FAST AND ROBUST IDENTIFICATION METHODS FOR JPEG IMAGES WITH VARIOUS COMPRESSION RATIOS

Fitri Arnia, Ikue IIZUKA, Masaaki FUJIYOSHI, and Hitoshi KIYA

Department of Electrical Engineering, Tokyo Metropolitan University
1-1 Minamiosawa Hachioji-shi, Tokyo 192–0397, Japan.

ABSTRACT

We propose two identification methods for JPEG-coded images. The purposes are to identify the images that are compressed from the same original image with various compression ratios in fast and robust manner. The first approach can avoid identification leakage or false negative (FN), and could result in a few false positives (FP). The second approach can avoid both FN and FP, with a slightly longer processing time. By combining the two schemes, a faster and a more perfect identification can be achieved, in which FN and FP can be avoided.

1. INTRODUCTION

A large volume of digital images are available recently, and many of them are compressed using JPEG (Joint Photographic Experts Group) standard [1], with various compression ratios. As a consequence, JPEG image identification is becoming increasingly important, and identification methods should require the following properties: (1) able to identify the images in short time, and (2) robust to JPEG compression in which JPEG images that are compressed from the same original image with various compression ratios can be identified without resulting in false negative (FN) and false positive (FP). FN or leakage refers to a failure of identifying images that are compressed from the same original image; and FP occurs when images compressed from different original images are also identified.

Although some methods have been introduced for image authentication and recognition [2, 3, 4], and retrieval or indexing [1, 5, 6, 7, 8], they are not sufficient to be used as identification means which fulfilling the requirement described previously. Because additional calculation for obtaining feature [1, 2, 3, 4, 5, 6, 7, 8] are required to achieve a higher identification accuracy. As a result, longer processing time might be required. In addition to that, those methods may result in a leakage and/or FP when accomplishing the identification between compressed images [2, 4, 5], since the effect of the JPEG compression had not been considered in advance.

In this paper, we propose two novel identification methods that can identify the images by considering two previous requirements, namely: without any additional effort for calculating features and without resulting in FN and FP. In the first method, the identification of two images is accomplished by comparing the positive and negative signs of quantized DCT (Discrete Cosine Transform) coefficients. These signs can be accessed without necessity of full JPEG decoding, thus a low processing time can be achieved. The second method utilizes the ratio of the quantized DCT coefficients of the two images and the ratio of their corresponding quantization tables. The first method is able to identify the images without FN, and could result in a few occurrences of FP. The second method can identify the images without FN and FP, and it requires a slightly higher processing time. By combining both approaches, a lower computation time and a more perfect identification can be achieved.

2. BACKGROUND

2.1. Image Identification Model

Let us consider there are two or more compressed images, which have different or the same compression ratios. Those images are originated from the same image and compressed by the same compression method. In this paper, the identification of those images is referred to as image identification. In other words, if the images do not originate from the same image, or are not compressed by the same compression method, they are unidentifiable from each other.

A simplified model of the image identification system is shown in Fig. 1. The system consists of three components, namely, a client (user), an image identifier, and a database. The database may contain various types of data, such as compressed images, parameter (feature) of the images, and image information (metadata). An identification is initialized by the user, by sending a query, which can be any kind of the data mentioned above to the image identifier. Then the image identifier checks the availability of the query in the database. Afterwards, if the query information is available, it can be directly sent or confirmed to the user. In this paper, the identification methods are focused on querying by image.

2.2. Applications

There are numerous applications for the previously mentioned identification model. Some examples are described in the following:
a. Security.
In a compressed image environment, it is important to identify any alterations in images caused by disturbances or alterations other than the compression itself. For instance, identifying the presence of malicious attacks, such as intentional cropping and addition or removal of objects [9].
b. Evaluation of Image Quality.
In image and video communications, a slight quality degradation due to compression noise is commonly accepted. However, the image quality degradation due to other causes, such as transmission and decoding errors, are usually unacceptable. A method to identify those errors in a fast and automatic way is required in such applications.
c. Evaluation of Image Validity.
Let us consider two images of the same scene, for example: chest X-ray images of two patients. Those images may have been labelled by name, date, or content description. However, this approach is very sensitive to human error, such as mislabelling. The mislabelled images can cause a misdiagnosis, which in turn could threaten the patient's life. Therefore, a more efficient and save method to guarantee the image validity is required.

