# Mediaprinting: Identifying Multimedia Content for Digital Rights Management

**Tiejun Huang, Yonghong Tian, and Wen Gao,** *Peking University*
**Jian Lu,** *Shanda Interactive Entertainment*

**Encryption and watermarking are the most common techniques used to protect copyrighted multimedia content, but both have many limitations. Mediaprinting offers a reproducible and reliable alternative for digital rights management and related applications on the Internet.**

The Internet is revolutionizing multimedia content distribution, offering users unprecedented opportunities to share digital images, audio, and video but also presenting major challenges for digital rights management (DRM).

The ease with which anyone can upload and download material inherently facilitates misuse, piracy, plagiarism, and misappropriation. In 2000, A&M Records and other leading record companies sued peer-to-peer music file-sharing site Napster for contributory and vicarious copyright infringement. Seven years later, Viacom took YouTube to court for "massive intentional copyright infringement." These and similar lawsuits highlight the importance of content protection and copyright management as the Internet evolves into a global multimedia distribution platform.

Technologically speaking, DRM refers to the technologies or systems that protect and enforce the rights associated with the use of digital content. Two proactive DRM approaches that have emerged in the past two decades are encrypting multimedia content to prevent unauthorized access and embedding watermarks for posterior authentication.[1] However, both approaches have many limitations, and neither encryption nor watermarking can help resolve the rights issues associated with the vast amount of content distributed by millions of Internet users.

*Mediaprinting* offers a retroactive but reproducible and reliable alternative approach for multimedia content identification and management on the Internet.

## CURRENT DRM APPROACHES

Encryption and watermarking are the two most common DRM techniques for protecting multimedia content.

### Encryption

As a fundamental information security technology, encryption is the process of scrambling confidential data into an unintelligible form. Providers can apply various encryption techniques to protect the confidentiality of, and prevent unauthorized access to, digital content. Multimedia encryption involves numerous technical complexities not encountered in encrypting text or other data.[1] In addition, this approach has several limitations.

**Lack of interoperability.** Typically, different DRM systems employ their own encryption and rights management techniques. This makes interoperation of these systems difficult. Moreover, DRM system vendors might refuse to disclose their systems' inner workings or license their technologies, resulting in competing and incompatible systems. For example, Apple's FairPlay system is incompatible with Microsoft's Windows Media system.

 Published by the IEEE Computer Society 0018-9162/10/$26.00 © 2010 IEEE

**Fair use and public availability restrictions.** DRM restricts fair use rights. For example, users cannot transfer and play content protected by encryption-based DRM on arbitrary devices. Moreover, encryption-based DRM hampers public availability of multimedia content even after copyright expiration.

**Deployment cost and complexity.** Encryption-based DRM is most effective in a closed-content system. In large, open environments like the Internet, encryption-based content distribution requires the deployment of costly and complex security mechanisms in a wide range of consumer devices. Furthermore, if an encryption system gets cracked, fixing the damage or upgrading the security infrastructure will incur additional cost.

### Watermarking

A digital watermark is a signal embedded in multimedia content. In addition to being perceptually invisible or inaudible to humans, watermarks should be statistically undetectable and resistant to any malicious attempts to remove them. In copyright protection applications, watermarks can carry information to assert the owner's copyright, licensing data for access control, or user-related information (such as a user's identity) to track illegal copy transfer. Researchers have developed different types of digital watermarks—for example, robust, semifragile, and fragile—to assist in DRM, but the technology still faces several fundamental challenges.

**Insufficient robustness.** Despite considerable efforts to develop watermarks resistant to content transformations such as JPEG compression, rotation, cropping, and additive noise, current watermarking techniques are not sufficiently robust for many DRM applications.

**Inevitable degradation of quality.** Multimedia content will inevitably degrade after watermarking. In general, it is easy to create either robust or imperceptible watermarks, but creating watermarks that have both qualities has proven to be quite difficult.[1]

**Incompleteness.** Even robust watermarking technology cannot authenticate ownership of multimedia content on its own, as anyone can embed watermarks in the content. That is to say, a third-party content registration and authentication authority is needed.

