Published in Lecture Notes in Computer Science (LNCS) 3568, 154-163, 2005
which should be used for any reference to this work

Efficient Compressed Domain Target Image Search and Retrieval

Javier Bracamonte, Michael Ansorge, Fausto Pellandini, and Pierre-André Farine

Institute of Microtechnology, University of Neuchâtel
Rue A.-L Breguet 2, 2000 Neuchâtel, Switzerland
{javier.bracamonte, michael.ansorge, fausto.pellandini, pierre-andre.farine}@unine.ch

Abstract. In this paper we introduce a low complexity and accurate technique for target image search and retrieval. This method, which operates directly in the compressed JPEG domain, addresses two of the CBIR challenges stated by The Benchathlon Network regarding the search of a specific image: finding out if an exact same image exists in a database, and identifying this occurrence even when the database image has been compressed with a different coding bit-rate. The proposed technique can be applied in feature-containing or featureless image collections, and thus it is also suitable to search for image copies that might exist on the Web for law enforcement of copyrighted material. The reported method exploits the fact that the phase of the Discrete Cosine Transform coefficients contains a significant amount of information of a transformed image. By processing only the phase part of these coefficients, a simple, fast, and accurate target image search and retrieval technique is achieved.

1 Introduction

Nowadays most image and video data are both stored and transmitted in compressed form. By processing these data directly in the compressed domain important savings can be made in terms of computational and memory requirements, processing speed, and power consumption. These savings come from the fact that it is no longer necessary to allocate resources to the computationally-intensive decompression modules and from the advantage that the amount of data to process is significantly less in the compressed domain. Multiple compressed-domain algorithms have been reported in the signal processing literature for a wide range of applications [1][2][3].

In this paper we address the issue of target image search and retrieval in the compressed domain, an issue listed by The Benchathlon Network among the multiple content-based image retrieval (CBIR) challenges [4].

Target search can be used by an image owner to track down with a Web crawler copy images that might be utilized elsewhere on the Internet without the proper rights. Specific target search can also be used as part of a more complex automated CBIR system for trademark image retrieval applications, where it would assist the verification of whether a potential new logo has been previously registered [5]. Additionally, target search can be of support for testing purposes [5], for the management of large image databases (e.g., to find duplicates), and is a particular option available in commercial image search, retrieval, and management systems [6][7].

Since the vast majority of images that populates the Web exist in compressed JPEG format, it is of large interest to save time, considering the huge number of existing images, that the target searching process takes place directly in the JPEG domain. A similar statement can be made regarding multi-million compressed-image databases. Additionally, given the lack of an associated feature-vector for most of the images available on the Web, the use of complex feature-based retrieval schemes would be computationally inefficient.

In this paper we introduce a low complexity and efficient technique for target search that copes with the issues reported in the previous paragraph. That is, it operates directly in the compressed JPEG domain, and it is applicable both to feature-containing and featureless image collections. The study of two different search cases will be discussed. First, when an exact copy of the target image is part of the database, and second, when the copy of the query in the database has been JPEG encoded with a higher compression ratio. The results of computational-complexity-reduction schemes aiming at speeding up the target search process will also be reported.

1.1 Previous Work

Previous papers have reported methods for indexing and retrieval in the compressed DCT domain. The algorithm reported in [8] 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 [9]. An energy histogram technique similar in concept to the pixel-domain color histogram method has been proposed in [10]. 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 [11] and [12] 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. A good synthesis on further indexing and retrieval techniques in the DCT-domain is reported in [13].

The previously cited papers deal with compressed-domain indexing/retrieval methods that are in general suitable for *similarity* queries in CBIR systems. Other studies more related to or specifically addressing the *exact image query* issue [14][15] have been proposed in [5][16][17][18][19]. The latter are all pixel-domain-based techniques and are rather suitable for feature-containing image collections.