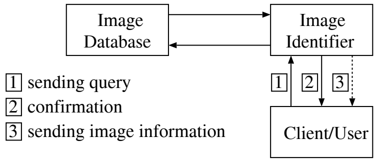

Fig. 1. Process of image identification

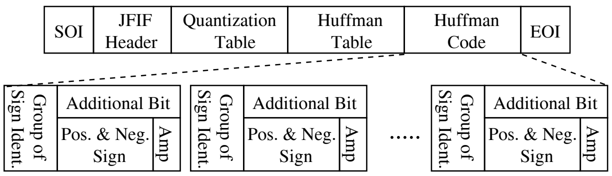

Fig. 2. JPEG Bitstream.

d. Image Information Retrieval.
In addition to image querying to obtain identical image, image querying to obtain image information (metadata) is comparably important. For the images, the metadata may include: photographer's name, image format, and date and time. The digital library is one area where metadata identification is important.

2.3. JPEG Bitstream Structure

The JPEG compresses images by a sequence of operations [10], namely (a) tiling the images into blocks of 8×8 pixels, (b) DCT transform, (c) quantization and (d) entropy encoding. The output of entropy encoder is JPEG bistream. Simplified structure of greyscale image bitstream is outlined in Fig. 2. The Start of Image (SOI) marker is the first marker in the JPEG stream. The JPEG File Interchange Format (JFIF) header is a minimal file format that enables the JPEG bitstream to be exchanged between a wide variety of platforms and applications. The next two entries are the tables that are required to decode the image, namely: quantization and Huffman tables. The Huffman code contains the image data bitstream. The bitstream can be arranged into groups referred to as groups of sign identifications. In each group, the MSB of additional bits represents the positive and negative signs of the DCT coefficient. Eventually, the End of Image (EOI) marker follows the last byte of the compressed data.

3. PROPOSED IDENTIFICATION METHODS

Two image identification methods and their features are described in this section. The JPEG images considered in this research are originated from the same image and compressed by the same method. The proposed methods utilize the following properties:
a. The first method
The positive and negative signs of the quantized DCT coefficients of the two images are equivalent in the corresponding location, even though their Q-factors (relating to compression ratios) are different. This condition does not apply for zero-valued coefficients.
b. The second method
The ratio of the quantized DCT coefficients of the two images in the corresponding location exists in a constant range based on the ratio of the quantization step sizes.

3.1. Notations and Terminologies

Several notations and terminologies used in the following sections are listed here.
• x represents an image. x can be: "A" for image A, "B" for image B, and "O" for the original image.
• M represents the number of constituent 8 × 8 block of an image. $0 \leq m \leq (M-1)$.
• N represents the number of DCT coefficient used for identification in each 8 × 8 block. $0 \leq N \leq 63$.
• The n^{th} DCT coefficient in the m^{th} block in image x is abbreviated as $c_x(m,n)$.
• The sign of the n^{th} DCT coefficient in the m^{th} block in image x is abbreviated as sgn $(c_x(m,n))$, and is derived based on the following equation:

$$\text{sgn}(a) = \begin{cases} -1, & a < 0 \\ 0, & a = 0 \\ 1, & a > 0. \end{cases} \tag{1}$$

• The quantization table for image x is represented as $\mathbf{q}_x = \{q_x(n) \mid n = 0, 1, \ldots, 63\}$.

3.2. First Identification Method

In the first method, the positive and negative signs of DCT coefficients are obtained by entropy-decoding the bitstream. Those signs of the AC coefficients are available in the bitstream and specified by the MSB of an additional bit group as shown in Fig. 2.

3.2.1. Identification Algorithm

Let us define image A as a JPEG-compressed image that is given by a user (a query image) and image D is a JPEG-compressed image that is a member of database $\mathbf{D}$, where $D \in \mathbf{D}$. The positive and negative signs of the quantized DCT coefficients of images A and D in the corresponding locations are compared, and the results are used to decide whether the images are compressed from the same original image. In each block m, not all 64 coefficients are required for comparison, the first N coefficients are sufficient. The first algorithm is accomplished according to the following steps:
1. Set the value of N.
2. Set $m := 0$.
3. Set $n := 0$.
4. For the m^{th} block in image A, extract the positive and negative signs sgn $(c_A(m,n))$ of the n^{th} quantized DCT coefficient $c_A(m,n)$. The sign is obtained by using Eq.(1). If sgn $(c_A(m,n)) = 0$, proceed to step 7.
5. For the m^{th} block in a database image D, extract the positive and negative signs sgn $(c_D(m,n))$ of the n^{th} quantized DCT coefficient $c_D(m,n)$. If sgn$(c_D(m,n)) = 0$, proceed to step 7.
6. If sgn $(c_A(m,n)) \neq$ sgn $(c_D(m,n))$, the algorithm decides that image A and B were not compressed from the same original image, and the process is halted. Otherwise, proceed to step 7.
7. Set $n := n + 1$.
8. If $n = N$, set $m := m + 1$, and proceed to step 9. Otherwise, continue to step 4.
9. If $m = M$, it is decided that image A has the same original image as image D. Otherwise, continue to step 3.

