HAL
archives-ouvertes.fr

Eff2 Videntifier: Identifying Pirated Videos in Real-Time

Kristleifur Daðason, Herwig Lejsek, Friðrik Ásmundsson, Björn Þór Jónsson, Laurent Amsaleg

To cite this version:

Kristleifur Daðason, Herwig Lejsek, Friðrik Ásmundsson, Björn Þór Jónsson, Laurent Amsaleg. Eff2 Videntifier: Identifying Pirated Videos in Real-Time. ACM International Conference on Multimedia, Sep 2007, Augsburg, Germany. 2007. <inria-00175874>

HAL Id: inria-00175874
https://hal.inria.fr/inria-00175874

Submitted on 1 Oct 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Eff^2 Videntifier: Identifying Pirated Videos in Real-Time

Kristleifur Dadason, Herwig Lejsek, Fridrik Ásmundsson, Björn Jónsson
Reykjavík University
Kringlan 1
IS-103 Reykjavík
Iceland
{kristleifur04, herwig, fridrik01, bjorn}@ru.is

Laurent Amsaleg
IRISA–CNRS
Campus de Beaulieu
35042 Rennes
France
laurent.amsaleg@irisa.fr

ABSTRACT

With the proliferation of high-speed internet access and the availability of cheap secondary storage, movie piracy has become a major problem. This demonstration paper describes the Eff^2 Videntifier, a content-based system for large-scale automatic copyright enforcement of videos. The paper briefly describes the database and image processing techniques underlying the system. It also describes our proposed demonstration, which realistically simulate scenarios of copyright violations of movies.

Categories and Subject Descriptors: H.2.4 [Database Systems]: Systems—*Multimedia Databases*

General Terms: Algorithms, Performance

Keywords: Copyright Protection, NV-Tree, Video Retrieval

1. INTRODUCTION

Publication and distribution of multimedia content is a large, high-revenue industry. Managing and tracking when and where multimedia content gets published has therefore become an integral part of the business of content providers.

With the proliferation of high-speed internet access and the availability of cheap secondary storage it has become very easy for home users to copy large amounts of data before massively sharing it over the web. That shared data typically contains much self-produced or free content, but in many cases also copyrighted material from third parties.

Despite the poor image quality, thousands of videos are uploaded every day to video-sharing sites such as YouTube. It is known that this material includes a significant percentage of copyrighted movies. Other file sharing platforms, such as eDonkey or BitTorrent, are also very popular, yet illegal, sources of copyrighted material. In 2005, a study conducted by the Motion Picture Association of America was published, which estimated that their members lost 2,3 billion US$ in sales due to video piracy over the Internet [1].

Due to the high risk of piracy, movie producers have tried many means of restricting illegal distribution of their material, albeit with very limited success. Video pirates have found ways to circumvent even the most clever protection mechanisms. In order to cover up their tracks, stolen (ripped) videos are typically compressed, modified and re-encoded, making them more suitable for easy downloading. Another very popular method for stealing videos is *camcording*, where pirates smuggle digital camcorders into a movie theatre and record what is projected on the screen. Once back home, that goes to the web.

Clearly, this environment calls for an automatic content-based video retrieval system. Such a system needs to be *effective* as it must cope with often severe attacks against the visual contents, and *efficient* as it must rapidly spot the original videos from a huge reference collection.

2. AUTOMATIC IDENTIFICATION

A fully automatic content-based video identification system builds a reference database of low-level features extracted from the videos to protect. Then, video streams to check are fed into the system to detect near-duplicate content. Those streams can originate from the web (via a robot), from a TV-broadcast or from a camera installed in front of a multimedia device; a service such as YouTube could even submit uploaded videos to the system.

As pointed out above, effectiveness and efficiency are crucial in order for a system to be usable in practice at a realistic scale. Efficiency means that it must be real-time, even when the reference database is very large or many incoming streams must be checked simultaneously (e.g., when monitoring many TV channels). Effectiveness means it must detect stolen material despite low visual quality due to severe alteration of the visual contents, such as downsampling, frame rate changing, border attacks, affine transformations, motion blurring, compression, etc. It must also avoid false positives while having only a few false negatives (e.g., missing extremely distorted videos might be acceptable).

Particular video copyright related applications have specific needs, however. Querying the system using small fractions of a video file found on the web might be sufficient to assess whether it is stolen material. Monitoring an incoming video stream to spot trailers or program boundaries requires a more detailed and continuous querying process. Furthermore, video copyright protection focuses for the most part on very recent videos. Popular movies typically make the major part of their profit in the first few months after publication. Nevertheless most movies are available online for download right after the premiere, sometimes even before. Supporting database updates is therefore key.

Copyright is held by the author/owner(s).
MM'07, September 23–28, 2007, Augsburg, Bavaria, Germany.
ACM 978-1-59593-701-8/07/0009.

