content-based comic identification
Comic fingerprinting
Global feature-based method
Comic copy detection

ABSTRACT

Copyright infringement has emerged as a significant issue in the growth of the e-book/comics market on account of illegal copying and distribution. It is therefore important to develop automated comic book identification techniques to prevent such problems. Fingerprinting methods have been typically used for multimedia identification; however previous fingerprinting methods are unsuitable for the identification of comic books, which may include several types of distortions, including geometric changes. In this paper, a new comic fingerprinting method is proposed based on a comparison of the average pixel intensity of sub-images, which are called pairwise patterns. In particular, for robust identification against such geometric distortions, circular patterns are newly proposed and evaluated after constructing a comic fingerprint database. The intra- and inter-distances of the features in a fingerprint are calculated to demonstrate its pairwise independence and robustness against various distortions. The result shows that the proposed fingerprinting method is more robust against various distortions, especially for rotational distortion, than those in previous methods. Moreover, owing to its low complexity, the method has potential advantages for commercial applications in real-time.[1]

1. Introduction

In recent years, hand-held devices such as smartphones, tablets, iPads and Kindle Fires have been leading the rapid spread of various forms of multimedia entertainment through audio, video and e-book/comics [1,2]. The development of portable devices continues to intensify the demand for readable and visual contents instead of using hard-copy versions of books. In addition, the growth of online file-sharing services based on web storage has made it possible for users to much more easily share their data with others. This has resulted in the increased conversion of paper books/comics to electronic versions. However, copyright violations from the illegal copying and distributing of e-book contents have become an important public issue, particularly for comics. To prevent this problem, it is necessary to develop an effective, less computationally complex technique of identifying illegally copied versions from originals.

To this end, content-based fingerprinting has become of interest to researchers and is receiving an increasing amount of attention [3,4]. The techniques generally assume that multimedia contents can be summarized using different features or fingerprints, which enables each issue to be uniquely characterized. Therefore, fingerprinting technology is exploited to identify unknown

* Corresponding author. Tel.: +82 2 2123 2767; fax: +82 2 313 2879.
E-mail addresses: tgoh27@yonsei.ac.kr (T. Oh), luckypensil@yonsei.ac.kr (N. Choi), tnyffx@yonsei.ac.kr (D. Kim), slee@yonsei.ac.kr (S. Lee).

[1] This research project was supported by Ministry of Culture, Sports and Tourism (MCST) and from Korea Copyright Commission in 2015. (2013-book_scan-9500).

http://dx.doi.org/10.1016/j.image.2015.07.009
0923-5965/

contents by measuring the distance between the query fingerprint and the fingerprints of all contents in the database. Content identification and copyright protection for file-sharing services are good examples of commercial applications through the utilization of fingerprinting [1]. For the legitimate distribution of web-based contents, a variety of multimedia fingerprinting methods for audio [5], image [6,7] and video [8–10] have been proposed; however, in the case of comic books, the research has only started recently [11].

Previous research on scanned comic books has mostly aimed to automatically extract the frames and text balloons [12]. The purpose of these studies has predominantly focused on how to display the comics to fit into the mobile screen by extracting and sorting frames from the comic images. For this purpose, conventional image processing algorithms have been employed such as line segmentation [13], labeling, and line segment clustering [14]. Conventional fingerprinting extraction stems from this comic frame detection [15] and aims to identify the original from the scanned version by measuring the size and position of the box frame [11]. However, even if this method can recognize the overall structure of comics from the top view, it exhibits less accuracy when the comic images are distorted in the distribution process. Moreover, extracting a fingerprint is time-consuming because each feature is identified through highly complex computation. In traditional multimedia identification systems, including music, video, and text, the server must provide a high computation ability to monitor all data and extract the fingerprints for identification. Therefore, some limitations exist in applying conventional fingerprinting techniques to comic fingerprinting in terms of reliability, robustness, and operation time.

In contrast to other media, each page of a comic book is composed of unique cartoon images and words; the sequence of frames is meaningful in the construction of the comic fingerprint. Image/video fingerprinting methods can be classified into two classes: local and global feature-based methods [16]. Local feature-based methods extract all descriptors of an image, with one descriptor representing the local region of interest (ROI). These methods are usually robust against various distortions; however they incur a high computational cost. On the other hand, global feature-based methods extract a vector representing the overall information of the image. Although they have a low computational cost, global methods are relatively vulnerable to distortion attacks. A comic image is typically drawn with a pen and is primarily described in terms of edges on a grayscale, whereas a regular image is described in terms of the shapes on a color scale. This characteristic of the limited expression of the image restricts the local descriptors in the comic image more than in a regular image. Furthermore, one volume of a comic book has a large number of pages, and each page has numerous local descriptors in the form of high-dimensional vectors. Accordingly, it is highly computationally complex to specify all local features from the book, which makes it inefficient for commercialized comic fingerprinting.

To satisfy the computational cost required for the commercialization of comic fingerprinting, the global features of image/video fingerprinting can be a good solution. To increase identification using only global features, the unique sequence of page fingerprints in comics could be an important clue for an identification decision. In addition, it is important to determine the feasible types of distortions, such as global, local, and geometric distortions, prior to system optimization, which usually occurs during the comic book distribution.

Many global features are generally extracted by summarizing the overall statistics of the sub-images in one image/frame and are used to construct a fingerprint based on the relativity of the sequence [7,8]. They tend to be tolerant of global distortion because the relativity in terms of the ordinal measure does not easily change. For local distortion such as cropping or logo insertion, however, the effectiveness of the relativity of the sub-images decreases. To overcome this local distortion, the authors calculated the rank values of two sub-blocks over various regions in one image/frame [8]. Moreover, illegal copy detection algorithms have been proposed by partitioning each image into specific patterns of regions.

The legacy ROI algorithm can be applied to copy detection by identifying objects partially located in the ROI [17]. However, immense processing time is required to figure out objects in the ROI, which is also exponentially increased in proportion to the image size. As an alternative approach, the correlation measurement can be an effective tool for quantifying the similarity between images [18]. This method provides robustness in evaluating image similarity when the statistics is not changed for a given geometric distortion. In contrast, when the statistics is changed, the image similarity is differently measured; it is therefore difficult to identify such distorted images as the original. Moreover, it is necessary to calculate correlation values of all images in order to identify the comic, which leads to high computational complexity.

In [3], the authors attempted to match distorted images with the originals by capturing features from pixel values of each video frame. However, because the method deals with the pixel value and position together, it is not robust to geometric distortion (e.g. a shift or rotation) in which the two factors are changed. In [19], a circular partition was utilized to capture the average pixel value of each partition and then, the differences between partitions were calculated. The average and difference values were used as features to detect replica from the original image. In applications, the authors attempted to identify the replica of an image of nature, architecture, streets, and so on. However, the pattern was not optimized to identify the comic. Furthermore, the authors of [19] did not propose robust patterns for distortion of translation.

In [10], the authors presented hierarchical symmetric difference features for video fingerprinting. The image/frame was divided into two sub-images of 38 patterns in the range of small to large regions, and the fingerprints were extracted based on a comparison utilizing the average pixel intensity of the two sub-images. Unlike conventional illegal copy detection by using global features, this algorithm has low computational complexity and is robust to local distortion. However, it demonstrates low performance on geometric distortions, in particular, when a