### Other limitations

In addition to the problems unique to encryption and watermarking, both DRM solutions are vulnerable to the so-called *analog hole*—that is, protected digital content can be recorded and copied through analog means, then redigitized and distributed to bypass the protection systems.

However, perhaps the biggest impediment to proactive techniques like encryption and watermarking is not their robustness but their coverage. For example, the same video can be distributed via DVD, satellite and cable broadcasting, online streaming, or digital download, to name only a few ways. Using encryption-based DRM to protect content in all forms and channels is practically impossible. Indeed, Steve Jobs attributes FairPlay's ineffectiveness to the coexistence of unprotected and protected music content. Similarly, most digital content on the Internet is not watermarked and therefore cannot be tracked or protected using this approach.

## MEDIAPRINTING FOR DRM

The limitations of both encryption and watermarking motivated the development of mediaprinting, a new DRM approach that attempts to retroactively protect copyrights by identifying multimedia content and checking whether it has been illegally distributed and shared on the Internet. Mediaprints are compact descriptors that, unlike extrinsic identifiers affixed to multimedia such as watermarks, or assigned identifiers such as International Standard Recording Code numbers for music, are extracted from the content. A mediaprint thus cannot be erased or faked because it can be always recomputed from the content. Unlike cryptographic hashes computed from binary data, which are extremely fragile and data-sensitive, mediaprints are robust (unchanging) across a wide range of modifications and transformations of the same content but sufficiently different for every unique content item.

**Mediaprints for content identification are analogous to fingerprints and voiceprints for personal identification.**

Mediaprints for content identification are analogous to fingerprints and voiceprints for personal identification. Different types of media have different types of mediaprints—thus, mediaprints for image, audio, and video content are called *imageprints*, *audioprints*, and *videoprints*, respectively. This concept can be extended to *docprints* for documents and *softwareprints* for source code. Mediaprints are also referred to as multimedia fingerprints, perceptual hashes, audio and visual signatures, and media DNA (for example, video DNA).

A mediaprint has at least two intrinsic properties:

- *Robustness.* A mediaprint is largely invariant for an original content item and its copies—it is not substantively altered by modifications such as editing operations or transformations such as transcoding and analog-to-digital conversion.