3.2.2. Features of First Identification Method

The main features of the first identification method are:
A. Robustness
The method is robust to JPEG compression because the quantization does not alter the signs of DCT coefficients. As a result, the images whose original image is the same but compression ratios are different still have equal pattern of DCT signs, regardless the zero-valued coefficients. Furthermore, because the method omits the zero-valued coefficients, the different number of zero-valued coefficients of the images does not effect the identification performance.
B. Low processing time
A shorter processing time can be achieved primarily due to (a) the positive and negative signs can be acquired by only entropy decoding the bitstream, invers quantization and invers DCT are not

required, (b) the algorithm stop the comparison in current image D and start comparison for a new image D from the database **D** if sgn $(c_A(m,n)) \neq$ sgn $(c_D(m,n))$, (c) instead of all 64 coefficients, N coefficients are sufficient to compare in each block, and (d) the algorithm omits the sign comparison if sgn $(c_A(m,n)) = 0$ or sgn $(c_B(m,n)) = 0$.

3.3. Second Identification Method

In the second identification method, the DC and AC coefficients in a quantized DCT domain are obtained by entropy-decoding the bitstream. The quantization table can be directly extracted from the bitstream.

3.3.1. Identification Algorithm

The second algorithm is accomplished according to the following steps:

1. Extract the quantization table $\mathbf{q}_A = \{q_A(n) \mid n = 0, 1, \ldots, 63\}$ of image A, and the quantization table $\mathbf{q}_B = \{q_B(n) \mid n = 0, 1, \ldots, 63\}$ of image B.

2. Set $m := 0$.

3. Set $n := 0$.

4. Then, entropy decode the n^{th} quantized DCT coefficient $c_A(m,n)$ of the m^{th} block in image A and the n^{th} quantized DCT coefficient $c_B(m,n)$ of the m^{th} block in image B.

5. If Eqs. (2) and (3) are fulfilled at the same time, proceed to step 7. If Eqs. (2) and (4) are satisfied at the same time, the algorithm decides that the images do not have the same original image, and the algorithm is stopped. Otherwise proceed to step 6.

$$q_A(n) < q_B(n) \tag{2}$$

$$|c_A(m,n)| > 0 \quad \& \quad |c_B(m,n)| = 0 \tag{3}$$

$$c_B(m,n) \neq 0 \quad \& \quad c_A(m,n) = 0. \tag{4}$$

where $\&$ stands for the logical and operation.

6. If Eq. (5), is satisfied, proceed to step 7. Otherwise, it means that images A and B do not have the same original image, and the identification is stopped.

$$\frac{c_A(m,n) - 0.5}{c_B(m,n) + 0.5} < \frac{q_B(n)}{q_A(n)} < \frac{c_A(m,n) + 0.5}{c_B(m,n) - 0.5}. \tag{5}$$

7. Set $n := n + 1$.

8. If $n \leq 63$, proceed to step 4. Otherwise, update $m := m + 1$ and continue to step 9.

9. If $m = M$, it is decided that image A and image B are compressed from the same original image, and the process is ended. Otherwise, continue to step 3.

3.3.2. Features of Second Identification Algorithm

The main features of the second identification method are:

A. Robustness

It turns out that two images compressed from the same original image but have different compression ratios are connected by Eq. (8), and not their compression ratios (Q-factors). Using the restriction formulated in Eq. (5), the compressed images that have the same origin are robustly identified. Moreover, disregarding zero-valued coefficients ($c_A(m,n) = 0$ or $c_B(m,n) = 0$), makes this scheme able to identify the images having different compression ratios.

B. Low processing time

The identification can be accomplished with a low processing time because it requires only the quantized DCT coefficients (invers quantization and invers DCT are excluded). Moreover, notice that many of the quantized DCT coefficients are zero-valued. Ignoring zero-valued coefficients can decrease the processing time.

