Published in Proceedings of the 6th COST 276 Workshop on
Information and Knowledge Management for Integrated Media Communication, 88-93, 2004
which should be used for any reference to this work

Low Complexity Image Matching in the Compressed Domain by using the DCT-phase

J. Bracamonte, M. Ansorge, F. Pellandini, and P.-A. Farine

Institute of Microtechnology, University of Neuchâtel Rue A.-L Breguet 2, 2000 Neuchâtel, Switzerland
Email: javier.bracamonte@unine.ch

Abstract

This paper presents a new, simple, and efficient algorithm for image matching in the compressed domain. The proposed technique operates directly on transformed frequency-domain data, which makes it a good choice when dealing with standard-compliant Discrete Cosine Transform (DCT)-based compressed images since it avoids the large computational overhead associated with the decoding stage. Furthermore, the algorithm itself features a very low computational complexity and a high matching accuracy. An example of the effectiveness of this algorithm is illustrated in the frame of a particular content-based image retrieval application. The proposed method exploits the fact that the phase of the coefficients of a DCT transformed image contains a significant amount of information of the original image. By processing only the phase part of the DCT coefficients, a simple and yet high matching accuracy method is achieved. The underlying principle can also be applied in other computer vision applications and in video processing in the compressed domain.

Categories and Subject Descriptors (according to ACM CCS): H.3.3 [Information Search and Retrieval]: Selection process

1. Introduction

Most digital images and video sequences today are stored and transmitted in compressed form. This generalized compressed status of visual information in current multimedia systems, promotes a large interest in the design and implementation of image and video processing algorithms that operate directly in the compressed domain. These algorithms present the advantage of avoiding the allocation of computational and electric power to the heavy decompression modules before executing the specific image processing application in the pixel-domain. Several compressed-domain-based techniques for still image applications have been reported in [1] [2] [3].

In this paper we address the issue of image matching in the compressed JPEG domain. The effectiveness of the proposed technique is illustrated in the frame of a particular content-based image retrieval (CBIR) [4] application, which is *exact match queries*. In effect, the multiple challenges in CBIR include the design of minimum retrieval time systems that recognize whether or not an identical copy of a query image is present in a database [5]. Furthermore, this presence should also be identified in case the existing copy has been encoded with a different bit rate, e.g., with a higher compression ratio [5].

1.1. Previous Work

Previous publications have reported methods for indexing and retrieval in the compressed DCT domain. The algorithm reported in [6] is based on the computation of the mean value μ, and the variance σ^2 of the DCT coefficients of each (8×8)-element basic block. By executing some vector-quantization-like process on the two-dimensional (2-D) (μ, σ) space, a 28-component vector is produced and used as the corresponding image feature. The same idea, but based on (4×4) blocks, has been reported in a previous paper [7]. An energy histogram technique similar in concept to the pixel-domain color histogram method has been proposed in [8]. The histogram is built by counting the number of times an energy level appears in the (8×8)-element blocks of DCT coefficients of a transformed image. Since most of the energy within such (8×8)-element blocks is generally distributed in the low frequency region, the proposed method reduces the computational complexity by selecting the DC and only few additional low frequency coefficients for creating the histogram. In [9] a procedure to speed up the generation of image features is reported; processing time is saved by adaptively selecting a reduced number of coefficients that are used as input to the Inverse DCT (IDCT) operation. Further indexing and retrieval techniques in the DCT-domain are reported in [10].

The previously cited papers deal with compressed-domain-based indexing methods that are in general suitable for *similarity* queries in CBIR systems. Other studies more specifically related to the *exact image query* issue have been proposed in [11] [12] [13] [14]. The latter are all pixel-domain-based techniques.

Additional methods concerning image matching in the JPEG domain in a non-CBIR application are found in [15] [16]. These techniques are applied in the matching of compressed photogrammetric image data. The authors claim having obtained only a very limited success, specially with the technique explored in [15]. The latter describes a compressed-domain symmetrical convolution, which is an extension of classic linear convolution used in signal processing. The second cited paper presents a compressed-domain least squares matching approach, which consists in taking a mathematic transform, the DCT in this case, of the least squares correlation method reported in [17].