1.2 Organization of the Paper

The remainder of this paper is organized as follows. Section 2 will present an overview of the basic principle underlying the image matching technique used by the proposed target search algorithm. The image matching procedure itself and the exact image search system will the described in Section 3. A small sample of the results of a large number of query tests will be introduced in Section 4, while Section 5 will present the results of some computational complexity reduction approaches. Finally, the conclusions will be stated in Section 6.

2 The DCT-Phase of Images

A study on the significance of the DCT-phase in images was reported in [20] where it is showed that the DCT-phase in spite of its reduced binary value $\{0, \pi\}$ conveys a significant amount of information on its associated image. An example from [20] is reproduced in Figure 1 and is briefly described below.

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 Inverse DCT (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 the binary-valued phase arrays and when 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) as described in the next paragraph.

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 Target Image Search Algorithm

Based on the reconstructed image results presented in the previous section, an image matching algorithm oriented to target image search was studied and implemented [21]. The rationale of this algorithm is that given the significant amount

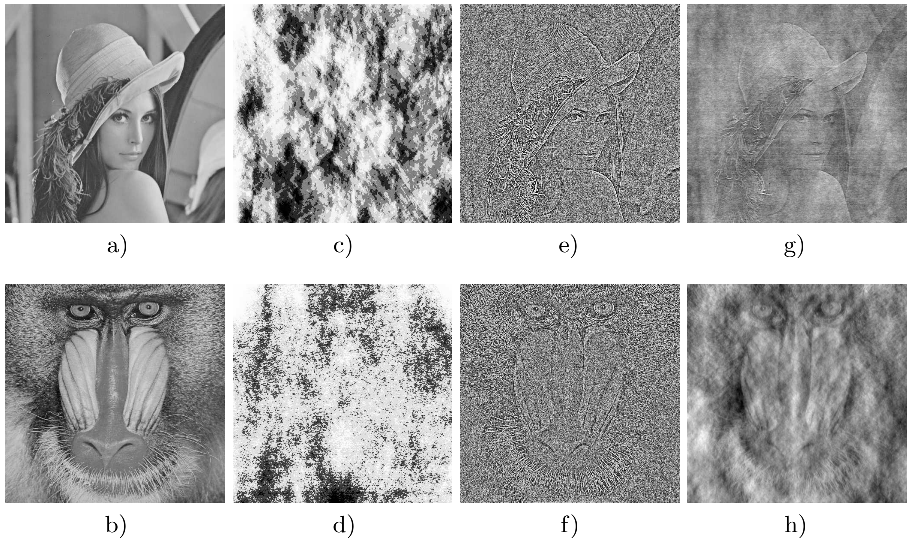

Fig. 1. Examples of the relevance of the DCT-phase in images [20]. Original images: (a) Lena; (b) Baboon. IDCT reconstructed images from: (c) DCT-magnitude of Lena with DCT-phase $\equiv 0$; (d) DCT-magnitude of Baboon with DCT-phase $\equiv 0$; (e) DCT-phase of Lena with DCT-Magnitude $\equiv 1$; (f) 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.

of information conveyed by the phase of the DCT-coefficients, a phase-only-processing scheme can provide a reliable metric of the correlation between two images, and thus be an efficient mean for specific target search. Since JPEG, currently and by far the most widely used image compression algorithm, is DCT-based, the resulting target search method suits perfectly for querying in databases composed of JPEG encoded images, and/or to explore the Web, where JPEG predominates as compressed image format.

An implementation scheme of the target search system is depicted in Figure 2. The image search space can be either a structured image collection or a streaming set of images from the Web. A partial entropy decoder followed by an elementary mapping unit are the only operations required to extract the DCT phase information from the files or bitstreams. The DCT-phase of the query and that of the database images are then compared by executing a correlation metric.