Table 1. Simulation conditions.

Image 1	Football video sequence (gray scale, 704 × 240, 8 bpp, 194 frame x 10)	
Image 2	Paris Video Sequence (gray scale, 352 × 288, 8 bpp, 60 frame x 10)	
Image 3	Claire Video Sequence (gray scale, 360 × 288, 8 bpp, 168 frame x 10)	
Image 4	Mobile Video Sequence (gray scale, 352 × 288, 8 bpp, 300 frame x 10)	
Q-factor	50, 200, 350, 500, 650 800, 950, 1100, 1250, 1400	
Coeffi cient	$N = \{4, 64\}$	
Query Frame	No. 31 for ”Football” No. 1 for ”Claire”	No. 30 for ”Paris” No. 80 for ”Mobile”

3.4. Combination of First and Second Methods

Let us consider image A as a JPEG-compressed image that is given by a user (the same as the one queried to the first algorithm) and images D are output of the first algorithm. The images D are then grouped in database **B** and may or may not contain false positives. If it is required to removed false positives occurrence, further identification can be performed by querying image A to the second algorithm to identify all images B compressed from the same original image, where $B \in$ **B**.

4. SIMULATION

To evaluate the performance of the proposed methods, several simulations were conducted. For comparison purposes, we incorporated two techniques in pixel domain, namely correlation and mean square error (MSE)[12].

4.1. Simulation Conditions

The simulation conditions are presented in Table. 1. Four video sequences, were used in the simulation. All frames of all sequences were compressed with 10 different quality factors (Q-factor). The original uncompressed versions were not included in the simulation. From each block m, the DCT coefficients of $N = 4$ and $N = 64$, which are obtained according to a zigzag scan, were used. Identification was accomplished by querying a compressed frame, for example: “Paris”frame No. 30, Q-factor 50, at a time. The same frame with other different compression ratios, for example: frame No. 30, $200 \leq$ Q-factor ≤ 1400, were subsequently queried to the system. The same procedures were also applied to the other sequences. Successful identification will identify all 10 images compressed from the same image. In following sections, “Football” frame No. 31 with Q-factor 50 will be referred to as “Football31-50”. The same notation was also applied to other frames with different Q-factors. The simulation is run on a computer with a 1.2 GHz processor, using a JPEG codec package from Portable Video Research Group at Stanford University [11].

4.2. First Algorithm

The results of some querying (Q-factors of 50 and 350) with N = 4 can be seen in Table 2. For ”Football” and ”Paris”, querying with all Q-factors resulted in a perfect identification. Specifically, all 10 images compressed from the same original image were identified. Example of frames that were perfectly identified is shown in Fig. 3. For ”Claire” and ”Mobile”, querying images with Q-factor ≥ 350 resulted in false positives. For both sequences, it is worth noting that there were no false negatives.

The method can identify the targeted images in short time. As an example, the total time required to identify 10 out of 1940 images with the query ”Football31-50”, N = 64, was approximately 0.5 seconds.

(a) "Paris30-50" (b) "Paris30-350"

Fig. 3. Example of frames used in simulations. For first and second methods, querying "Paris" whose Q-factors are between 50 and 1400 resulted in perfect identification, without false positives nor false negatives.

Table 2. Results of querying frames to first method (N = 4). For each Q-factor: column "a" represents "Football", "b" represents "Paris", "c" represents "Claire", and "d" represents "Mobile".

Q-factor	50				350			
Frame	a	b	c	d	a	b	c	d
Total Retrieved Images	10	10	10	10	10	10	15	14
Same Original Images	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	0	0	5	4
False Negatives	0	0	0	0	0	0	0	0

4.3. Second Algorithm

The querying to the second algorithm resulted in neither leakage nor false positives. In other words, only images compressed from the same original image were identified. A slightly higher processing time was required to identify the images. As an illustration, the total time required to identify 10 out of 1940 images with the query "Football31-50", N= 64, was approximately 8 seconds.

4.4. Combination of First and Second Algorithms

Some querying of "Claire" and "Mobile" frames to the first algorithm resulted in false positives. For these frames, the total retrieved images for Q-factor $\geq$ 350 were then collected to form a smaller-sized database **B**. The same query images A of "Claire" (and "Mobile") as used in the first algorithm are then input to the second algorithm to identify images in the database **B**. All the false positives of "Claire" and "Mobile" were removed by the second algorithm. Required processing time of this combined approach would be insignificant since total images in the database **B** are commonly a few in number as compared to the total images in the database **D**.