In this paper we introduce a novel, low complexity, image matching technique that is based on the processing of the phase of DCT coefficients. Besides its reduced computational requirements, the proposed algorithm produces an excellent matching accuracy as it will be demonstrated in the following sections.

1.2. Organization of the paper

The remainder of this paper is organized as follows. Section 2 recalls the underlying principle that motivated the study of the proposed matching technique. Section 3 describes the compressed domain matching algorithm and introduces one low complexity implementation scheme. Section 4 presents the results obtained in the frame of an exact match image retrieval application. Section 5 summarizes the computational resources required by the matching algorithm. Finally, Section 6 states the conclusions.

2. The DCT-phase of Images

A study on the significance of the DCT-phase in images was reported in [18] where it is showed that the DCT-phase in spite of its reduced binary value $\{0, \pi\}$ conveys a significant amount of information of its associated image. An example given in [18] is reproduced in Figure 1 and is briefly described in the following paragraph.

Figures 1(a) and 1(b) show the test images Lena and Baboon, both monochrome and with a spatial resolution of (512×512) pixels. By applying a 512-point 2-D DCT over these images, two sets of transformed coefficients are obtained. Figures 1(c) and 1(d) show the reconstruction back into the spatial domain after an IDCT has been applied over the magnitude array of the two sets of transform coefficients and when the corresponding phase values were all forced to zero. Figures 1(e) and 1(f) show the reconstruction when the IDCT is applied over only the original binary-valued phase arrays, while the value of the magnitudes was set to one. These last two figures put in evidence the high amount of information conveyed by the DCT-phase, which is further emphasized in Figures 1(g) and 1(h).