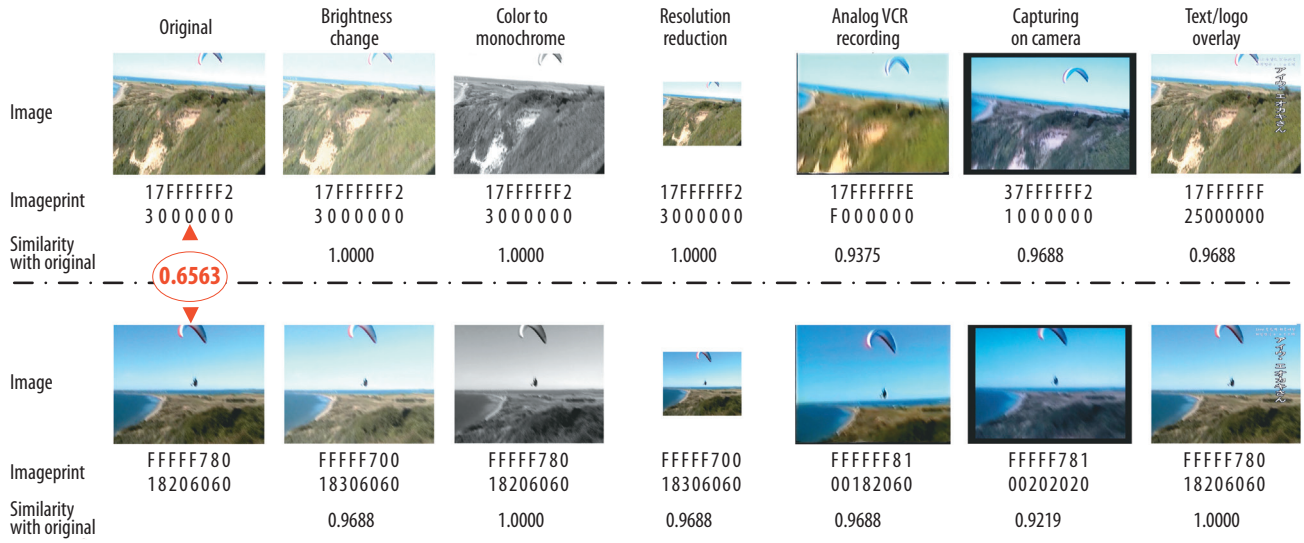


**Figure 1.** Robustness and uniqueness of mediaprints. This example shows 64-bit DCT-based imageprints extracted from two original images and six transformed copies. Despite having a very similar visual appearance, the two original images have distinct imageprints. The imageprints are largely unchanged for the transformed copies of the same original image.

- *Uniqueness.* Mediaprints extracted from different original media items are significantly different. In other words, mediaprints can accurately distinguish different media items.

Figure 1 illustrates these properties. The example shows 64-bit DCT (discrete cosine transform)-based imageprints extracted from two original images and six transformed copies. Despite having a very similar visual appearance, the two original images have distinct imageprints (a similarity of 0.6563, calculated using Hamming distance). Note that the imageprints are largely unchanged for the transformed copies of the same original image.

The Moving Picture Experts Group has specified other intrinsic properties for MPEG visual signatures,[2] including fast matching, fast extraction, compactness, nonalteration, self-containment, and coding agnosticism. For example, the nonalteration property signifies that visual signatures are extracted and measured without altering the content.

Mediaprinting-based DRM systems typically include two major processes. In the *registration* process, the system extracts media prints from copyrighted media content and stores it in a database along with metadata—for example, the item's title, ownership information, production and release dates, and locations—and rules specified by content owners as to what actions to take when an unauthorized copy of reference content is identified.[3] In the *identification* process, the system extracts the mediaprint of a given query item and then compares it with all mediaprints in the database to determine whether it matches a registered item. If the matching result indicates that the query item is an unauthorized copy, then the content owner will take the specified action.

## RESEARCH QUESTIONS

As Figure 2 shows, mediaprinting research centers on three main questions.

*How do we extract mediaprints that are robust across a wide range of transformations?* We can easily recall a movie we saw many years ago when it is replayed on a different device, such as a TV or DVD player, or in different environmental conditions, such as a brightly lit living room versus a dark theater. We can also immediately recall a song just by catching the melody. In these cases, the representation of the images or music in our mind should be the same, or at least very similar, and can serve as the ideal "mediaprint."

For DRM and other applications, extracting unique mediaprints is a key challenge. Mediaprint extraction is similar to feature extraction in content-based retrieval in that both aim to describe media content as a multidimensional space of features, such as color, texture, shape, and motion. However, mediaprint extraction is more difficult because mediaprints must represent images, audio, and video as unique entities.

*How do we efficiently match mediaprints in a large-scale database?* Because mediaprints are not identical for different versions of the same content item, mediaprint matching is not a simple database table lookup—it is a similarity search problem in high-dimensional space.[3] The enormous amount of copyrighted content calls for algorithms that

