Image and Vision Computing 76 (2018) 48–63

Contents lists available at ScienceDirect

Image and Vision Computing

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

A brand new application of visual-audio fingerprints: Estimating the position of the pirate in a theater - A case study☆

R. Roopalakshmi

Alva's Institute of Engineering and Technology, India

ARTICLE INFO

Article history:
Received 21 June 2016
Received in revised form 22 May 2018
Accepted 5 June 2018
Available online xxxx

Keywords:
Camcorder piracy
Content-Based video Copy Detection (CBCD)
Visual fingerprints
Audio signatures
SURF
Projective geometry
Geometric distortions

ABSTRACT

Combating against camcorder piracy requires identification of the theater and show time information, followed by the estimation of camcorder location in a theater from which an illegal recording was made, in order to find out the pirate and limit the number of pirate suspects. State-of-the-art pirate position estimation frameworks employ watermarking techniques to approximate the position of the pirate in a movie theater. However, watermarks are fragile in nature and involve complex procedures, which may damage the video content. To solve this, a novel forensic tracking framework, which exploits visual-audio fingerprints for estimating the location of the pirate in a theater without embedding digital watermarks is presented. Precisely, the proposed framework first spatio-temporally aligns the source movie and captured video contents, then estimates the geometric distortions and consequently derives the illegal capture location in a theater by exploiting multimodal features. The case study results in the form of sophisticated In-theater experiments prove that, it is certainly possible to estimate the illegal capture location in a theater with a mean absolute error of (38.25, 22.45, 11.11) cm, by employing multimodal fingerprints. In this way, the proposed article demonstrates a brand-new application of video fingerprinting for investigating the location of illegal camcorder capture in a theater, which is applicable for digital cinema applications.

1. Introduction

Due to the exponential growth of Internet technologies, the downloading and distribution of illegal movies is unprecedented, which in turn causes huge piracy and copyright issues. As a result of this movie piracy, worldwide motion picture industry including theaters and film distributors are losing billions of dollars from the past few years. For instance, the latest Canadian Motion Picture Distributors Association (CMPDA)-2011 [1] report alarms that, 133 million pirated movies are watched in Canada in 2010. This report also estimates the total loss to Canadian economy in 2010 as C$895 million due to movie piracy. Moreover, according to Motion Picture Association (MPA), over 90% of the pirated versions of newly released movies are created by camcorder piracy, which occurs due to illegal camcorder captures in theaters [2]. Thanks to the sophisticated camcorder models, camcorder piracy has raised as a serious issue for the movie industry over the past few years, which needs to be solved.

On the other hand, digital cinema system is introduced to uniformly project and distribute motion pictures and also to protect digital cinema. Strictly speaking, Digital Cinema Initiatives (DCI) is the entity, created to establish the technical specifications and requirements for mastering, distributing as well as theatrical playback of digital cinema content [3]. DCI defines a forensic watermarking system for copyright protection and also specifies that, the payload of the forensic watermark should contain time stamp and the theater information of movie playback. However, as per the requirements for protecting digital cinema, detecting the theater and time stamp information is not sufficient, it is necessary to identify the pirate so that the number of piracy suspects is restricted. Once the pirate is identified, then the next step is bringing the pirate to justice.

A scenario shown in Fig. 1 is considered for the purpose of identifying the movie pirate, which is defined as follows: (1) The pirate illegally records a movie using a camcorder in a theater and uploads the pirated content to the Internet. (2) In any anti-piracy strategy copy identification is the first step; hence, here Content-Based video Copy Detection (CBCD) approach is employed, which detects the best matching master video for the given pirated clip. (3) A conventional watermarking system such as [4] is used to determine the theater and the show time at which the illegal camcorder captures were made. (4) Then our position estimation framework estimates the

☆ This paper has been recommended for acceptance by Matthew Turk.

E-mail address: drroopalakshmir@aiet.org.in.

