Generating Robust Digital Signature for Image/Video Authentication

Ching-Yung Lin and Shih-Fu Chang

Department of Electrical Engineering
and New Media Technology Center
Columbia University
New York, NY 10027, USA

{cylin, sfchang}@ctr.columbia.edu

ABSTRACT

Image/video authentication techniques protect the recipients against malicious forgery. In this paper, we describe an image authentication technique that verifies the originality of the received images. The authentication signature can distinguish content-changing manipulations (such as pixel replacing) from content-preserving manipulations (such as JPEG compression). We also propose a video authentication method that generates robust signatures for compressed video. The signatures can survive some of the transcoding process of MPEG.

KEYWORDS

authentication, watermark, digital signature, manipulation, transcoding

1. INTRODUCTION

The concept of content-based image/video authentication builds upon the increasing need for trustworthy digital multimedia data in commerce, industry, defense, *etc*. Digital media become popular in the past few years partly because of their efficiency of manipulation. Editing or modifying the content of a digital image or video can be done efficiently and seamlessly. However, these advantages decrease the credibility of digital data. To ensure trustworthiness, content-based image/video authentication techniques are needed for verifying the originality of video content and preventing forgery [1]. Observers require them to verify either the "*reality*" of images/videos of natural events or the "*intactness*" of artificial images/videos such as motion pictures, film, *etc*.

The proof of the "*reality*" of a video clip or an image can be provided only by the digital camera that took the shot. Similarly, the proof of the "*intactness*" of a received image/video should be provided by the producer. A signature, which conveys the identification of the camera or the producer and is relative to the contents, can be the proof. Image/video authentication techniques are based on two methods: embedded watermark and external digital signature. Embedding a watermark in the image/video is equivalent to signing a specific digital producer identification (signature) on the content of images/videos [2,3]. Once the image/video is manipulated, this watermark will be destroyed such that the authenticator can examine it to verify the originality of contents. Another approach generates a content-based digital signature which includes the important information of contents and the exclusive producer identification [4-10]. The signature is generated by a producer-specific private key such that it can not be forged. Therefore, the authenticator can verify a received image/video by examining whether its contents match the information conveyed in the signature [4].

Today, most digital multimedia data are stored or distributed in compressed form. Moreover, to satisfy the various needs of broadcasting, storage and transmission, some transcoding of compressed digital images/videos may be required [11,12]. For instance, digital video clips are usually shot and stored in the compressed format with a pre-determined bit-rate. But the final distributed bit rate of them may be different. Another example is that digital images shot and stored in one format may need to be distributed in different formats. These transcoding processes change the pixel values of the digital image/video but not its content. Therefore, these processes should not alter the authenticity of the data.

Robustness is an important concern in developing multimedia authentication techniques. Without robustness, an authentication method can only *verify* the images/videos at the final stage of transcoding processes, but not *authenticate* them. In other words, unless we trust all the transcoders in the processes, the "*reality*" or the "*intactness*" of the multimedia data cannot be proven without robust signatures.

Robustness consideration for authentication is different from that for general watermarking techniques [13-15]. Watermarks used for copyright protection are expected to be robust to most manipulations. But authentication signatures are expected to survive only acceptable transcoding or compression and reject other manipulations.

Of the two authentication methods, the embedded watermarking method is more convenient but usually does not work well with lossy compression. The watermarks are either too fragile for compression or too flexible for manipulations. In other words, a watermarking method that can reliably distinguish compression from other manipulations still has not been found. The external signature method is not as efficient because anyone who needs to authenticate the received image/video has to request the source to provide the signature. But since the signatures remain untouched when the pixel values of the images/videos are changed, they provide a better prospect for achieving robustness.

In this paper, we describe an effective technique for content-based image/video authentication that is based on the robust authentication signature we proposed in [8-10]. This signature can survive JPEG compression, because the content-based information included in the signatures is invariant before and after JPEG compression. The proposed video authentication signature is also robust to some of the transcoding process of MPEG.

Section 2 describes the proposed robust image authentication system and its characteristics. Section 3 shows the process of generating robust signatures. Section 4 describes the authenticator. In Section 5, we describe two methods to enhance the performance of the authentication system. Section 6 shows the robustness of this robust digital signature. In Section 7, we show some experimental results of the image authentication system. Section 8 describes the common transcoding processes of MPEG compressed videos and a robust video authentication system. We present a brief conclusion in Section 9.