3.1 Correlation Metric

Referring to Figure 2 for a given compressed query image Q with a horizontal and vertical pixel resolution of W and H respectively, the output of the mapping unit is a ternary-valued $\{-1, 0, +1\}$ DCT-phase matrix θ_Q of $(W \times H)$ elements [21]. In accordance with the (8×8)-element block-based processing of

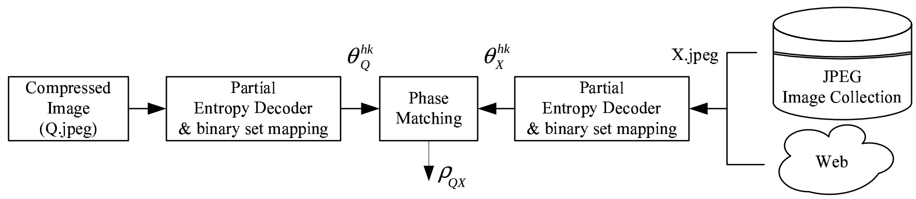

Fig. 2. Image matching scheme for the target search algorithm.

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 subblocks that compose the complete $(W \times H)$-element array, and 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}.

Among the multiple evaluated metrics to estimate the correlation between the images Q and X, one that proved efficient for the exact matching application was:

$$\rho_{QX} = \sum_{hk} \sum_{ij} \theta_Q^{hk}(i,j) \theta_X^{hk}(i,j) \tag{1}$$

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

$$\rho_{QXn} = \frac{1}{\alpha W H} \sum_{m} \rho_{QX_m} \tag{2}$$

where the index m iterates over the resulting sum in Equation (1) for each of the Y, Cb, and Cr bands, in case 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, in which case $\alpha = 1$, or to a color image, where $\alpha = 1.5$ due to JPEG's 4:2:0 chroma subsampling ratio.

4 Results

The target search algorithm was intensively tested by submitting a large number of exact match queries to a database of 6'800 color images. This database corresponds to a subset of all the (128×96)-pixel images from the 10'000-element Corel image collection that is available in [22].

A sample of these target search results is shown in Figure 3. The image on the left column of each set represents simultaneously (to save space and avoid redundancy) the query, i.e., the target searched image, and its perfect matching image found in the 6'800-image database. For illustration purposes, the images on the right column of each set correspond to the second best matching image. The same identification number given to the pictures in [22] is indicated as

Fig. 3. Results of the target specific search. The captions indicate the image identification number (ID No.) along with its normalized correlation score obtained with Equation (2).

caption of each image, along with the normalized correlation measure obtained with Equation (2).

Given the main goal of target search, i.e., to identify the existence of an exact copy of the query image, the visual content of the second best matching image might not be relevant. In Figure 3 it is very important however to remark upon the high discriminating power of the proposed technique by noting the dramatic difference of the correlation values, ρ, for the exact matching image (found target) ρ_M, and that of the second best matching picture ρ_{SB}.

The accuracy and the robustness of the target search system was substantiated by the true positive outcome for all the large number of launched queries.

Table 1. Results of target search with copies encoded with higher compression ratios.

Query Image (ID No.)	CR of the query image (CR1)	CR of the found target (CR2)	Correlation score of the found target (%)	Correlation score of the 2nd best match (%)	CR2/CR1
1255	6.38	38.12	1.61	1.57	5.9
1329	20.30	39.47	1.69	1.67	1.9
1360	13.24	39.77	1.57	1.56	3.0
2355	14.12	31.59	1.67	1.65	2.2
5555	6.62	9.33	27.03	23.00	1.4
6555	8.35	14.97	12.23	11.87	1.7
7234	7.43	38.28	1.57	1.56	5.2
8000	16.21	35.83	2.26	2.16	2.2
8111	10.47	37.81	1.73	1.72	3.6
8223	9.85	37.39	2.05	2.02	3.7

The high selectivity power was equally reconfirmed for *all* the launched queries even in cases where the database contains very similar images that are, for a human observer, visually identical to the target, e.g., when querying with image 5555.jpg.

4.1 Target Search of Higher-Compressed Image Copies

This section presents the results concerning the robustness of the target search algorithm to identify copies when the latter have been encoded with a higher compression ratio (CR) than the query image. The goal was to find out how much the image copy can be compressed and still be correctly identified as a replica of the query in a database. The results of this study for the same set of images in Figure 3 are given in Table 1.