https://doi.org/10.1016/j.imavis.2018.06.001
0262-8856/

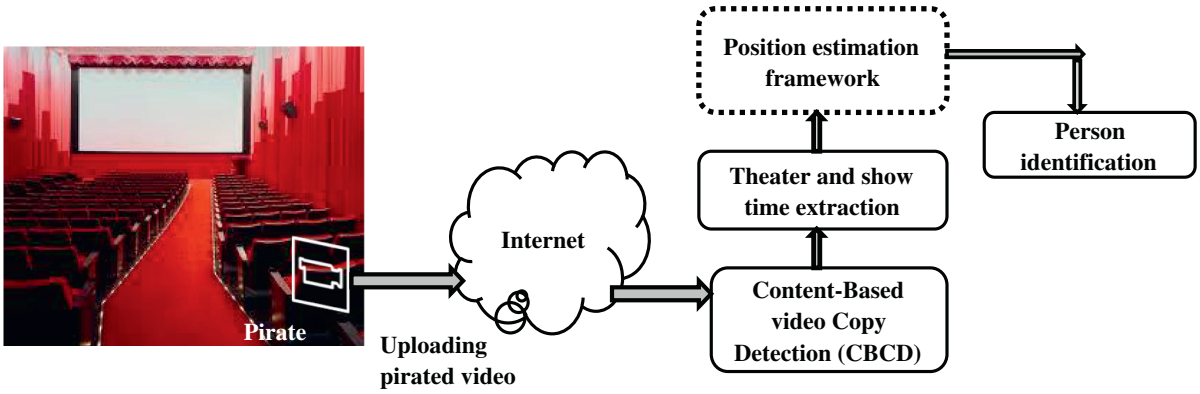

Fig. 1. Scenario for identifying a movie pirate.

position of the pirate in the theater in terms of specific seat information. Specifically, after the copy detection task, spatio-temporal alignments and estimation of distortion model between pirate and master video contents are prerequisites, in order to approximate the capture location in a theater. In order to achieve this, the proposed position estimation framework is implemented in 3 stages: i) It computes spatio-temporal frame alignments of the source movie and pirated video contents by exploiting visual-audio fingerprints; ii) Then it estimates the geometric distortions in the pirated video in terms of the projective matrix; iii) Consequently, the camcorder optical axis to the screen perpendicular is determined by redefining the theater projective geometry and eventually the position of the pirate in the theater is estimated. (5) Finally, an electronic ticketing system may be used to identify the exact person who illegally captured the movie. In this way, the proposed position estimation framework restricts the number of piracy suspects and helps to identify the pirate. This paper focuses on the position estimation framework shown in Fig. 1 with dotted lines, which is an essential component of this scenario.

2. Related work

This section details the state-of-the-art techniques related to all the three stages of proposed position estimation framework namely spatio-temporal frame alignment, geometric distortion estimation and pirate position estimation as given below.

2.1. Spatio-temporal frame alignments

In literature, several papers are proposed to address the frame alignments of source and pirated video sequences. For instance, in 2003, Delannay et al. [5] presented a temporal registration scheme, by matching key frames of two video contents. Cheng [6] introduced a dynamic programming based scheme for the temporal alignment of source and captured video sequences. Though this method provides accurate frame alignments, yet it is much affected by distortions such as noise addition. Cheng and Isnardi [7] proposed a spatial, temporal and histogram registration algorithm for detecting forensic watermark information by integrating contextual costs. Chupeau et al. [8] designed a temporal registration scheme using color histograms in order to match the illegal video and source movie contents with the help of dynamic programming. Due to the global descriptive nature of histograms, this scheme achieves poor results for region-based distortions.

Baudry et al. [9] employed global and local features for matching source and captured video contents. Although this framework achieves better accuracy, yet it performs poor for low motion frames and distortions such as subtitles and logo insertions. Chen et al. [10] proposed a video matching scheme by employing video fingerprints based on temporal ordinal measurements. Lee et al. [11] presented a frame matching algorithm based on dynamic programming by utilizing effective matching cost function, which reduces matching errors. However, this method performs better for few types of transformations such as compression, frame insertions, shuffle and frame removals. Recently, Baudry et al. [12] presented a temporal registration scheme for video copies, in which fingerprints are computed based on hierarchical encoding of the wavelet coefficients. However, this scheme suffers due to its cost in terms of memory and CPU.

To summarize, previous research on video copy frame alignment utilize only visual features for aligning the source and pirated movie sequences [5-12], while no effort is made to employ acoustic features. However, audio is an essential information source of a video, which is less affected in most of the illegal camcorder captures compared to the visual content [13]. Hence, promising registration schemes employing visual-audio fingerprints are needed in order to achieve accurate frame-to-frame alignments of the pirated video with the master video content.

2.2. Geometric distortion estimation

There are only a few papers that focus on estimating the geometric distortions in illegal contents. Delannay et al. [14] presented a scheme for estimating and compensating the geometric distortions in pirated contents which occur due to handy cam attacks. The authors employed displacement vectors to estimate the distortions and utilized 12-parameter bilinear transformation model for compensating the distortions. Though this method is useful in digital cinema applications, yet, the performance of this method is sensitive to the watermark embedding algorithm. Chupeau et al. [15] introduced a method for estimating and compensating the geometric distortions in video copies. The authors employed a visual descriptor based on luminance values to generate the temporal profile of two video contents.

Existing attempts on geometric distortion estimation are exploiting only visual signatures of videos and addressing the distortions in watermarked video contents [14,15]. However, it is important to note that all copyrighted contents are not watermarked [16]. Hence, promising frameworks utilizing multimodal features are required to