2. Image Authentication System

The proposed method is shown in Figure 1. Our method uses a concept similar to that of the digital signature method proposed by Friedman [4], but their technique doesn't survive lossy compression. A signature and an image are generated at the same time. The signature is an encrypted form of the feature codes or

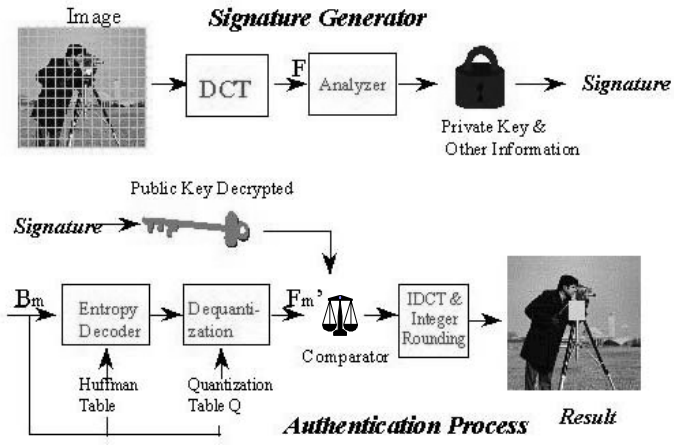

Figure 1: Signature Generator and Image Authentication Process

hashes of this image, and it is stored separately. Once a user needs to authenticate the image he receives, he should decrypt this signature and compare the feature codes (or hash values) of this image to their corresponding values in the original signature. If they match, this image can be claimed to be "authentic". The most important difference between our method and Friedman's "trustworthy camera" is that we use invariance properties in JPEG lossy compression as robust feature codes instead of using hashes of the raw images.

3. Signature Generation

The generation of a signature can be divided into two parts: feature extraction and feature encryption. Feature extraction is the core problem of this paper. From the compression process of JPEG, we have found that some quantitative invariants or predictable properties can be extracted.

Because all DCT coefficient matrices are divided by the same quantization table in the JPEG compression process, the relationship between two DCT coefficients of the same coordinate position should remain the same after the quantization process. Furthermore, due to the rounding effect after quantization, the relationship of the two may be the same or become equal. For instance, if one coefficient $Fp(n)$ in the position n of block p is larger than the other coefficient $Fq(n)$ in the position n of block q, then after compression, their relationship, $Fp'(n) \geq Fq'(n)$, where $Fp'(n) = Integer\ Round\ (Fp(n)/Q) \cdot Q$ and $Fq'(n) = Integer\ Round\ (Fq(n)/Q) \cdot Q$, is guaranteed. It can be

Theorem 1:

- if $Fp(n) > Fq(n)$ then $Fp'(n) \geq Fq'(n)$,
- if $Fp(n) < Fq(n)$ then $Fp'(n) \leq Fq'(n)$,
- if $Fp(n) = Fq(n)$ then $Fp'(n) = Fq'(n)$.

summarized as Theorem 1:

This property holds for any number of decoding and re-encoding processes.

The signature generation process is as follows: Each 8x8 block of an image captured directly by a digital camera, a digital camcorder, or computer graphic software is transformed to the DCT coefficients, and sent to the image analyzer. The feature codes are generated according to two controllable parameters in the analyzer: mapping function, *W*, and selected positions, *b*, in the DCT domain. Given a block *p* in an image, the mapping function is used for selecting the other block to form a block pair, *i.e.*, $q = W(p)$. A coefficient position set, *b*, is used to indicate which positions in a 8x8 block are selected. The feature codes of the image records the relationship of the difference value, $Fp(n)$-$Fq(n)$, and *zero*, at the *b* selected positions. If the difference is larger than or equal to *zero*, a bit *1* is represented; otherwise, a bit *0* is recorded. This process is applied to all blocks to ensure the whole image is protected. (*i.e.*, each block has to be, at least, in a block pair.) In the last step, the feature codes are encrypted with a private key by using the Public Key Encryption method [4]. More detailed descriptions of the signature generation process are in [10].