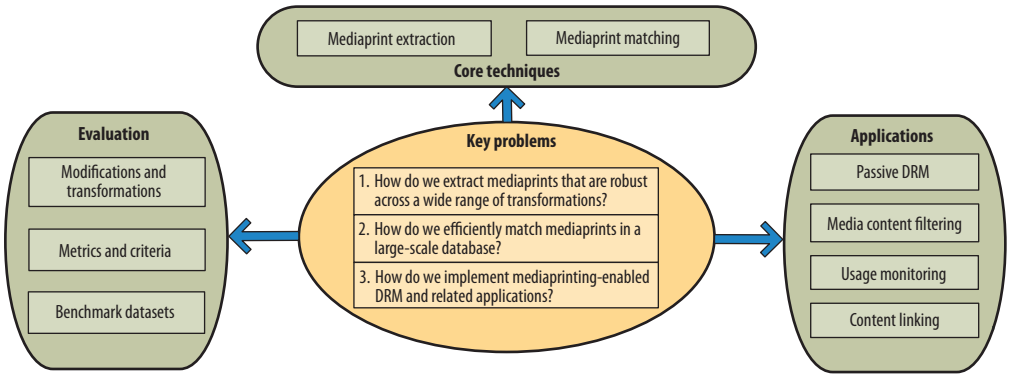


**Figure 2**. Mediaprinting research centers on three main questions.

- efficiently index mediaprints to enable fast searching in a very large database,
- assess the similarity between a pair of mediaprints, and
- measure and evaluate mediaprint matching techniques—for example, for accuracy and speed.

*How do we implement mediaprinting-enabled DRM and related applications?* This involves building a large-scale, accessible mediaprint database as well as combining mediaprinting with other DRM approaches such as encryption and watermarking.

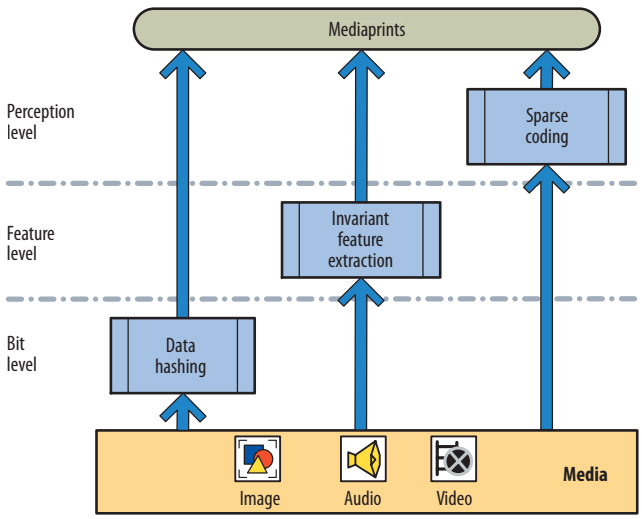


**Figure 3.** Mediaprint extraction. At the bit level, data-hashing functions can generate a fixed-length string of bits as a mediaprint. At the feature level, algorithms can extract invariant features from the spatial, temporal, or frequency domain and convert these features into mediaprints. At the perception level, sparse coding can simulate the physiological system that generates a compact expression for a media item.

## MEDIAPRINT EXTRACTION

Mediaprints can be extracted from content in various ways, as Figure 3 shows. At the binary data level, one approach is to treat the content as a bitstream and use data-hashing functions to generate a fixed-length string of bits as its mediaprint. A more sophisticated approach is to extract invariant features from the content in the spatial, temporal, or frequency domain, or use a combination of features from all three domains, and then convert these features into mediaprints; the task then becomes finding such invariant features. Because the human brain stores multimedia content in a sparse way, the best mediaprinting approach might be to use sparse coding to simulate the physiological system that generates a compact expression (mental imagery) for a media item.

Mediaprint extraction techniques fall into two categories. *Feature-based* techniques generate a mediaprint by extracting physically meaningful features from the content that characterize certain aspects of its uniqueness. Unlike pattern-recognition applications, which use features that are robust in a particular category, features for mediaprinting should be robust to content distortions—for example, transformation-invariant and visually salient. The features can either be directly extracted from the content or obtained using feature transforms such as dimension reduction. *Process-based* techniques generate mediaprints from content directly via a linear or nonlinear mapping function—typically an artificial neural network that identifies the content being viewed or listened to by functionally simulating the human auditory or visual process with sparse coding.