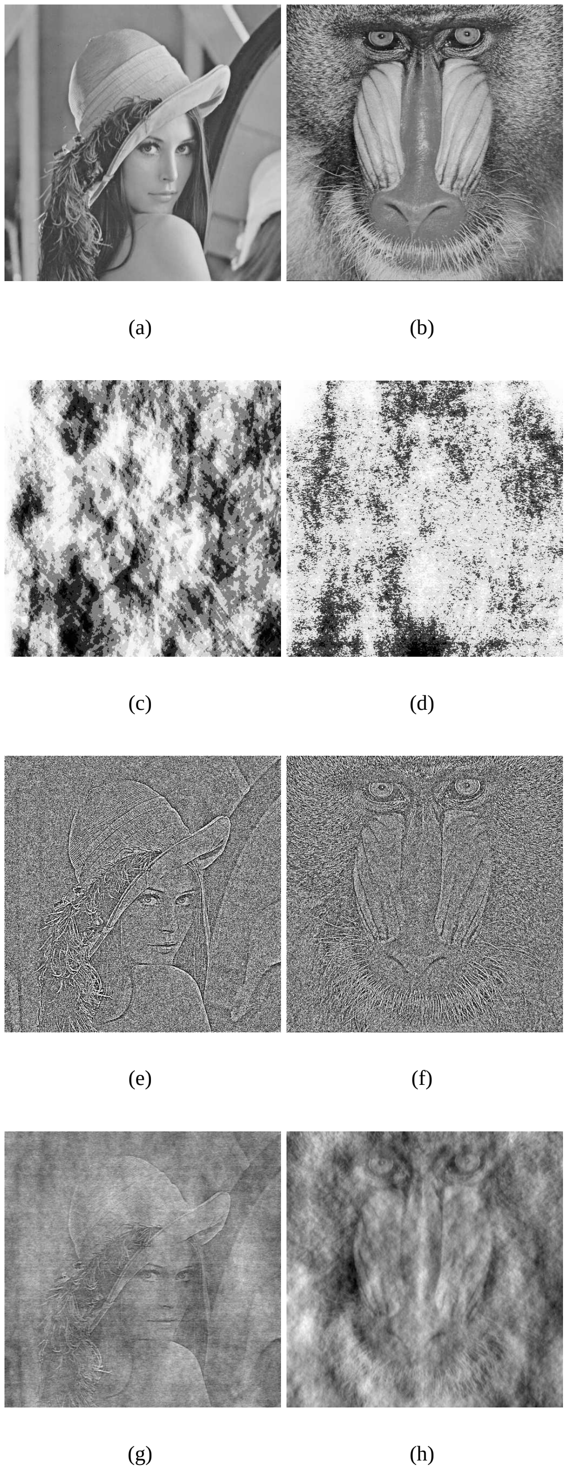


Figure 1: *Examples of the relevance of the DCT-phase in images [18]. Original images: (a) Lena; (b) Baboon. IDCT reconstructed images from: (c) DCT-magnitude of Lena with DCT-phase* $\equiv 0$*; (d) DCT-magnitude of Baboon with DCT-phase* $\equiv 0$*; (e) DCT-phase of Lena with DCT-Magnitude* $\equiv 1$*; (f) DCT-phase of Baboon with DCT-magnitude* $\equiv 1$*; (g) DCT-phase of Lena with DCT-magnitude of Baboon; (h) DCT-phase of Baboon with DCT-magnitude of Lena.*

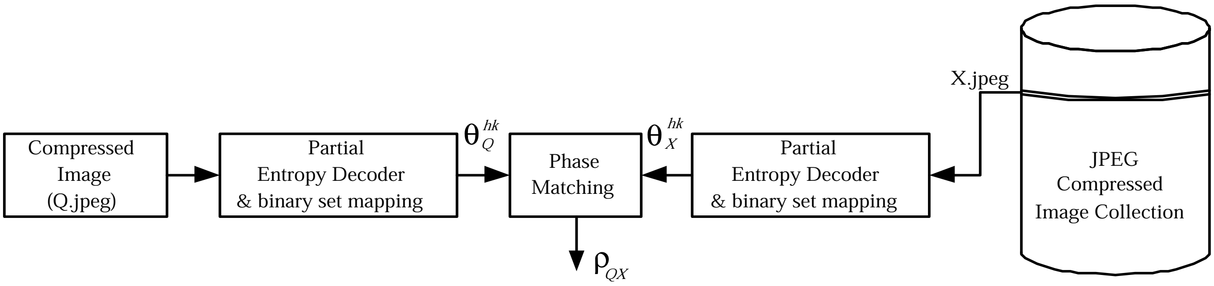


Figure 2: *DCT-phase matching scheme.*

The reconstructed image in Figure 1(g) is the result of the IDCT when applied on the magnitude of the DCT coefficients of Baboon combined with the DCT-phase of Lena; the result of the alternative magnitude-phase combination is shown in Figure 1(h). It is clear from these images that the DCT-phase prevails over the magnitude in this reconstruction process. It is remarked that in order to highlight the content of the reconstructed images in Figures 1(c) to 1(f), the result of the IDCT was normalized to the range [0, 255], and then contrast enhanced by histogram equalization.

3. Image Matching Scheme

Considering the results of the reconstructed images in the previous section, an image matching algorithm was studied and implemented. The underlying rationale of the algorithm is that given the significant amount of information conveyed by the phase, a phase-only matching scheme can provide a reliable metric of correlation between two images. Since the phase lies in the DCT domain, the method is inherently suitable for JPEG compressed images.

Unlike the Fourier phase, which is a real-valued number that requires expensive (floating-point) arithmetic and memory for its processing, the phase of a DCT coefficient is simply represented with the binary set $\{0, \pi\}$. This makes DCT-phase-based processing algorithms attractive from a computational complexity perspective.

A general implementation scheme for the matching of two JPEG compressed images is shown in Figure 2 with related details being described in this section and in Section 3.2. A partial entropy decoding is all that is required to extract the phase information from the file or bitstream in order to proceed with the image correlation operation. The phase values produced by the partial entropy decoder belong to the binary set $\{0, 1\}$, which represent respectively, negative and positive DCT coefficients, considering that in JPEG zero-valued DCT coefficients are not explicitly entropy coded, except for the DC values. Depending on the nature of the correlation operations, it might be useful to map this binary set into an alternative application-dependent and/or implementation-friendly binary or ternary set, as will be illustrated below.