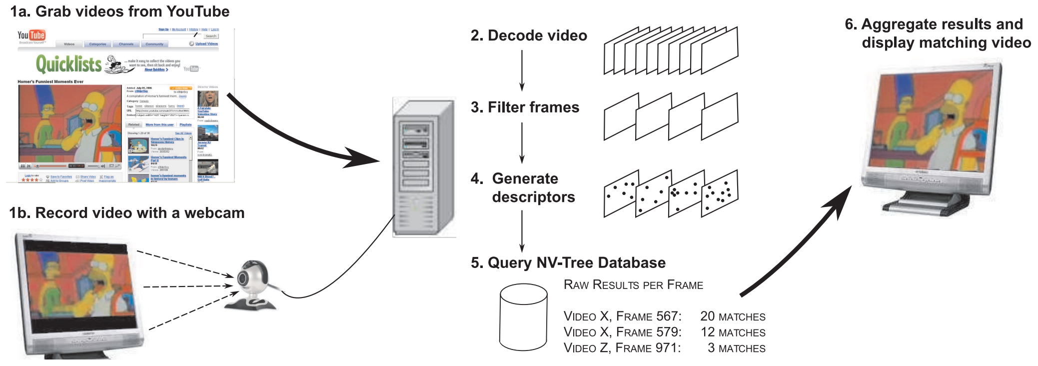

Figure 1: The Eff^2 Videntifier video monitoring prototype

3. THE Eff^2 VIDENTIFIER

With the Eff^2 Videntifer we show a system that fulfills all the requirements stated above. The Eff^2 Videntifer is built on top of a novel high dimensional database technology, the NV-Tree [2], which is able to perform accurate approximate high-dimensional nearest neighbour queries in constant time, independent of the size of the video descriptor collection. The NV-Tree also copes well with updates and its design allows for easy distribution of queries across several computers, which increases system throughput linearly [3].

Video frames are described using the Eff^2 image descriptors. The Eff^2 descriptors are of the SIFT family [4], but have been shown to have much better recognition power for very large collections [5], making the Eff^2 Videntifer very effective at detecting pirated movies.

4. DEMONSTRATION SCENARIOS

We have set up a database containing about 100 million video descriptors from more than 100 full-time videos (about 150 hours of video). Figure 1 shows the configuration of the demonstration system, which provides two related scenarios.

In a first scenario (1a) we will demonstrate a copyright protection application which continuously downloads videos from the net and checks them against our collection. We will show that Videntifier is able to detect pirated copies of video, even in case of strong modifications. Among them are cropping, blurring, brightness and contrast changes, reencoding, camcorder attacks and even compression to 56Kbps quality.

Videos to check are first decomposed into frames. Then, every 12th frame is used to compute up to 400 Eff^2 descriptors. These descriptors are then used as queries and the NV-Tree database is probed. The results are aggregated in a two step process; first at a frame level, and then at a video level. Some false matches are filtered using a sequence aware quality metric taking into account that successive query frames should yield successive matches.

As a second scenario (1b) we will show a video monitoring application. This scenario differs from the previous one as many more frames (overall and per second) must be checked. A decent web cam is used to record a movie from a computer screen, adding several interesting artifacts to the recorded video. First the camera adds lense distortion, noise and saturation changes to the recording. Second, the position of the camera plays an important role, since it introduces affine distortions, small rotations, cropping or bordering effects to the video. Finally, since webcams do not record at full frame rate, significant motion blur is introduced for fast moving scenes. We will demonstrate that our system is nonetheless capable of detecting the correct reference movie from the database in real-time and in as little as 5 seconds.

5. RELATED WORK

The closest competitor to the Eff^2 Videntifier is a real-time TV monitoring system proposed by Joly et al. [6]. That work includes experiments using significantly larger collections of TV recordings (40,000 hours). The system still showed real-time performance at that scale due to strong filtering of the video data yielding about 10 times fewer descriptors per hour of video. When used on descriptor collections of comparable size, however, our system handles more difficult modifications and reports significantly better response time than their system.

6. ACKNOWLEDGEMENTS

This work was partly supported by an Icelandic Research Fund grant 60036021, INRIA Eff^2 Associate Teams grant, and an ÉGIDE Jules Verne travel grant.

7. REFERENCES

[1] L.E.K. Consulting The Cost of Movie Piracy. *http://www.mpaa.org/2006_05_03leksumm.pdf*

[2] H. Lejsek, F. H. Ásmundsson, B. Th. Jónsson and L. Amsaleg. NV-Tree: An efficient disk-based index for approximate search in very large high-dimensional collections. *Tech. Rep. RUTR-CS07001, Reykjavík University*, 2007.

[3] F. H. Ásmundsson. NV-Network: A Distributed Architecture for High Throughput Image Retrieval. *Master's thesis, Reykjavík University*, August 2006.

[4] D. G. Lowe. Distinctive Image Features from Scale-invariant Keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[5] H. Lejsek, F. H. Ásmundsson, B. Th. Jónsson and L. Amsaleg. Scalability of Local Image Descriptors: A Comparative Study. *Proc. ACM Multimedia Conference*, Santa Barbara, CA, USA, 2006

[6] A. Joly, O. Buisson and C. Frélicot. Content-Based Copy Detection using Distortion-Based Probabilistic Similarity Search. *IEEE Transactions on Multimedia*, 9(2):293–306, 2007.