4. Authentication Process

The procedure of authentication is also shown in Fig. 1. Given a signature derived from the original image and a JPEG compressed image bitstream, *Bm*, for authentication, at the first step, we have to decrypt the signature and reconstruct DCT coefficients from *Bm*. Because the feature codes decrypted from the signature record the relationship of the difference values and *zero*, they indicate the sign of the difference of DCT coefficients, despite the changes of the coefficients incurred by lossy JPEG compression. If these constraints are not satisfied, we can claim that this image has been manipulated by another method.

5. Performance Enhancement

5.1 Tolerance bound for recompressing noise

Rounding noises may be added during the JPEG compression process and they may cause false alarm. In practice, computer software and hardware calculate the DCT with finite precision. Because the error may accumulate throughout the multiple recompression processes, we have to introduce some tolerance bounds to prevent the authenticator from reporting some *false alarm* in the accepted recompression process. If we assign a tolerance bound, τ, to the authentication system, then the following property,

- if $Fp(n) \geq Fq(n)$ then $Fp'(n) - Fq'(n) \geq -\tau$,
- if $Fp(n) < Fq(n)$ then $Fp'(n) - Fq'(n) \leq \tau$,

should be considered as acceptable value changes in the authenticator.

5.2 Multi-layer feature codes

Given two DCT coefficients at the same positions of two blocks, not only their relationship after compression is constrained, but also the range of their difference after compression is limited. Defining *Qp* and *Qq* as the quantization matrix of the block *p* and *q*, respectively, the following theorem must be satisfied:

Theorem 2:

- if $Fp(n)-Fq(n) \geq k$ then
  $Fp'(n)-Fq'(n) \geq k - 1/2 \cdot (Qp(n)+Qq(n))$,
- if $Fp(n)-Fq(n) < k$ then
  $Fp'(n)-Fq'(n) \leq k + 1/2 \cdot (Qp(n)+Qq(n))$

Applying Theorem 2, we can use multi-layer feature codes to protect the DCT difference values within more precise ranges. For instance, the r-th layer feature codes record the relationship of the difference value, $Fp(n)$-$Fq(n)$, and a threshold, k_r. Therefore, they indicate the possible ranges of the difference of DCT coefficients, which will be tested in the authenticator.

6. Robustness

The feature codes generated in the Section 3 are based on the characteristics of JPEG compression. With the robust digital signature generated from these feature codes, images may be compressed and decompressed several times and still considered as authentic.

In some practical applications, some other manipulations are also considered acceptable, such as intensity enhancement, scaling, cropping, file format transformation, *etc*. These acceptable manipulations can be either pre-determined by the signature generator with special consideration on the controllable parameters, or decided by the authenticator with case-dependent tolerance bound. The methods for achieving robustness to these manipulations are discussed as follows:

- **Intensity enhancement:**
  If a constant intensity change is applied to the whole image, it only changes the DC values of all the 8x8 DCT blocks. Because the authenticator compares the difference of DCT coefficients, this manipulation will be considered as acceptable. On the other hand, if the authenticator wants to reject it or limit the range of change, we can include the mean value of all DC coefficients in the signature such that the authenticator can reject large intensity changes.
- **Cropping:**

In most situations, cropping only selects a part of the image, such that it may introduce a different visual meaning to the cropped image. However, if this manipulation is allowed in some situations, we can design a robust signature with carefully selected mapping function. For instance, we can select block pairs from adjacent blocks. Then, the feature codes of those cropped blocks can be found in the original signature. In practical situations, the cropped image has to provide its related location on the original image to the authenticator. Because the origin point of the cropped image may not be at the grid points of the original image, (*i.e.*, each 8x8 block in the cropped image may cover parts of four 8x8 original blocks), the authenticator can only verify the cropped image excluding its boundary pixels. In this case, the recompressing process may introduce different variations of pixels, from recompressing the original image. Therefore, some tolerance may be needed in this situation.

