

A Robust Identification Scheme for JPEG XR Images with Various Compression Ratios

Hiroyuki Kobayashi¹(✉), Shoko Imaizumi², and Hitoshi Kiya³

¹ Tokyo Metropolitan College of Industrial Technology, Shinagawa,
Tokyo 140-0011, Japan
hkob@metro-cit.ac.jp

² Chiba University, Chiba 263-8522, Japan
shoko@chiba-u.jp

³ Tokyo Metropolitan University, Hino, Tokyo 191-0065, Japan
kiya@sd.tmu.ac.jp

Abstract. A robust scheme for identifying JPEG XR coded images is proposed in this paper. The aim is to identify the images that are generated from the same original image under various compression ratios. The proposed scheme is robust against a difference in compression ratios, and does not produce false negative matches in any compression ratio. A new property of the positive and negative signs of lapped biorthogonal transform coefficients is considered to robustly identify the images. The experimental results show the proposed scheme is effective for not only still images, but also video sequences in terms of the querying such as false positive, false negative and true positive matches.

Keywords: Image identification · JPEG XR · Robustness

1 Introduction

The use of digital images and video sequences has greatly increased recently because of the rapid growth of the Internet and multimedia systems. It is often necessary to identify a certain image in a database that has a large number of images in various types of the applications of digital images/videos, such as security and evaluation of image validity. The image database generally consists of images in a compressed form to reduce the amount of data. Several international standards for searching for images/videos have been developed [1–4] in connection to this. “Identification” in this work is defined as the operation of finding an image that is identical to a given original image from an image database. In this paper, a robust scheme for identifying JPEG XR coded images is proposed.

JPEG XR is an image coding standard from the JPEG committee [5, 6]. It allows lossy and lossless coding for still images and videos. It supports not only images with 8 bits but also images with over 8 bits and floating point representation. Thus, it can support various kinds of images including high dynamic range(HDR) images [7–9] for a new generation of digital cameras. Therefore the proposed scheme is widely available for identifying many kinds of images.



Fig. 1. Process of image identification

So far, several schemes have been developed for identifying compressed images [10–19]. The schemes described in [13–17] are for the JPEG standard, and the schemes in [17–19] are for the JPEG 2000 standard, where some properties of transform coefficients i.e. DCT(Discrete Cosine Transform) and DWT(Discrete Wavelet Transform) coefficients, play an important role for image identification. In addition, they have been extended to image identification schemes in the encrypted domain securely to operate images/videos [20,21]. However, there is still none for the JPEG XR standard. Moreover the previous schemes are not available for JPEG XR images, because JPEG XR is the only image coding standard that uses a lapped biorthogonal transform(LBT), which is different from DCT and DWT [22].

Because of this situation, a scheme for identifying JPEG XR coded images is considered in this paper. The aim of the proposed scheme is to identify JPEG XR images that are generated from the same original image under various compression ratios. The proposed scheme does not produce false negative matches in any compression ratio. A new property of the positive and negative signs of LBT coefficients is utilized to identify the images. The experimental results shows the proposed scheme is effective for not only still images, but also video sequences in terms of the retrieval performance such as false positive, false negative and true positive matches.

2 Background

2.1 Image Identification Model

Let us consider that 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. This paper focuses on querying some properties of JPEG XR images.

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 image 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.

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 a 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.

2.3 JPEG XR

JPEG XR is an image coding standard from the JPEG committee. It allows lossy and lossless coding for still images and videos. It supports not only fixed point representation but also floating point representation. Thus, it can support various kinds of images including HDR images for a new generation of digital cameras.

The block diagram of JPEG XR encoding is illustrated in Fig. 2. JPEG XR is based on a block transform design, and it uses some of the same high level building blocks as in most image compression schemes, such as color conversion, spatial transformation, scalar quantization, coefficient scanning, and entropy coding. The encoding consists of the following basic steps:

- (1) Performing a color conversion.
- (2) Dividing an image into non-overlapped consecutive 16×16 blocks, called macro block, and then each macro block into consecutive 4×4 blocks, called block (see Fig. 3(a)).