Different media types generally require different mediaprint extraction algorithms. There are three basic approaches to imageprinting. One is to extract features describing the entire image to generate imageprints. Although this approach performs well in many cases, it is less robust to local modifications such as cropping, embedding, and combining. An alternative is to extract local features from the image to generate imageprints—for example, to partition an image into several blocks (or

regions) and then extract features from these blocks. A third, keypoint-based, approach has recently attracted interest. For example, Vishal Monga and Brian L. Evans have proposed an image perceptual hashing algorithm using visually significant feature points.[4]

Extraction algorithms for audio and video differ from those for images because they can take into account the temporal correlation of neighboring frames to generate more robust mediaprints. Audioprinting algorithms use features originally designed for content-based audio retrieval such as Mel-frequency cepstral coefficients (MFCCs), mean energy, normalized spectral sub-band moments, and audio spectrum flatness (ASF) in MPEG audio signatures.

Videoprinting algorithms use one of two approaches. The first employs 3D data transforms (spatiotemporal DCT) to extract a global descriptor of a video clip. However, this is difficult for partial content matching—namely, to determine whether one segment of a query clip matches a certain segment of a reference. Another approach is to employ imageprinting methods on key frames, then assemble the corresponding frameprints to form the videoprint.[5] Using spatial or temporal information such as the difference between or correlation of neighboring frames[6] can generate more robust and discriminable videoprints.

**Efficient indexing and search techniques are needed to enable rapid mediaprint matching in large, continuously expanding reference databases.**

Various combinations of video transformations (for example, pattern or picture insertions and simulated camcording) and audio transformations (for example, mp3 compression and adding speech) can change a video clip. In such cases, videoprints and audioprints should be extracted independently from the video and audio tracks and then aligned to accurately and efficiently identify near-duplicate video copies from a large collection of video clips.

## MEDIAPRINT MATCHING

There are two query scenarios in mediaprint matching: in *direct matching*, the system determines whether the query item matches an entire mediaprint in the database; in *partial matching*, the system determines whether a segment of the query item matches a segment of one or more mediaprints in the database.[2] Mediaprint matching is a similarity search problem, and researchers use various distance metrics to measure the degree of similarity between two mediaprints, including Hamming distance, Euclidean distance, Manhattan distance, and bit error rate (BER).[3]

Efficient indexing and search techniques are needed to enable rapid mediaprint matching in large, continuously expanding reference databases. In most cases, a well-designed approximate search can find a best match in a fraction of the time required for an exhaustive search, which is clearly not scalable for practical applications.[3]

During the past two decades, researchers have developed numerous nearest-neighbor search techniques for high-dimensional datasets. *Locality-sensitive hashing*, a widely used algorithm for fast mediaprint matching, is based on the simple idea that, if two points are close together, they will remain so after a "projection" operation.[7] The goal of LSH is to hash a large reference database into a much-smaller-size bucket of match candidates, then use a linear, exhaustive search to find the points in the bucket that are closest to the query point. The challenge is to devise functions that hash the close points into the same bucket with high probability.

To address this issue, Kave Eshghi and Shyamsundar Rajaram proposed a new class of LSH functions for cosine similarity based on concomitants that capture the relation between the order statistics of $X$ and $Y$.[8] Brian Kulis and Kristen Grauman generalized LSH to accommodate arbitrary kernel functions, making it possible to preserve the algorithm's advantage of sublinear time search for a wide class of useful similarity functions.[9] Shumeet Baluja and Michelle Covell proposed a "learning to hash" technique that uses machine learning methods and training data to devise a hashing system that adapts to the identification task and data, resulting in a more compact hash bucket that contains significantly fewer candidates to be compared with a linear search.[10]

Though promising, hash-based approximate search can lead to low recall rates. To boost search quality, Yin-Hsi Kuo and colleagues proposed two novel strategies: *intra-expansion* to increase the number of target feature points similar to those in the query, and *inter-expansion* to mine feature points that co-occur with the search targets but are not present in the query.[11] Another problem with LSH is that large collections require a large main memory to store the hash tables and avoid frequent disk accesses. Herwig Lejsek and colleagues recently proposed the NV-tree as an efficient disk-based data structure that can give good approximate answers to nearest-neighbor queries with a single disk operation, even for very large collections of high-dimensional data.[12]

