PAPER *Special Section on Papers Selected from ITC-CSCC 2005*

Fast Image Identification Methods for JPEG Images with Different Compression Ratios

Fitri Arnia[†a), *Nonmember*, **Ikue IIZUKA**[†], *Student Member*, **Masaaki FUJIYOSHI**[††], *and* **Hitoshi KIYA**[††], *Members*

SUMMARY Two schemes for fast identification of JPEG coded images are proposed in this paper. The aim is to identify the JPEG images that are generated from the same original image and have equivalent or different compression ratios. Fast identification can be achieved since the schemes work on the quantized Discrete Cosine Transform (DCT) domain. It is not required to inverse the quantization and the DCT. Moreover, only a few coefficients are commonly required for identification. The first approach can avoid identification leakage or false negative (FN), and probably result in a few false positives (FP). The second approach can avoid both FN and FP, with a slightly higher 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.

key words: *image identification, fast identification, compression ratio, image database*

1. Introduction

A large amount of digital images are recently extensively available, and many of them are compressed by variously different compression ratios before being stored in a database. These things make research in image identification increasingly more important. In this paper, we focus on research of image identification system that is applied to database and having the following features: (1) it can identify all targeted images without leakage and (2) can identify images without any additional calculation to extract features. The system is suitable for applications such as image security and fast retrieval and identification through the Internet. In this case, the aim of the identification is to identify the images that are compressed using Joint Photographic Experts Group (JPEG) standard [18] from the same original image, and have different or equivalent compression ratios.

Although some works have been proposed for image authentication [1]–[4], recognition [5], and retrieval or indexing [6]–[14], they are not sufficient to be used as identification means which fulfilling the requirement of the system described above. Because additional calculation for obtaining feature [1]–[14] or many features [4], [6], [7] are required to achieve a higher identification accuracy. Those methods may also result in a leakage when accomplishing the indentification between compressed images [1], [5]–[9], [13], since the effect of the compression had not been considered in advance.

In this paper, we propose two image identification methods that can identify the images without leakage and without any additional effort for calculating features [15]–[17]. In the first method, the identification of the two images is accomplished by comparing the positive and negative signs of the quantized Discrete Cosine Transform (DCT) coefficient. These signs can be accessed from the bitstream, without necessity of full entropy 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 table. The first method is able to identify the images without identification leakage or false negative (FN), with only a few occurrences of identification addendum or false positive (FP). The second method is also able to prevent leakage, and in addition, it can avoid FP. It requires a slightly higher processing time. By combining both approaches, a lower computation time and a more perfect identification can be achieved, in which FN and FP can be avoided.

The rest of this paper is organized as follows. Section 2 describes the backgrounds of the proposed methods, including the image identification model and its potential applications, JPEG encoding procedures and JPEG bitstream structures. Section 3 illustrates our proposed methods. Finally, the results of simulations and conclusions are presented in Sects. 4 and 5 respectively.

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. The purpose of the identification is to robustly identify the compressed images at a bitstream level.

Manuscript received September 9, 2005.
Manuscript revised December 16, 2005.
Final manuscript received February 1, 2006.

[†]The authors are with the Department of Electrical Engineering, Tokyo Metropolitan University, Hachioji-shi, 192–0397 Japan.
[††]The authors are with the Department of Information and Communication Systems Engineering, Tokyo Metropolitan University, Hachioji-shi, 192–0397 Japan.
a) E-mail: arnia@isys.eei.metro-u.ac.jp

DOI: 10.1093/ietfec/e89–a.6.1585

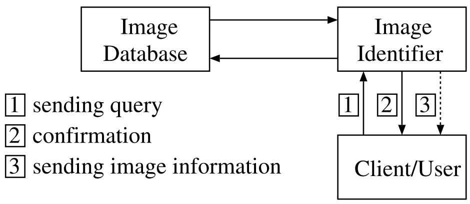


Fig. 1 Process of image identification.

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, or the addition or removal of objects [15].

b. Detection of Errors in Images
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.