- **Scaling:**
  Scaling is a common operation on the images, which is accepted in many situations. For instance, a scanner may scan an image with a high resolution. This image may be down-sampled to an appropriate size later. In the scaling cases, the signature generator has to record the original size of the image. An authenticator can re-scale this scaled image to its original size before general authentication processes. Because the DCT transformations are linear and the difference in the pixel values of the original and the re-scaled image should not be too great, there will be no large changes in the DCT coefficients. Similar to the general recompression noise, these changes can be also considered as some kinds of noise that can be solved by allowing larger tolerance values in the authenticator.
- **Format transformation with other lossy compressions:**
  Other lossy compressions such as wavelet-based methods or color space decimation methods can be considered as introducing noises to the original image. Similarly, we can use larger tolerances in the authenticator to allow these lossy compressions.
- **Filtering and other operations:**
  Filtering, such as low-pass filtering and edge enhancement, may probably change more visual meaning of images. The authenticator would be hard to deal with these operations. However, if the change in pixel values is not too great, we can still consider them as some kind of noise and use larger tolerance values. This method can also be applied to other operations.

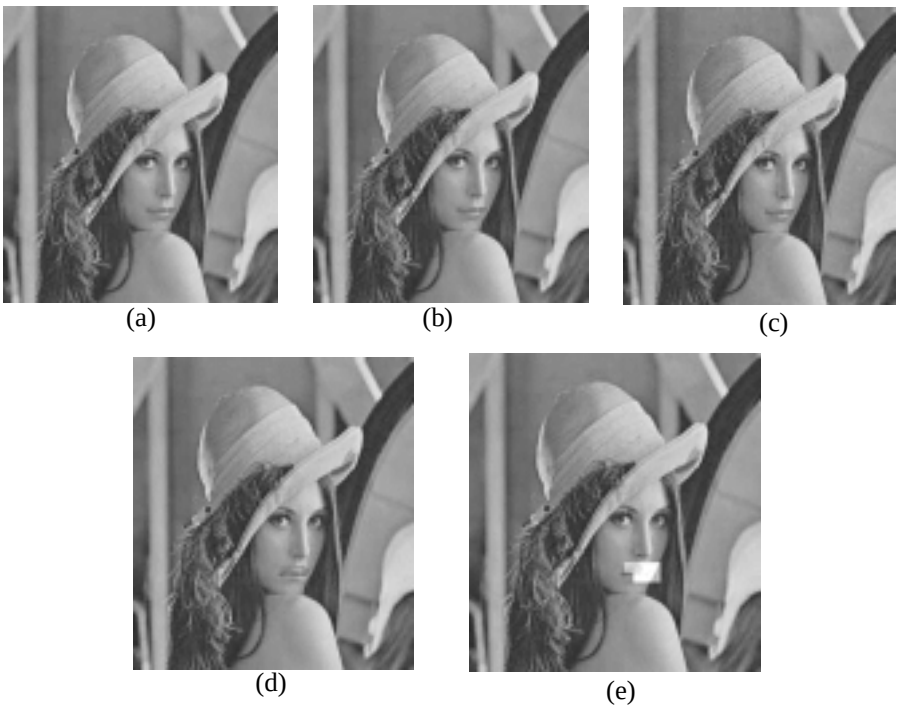

Figure 2: Experimental Results: (a) original image, (b) 9:1 JPEG compressed, (c) 9:1 JPEG recompressed from a 6:1 compressed image, (d) manipulated image, (e) authentication result of the manipulated image.

7. Experimental Results

The 'Lenna' image is compressed with a compression ratio of 9:1. The authentication signature is generated based on the original image. The compressed bitstream is sent to the system for authentication. As predicted, the authenti-cator will verify the compressed image as authentic and decompress this image perfectly. The authentication result is shown in Fig. 2(b).

The original image is compressed with a compression ratio 6:1. Then, this image is decompressed by Photoshop 3.0, rounded to integral values, and recompressed into an image with compression ratio 9:1. In this case, the recompression process (9:1) does not trigger the manipulation detector and the final compressed image is still verified as authentic. The final decoded image is similar to Fig. 2(c).

In the third experiment, we flipped the mouth area of the image. It is shown in Fig. 2(d), with its authentication result shown in Fig. 2(e). It can be clearly shown that the manipulated part has been detected as fake and highlighted by the authenticator.