3.1. Phase representation of quantized coefficients

The reconstructing IDCT operator that produces the results shown in Figures 1(c) to 1(h) was applied over non-quantized DCT coefficients. For the images Lena and Baboon in Figure 1, there were many DCT coefficients with a magnitude very close to zero, but none was an exact zero (in the study referred to in [18], the coefficients were represented as real numbers and a high resolution arithmetic was used in the software implementation of the DCT and IDCT). Thus, formally, the DCT-phase was completely represented with the binary set $\{0, \pi\}$.

Lossy image compression is in general characterized by executing some sort of quantization. The application of the latter forces many coefficients to become zero. In some cases, for practical purposes, it might be useful to associate the zero-valued coefficients with positive coefficients, and thus, the original binary-valued status of the DCT phase remains unchanged. Alternatively, when the effects of the quantization should be emphasized, one can consider extending the previous binary set to a more accurate ternary set, that would include a particular implementation-dependent symbol to represent the phase of zero-valued coefficients. In the correlation process in Figure 2, the latter choice was indeed used since it produced a higher matching accuracy while also contributing to find matching images, even when they have been encoded with different compression ratios. Thus, as indicated in Figure 2, the output binary set $\{0, 1\}$ produced by the partial entropy decoder was further mapped to the ternary set $\{-1, 0, 1\}$, which represents respectively, negative, zero-valued, and positive DCT coefficients. The mapping operation uses information from the variable length code *VLC1* (Section 5) to identify zero-valued coefficients.

3.2. Correlation metric

Once the ternary representation of the phase is available, multiple metrics can be implemented to determine the correlation of two images. In this paper we describe an example of a simple correlation metric that proved to be effective in the context of exact image retrieval.

Referring to Figure 2 and Figure 3, for a given compressed image Q with a horizontal and vertical pixel resolution of W and H respectively, the output of the mapping unit is a ternary-valued DCT-phase matrix θ_Q of $(W \times H)$ elements. In accordance with the (8×8)-element block-based processing of JPEG, this matrix can also be expressed as θ_Q^{hk}, where the indexes h and k identify the corresponding (8×8)-element DCT-phase subblock that compose the complete

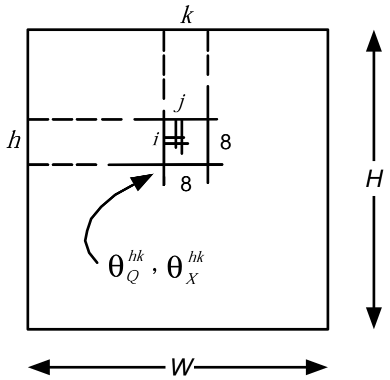


Figure 3: *Indexing scheme used for the DCT-phase matrices θ_Q and θ_X.*

$(W \times H)$-element array, where $h = 0,1,2,\ldots,(H/8)-1$, and $k = 0,1,2,\ldots,(W/8)-1$. By following the same notation, the DCT-phase array of the image X can be expressed as θ_X^{hk}.

A simple measure of the correlation between the images Q and X can be defined as:

$$\rho_{QX} = \sum_{hk}\sum_{ij}\theta_Q^{hk}\theta_X^{hk} \qquad (1)$$

where, $i, j = 0,1,2,\ldots,7$, represent the row and column indexes within an (8×8)-element block as illustrated in Figure 3. In normalized form, the previous correlation function can be expressed as:

$$\rho_{QXn} = \frac{1}{\alpha WH}\sum_{ycc}\rho_{QX_{ycc}} \qquad (2)$$

The index *ycc* iterates over the resulting sum in Equation (1) for each of the Y, Cb, and Cr bands, when they are all available. Accordingly, the value α is used to adjust the normalization factor $(W \cdot H)$ depending on whether the compressed data corresponds to a monochrome image, with $\alpha = 1$, or to a color image, in which case $\alpha = 1.5$ due to JPEG's 4:2:0 chroma subsampling ratio. This variable can also be used to adjust the normalization factor, in case the upper limits of the inner sum indexes, i, j, are selected such as to reduce the number of matchings within the basic (8×8)-element blocks (e.g., when discarding the contribution of the highest frequency coefficients for computational savings purposes).