The CRs of the *original* query images in JPEG format are shown in the second column of Table 1 in which the variable coding bit-rate nature JPEG is put in evidence. In effect, for a similar image reconstruction quality, as it is fairly the case for all the images in the database, the CR obtained with JPEG varies in function of the image-complexity content of the pictures.

The third column of Table 1 shows the maximum CR that can be applied to the target's image copy in the database so that the search procedure continues to produce true positive results. In other words, this is the maximum CR for which the target's copy produces a correlation score (which is shown in the fourth column) that is still the highest for all the images in the database. For comparison, the second best score previously reported in Figure 3 is shown in the fifth column.

It is noticeable from columns three and six the excellent robustness of the search target algorithm to identify higher-compressed versions of the queried im-

Table 2. Results of the computational complexity reductions approaches.

Query image (ID No.)	Ratio of correlation scores: found target / second best match						
	Y	Y_{32}	Y_{16}	Y_8	Y_4	Y_2	Y_1
1255	27.57	20.11	16.37	9.45	5.98	3.72	3.08
1329	4.61	4.53	4.18	3.29	2.44	1.60	1.47
1360	10.48	10.15	8.96	5.52	3.72	2.65	2.08
2355	9.85	9.51	7.75	4.93	3.07	2.29	1.74
5555	1.73	1.56	1.39	1.27	1.22	1.22	1.18
6555	2.15	2.23	2.46	2.45	2.24	2.10	1.87
7234	18.87	16.40	11.57	8.55	6.31	4.24	3.30
8000	6.48	6.26	4.76	3.60	2.23	1.80	1.59
8111	12.61	11.06	8.44	5.75	3.91	3.27	2.45
8223	12.82	10.56	7.62	5.01	3.39	2.74	2.78

ages. In effect, the search algorithm continues to produce true positive outcomes even with replicas that have undergone a dramatic increase of CR.

5 Computational Complexity Reduction Schemes

It can be easily confirmed from Figure 2 and Equation (2) that the amount of computational resources needed by the target search algorithm is quite low. These requirements are limited to three logic units which carry out respectively: the partial Huffman decoding, the mapping unit, and the phase comparison in the correlation unit, the latter being completed with an accumulator.

Beyond these very limited hardware/software implementation requirements, an additional dramatic reduction of the computational complexity can be achieved at the algorithm level. This is possible by lowering the amount of processed data, by excluding both the image's color bands and the contribution of the high frequency DCT-coefficients in the DCT-phase correlation operation as reported in the next paragraphs.

Table 2 shows the results of different computational complexity reduction schemes for the same set of images shown before. The column identified as Y corresponds to the results when only the luminance data is considered in the computation of Equation (2). Then, Y_N, indicated in the remaining columns, corresponds to the scheme in which only the first N coefficients (when ordered in JPEG's typical zig-zag scanning pattern [23]) of each 64-element block of the luminance are considered in the computation of Equation (1). Thus, Y_1 for example, corresponds to the case where only the (DPCM encoded) DC coefficients of each luminance block are evaluated.

The entries given in Table 2 correspond to the ratio of the correlation scores ρ_M/ρ_{SB}. It can be noticed that the value of these entries are all greater than 1.0, a fact that was consistently observed for all the launched queries in this study.

This demonstrates the robustness of the target search system when using one of these computational complexity reduction schemes, since true positive results continue to be produced. When evaluating the data in Table 2, it is important to recall that for a given query image the second best image is not necessarily the same image when passing from column Y_N to column $Y_{N/2}$.

6 Conclusions

This paper introduced a new, simple, and efficient target search algorithm that operates directly in the JPEG domain, along with computational complexity reduction schemes. The presented results showed that the proposed technique features a high selectivity power and a very good robustness to changes in the compression ratio of the matching images. The algorithm is based on the exploitation of the rich information conveyed by the phase component of DCT coefficients. The method can thus be extended to related video applications in which the bitstream has been generated with any of the ISO or ITU DCT-based video compression standards such as MPEG-x or H.26x.