## EVALUATION AND STANDARDIZATION

As part of the standardization process for mediaprinting tools, MPEG began evaluating image and video signatures in 2007.[2] That same year, the Motion Picture Association of America (MPAA) organized an industry-wide content-

## USING MEDIAPRINTING TO PROTECT OLYMPIC CONTENT

The 2008 Beijing Summer Olympics set a milestone in Olympic media content distribution and protection. For the first time, the International Olympic Committee sold media rights separately for over-the-air TV broadcasting and new media (including Internet and mobile) distribution. The IOC established strict requirements for rights-holding broadcasters (RHBs) to implement satisfactory antipiracy and content security measures. It also set up a special Internet monitoring program.

To protect Olympic video content, China Central Television and CCTV.com, the RHBs for TV and new media in the host country, partnered with Vobile, a California company that has developed VideoDNA, a videoprinting technology.[1] CCTV used VideoDNA to extract videoprints from the live feeds of Olympic events and deployed its VideoTracker system, shown in Figure A, to monitor online distribution of video content.

Over the course of the 2008 Games, VideoTracker ingested videoprints of 929 live events, monitored 312 online sites, and identified 4,364 infringements. Bloomberg Businessweek called it "a surprise victory for the broadcasters in the antipiracy Olympics."[2]

### References


1. P. Burrows, "Video Piracy's Olympic Showdown," 29 May 2008, *Bloomberg Businessweek*; www.businessweek.com/magazine/content/08_23/b4087073685542.htm.
2. P. Burrows, "A Surprise Victory for the Broadcasters in the Anti-Piracy Olympics," 12 Sept. 2008, Bloomberg Businessweek; www.businessweek.com/the_thread/techbeat/archives/2008/09/a_surprise_victory_for_the_broadcasters_in_the_anti-piracy_olympics.html.


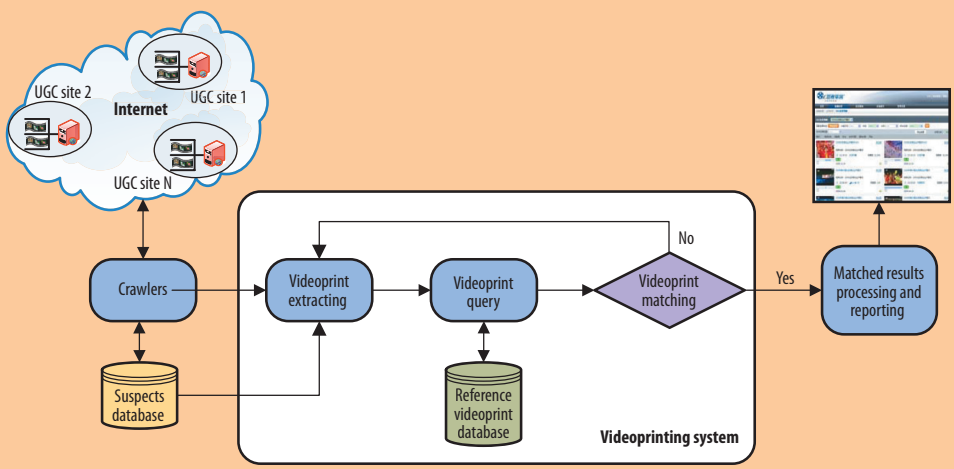


**Figure A. Vobile's VideoTracker system, which consists of Web crawlers looking for infringement suspects, the VideoDNA videoprinting system, and a Web interface to report and update tracking results, successfully monitored online distribution of video content from the 2008 Beijing Summer Olympics.**

recognition effort focusing on mediaprinting. And in 2008, TRECVID initiated a content-based copy detection (CBCD) scheme.