8. Video Authentication System

Similar to the image authentication system, a video authentication signature has to be robust to the transcoding processes. Regardless of the format transformation between different compression standards (such as MPEG-1, MPEG-2, H.261 and H.263), five transcoding processes may be applied to the compressed video [16,17]:

1. Dynamic Rate Shaping [18,19]: A real-time rate-control scheme in the compressed domain. This technique sets dynamic control points to drop the high-frequency DCT coefficients on each 8x8 block in a macroblock. Motion vectors are not changed.
2. Rate Control without Drift Error Correction [20,21]: This technique is also applied in the compressed domain. DCT coefficients are re-quantized to satisfy different bit-rate constraint. Motion vectors are not changed.
3. Rate Control with Drift Error Correction [16]: This technique improves the video quality, but it needs more computations. DCT coefficients of the residue of intercoded blocks are changed to satisfy the change of the re-quantized intracoded blocks. Motion vectors are not changed in this case.
4. Transcoding with Mostly Consistent Frame Types [16,17,23]: The frame types (I, P and B), are kept unchanged in each generation. It may be used in creating a new sequence by cutting and pasting several video segments with consistent GOP units within each segment except the frames at the boundary.
5. Transcoding with Inconsistent Frame Types [16]: In some editing process, the compressed videos are transformed to the uncompressed bitstreams which are then re-encoded. The GOP structures of frames and the motion vectors may change in this case.

Video authentication signatures can be generated for different situations. For instance, to generate a signature that is robust to situations 1, 2 and 4, we can use the DCT coefficients of the luminance and chromatic matrices in each macroblock to generate the comparison pairs. Since the *quantization_scale* is specified for each macro-block [25], the relative relationships of the coefficients are invariant during transcoding. Therefore, similar to the signature generation process of images, we can use them to generate the feature codes. If a more flexible choice of comparison pair is necessary, the authentication system can generate signatures based on the criteria we have proposed in [9,10]. It should be noted that, in situation 4, the frames in the boundary of video segmentations cannot be verified by this method.

Because the drift error correction process changes the DCT coefficient values, statistical models of the changes can be used to provide tolerance bounds for the coefficient relationships, similar to that described in [10].

Situation 5 poses the most challenging case for authentication. The GOP structure in the video is changed and so is the relationship of DCT coefficients among blocks. The design scheme for generating a robust signature in this situation is still under study.

A more detailed description of the content-based video authentication techniques will be shown in [26].

9. CONCLUSION

In this paper, we have described a method for robust image/video authentication. Robust signatures can distinguish the JPEG lossy baseline compression from other malicious manipulations for images, and the Rate-Control Coding from other manipulations for compressed videos. Our analytic and empirical performance analyses have shown the effectiveness of the image authentication system and presented a possible direction for further video authentication research.

10. REFERENCES

[1] Bearman, D., and Trant, J. *Authenticity of Digital Resources: Towards a Statement of Requirements in the Research Process*. D-Lib Magazine, June 1998.

[2] Yeung, M. and Mintzer, F. *An Invisible Watermarking Technique for Image Verification.* Proc. Of ICIP, Santa Barbara, CA, USA, Oct. 1997.

[3] Lin, C.-Y. and Chang, S.-F. *A Watermark-Based Robust Image Authentication Method Using Wavelets.* ADVENT Report, Columbia University, Apr. 1998. http://www.ctr.columbia.edu/~cylin/pub/a98wav.doc

[4] Friedman, G.L. *The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image*. IEEE Trans. on Consumer Electronics, Vol.39, No.4, pp.905-910, Nov. 1993.

[5] Quisquater, J.-J., Macq, B., Joye, M., Degand, N. and Bernard, A. *Practical Solution to Authentication of Images with a Secure Camera.* SPIE Storage and Retrieval for Image and Video Databases, San Jose, CA, USA, Feb. 1997. pp.290-297.

[6] Gennaro, R., and Rohatgi, P. *How to Sign Digital Streams.* CRYPTO '97, Santa Barbara, CA, USA, August 1997, pp.180-197.

[7] Gennaro, R., Krawczyk, H. and Rabin, T. *RSA-based Undeniable Signatures.* CRYPTO '97, Santa Barbara, CA, USA, August 1997, pp. 132-149.

[8] Lin, C.-Y. and Chang, S.-F. *A Robust Image Authentication Method Surviving JPEG Lossy Compression*. SPIE Storage and Retrieval for Image and Video Databases, San Jose,CA, USA, Jan. 1998.