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.

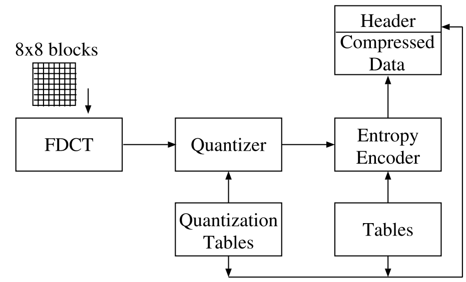


Fig. 2 Processing steps for JPEG DCT-based encoder.

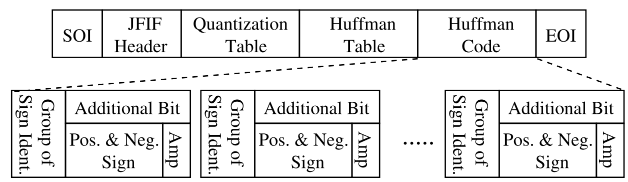


Fig. 3 JPEG bitstream.

2.3 JPEG Encoding Algorithm

The JPEG encoding procedure [18] is shown in Fig. 2 and can be summarized as follows: The original image is tiled into 8×8 blocks, in which each pixel value is shifted from an unsigned to a signed integer. Then a two-dimensional Forward Discrete Cosine Transform (FDCT) function is applied to each block, producing 1 DC and 63 AC coefficients. Quantization is then applied to each coefficient value. The quantization step size controls the compression ratio of an image. The compression ratio relates to a term known as a Q-factor. Finally, all the coefficients are zigzag ordered followed by entropy (Huffman or arithmetic) encoding. The output of the entropy encoder and some additional information represents the JPEG bitstream image.

2.4 JPEG Bitstream Structure

Figure 3 outlines the structure of an interchange format for a JPEG bitstream of a gray scale image. 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 and it contains the specifications of all the tables to decode the image. 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 is divided 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 are different. However, 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 nth DCT coefficient in the mth block in image x is abbreviated as $c_x(m,n)$.
- The sign of the nth DCT coefficient in the mth block in image x is abbreviated as $\mathrm{sgn}\,(c_x(m,n))$, and is derived based on the following equation:

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

- The quantization table for image x is represented as $\mathbf{q}_x = \left\{ q_x(n) \,\middle|\, n = 0, 1, \ldots, 63 \right\}$.

3.2 First Identification Method

In the first method, the positive and negative signs of a DC coefficient are obtained by entropy-decoding a different value of the DC coefficients from two consecutive blocks. However, the positive and negative signs of the AC coefficients can be obtained directly from the bitstream after installing the Huffman table. This is possible because those signs are already available in the bitstream and specified by the MSB of an additional bit group as shown in Fig. 3.

3.2.1 Identification Algorithm

Let us define image A as a JPEG-compressed image that is given by an 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 mth block in image A, extract the positive and negative signs $\mathrm{sgn}\,(c_A(m,n))$ of the nth quantized DCT coefficient $c_A(m,n)$. The sign is obtained by using Eq. (1). If $\mathrm{sgn}\,(c_A(m,n)) = 0$, proceed to step 7.
5. For the mth block in a database image D, extract the positive and negative signs $\mathrm{sgn}\,(c_D(m,n))$ of the nth quantized DCT coefficient $c_D(m,n)$. If $\mathrm{sgn}(c_D(m,n)) = 0$, proceed to step 7.
6. If $\mathrm{sgn}\,(c_A(m,n)) \neq \mathrm{sgn}\,(c_D(m,n))$, 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.

In the first identification, if $\mathrm{sgn}\,(c_A(m,n)) = 0$ or $\mathrm{sgn}\,(c_D(m,n)) = 0$, the comparison is ignored. This action provides some contributions to speed and robustness of the scheme, and will be discussed in Sect. 3.2.2.