All of these efforts used sample datasets consisting of content that had undergone various types of modifications to different degrees such as light, medium, and heavy. The modifications consisted of

- coding format changes, such as transcoding;
- editing operations—for example, the deletion or insertion of frames, the overlay of text or graphic patterns such as logos, and various types of image processing;
- quality changes—for example, the addition of noise, analog VCR recording, and camcording; and
- combinations of these changes.

The evaluation metrics included false alarm rate (FAR) and miss alarm rate (MAR), or equivalently precision and recall. TRECVID also used an overall measure—minimal normalized detection cost rate (DCR)—to evaluate detection effectiveness.

Expanding on these initiatives, MPEG and other standards organizations are working to develop industry standards for mediaprinting.

## OTHER MEDIAPRINTING APPLICATIONS

Beyond DRM, mediaprinting can be used for media content filtering, usage monitoring, and content linking.

Content-sharing sites such as Flickr and YouTube could employ mediaprinting to prevent users from uploading copyrighted material. When a user uploads a video clip, for example, the site could extract its videoprint, compare it to all copyrighted assets in a reference database, and, depending on the query match result, either publish the item to the Web or remove it according to the content owners' rules.

Mediaprinting already offers an effective way to monitor media usage. Content providers have successfully used

it to track high-valued copyrighted video ranging from Hollywood blockbusters to coverage of the 2008 Beijing Summer Olympics, as detailed in the "Using Mediaprinting to Protect Olympic Content" sidebar. In addition, copyright owners could use mediaprinting to collect royalties for their content, advertisers could use it to audit the airing of their commercials in paid time slots by a broadcast network, and brand holders could use it to detect plagiarism and misappropriation of their registered brands.

A wide range of content-linking applications such as contextual advertising and content-based retrieval would also benefit from mediaprinting. For example, mediaprint-based content identification can recognize exactly what content users are consuming, leading to more relevant advertisements.[3] By employing videoprints, advertisers could identify the most effective ad clips on different TV channels to enhance content-based video browsing and retrieval.

Mediaprinting could also play a fundamental role in semantically organizing digital items such as images, music, and webpages in cyberspace. For example, mediaprinting could add a new layer to the current manually linked Web by semantically linking near-duplicate or similar media content on different sites. The new link layer would improve the relevance performance of search engines and enable many interesting new applications.

Despite significant progress in mediaprinting technologies in recent years, continuing research and development are needed to provide robust multimedia content identification and management on the Internet.

Designing robust mediaprints that can accurately identify multimedia content is an important challenge. Many current mediaprint designs have crossed the bar of being "good enough" to use in real-world applications, but there is a need to reduce false positive and negative rates in content identification, particularly in large mediaprint databases.

In practice, some mediaprints are robust against certain types of distortions in media content but vulnerable to other types of distortions, and vice versa. Combining a set of mediaprints that complement each other could enhance robustness and discriminability.[3] An example is using classifier ensembling to improve pattern classification accuracy. We validated this approach during the 2010 TRECVID-CBCD contest, where we combined various types of imageprints and audioprints to accurately detect near-duplicate video copies from a large collection.

As providers distribute more copyrighted content on the Internet, scalability becomes critical for mediaprint matching. More research is needed on accurate and efficient mediaprint indexing that enables fast search as well as on-search optimization that optimally trades off accuracy and speed.

Researchers also must address many practical issues related to mediaprint systems and workflows, such as constructing and managing a large-scale, continuously expanding universal reference mediaprint database of copyrighted content and more effectively associating mediaprints with metadata and owner-specified copyright violation rules. C

## Acknowledgments

The work described in this article was supported by grants from the Chinese National Natural Science Foundation under contract No. 60973055 and No. 90820003, and a grant from National Key Technologies R&D Program of China under contract No. 2009BAH51B01.