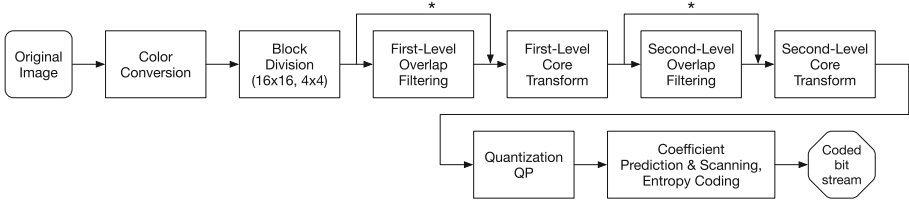


Fig. 2. Basic block diagram of JPEG XR encoding (* There are three modes)

- (3) Applying two basic operators i.e. core transform and optional overlap filtering to the blocks, where the operators are hierarchically executed twice shown in Fig. 3(b).
- (4) Applying a coefficient quantization approach controlled by quantization parameters (QPs).
- (5) Executing adaptive coefficient scanning to convert the two-dimensional array transform coefficients within a block into a one-dimensional vector to be encoded. Finally, the coefficients are entropy encoded.

In step (3), one temporally DC coefficient and 15 HP coefficients are obtained for each block by the 1st-level core transform, and 16 temporally DC coefficients are gathered from each macro block as shown in Fig. 3(b). The 2nd-level core transform is then applied to them. As a result, one DC coefficient, 15 LP coefficients and 15×16 HP coefficients are calculated for each macro block, where core transform, referred to as lapped biorthogonal transform (LBT), is common between two levels. Therefore, the transform coefficients are often called LBT coefficients, which consist of DC, LP and HP ones.

The overlap filtering may be used to reduce blocking artifacts. JPEG XR has three overlapping-modes. When mode 0 is chosen, no overlap filtering is performed. Otherwise, only the 1st-level overlap filtering is performed for mode 1, and both filtering operations are done for mode 2.

3 Proposed Identification Scheme

The aim of the proposed scheme is to identify JPEG XR images that are generated from the same original image under various compression ratios. The proposed scheme does not produce false negative matches in any compression ratio. A new property of the positive and negative signs of LBT coefficients is utilized to identify the images.

3.1 Notation and Terminologies

Several notations and terminologies used in the following sections are listed here.

- x represents an image. x can be “ Q ” for image Q , “ D ” for image D and “ O ” for the original image, where all images have the same size.
- B represents the number of blocks in an image.