3.2.2 Features of First Identification Method

The main features of the first identification method are:

A. Robustness
The algorithm is robust in identifying the images with different compression ratios, if the images are compressed from the same original image and coded by the same method. The robustness is analyzed as follows:

   a. The quantization process is simply a scaling operation, and thus does not alter the DCT coefficient signs of an image.
   b. If $c_A(m,n) = |c_O(m,n)/q_A(n)| < 0.5$, even in the case of $\mathrm{sgn}\,(c_O(m,n)) = \pm 1$, the sign becomes zero, i.e., $\mathrm{sgn}\,(c_A(m,n)) = 0$. These procedures increase the number of zero coefficients, and can

contribute to the robustness aspect of the method, as will be briefly explained in part c.

c. Omitting zero coefficients. If $\mathrm{sgn}(c_A(m,n)) = 0$, or $\mathrm{sgn}(c_B(m,n)) = 0$, the comparison is omitted. Therefore, the images that have the same origin, but have been compressed by different compression ratios, thus having different numbers of zero-valued coefficients, will be successfully identified.

B. Low processing time
A shorter processing time can be achieved by considering the following aspects:

a. The positive and negative signs can be acquired straightforwardly, without fully decoding the entropy of the AC coefficients, because the MSB of the additional bit group that composes the sign word of the AC coefficient shows the positive and negative signs of the AC coefficients. This results in a lower processing time as compared to when a complete JPEG decoding processing steps, which include entropy decoding, reverse-quantization, and reverse-DCT are required to be accomplished.
b. Since 98% of the DCT coefficients are AC coefficients, the identification is further accelerated.
c. If $\mathrm{sgn}(c_A(m,n)) \neq \mathrm{sgn}(c_D(m,n))$, the comparison between the coefficients of the following blocks of images A and B are omitted, and the algorithm proceeds to the next image D in database **D**.
d. In each block, not all 64 coefficients are required to compare, but only N coefficients according to a zigzag scanning.
e. If $\mathrm{sgn}(c_A(m,n)) = 0$ or $\mathrm{sgn}(c_B(m,n)) = 0$, the comparison is omitted.

It is worth noting that the first algorithm does not cause leakages (false negatives), for not identifying the coded images compressed from the same original image, although there is possibility of false positives occurrences for identifying the images compressed from different original images.

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 nth quantized DCT coefficient $c_A(m,n)$ of the mth block in image A and the nth quantized DCT coefficient $c_B(m,n)$ of the mth 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
The algorithm is robust in identifying the images with different compression ratios, if the images are compressed from the same original image and coded by the same method. The robustness can be obtained by considering the following aspects:

- The original coefficient $c_O(m,n)$ is dividing by its corresponding quantization step size and then rounded to the nearest integer to obtain an integer number $c_A(m,n)$. If Eqs. (6) and (7) are fulfilled, it is decided that image A and image B are compressed from the same original image. From Eqs. (6) and (7), a more compact representation is derived and written in Eq. (5). Using the restriction formulated in Eq. (5), the compressed images that have the same origin are robustly identified.

$$c_A(m,n) - 0.5 \leq \frac{c_O(m,n)}{q_A(n)} < c_A(m,n) + 0.5 \tag{6}$$

$$c_B(m,n) - 0.5 \leq \frac{c_O(m,n)}{q_B(n)} < c_B(m,n) + 0.5. \tag{7}$$

- The algorithm ignores the coefficients when $c_A(m,n) = 0$ or $c_B(m,n) = 0$. Similar to the reason for the robustness of the first identification method, disregarding zero-valued coefficients makes this scheme able to identify the images having different compression ratios.

B. Low processing time
The identification requires a low processing time. Some analysis are as follows:

a. The method requires only the quantized DCT coefficients. Therefore, time required for inverting the quantization and the DCT are excluded.
b. The algorithm ignores the coefficient comparison when $c_A(m,n) = 0$ or $c_B(m,n) = 0$. Notice that the second algorithm is applied to the quantized DCT coefficients, where many of them are zero-valued, thus ignoring zero-valued coefficients can decrease the processing time.