In words, the correlation variable ρ_{QX} in Equation (1) is incremented each time the phase components of two non-zero-valued DCT coefficients match, and decremented when they do not. Zero-valued coefficients do not contribute, in the current case studied, to the correlation sum.

4. Results

The performance of the DCT-phase correlation method in exact matching retrieval and using the selected metric above was tested by running a large number of exact queries on a database of 6'800 color images. More specifically, this database corresponds to a subset of all the (128×96)-pixel images from the 10'000-image Corel image collection that is available in [19].

Figure 4 shows the results of some of the launched queries. The images on the left column represent simultaneously, both the image that was used as the query image and the exact matching image that was found in the database. The images on the right column represent for each row the second best matching images among the 6'800 pictures. The same identification number of the pictures given in [19] is indicated in the caption of the images along with the corresponding normalized correlation measure obtained with Equation (2). Both luminance and chrominance values were considered for computing this correlation function.

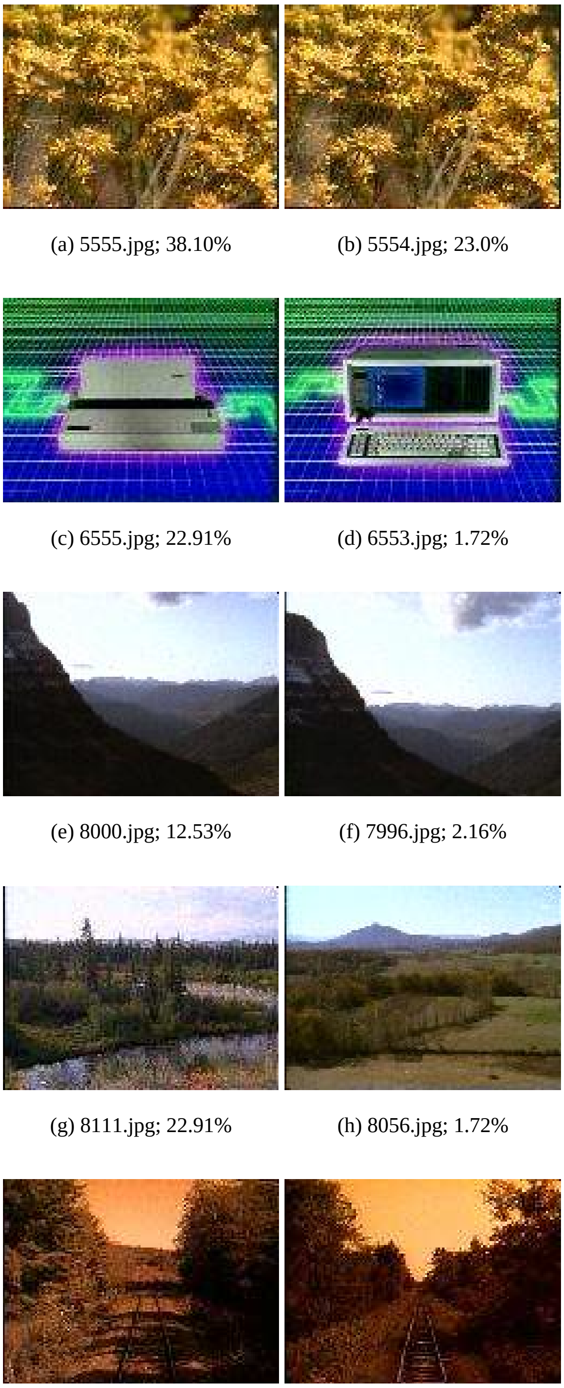
It is worth highlighting the significant distance between the correlation values of the exact matching image and the second best result. This was the case for *all* the launched queries carried out in this study, which demonstrates the selectivity power of the proposed method to match the *right* image. In this study, the highest correlation value obtained among all the images identified the exact copy of the query found in the 6'800 images database.