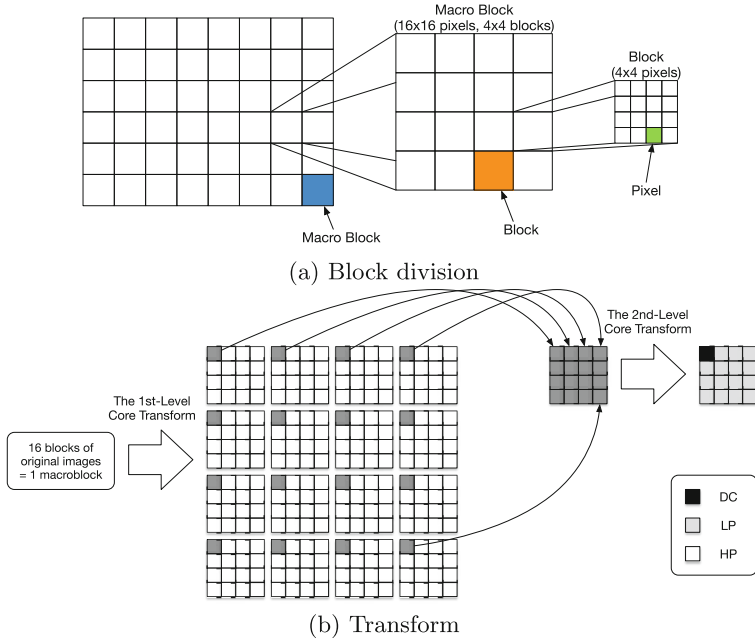


Fig. 3. Lapped biorthogonal transform used in JPEG XR

- M represents the number of macroblocks in an image.
- N represents the number of coefficients in a 4×4 core transform, and the number of blocks in a macroblock, where $N = 16$.
- $DC_x(m)$ represents the DC coefficient of the m^{th} macroblock in image x , where $0 \leq m < M$.
- $LP_x(m, n)$ represents the n^{th} LP coefficient of the m^{th} macroblock in image x , where $0 \leq m < M, 1 \leq n < N$.
- $HP_x(b, n)$ represents the n^{th} HP coefficient of the b^{th} block in image x , where $0 \leq b < B, 1 \leq n < N$.
- P represents the number of all coefficients in an image, where $P = MN + B(N - 1)$.
- $\text{sgn}(c)$ represents the sign of a real value c as

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

- $C_x(k)$ represents LBT coefficients sequence given by

$$C_x(k) = \begin{cases} DC_x(k), & 0 \leq k < M, \\ LP_x(m, n), & M \leq k < MN, \\ & m = \text{mod}(k, M), n = \lfloor (k - M)/M \rfloor, \\ HP_x(b, n), & MN \leq k < P, \\ & b = \text{mod}(k - MN, B), n = \lfloor (k - MN)/B \rfloor + 1. \end{cases} \quad (2)$$

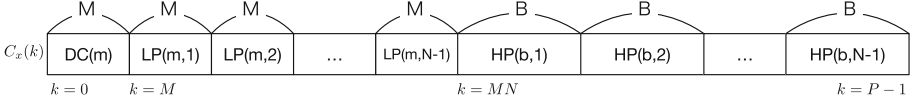


Fig. 4. LBT coefficients sequence $C_x(k)$

where $\text{mod}(x, d)$ denotes the remainder when x is divided by d , and $\lfloor x \rfloor$ denotes the integer part of x . The length of $C_x(k)$ is $P = MN + B(N - 1)$, (see Fig. 4).

3.2 Identification Scheme

The proposed scheme focuses on the positive and negative signs of LBT coefficients, which can be obtained by entropy-decoding from JPEG XR bit streams. It is verified that quantized LBT coefficients have the following property.

- When images Q and D_i are generated from the same original image O , the positive and negative signs of LBT coefficients of the two images are equivalent in the corresponding location, even though quantization parameters (QPs) are different. Namely, the relation is given as

$$\text{sgn}(C_Q(k)) = \text{sgn}(C_{D_i}(k)), (0 \leq k < P) , \tag{3}$$

where this property does not apply in zero-value coefficients.

The above property, which can be theoretically explained, is illustrated in Fig. 5. Figure 5(a) and (b) are examples of quantized LBT coefficients of images Q and D_1 that are generated from the same original image O . It is confirmed that the positive and negative signs of LBT coefficients of the two images are equivalent in the corresponding location, except for the case in zero-value coefficients. On the other hand, image D_2 in Fig. 5(c) that is generated from the other original image, does not have the same signs as those in Fig. 5(a). In this manner, there is no guarantee that two images generated from different original images have the same signs. Note that the number of zero-value coefficients depend on quantization parameters (QPs).

-349	0	4	-1
4	2	-6	7
-1	-3	-3	-4
0	3	3	-1

-116	0	1	0
1	1	-2	2
0	-1	-1	-1
0	1	1	0

-29	-15	7	-8
1	1	7	-25
-8	10	-2	-2
-24	7	6	-1

(a) Image Q ($QP = 10$) (b) Image D_1 ($QP = 30$) (c) Image D_2 ($QP = 10$)

Fig. 5. Examples of LBT (DC and LP) coefficients. Image D_1 has the same signs as image Q except for zero-value coefficients

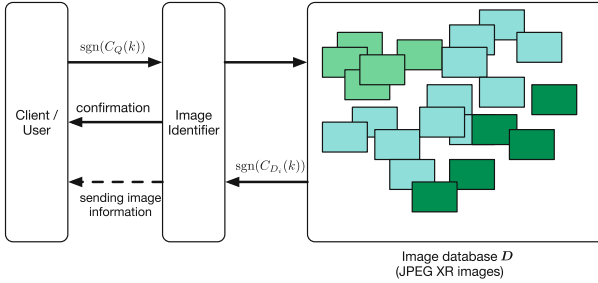


Fig. 6. Image identification for JPEG XR images

Let us define image Q as a JPEG XR coded image that is given by user (a query image) and image D_i is a JPEG XR image that is given from a database D , where $D_i \in D$ (see Fig. 6). The positive and negative signs of the quantized LBT coefficients of the images Q and D_i in the corresponding locations are compared, and the results are used to decide whether the images are compressed from the same original image.