3.4 Combined Method

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 image D is output of the first algorithm. The images D are then grouped in database $\mathbf{B}$ and may or may not contain false positives. If it is suspected that false positives have occurred, further identification can be performed by querying image A to the second algorithm to identify images B, where $B \in \mathbf{B}$.

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) [20].

4.1 Simulation Conditions

The simulation conditions are presented in Table 1. Two video sequences, "Football" and "Claire" were used in the simulation. "Football" is a class of images with many object movements between subsequent frames. "Claire" is vice versa.

Originally, there were 194 uncompressed frames for "Football" and 168 for "Claire." All frames of both sequences were compressed with 10 different quality factors (Q-factor), which were in a range of $50, 200, ... \leq$ Q-factor $\leq ..., 1250, 1400$, resulting in 1940 and 1680 images respectively. The original uncompressed versions were not included in the simulation. From each block m, the DCT coefficients of $N = 4$ and $N = 64$ that were obtained according to a zigzag scan were used.

Table 1 Simulation conditions.

Image 1	Football video sequence (gray scale, 704×240, 8 bpp, 194 frame $\times$ 10)
Image 2	Claire video sequence (gray scale, 360×288, 8 bpp, 168 frame $\times$ 10)
Q-factor	50, 200, 350, 500, 650 800, 950, 1100, 1250, 1400
Coefficient	$N = \{4, 64\}$
Query Frame	No. 31 for "Football" No. 1 for "Claire"

Identification was accomplished by querying a compressed frame, for example: "Football" frame No. 31, Q-factor 50, at a time. The same image with other different compression ratios, for example: frame No. 31, $200 \leq$ Q-factor ≤ 1400, are then queried to the system. The same procedures were also applied to "Claire" image. 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 "Football 31-50." The same notation was also applied to other frames with different Q-factors.

The simulation was run on a PC, with a 1.2 GHz processor and a main memory of 512 Mbytes. We implemented the proposed methods using a JPEG codec package from Portable Video Research Group at Stanford University [19].

4.2 First Method

Querying results of the first method can be seen in Tables 2(a) and (b). For "Football," querying with all Q-factors resulted in a perfect identification. Specifically, only the images compressed from the same original image were identified. However, for "Claire," querying images with Q-factor ≥ 350 resulted in false positives. For both sequences, it is worth noting that there were no false negatives.

In the first method, the images compressed from the same original image with different compression ratios (Q-factor) will have the same pattern of DCT coefficients signs, regardless zero-valued coefficients. Because the first method identifies the images by comparing these signs, and ignoring zero-valued coefficients, the leakages can be avoided.

Based on our observation, pattern of DCT coefficient signs of different images can eventually be the same, regardless zero-valued coefficients, because of severe JPEG compression. As the first scheme identifies the images by comparing these signs, a few false positives may occur.

Some examples of the frames used in the simulations can be seen in Figs. 4(a) and (b) and Figs. 5(a) and (b) respectively. The first image group shows "Football" and the latter shows "Claire."

The total time required to identify one image comprises: (1) time to install the Huffman tables, (2) time to entropy decode the bitstream, and at the same time extract the positive and negative signs of DC coefficient and several AC coefficients, and (3) time to compare the signs. As an illustration, the total time required to identify 10 out of 1940 images with the query "Football 31-50," $N = 64$, was approximately 0.5 seconds.

Table 2 Result of querying frames to first identification method. Leakages and false positives did not occur for "Football," but some false positives appeared in "Claire" sequences.