4.5. Pixel Domain Identification

In general, querying the images using both pixel correlation and MSE resulted in many leakages and false positives. An example of querying results of "Claire1-50" using correlation is shown in Table 3. As can be seen, among the images with the first 10 highest correlation values, only 4 images compressed from the same original image were identified, 6 images were false positives and 6 remaining images were false negatives (not included in Table 3 due to their correlation ranks are far below 10).

It is apparent that although pixel correlation and MSE methods are as considerably simple as the proposed methods (regardless the domain of identification), the proposed methods provided better identification results.

5. CONCLUSION

We have proposed two novel schemes for identifying JPEG-coded images in fast and robust manner. The first approach was able to avoid identification leakages or false negatives and could results in a few false positives. The second approach can avoid both false positives and false negatives, with a slightly higher processing time. By combining both methods, processing time can be reduced and a more perfect identification can be achieved. An extension of the proposed methods to other compression standards such as MPEG and JPEG2000 are considered as future works.

Table 3. Results of querying using correlation. Query: "Claire1-50". False positives and false negatives occurred.

Rank	Retrieved Image	Corr. Value	Note
1	Claire1-50	1	Identified
2	Claire0-50	0.9971	False Positive
3	Claire1-200	0.9966	Identified
4	Claire2-50	0.9948	False Positive
5	Claire0-200	0.9946	False Positive
6	Claire1-350	0.9937	Identified
7	Claire2-200	0.9925	False Positive
8	Claire0-350	0.9915	False Positive
9	Claire3-50	0.9910	False Positive
10	Claire1-500	0.9896	Identified

ACKNOWLEDGMENTS

We would like to thank Mr. Takahiro Fukuhara of Sony Corporation for his useful advice.

6. REFERENCES

[1] A.R. McIntyre and M.I. Heywood, "Exploring content-based image indexing techniques in compressed domain", in Proc. of Canadian Conf. Elect. and Comp. Eng., IEEE, 2002, Vol.2, pp.957-962.

[2] H. Kobayashi and H. Kiya, "Robust Image Authentication Using Hash Function," in Proc. of Region 10 Conference Volume A (TENCON), IEEE, 2004, vol.1, pp. 435-438.

[3] J.J. Eggers and B. Girod, "Blind Watermarking Applied to Image Authentication," in Proc. of Intl. Conf. Acoustics, Speech, and Signal Proc., IEEE, 2001, Vol.3, pp. 1997-1980.

[4] J. Sun, Q. Zhuo, and W. Wang, "The Study of Sparse Features in Image Recognition", in Proc. of Image Extraction, Segmentation and Recognition, SPIE, 2001, Vol. 4550, pp. 79-84.

[5] W. Niblack, R. Barber, W. Equitz, M. Flickner, E. Glassman, D. Petkovic, P. Yanker, C. Faloutsos and G. Taubin "The QBIC Project: Querying images by content using color, texture and shape", Proc. SPIE Intl. Soc. Opt. Eng., Vol. 1908, pp. 173-187, 1993.

[6] M.Shneier and M.Abdel-Mottaleb, "Exploiting the JPEG compression scheme for image retrieval", IEEE Trans. Pattern Anal. Mach. Intell., Vol.18, no.8, pp.849-853, Aug.1996.

[7] J. Jiang, A. Amstrong, G.C. Feng "Web-based image indexing and retrieval in JPEG compressed domain", Multimedia System, Vol.9, pp.424-432, 2004.

[8] M. K. Mandal and C. Liu, "Efficient Image Indexing Techniques in the JPEG2000 Domain", Journal of Electronic Imaging, Vol.13, No.1, pp.182-187, January 2004.

[9] Fitri Arnia, I. Iizuka, Y. Nakano, A. Takagi, and H. Kiya, "Compressed Domain Based Image Authentication for JPEG Coded Images", in Proc. of ITC-CSCC , July 2005, Vol. 4, pp. 1465-1466.

[10] Information Technology - Digital Compression and Coding of Continuous-Tone Still Images - Requirement and Guidelines, ISO/IEC International Standard 10918-1, 1992.

[11] PVRG(Portable Video Research Group)- JPEG CODEC, ver.1.2 (1994)

[12] Z.Wang and E.P.Simoncelli, "Translation Insensitive Image Similarity in Complex Wavelet Domain", Proc. IEEE Intl. Conf. Acoustic, Speech & Signal Proc., March 2005.