When compressed image Q and image D_i ($i = 1, 2, \dots$) are compared, the identification algorithm is accomplished according to the following steps.

- (a) Set the value of L , where L is the number of LBT coefficients used for identification ($1 \leq L \leq P$).
- (b) Set $k := 0$.
- (c) For the k^{th} coefficients A , extract the positive and negative signs. If $\text{sgn}(C_A(k)) = 0$, proceed to step (e).
- (d) If $\text{sgn}(C_Q(k)) \neq \text{sgn}(C_{D_i}(k))$, the algorithm decides that image Q and D_i were not compressed from the same original image, and the process is halted. Otherwise, proceed to step (e)
- (e) Set $k := k + 1$.
- (f) If $k = L$, it is decided that image Q has the same original image as image D_i . Otherwise, continue to step (c).

When $L = M$ is chosen, only DC coefficients are used for identification. Otherwise, DC and LP coefficients are used for $L = MN$, and all LBT coefficients are done for $L = P$, respectively.

4 Simulation

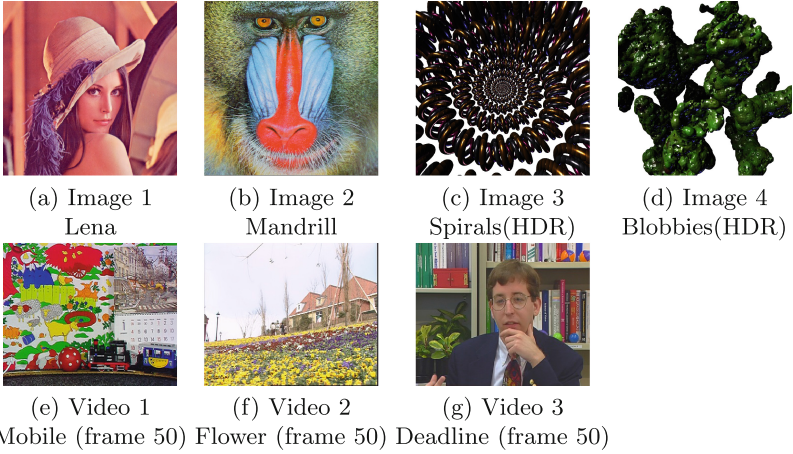
To evaluate the performance of the proposed scheme, several simulations are conducted.

4.1 Simulation Conditions

The simulation conditions are presented in Table 1. Two still images with 8×3 bpp(bit per pixel), two still HDR images with the OpenEXR format

Table 1. Simulation conditions

Image 1	Lena still image (color 512×512 , 8×3 bpp)
Image 2	Mandrill still image (color 512×512 , 8×3 bpp)
Image 3 (OpenEXR format)	Spirals still image (color 1040×1040 , 16×4 bpp)
Image 4 (OpenEXR format)	Blobbies still image (color 1040×1040 , 16×4 bpp)
Video 1	Mobile video sequence (color 352×288 , 8×3 bpp, 100frames)
Video 2	Flower video sequence (color 352×288 , 8×3 bpp, 100frames)
Video 3	Deadline video sequence (color 352×288 , 8×3 bpp, 100frames)
<i>OM</i> (Overlapping mode)	0(non overlapping), 1(HP only), 2(LP &HP)
<i>QP</i>	10, 20, 30, 40, 50, 60, 70, 80, 90
The number of LBP coefficients	$L = M$ (DC only), MN (DC and LP), P (all)

**Fig. 7.** Images and videos used in simulation

(16×4 bpp) [23] and three video sequences, i.e. “Mobile”, “Flower” and “Deadline” were used in the simulation (Fig. 7). “Mobile” and “Flower” are in a class of images with large object movements between subsequent frames. “Deadline” is vice versa. All images were compressed with 9 different quantization parameters (QP). In the following section, for example, “Mobile” frame No.5 with $QP = 10$ will be referred to as “Mobile5-10”. The JPEG XR reference software 1.8 [24] was used in the simulation. The simulation was run on a PC, with a 2.7 GHz processor and a main memory of 16 Gbytes.