(a). Query: "Football" sequence, frame No. 31. Number of images in database **D**: 1940

Q-factor	50		200		350		500		650		800		950		1100		1250		1400	
N	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64
Total Retrieved Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Same Original Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
False Negatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b). Query: "Claire" sequence, frame No. 1. Number of images in database **D**: 1680

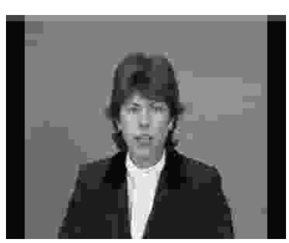
Q-factor	50		200		350		500		650		800		950		1100		1250		1400	
N	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64
Total Retrieved Images	10	10	10	10	15	10	52	35	62	43	63	43	63	43	63	43	63	43	63	43
Same Original Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	5	0	42	25	52	33	53	33	53	33	53	33	53	33	53	33
False Negatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(a) Football Frame No. 31 with Q-factor 50

(b) Football Frame No. 31 with Q-factor 1400

Fig. 4 Examples of "Football" frame with two different Q-factors. Querying frames whose Q-factors are between 50 and 1400 resulted in perfect identification, without false positives nor false negatives.

(a) Claire Frame No. 1 with Q-factor 50



(b) Claire Frame No. 1 with Q-factor 500

Fig. 5 Example of "Claire" frames with two different Q-factors. Querying frames whose Q-factors are between 50 and 1400 did not resulted in perfect identification. Querying the frames with Q-factor ≥ 350 resulted in false positives, but not false negatives.

4.3 Second Method

Simulation results of the second algorithm is provided in Tables 3(a) and (b). For both "Football" and "Claire," neither leakages nor false positives occurred. However, a slightly higher processing time was required to identify the images.

Severe JPEG compression on different images can produce images with the same pattern of DCT coefficient signs. However, DCT coefficient values of those images remain unique. These values are then evaluated by Eq. (5) in Sect. 3.3.1. Equation (5) guarantees that all images compressed from the same image will be identified, and other images will not be identified (rejected). Therefore, in the second algorithm, there were no leakage and false positive occurrences.

To accomplish an identification, the total time required comprises: (1) time to install the Huffman and the quantization tables, (2) time to decode the DC and some AC coefficients, and (3) time to perform comparison using the tables and the entropy-decoded coefficients. As an example, the total time required to identify 10 out of 1940 images with the query "Football 31-50," $N = 64$, was approximately 8 seconds.

4.4 Combined Method

As presented in Sect. 4.2, some querying of "Claire" images to the first algorithm resulted in false positives. For instance, when querying "Claire 1-500" (Fig. 5(b)) to the first method, the number of false positive images with $N = 64$ was 25 (Table 2(b)). Figures 6(a) and (b) show two images that were false positively identified.

The total retrieved images from the first method, which comprised images with the same original image and the false positively identified images, were then collected to form a smaller-sized database **B**. In the experiment, the database **B** consisted of all the retrieved images for Q-factor ≥ 350 (Table 2(b)). The same query images A of "Claire" as used in the first algorithm were then input to the second algorithm to identify images in the database **B**. The results of the identification can be seen in Table 4, all false positives have been removed.

Because false positive images are commonly a few in number as compared to the total images, the required processing time would be insignificant.

Table 3 Result of querying frames to second identification method. Leakages and false positives did not occur for both "Football" and "Claire" sequences.

(a). Query: "Football" sequence, frame No. 31. Number of images in database **D**: 1940

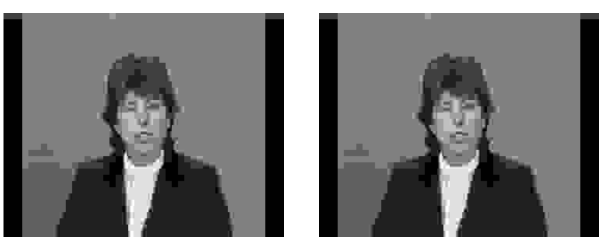
Q-factor	50		200		350		500		650		800		950		1100		1250		1400	
N	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64
Total Retrieved Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Same Original Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
False Negatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

(b). Query: "Claire" sequence, frame No. 1. Number of images in database **D**: 1680

Q-factor	50		200		350		500		650		800		950		1100		1250		1400	
N	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64
Total Retrieved Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Same Original Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
False Negatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4 Result of querying output of first algorithm of "Claire" sequence to second algorithm. Both false positives and false negatives were removed.