This technique can be used straightforwardly at its current status in those applications in which the copies of the target image have not been purposely altered, for example, in database management, in CBIR testing, in searching on the Web for non-authorized use of images from users unaware of the intellectual property status of the utilized images, or as an embedded option in trademark registration systems, to name a few.

This target search system is obviously open to multiple optimizations in order to render it robust against other kinds of image modifications. This is currently the object of further study.

Acknowledgements

Dr. Stéphane Marchand-Maillet is thankfully acknowledged for the interesting discussions regarding CBIR issues and the content of this paper. This work was partially supported by the Swiss State Secretariat for Education and Research under Grant SER C00.0105 (COST 276 Research project).

References

1. Smith B.C., Rowe L.A.: Compressed domain processing of JPEG-encoded images, Real-Time Imaging J., 2(1):3–17, 1996.
2. Mandal M.K., Idris F., Panchanathan S.: 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. Wong P.H.W., Au O.C.: A blind watermarking technique in JPEG compressed domain, Proc. IEEE Int'l Conf. on Image Processing (ICIP02), 3:497–500, Sept. 2002.

4. The Benchathlon Network: CBIR Challenges – http://www.benchathlon.net/resources/challenges.html
5. Cox I.J., Miller M.L., Minka T.P., Papathomas T.V., Yianilos P.N.: The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments, IEEE Trans. on Image Processing, 9(1):20–37, Jan. 2000.
6. Image-Seeker^{TM}– http://www.ltutech.com
7. Acdsee^{TM}– http://www.acdsee.com
8. Feng G., Jiang J.: 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.
9. Smith J.R., Chang S.-F.: 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.
10. Lay J.A., Guan L.: 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.
11. Armstrong A., Jiang J.: An efficient image indexing algorithm in JPEG compressed domain, Proc. IEEE Int'l Conf. on Consumer Electronics (ICCE01), pp. 25–30, June 2001.
12. Jiang J., Armstrong A., Feng G.C.: Web-based image indexing and retrieval in JPEG compressed domain, Multimedia Systems, 9:424–432, 2004.
13. Panchanathan S.: Compressed or progressive image search, Chapter 16 in Image Databases: Search and Retrieval of Digital Imagery, Eds. Castelli V., Bergman L., John Wiley & Sons, USA, 2002.
14. Bimbo A. del: Visual Information Retrieval, Morgan Kaufman Publ., USA, 1999.
15. Manolopoulos Y., Theodoridis Y., Esotras V.J.: Advanced database indexing, Kluwer Academia Publishers, USA, 1999.
16. Johnson N.F.: 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.
17. Guru D.S., Punitha P.:, An invariant scheme for exact match retrieval of symbolic images based upon principal component analysis, Pattern Recognition Letters, 25(1):73–86, Jan. 2004.
18. Chang C.-C., Wu T.-C.: An exact match retrieval scheme based upon principal component analysis, Pattern Recognition Letters, 16(5):465–470, May 1995.
19. Bosch P., Ballegooij A. van, Vries A. de, Kersten M.: Exact matching in image databases, Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME01), pp. 513–516, Aug. 2001.
20. Bracamonte J.: The DCT-phase of images and its applications, Technical Report IMT No. 451 PE 01/04, Institute of Microtechnology, University of Neuchâtel, Switzerland, Jan. 2004.
21. Bracamonte J., Ansorge M., Pellandini F., Farine P.-A.: Low complexity image matching in the compressed domain by using the DCT-phase, Proc. of the 6th COST 276 Workshop on Information and Knowledge Management for Integrated Media Communications, Thessaloniki, Greece, pp. 88–93, May, 2004.
22. Wang J.: Downloads/Related Links – http://wang.ist.psu.edu/docs/related
23. Pennebaker W.B., Mitchel J.L.: JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, USA, 1993.