4.2 Evaluation for Still Images

Four still images including HDR ones were compressed with nine different quantization parameters (QP s) shown in Table 1 to generate 36 compressed images, of which four images with $QP = 50$ were in the database \mathbf{D} , and $4 \times 9 = 36$ compressed images were used as a query image. The original uncompressed versions were not included in the simulation. Identification was accomplished by querying a compressed image.

Querying results for still images are shown in Table 2. Table 2 summarizes the number of true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN) matches. Besides, the table shows the false-positive-rate (FPR) and true-positive-rate (TPR) [25], defined by

$$FPR = FP / (FP + TN) , \quad (4)$$

$$TPR = TP / (TP + FN) . \quad (5)$$

Moreover, the F_1 -score (F_1) [25] is known to be one measure used in the field of information retrieval for measuring the performance of search, document classification, and query classification. A higher F_1 -score means better performance. The value F_1 is given by

$$F_1 = \frac{2}{1/\text{precision} + 1/\text{recall}} , \quad (6)$$

$$\text{precision} = \frac{TP}{TP + FP} , \quad (7)$$

$$\text{recall} = \frac{TP}{TP + FN} . \quad (8)$$

It is confirmed that there were not any false positive and false negative matches, under all overlapping modes (OM) and any compression ratios. In other words, querying with all QP s resulted in a perfect identification for all images.

Table 2. Querying results for still images

OM	L	TP	TN	FP	FN	FPR (%)	TPR (%)	F_1 score
0 (non overlapping)	M (DC only)	36	108	0	0	100	0	1.0
	MN (DC and LP)	36	108	0	0	100	0	1.0
	P (all)	36	108	0	0	100	0	1.0
1 (HP only)	M (DC only)	36	108	0	0	100	0	1.0
	(DC and LP)	36	108	0	0	100	0	1.0
	P (all)	36	108	0	0	100	0	1.0
2 (LP and HP)	M (DC only)	36	108	0	0	100	0	1.0
	MN (DC and LP)	36	108	0	0	100	0	1.0
	P (all)	36	108	0	0	100	0	1.0

4.3 Evaluation for Videos

The three video sequences shown in Table 1 were used to confirm the effectiveness of the proposed scheme. Originally, there were 100 uncompressed frames for each video sequence. All video frames were compressed with three different quantization parameters i.e. $QP = 10, 50$ and 90 . As a result, 300 compressed frames were generated from each sequence, and 900 compressed frames were used in total in the simulation. Three video sequences with $QP = 50$ i.e. 300 frames in total were in the database \mathbf{D} , and all compressed frames i.e. 900 frames were used as a query image. The original uncompressed versions were not included in the simulation. Therefore, 900×300 combinations were carried out to evaluate the proposed scheme.

Querying results for videos are shown in Table 3. From the results, it is confirmed that a larger L proves higher recognition accuracy and a smaller QP also gives higher one, because these conditions enable to supply a large number of the positive and negative signs of LBT coefficients to image identifier. In particular, for $L = P$, querying with all QPs resulted in a perfect identification for all video sequences. The performance trends can be reconfirmed via F_1 -scores as shown in Fig. 8. Besides, compared to “Mobile” and “Flower”, F_1 -scores decrease

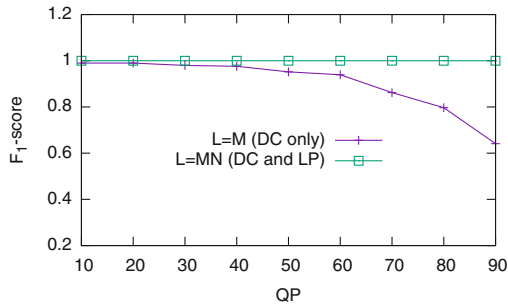


Fig. 8. F_1 -scores for experimental results (Flower)