Q-factor	50		200		350		500		650		800		950		1100		1250		1400	
N	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64	4	64
Total Retrieved Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
Same Original Images	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
False Positives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
False Negatives	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0



(a) Claire Frame No. 21, with Q-factor 1250 (b) Claire Frame No. 22 with Q-factor 1100

Fig. 6 Examples of false positively identified images of first scheme, but successfully identified by second method. Query frame No. 1, Q-factor = 500, N = 64.

4.5 Pixel Domain Identification

A. Pixel Correlation

Pixel correlation of images was calculated using Eq. (8). Here, r is a correlation coefficient, X_i and Y_i are pixels of image X and Y at location i respectively.

$$r = \frac{\sum(X_i - \bar{X}).(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2 . \sum(Y_i - \bar{Y})^2}} \tag{8}$$

$r = 0$ indicates no linear relationship between image X and Y, and $r = 1$ suggests a strong linear relationship between image X and Y.

When querying "Football" frame No. 31 with all Q-factors, all JPEG images compressed from the same original images were identified. "Football" is a class of images that has many object movements (differences) between subsequent frames, therefore correlation of two subsequent frames was weaker than correlation between the same image with different compression ratios.

Table 5 Result of querying using correlation. Query: "Claire" sequence, frame No. 1, Q-factor 50. False positive (FP) and false negative (FN) occurred.

Rank	Retrieved Image	Corr. Value	Note
1	Claire1-50	1	Identified
2	Claire0-50	0.9971	FP
3	Claire1-200	0.9966	Identified
4	Claire2-50	0.9948	FP
5	Claire0-200	0.9946	FP
6	Claire1-350	0.9937	Identified
7	Claire2-200	0.9925	FP
8	Claire0-350	0.9915	FP
9	Claire3-50	0.9910	FP
10	Claire1-500	0.9896	Identified
21	Claire1-650	0.9834	FN
30	Claire1-800	0.9834	FN
39	Claire1-950	0.9800	FN
50	Claire1-1100	0.9791	FN
52	Claire1-1250	0.9788	FN
53	Claire1-1400	0.9788	FN

However, when querying "Claire" frame No. 1 with all Q-factors, many leakages and false positives occurred. Querying results of "Claire 1-50" can be seen in Table 5. Among 10 images that have the highest correlation values, it turned out that only 4 images compressed from the same image (those which marked "identified"). There were 6 other images that originated from different images (marked as FP) and 6 leakage images (marked as FN). Figure 7(a) shows "Claire 1-50," the query image, and Fig. 7(b) shows "Claire 0-50," one of the false positive images. Successful identification will include all 10 images with different Q-factors in the first 10 identification rank. Note that the ranks of leakage

(a) Claire Frame No. 1, with Q-factor 50 (b) Claire Frame No. 0 with Q-factor 50

Fig. 7 An example of false positive image identified by pixel correlation. (a) query image and (b) false positive image.

Table 6 Result of querying using MSE. Query: “Claire” sequence, frame No. 1, Q-factor 50. False positive (FP) and false negative (FN) occurred.

Rank	Retrieved Image	MSE Value	Note
1	Claire1-50	0	Identified
2	Claire0-50	16.59	FP
3	Claire1-200	19.41	Identified
4	Claire2-50	29.57	FP
5	Claire0-200	30.78	FP
6	Claire1-350	35.76	Identified
7	Claire2-200	42.68	FP
8	Claire0-350	48.09	FP
9	Claire3-50	51.33	FP
10	Claire2-350	59.51	FP
13	Claire1-500	65.78	FN
20	Claire1-650	93.00	FN
48	Claire1-800	119.26	FN
61	Claire1-950	125.46	FN
67	Claire1-1100	130.74	FN
71	Claire1-1250	132.01	FN
72	Claire1-1400	132.12	FN

images were far below 10.