## References

1. W. Zeng, H. Yu, and C.-Y. Lin, eds., *Multimedia Security Technologies for Digital Rights Management,* Academic Press, 2006.
2. M. Bober and P. Brasnett, "MPEG-7 Visual Signature Tools," *Proc. 2009 IEEE Int'l Conf. Multimedia and Expo* (ICME 09), IEEE Press, 2009, pp. 1540-1543.
3. J. Lu, "Video Fingerprinting for Copy Identification: From Research to Industry Applications," *Proc. SPIE,* vol. 7254, 2009; http://159.226.42.40/jiaoxue-MMF/2009/VideoFingerprinting_SPIE-MFS09.pdf.
4. V. Monga and B.L. Evans, "Perceptual Image Hashing via Feature Points: Performance Evaluation and Tradeoffs," *IEEE Trans. Image Processing,* vol. 15, no. 11, 2006, pp. 3452-3465.
5. S. Lee and C.D. Yoo, "Robust Video Fingerprinting for Content-Based Video Identification," *IEEE Trans. Circuits and Systems for Video Technology,* vol. 18, no. 7, 2008, pp. 983-988.
6. J. Oostveen, T. Kalker, and J. Haitsma, "Feature Extraction and a Database Strategy for Video Fingerprinting," *Proc. 5th Int'l Conf. Recent Advances in Visual Information Systems* (VISUAL 02), LNCS 2314, Springer, 2002, pp. 117-128.
7. Z. Yang, W.T. Ooi, and Q. Sun, "Hierarchical, Non-Uniform Locality Sensitive Hashing and Its Application to Video Identification," *Proc. 2004 IEEE Int'l Conf. Multimedia and Expo* (ICME 04), vol. 1, IEEE Press, 2004, pp. 743-746.
8. K. Eshghi and S. Rajaram, "Locality Sensitive Hash Functions Based on Concomitant Rank Order Statistics," *Proc. 14th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining* (KDD 08), ACM Press, 2008, pp. 221-229.
9. B. Kulis and K. Grauman, "Kernelized Locality-Sensitive Hashing for Scalable Image Search," *Proc. 12th IEEE Int'l Conf. Computer Vision* (ICCV 09), IEEE Press, 2009, pp. 2130-2137.

10. S. Baluja and M. Covell, "Learning to Hash: Forgiving Hash Functions and Applications," *Data Mining and Knowledge Discovery,* vol. 17, no. 3, 2008, pp. 402-430.
11. Y.-H. Kuo et al., "Query Expansion for Hash-Based Image Object Retrieval," *Proc. 17th ACM Int'l Conf. Multimedia* (MM 09), ACM Press, 2009, pp. 65-74.
12. H. Lejsek et al., "NV-Tree: An Efficient Disk-Based Index for Approximate Search in Very Large High-Dimensional Collections," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 31, no. 5, 2009, pp. 869-883.

***Tiejun Huang*** *is a professor in the School of Electrical Engineering and Computer Science (EE & CS), and deputy director of the National Engineering Laboratory for Video Technology, at Peking University, China. His research interests include image understanding, video coding, digital libraries, and digital copyright management. Huang received a PhD in pattern recognition and intelligent systems from Huazhong University of Science and Technology. He is a member of IEEE and the ACM. Contact him at tjhuang@pku.edu.cn.*

***Yonghong Tian****, the corresponding author for this article, is an associate professor in the School of EE & CS at Peking University. His research interests include machine learning and multimedia content analysis, retrieval, and copyright management. Tian received a PhD in computer applications from the Institute of Computing Technology, Chinese Academy of Sciences. He is a senior member of IEEE. Contact him at yhtian@pku.edu.cn.*

***Wen Gao*** *is a professor in the School of EE & CS, and director of the National Engineering Laboratory for Video Technology, at Peking University. His research extends to all fields of digital media technology. Gao received a PhD in electronic engineering from the University of Tokyo. He is an IEEE Fellow. Contact him at wgao@pku.edu.cn.*

***Jian Lu*** *is vice president of multimedia technology at Shanda Interactive Entertainment and an adjunct professor at the National Engineering Laboratory for Video Technology at Peking University. His research interests include multimedia content identification and search, and copyright management. Lu received a PhD in electrical engineering from Dartmouth College. He is a senior member of IEEE. Contact him at jian@computer.org.*