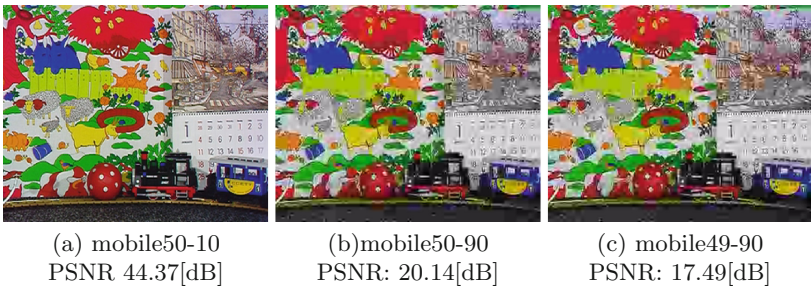


Fig. 9. Compressed frames with different QPs (Mobile frame 50 and 49)

Table 3. Querying results for video images (Overlapping mode = 1)

Video	QP	$L = M(\text{DC only})$						$L = MN(\text{DC \& LP})$					
		TP	TN	FP	FN	FPR (%)	F_1	TP	TN	FP	FN	FPR (%)	F_1
Mobile	10	100	29898	2	0	0.01	0.990	100	29900	0	0	0.00	1.000
	50	100	29898	2	0	0.01	0.990	100	29900	0	0	0.00	1.000
	90	100	29896	4	0	0.01	0.980	100	29900	0	0	0.00	1.000
Flower	10	100	29895	5	0	0.02	0.976	100	29900	0	0	0.00	1.000
	50	100	29890	10	0	0.03	0.952	100	29900	0	0	0.00	1.000
	90	100	29788	112	0	0.37	0.641	100	29900	0	0	0.00	1.000
Deadline	10	100	29888	12	0	0.04	0.943	100	29900	0	0	0.00	1.000
	50	100	29874	26	0	0.09	0.885	100	29900	0	0	0.00	1.000
	90	100	29763	137	0	0.46	0.593	100	29893	7	0	0.02	0.966

for “Deadline”, since it does not include large objective movements between subsequent frames. For all conditions, it is worth noting that there were no false negatives.

Figure 9 shows examples of compressed frames with different QPs , where PSNR(Peak Signal to Noise Ratio) is a measure of image quality. From these examples, it is shown that the successive frames are very similar and moreover compressed frames include large amount of quantization noise in general. The proposed scheme enables to detect the slight difference between frames, even though there is such a situation.

5 Conclusion

A novel scheme for identifying JPEG XR images in the compressed domain has been proposed in this paper. The conventional schemes for compressed images are not available for JPEG XR images, due to the use of a LBT. A new property of the positive and negative signs of LBT coefficients has been considered robustly to identify the images. The proposed scheme does not produce false negative matches in any compression ratio. The experimental results have showed the proposed scheme is effective for not only still images, but also video sequences in terms of the retrieval performance such as false positive, false negative and true positive matches. In particular, in the case of using DC and LP coefficients, i.e. $L = MN$, querying with all QPs resulted in a near-perfect identification for all images and videos. The proposed scheme will be extended to a identification scheme in the encrypted domain as a future work.

References

1. Information technology - JPSearch - Part 1: System framework and components. International Standard ISO/IEC TR-24800-1 (2007)
2. Compact descriptors for visual search: Applications and use scenarios. ISO/IEC JTC1/ SC29/WG11/N11529 (2010)