“Claire” is a class of images with only a few changes between subsequent frames (for instance, only a slight head movement). Therefore, because of compression, correlation of two images compressed from the same original images can be weaker than correlation of two subsequent frames with equal compression ratios. For example, “Claire 1-50” has a stronger correlation with “Claire 0-50,” than to “Claire 1-200.”

B. Mean Square Error

Mean square error (MSE) equation is given in Eq. (9). Here, m is total number of pixel in the image, X and T are the database and query images respectively.

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(X_i - T_i)^2 \qquad (9)$$

If the query and the database image are equal, there will be no error, therefore, MSE value is zero. Otherwise, the MSE results a positive real number that can be thought as “error distance” of an image in the database from the query image. The greater the number, the greater the error, thus the further the distance.

In our simulation, querying results of MSE showed the same tendency as the pixel correlation technique. When querying “Football” frame No. 31 with all Q-factors, there was no leakages nor false positives. All 10 targeted images were retrieved.

However, when querying “Claire” frame No.1, there were many false positives and false negatives. For example, querying results of “Claire 1-50” was summarized in Table 6. Only 3 images were identified, 7 frames were false positives and 7 images were leakage images.

As a summary, the correlation and MSE techniques in pixel domain failed to identify all the expected images. In addition, a complete JPEG-decoding is required, therefore longer processing time is needed.

5. Conclusions

We have proposed two novel schemes for identifying JPEG coded images. These schemes require low processing time, since feature calculation was not needed and they work on the quantized DCT domain. Moreover, only several first DCT coefficients are commonly required from each image. The first approach is able to avoid identification leakages or false negatives and results in only 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 perfect identification can be achieved.

An extension of the proposed methods to other compression standards such as MPEG and JPEG 2000 are considered as future works. Furthermore, a development of the proposed schemes as image retrieval methods that based on similarity measure is also considered.

References

[1] H. Kobayashi and H. Kiya, “Robust image authentication using hash function,” Proc. 2004 IEEE TENCON, vol.1, pp.435–438, Chiang Mai, Thailand, Nov. 2004.
[2] J.J. Eggers and B. Girod, “Blind watermarking applied to image authentication,” Proc. IEEE ICASSP, vol.3, pp.1997–1980, May 2001.
[3] Q. Shun and S.-F. Chang, “Semi fragile image authentication using generic wavelet domain features and ECC,” Proc. IEEE ICIP, vol.2, pp.901–904, Sept. 2002.
[4] A. Datta, N.D.V. Lobo, and J.J. Leeson, “Novel feature vector for image authentication,” Proc. IEEE ICME, vol.2, pp.221–224, July 2003.
[5] J. Sun, Q. Zhuo, and W. Wang, “Study of sparse features in image recognition,” Proc. SPIE, vol.4550, pp.79–84, Sept. 2001.
[6] 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, vol.1908, pp.173–187, April 1993.
[7] I. Kompatsiaris, E. Triantafilow, and M.G. Srintzis, “Region-based color image indexing and retrieval,” Proc. IEEE ICIP, vol.1, pp.658–661, Oct. 2001.
[8] B.S. Manjunath and W.Y. Ma, “Texture features for browsing and retrieval of image data,” IEEE Trans. Pattern Anal. Mach. Intell., vol.18, no.8, pp.837–842, Aug. 1996.
[9] F. Mahmoudi, J. Shanbehzadeh, A.-M. Eftekhari-Moghadam, and H. Soltanian-Zadeh, “A new non segmentation shape based image indexingmethod,” Proc. IEEE ICASSP, vol.3, pp.17–20, April 2003.

[10] 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.

[11] J. Jiang, A. Armstrong, and G.C. Feng, "Web-based image indexing and retrieval in JPEG compressed domain," Multimedia Syst., vol.9, pp.424–432, 2004.

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

[13] M.K. Mandal and T. Aboulnasr, "Fast wavelet histogram techniques for image indexing," Proc. Comp. Vis. and Image Understanding, vol.75, pp.99–110, July/Aug. 1999.