[9] Lin, C.-Y. and Chang, S.-F. *An Image Authenticator Surviving DCT-based Variable Quantization Table Compression.* CU/CTR Technical Report 490-98-24, Nov. 1997.

[10] Lin, C.-Y. and Chang, S.-F. *A Robust Image Authentication Method Distinguishing JPEG Compression from Malicious Manipulation*. CU/CTR Technical Report 486-97-19, Dec. 1997. http://www.ctr.columbia.edu/~cylin/pub/authpaper.ps

[11] Wells, N.D. *The Atlantic Project: Models for programme production and distribution.* Proceedings of the European Conference on Multimedia Applications Services and Techniques (ECMAST 96), Louvaine-la-Neuve, Belgium, May 1996, pp. 243-253.

[12] Brightwell, P.J., Dancer, S.J. and Knee, M.J. *Flexible Switching and Editing of MPEG-2 Video Bitstreams.* International Broadcasting Convention (IBC 97), Amsterdam, Sep. 1996, pp. 547-552.

[13] Cox, I., Kilian, J., Leighton, T., and Shamoon, T. *Secure Spread Spectrum Watermarking for Multimedia.* NEC Research Institute Technical Report, 95-10, 1995.

[14] Braudaway, G.W., Magerlein, K.A. and Mintzer, F. *Protecting Publicly-Available Images with a Visible Image Watermark.* IBM Research Division, T.J. Watson Research Center, Technical Report 96A000248, 1996.

[15] Meng, J. and Chang, S.-F. *Embedding Visible Watermarks in the Compressed Domain.* IEEE International Conference on Image Processing (ICIP 98), Chicago, IL, USA, Oct. 1998.

[16] Tudor, P.N. and Werner, O.H. *Real-Time Transcoding of MPEG-2 Video Bit Streams.* International Broadcasting Convention (IBC 97), Amsterdam, Netherlands, Sep. 1997, pp. 286-301.

[17] Werner, O.H. *Generic Quantiser for Transcoding of Hybrid Video.* Proceedings of the 1997 Picture Coding Symposium, Berlin, Germany, Sep 1997.

[18] Eleftheriadis, A. and Anastassiou, D. *Constrained and General Dynamic Rate Shaping of Compressed Digital Video.* Proceedings of the 2nd IEEE International Conference on Image Processing (ICIP 95), Arlington, VA, USA, Oct. 1995.

[19] Jacobs, S. and Eleftheriadis, A. *Straming Video using Dynamic Rate Shaping and TCP Flow Control.* Visual Communication and Image Representation Journal, Jan. 1998.

[20] Viscito, E. and Gonzales, C. *A Video Compression Algorithm with Adaptive Bit Allocation and Quantization.* SPIE Vol. 1605 Visual Communications and Image Processing '91.

[21] Ding, W. and Liu, B. *Rate Control of MPEG Video Coding and Recording by Rate-Quantization Modeling.* IEEE Trans. on Circuits and Systems for Video Technology, Vol. 6, No. 1, pp.12-19, Feb. 1996.

[22] Meng, J. and Chang, S.-F. *Tools for Compressed-Domain Video Indexing and Editing.* SPIE Conference on Storage and Retrieval for Image and Video Database, Vol. 2670, San Jose, CA, USA, Feb. 1996.

[23] Meng, J. and Chang, S.-F. *CVEPS – A Compressed Video Editing and Parsing System.* Proceedings of ACM Multimedia 96, Boston, MA, USA, Nov. 1996.

[24] Chang, S.-F. and Messerschmitt, D. G. *Manipulation and Compositing of MC-DCT Compressed Video.* IEEE Journal of Selected Areas in Communications, Vol. 13, No. 1, pp.1-11, Jan. 1995.

[25] Haskell, B.G., Puri, A. and Netravali, A.N. *Digital Video: An Introduction to MPEG-2.* Chapman and Hall, 1997.

[26] Lin, C.-Y. and Chang, S. F. *Issues and Solutions for Authenticating MPEG Video.* SPIE Storage and Retrieval for Image and Video Databases, San Jose, CA, USA, Jan. 1999.