5. Computational Complexity

The computational complexity of the matching algorithm is easily deduced from Figure 2 and from Equation (1). In JPEG the entropy code of the quantized, zigzag reordered, and runlength coded DCT coefficients is composed of the concatenation of two variable length codes: *VLC1* and *VLC2* [20]. The binary code *VLC1* is a regular Huffman code that has been properly defined in a given Table. This code is chosen in function of the magnitude range of the quantized coefficient and on the number of preceding zero-valued coefficients. *VLC2* is a code associated with the binary representation of the DCT coefficient. The very first bit of *VLC2* determines the phase of the quantized DCT coefficient. Thus, after the detection of each *VLC1*, reading an extra bit from the JPEG file or bitstream produces the phase array. For simplification, the decoding description above omits the details of handling particular and infrequent irregularities in the processing pipeline.

The just described phase generation process is executed by using low complexity logic functions. Expensive arithmetic operations are not involved. The speed of the phase extraction process can vary depending on whether the *VLC1* detection is made by scanning multiple bits in parallel or one single bit in each clock cycle.

A second simpler logic unit is used to map, as described in Section 3.1, the binary set $\{0,1\}$ produced by the partial entropy decoder into the set $\{-1,0,1\}$. The latter values are stored in a memory array, whose size depends on the spatial resolution $(W\times H)$ of the compressed image. The correlation unit can be executed with an adder/accumulator and an additional logic block composed of a few IF and AND statements/operators that carry out the comparison of the two ternary-valued phase symbols θ_Q^{hk} and θ_X^{hk}; alternatively, instead of a logic block, a minimum complexity two-bit mul-



(a) 5555.jpg; 38.10% (b) 5554.jpg; 23.0%

(c) 6555.jpg; 22.91% (d) 6553.jpg; 1.72%

(e) 8000.jpg; 12.53% (f) 7996.jpg; 2.16%

(g) 8111.jpg; 22.91% (h) 8056.jpg; 1.72%

(i) 8223.jpg; 24.81% (j) 8236.jpg; 2.02%

Figure 4: *Results of the exact match queries. The left column shows simultaneously both the queries and their exact matching image found in the database. The right column shows the second best matching result. The identification number of the images and their related normalized correlation measure obtained from Equation (2) are indicated.*

tiplier could also be used as explicitly expressed in Equation (1). The results presented in Section 4 are for the case where the i, j indexes address the full 64 coefficients in an (8×8)-element block. As it is well known, and regularly applied in many algorithms, the complexity can be reduced by discarding the contribution of the highest frequency coefficients.

It is important to highlight that with respect to a generic CBIR system, a set of pre-computed image features is not present in the basic scheme shown in Figure 2. In effect, the low complexity of the matching algorithm allows the feature vector, i.e., in this case the DCT-phase of the image, to be generated in real time. All the same this matching process can be largely accelerated, in exchange of memory, if a phase-based image feature is pre-calculated. This is currently the object of further study.

6. Conclusions

Be it on the Internet, in databases, on consumer electronics and particularly in the digital still image industry, JPEG [20] is currently, and by far, the most popular compressed image format. In addition to its industry standard status, one of its main advantages is its modest computational complexity and its good image encoding quality at medium and especially at higher bit rates.

This paper introduced a new, simple, and efficient low complexity image matching technique that operates directly in the compressed JPEG domain. The underlying method is based on the exploitation of the rich information conveyed by the phase component of DCT coefficients. The effectiveness of the proposed technique has been illustrated in the frame of the particular exact matching retrieval application. The results of the study carried out demonstrate the high matching accuracy of the algorithm, as confirmed by the systematic true positive outcome obtained for *all* the launched queries. Furthermore, the proposed method features a high discrimination capability to find the right image, as reflected by the significant distance between the correlation values obtained for the exact matching image and the second best matching one. The processing requirements of the matching algorithm are particularly low, being fulfilled with the use of three different and simple logic units, and one adder/accumulator.

The image matching approach presented in this paper represents a first conceptual scheme that is open to multiple extensions and optimizations for rendering it more versatile in CBIR and in other applications. For example, given that most ISO and ITU video compression standards, MPEG-x and H.26x, are all DCT-based, the underlying principle of the proposed technique could also be applied in different kinds of correlation processes of video frames in the compressed domain.