[14] M.K. Mandal and C. Liu, "Efficient image indexing techniques in the JPEG2000 domain," J. Electronic Imaging, vol.13, no.1, pp.182–187, Jan. 2004.

[15] F. Arnia, I. Iizuka, Y. Nakano, A. Takagi, and H. Kiya, "Compressed domain based image authentication for JPEG coded images," Proc. 2005 ITC-CSCC, vol.4, pp.1465–1466, Jeju, Korea, July 2005.

[16] I. Iizuka, F. Arnia, M. Fujiyoshi, and H. Kiya, "Fast image identification for coded images with different compression ratio," IEICE Technical Report, SIP2005-41, June 2005.

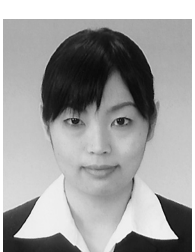
[17] Y. Nakano, A. Takagi, and H. Kiya, "A method of stream base fast image retrieval for JPEG coded images," ITE Tech. Rep., no.ME2005-55, pp.21–24, Feb. 2005.

[18] Information technology—Digital compression and coding of continuous-tone still images—Requirement and guidelines, ISO/IEC International Standard 10918-1, 1992.

[19] JPEG CODEC, Portable Video Research Group at Stanford Univ., ver.1.2, 1994.

[20] Z. Wang and E.P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," Proc. IEEE ICASSP, vol.2, pp.573–576, March 2005.

Fitri Arnia was born in Kisaran, North Sumatera Province, Indonesia, on November 12, 1973. She received the B.E. degree in electrical engineering from North Sumatera University, Medan, Indonesia, in 1997 and M.Eng.Sc. degree in electrical engineering from University of New South Wales, Sydney, Australia, in 2004. She is currently a candidate for D.E. degree at Tokyo Metropolitan University, Tokyo, Japan, and on study leave from Electrical Engineering Department, Syiah Kuala University, Banda Aceh, Indonesia. Her research interests are signal and image processing. She is a student member of the IEEE.



Ikue Iizuka was born in Saitama, Japan, on August 25, 1979. She received the B.E. degree in electrical engineering from Tokyo Metropolitan University (TMU), Tokyo, Japan, in 2005. Currently, she is a master course student at TMU. Her research interest is image processing.

Masaaki Fujiyoshi received his B.Arts, M.Eng., and Ph.D. degrees from Saitama University, Japan, in 1995, 1997, and 2001, respectively. Since 2001, he was with the Graduate School of Engineering at Tokyo Metropolitan University, and he has been a Research Associate of Information and Communication Systems Engineering at Tokyo Metropolitan University from 2005. His research interests include image processing, secure communications, and spread spectrum communications. Dr. Fujiyoshi is a Member of the IEEE and the ITE (Institute of Image Information and Television Engineers). He received the Young Engineer Award from the IEICE in 2001.

Hitoshi Kiya received his B.E. and M.E. degrees from Nagaoka University of Technology, Japan, in 1980 and 1982, respectively, and his D.E. degree from Tokyo Metropolitan University, Japan, in 1987. In 1982, he joined Tokyo Metropolitan University, and he is currently a Professor of Information and Communication Systems Engineering, Faculty of System Design. He was a Visiting Fellow at the University of Sydney, Australia, from Oct. 1995 to March 1996. His research interests include digital signal processing, multirate systems, adaptive filtering, image processing, and security for multimedia. Prof. Kiya has served as an Associate Editor for the IEICE Trans. Fundamentals (Japanese Edition) and the IEEE Trans. Signal Processing from 1998 to 2002 and from 1998 to 2000, respectively. He is a vice chair of the Signal Processing Committee of the IEICE, and the chair of the Media Engineering Committee of the ITE (Institute of Image Information and Television Engineers). He is also a Member of the ITE and the IIEEJ (Institute of Image Electronics Engineers of Japan), and a Senior Member of the IEEE.