3. Compact descriptors for visual search: Context and objectives. ISO/IEC JTC1/SC29/WG11/N11530 (2010)
4. Compact descriptors for visual search: Requirements. ISO/IEC JTC1/SC29/WG11/N11531 (2010)
5. Rec, ITU.-T., T.832: Information technology - JPEG XR image coding system - Image coding specification. <http://www.itu.int/rec/T-REC-T.832>
6. Dufaux, F., Sullivan, G., Ebrahimi, T.: The JPEG XR image coding standard [Standards in a Nutshell]. *IEEE Signal Process. Mag.* **26**(6), 195–204 (2009)
7. Dobashi, T., Tashiro, A., Iwahashi, M., Kiya, H.: A fixed-point implementation of tone mapping operation for HDR images expressed in floating-point format. *APSIPA Trans. Signal Inf. Process.* **3**(e11), 1–11 (2014)
8. Iwahashi, M., Kiya, H.: Two layer lossless coding of HDR images. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1340–1344 (2013)
9. Reinhard, E., Ward, G., Pattanaik, S., Debevec, P., Heidrich, W., Myszkowski, K.: *High Dynamic Range Imaging - Acquisition, Display and Image Based Lighting*. Morgan Kaufmann, Burlington (2010)
10. Mandal, M.K., Liu, C.: Efficient image indexing techniques in the JPEG2000 domain. *J. Electron. Imaging* **13**(1), 182–187 (2004)
11. Uchida, Y., Sakazawa, S.: D-12-93 near-duplicate video detection considering temporal burstiness of local features. In: *Proceedings of the IEICE General Conference 2011* (2011)
12. McIntyre, A.R., Heywood, M.I.: Exploring content-based image indexing techniques in compressed domain. In: *Proceedings of the IEEE Canadian Conference on Electrical & Computer Engineering*, vol. 2, pp. 957–962 (2002)
13. Arnia, F., Iizuka, I., Fujiyoshi, M., Kiya, H.: Fast and robust identification methods for JPEG images with various compression ratios. In: *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, vol. II, no. IMDSP-P4.6 (2006)
14. Arnia, F., Iizuka, I., Fujiyoshi, M., Kiya, H.: “Fast Method for Joint Retrieval and Identification of JPEG Coded Images Based on DCT Sign,” *Proc. IEEE International Conference on Image Processing*, vol. II, no. MP-P1.10, pp. 229–232 (2007)
15. Jiang, J., Armstrong, A., Feng, G.C.: Web-based image indexing and retrieval in JPEG compressed domain. *Multimedia Syst.* **9**, 424–432 (2004)
16. Shneier, M., Abdel-Mottaleb, M.: Exploiting the JPEG compression scheme for image retrieval. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**(8), 849–853 (1996)
17. Cheng, K.O., Law, N.F., Siu, W.C.: A fast approach for identifying similar features in retrieval of JPEG and JPEG 2000 images. In: *Proceedings of APSIPA ASC 2009* (2009)
18. Watanabe, O., Iida, T., Fukuhara, T., Kiya, H.: Identification of JPEG 2000 Images in encrypted domain for digital cinema. In: *Proceedings of the IEEE International Conference on Image Processing*, no. MA.PJ.PJ8, pp. 2065–2068 (2009)
19. Watanabe, O., Fukuhara, T., Kiya, H.: Fast identification of JPEG 2000 images for digital cinema profiles. In: *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, no. IVMS-P-L4.6, pp. 881–884 (2011)
20. Dobashi, T., Watanabe, O., Fukuhara, T., Kiya, H.: Hash-Based Identification of JPEG 2000 images in encrypted domain. In: *Proceedings of the IEEE International Symposium on Intelligent Signal Processing and Communication Systems*, no. D2.4, pp. 469–472. (2012)

21. Watanabe, O., Fukuhara, T., Kiya, H.: Codestream-based identification of JPEG 2000 images with different coding parameters. *IEICE Trans. Inf. Sys.* **E95-D**(4), 1120–1129 (2012)
22. Taubman, D.: High performance scalable image compression with EBCOT. *IEEE Trans. Image Process.* **9**(7), 1158–1170 (2000)
23. Bogart, R., Kainz, F., Hess, D.: The OpenEXR image file format. In: *Proceedings of the ACM SIGGRAPH Technical Sketches*, San Diego, CA, USA (2003)
24. ITU-T Rec. T.835: Information technology - JPEG XR image coding system - Reference software. <http://www.itu.int/rec/T-REC-T.835>
25. Fawcett, T.: An introduction to ROC analysis. *Pattern Recogn. Lett.* **27**(8), 861–874 (2006)