Acknowledgements

This work was partially supported by the Swiss Federal Office for Education and Science under Grant OFES C00.0105 (COST 276 Research Project).

References

[1] B.C. Smith, L.A. Rowe, "Compressed Domain Processing of JPEG-encoded Images", *Real-Time Imaging J.*, **2**(1):3–17, 1996.

[2] M.K. Mandal, F. Idris, S. Panchanathan, "A Critical Evaluation of Image and Video Indexing Techniques in the Compressed Domain", *Image and Vision Computing Journal*, Special issue on Content Based Image Indexing, **17**(7):513–529, May 1999.

[3] P.H.W Wong, O.C. Au, "A blind watermarking technique in JPEG compressed domain", *Proc. IEEE Int'l Conf. on Image Processing (ICIP02)*, **3**:497–500, Sept. 2002.

[4] A. del Bimbo, *Visual Information Retrieval*, Morgan Kaufman Publ., USA, 1999.

[5] The Benchathlon Network: *CBIR Challenges*, http://www.benchathlon.net/resources/challenges.html

[6] G. Feng, J. Jiang, "JPEG image retrieval based on features from the DCT domain", *Proc. Int'l Conf. on Image and Video Retrieval (CIVR02)*, Lecture Notes in Computer Science, **2383**:120–128, Springer, July 2002.

[7] J.R. Smith and S.-F. Chang, "Transform features for texture classification and discrimination in large image databases", *Proc. IEEE Int'l Conf. on Image Processing (ICIP94)*, **3**:407–411, Nov. 1994.

[8] J.A. Lay, L. Guan, "Image Retrieval Based on Energy Histograms of the Low Frequency DCT Coefficients", *Proc. IEEE Int'l Conf. on Acoustics, Speech, and Signal Processing (ICASSP99)*, **6**:3009–3012, March 1999.

[9] A. Armstrong, J. Jiang, "An Efficient Image Indexing Algorithm in JPEG Compressed Domain", *Proc. IEEE Int'l Conf. on Consumer Electronics (ICCE01)*, pp. 350–351, June 2001.

[10] S. Panchanathan, "Compressed or Progressive Image Search", Chapter 16 in *Image Databases: Search and Retrieval of Digital Imagery*, Eds. V. Castelli, L. Bergman, John Wiley & Sons, USA, 2002.

[11] N.F. Johnson, *In Search of the Right Image: Recognition and Tracking of Images in Image Databases, Collections, and the Internet*, Technical Report, CSIS-TR-99-05-NFS, Center for Secure Information Systems, George Mason University, Fairfax, VA, USA, April 1999.

[12] D.S. Guru, P. Punitha, "An Invariant Scheme for Exact Match Retrieval of Symbolic Images Based upon Principal Component Analysis", *Pattern Recognition Letters*, **25**(1):73–86, Jan. 2004.

[13] C.-C. Chang, T.-C Wu, "An exact match retrieval scheme based upon principal component analysis", *Pattern Recognition Letters*, **16**(5):465–470, May 1995.

[14] P. Bosch, A. van Ballegooij, A. de Vries, M. Kersten, "Exact matching in image databases", *Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME01)*, pp. 513–516, August 2001.

[15] R. Reeves, K. Kubik, "Compressed domain image matching using symmetric convolution", *Proc. IEEE Annual Conf. on Speech and Image Technologies for Computing and Telecommunications (TENCON97)*, **2**:543–546, Dec. 1997.

[16] R. Reeves, K. Kubik, "Least squares matching in the transform domain", *International Archives of Photogrammetry and Remote Sensing*, **32**(3/1):168–176, July 1998.

[17] F. Ackermann, "Digital image correlation: Performance and potential application in photogrammetry", *Phogrammetric Record*, **11**(64):429-439, Oct. 1984.

[18] J. Bracamonte, *"The DCT-phase of images and its applications,"* Technical Report IMT No. 451 PE 01/04, Institute of Microtechnology, University of Neuchâtel, Switzerland, January, 2004.

[19] J. Wang, *Related Links*, http://wang.ist.psu.edu/

[20] W.B. Pennebaker, J.L. Mitchel, *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, USA, 1993.