RETRIEVAL OF OCCLUDED IMAGES USING DCT PHASE AND REGION MERGING

P. Ayyalasomayajula, S. Grassi, P.-A.Farine

Institute of Microengineering, Electronics and Signal Processing Laboratory, Ecole Polytechnique Fédérale de Lausanne, EPFL STI IMT ESPLAB, Rue A.-L. Breguet 2, 2000, Neuchâtel, Switzerland

ABSTRACT

In this paper we present an efficient method for Content Based Image Retrieval (CBIR) of occluded images using DCT-phase. The proposed method utilizes a novel correlation metric for ternary-valued DCT-phase, as well as a region merging method to reconstruct the non-occluded regions in the retrieved image. The proposed image retrieval method showed good performance when tested with different datasets containing reference images, occluded images, fused images and images with different JPEG compression ratios. Experimental evaluation also showed that the proposed image retrieval method performs better than current state of the art DCT-phase based image retrieval methods while retrieving not only occluded images but also reference images, fused images and images with different JPEG compression ratios.

Index Terms— DCT phase, CBIR, region merging, occluded image retrieval, fused image retrieval

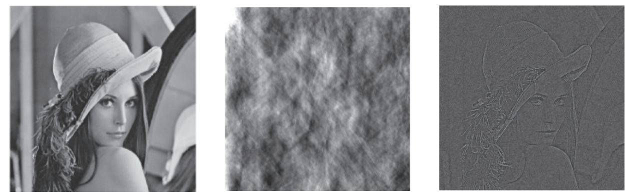
1. INTRODUCTION

As the number and size of digital image databases are increasing, the development of image retrieval systems is becoming increasingly important. Among them, Content Based Image Retrieval (CBIR) has gained lot of interest from researchers starting from early 1990's [1]. Nowadays images are usually stored and transmitted in compressed form. By processing them directly in the compressed domain, important savings in terms of computation, memory requirements and processing speed can be made. Multiple compressed-domain image retrieval algorithms are proposed in the literature [2].

Discrete Cosine Transform (DCT) is an essential processing block of JPEG compression. As the DCT-phase can be easily derived from JPEG compressed image with low complexity calculations, many methods have been proposed in literature [3],[4],[5],[6],[7] which use DCT-phase data to match and retrieve images from databases. DCT-phase based image retrieval is well suited for efficient implementation in embedded systems and especially in handheld devices [8]. These DCT-phase based methods generally work well with clean images. However, a key challenge is to keep their good performance in case of distortions or occlusions.

In this paper we propose a DCT-phase only image retrieval method for occluded and fused [9] images. This proposed method utilizes a novel correlation metric for ternary-valued DCT-phase. A new region merging method is also introduced, which is used to reconstruct the non-occluded regions of the image.

The paper is organized as follows. A brief introduction to DCT-phase and its ternary-valued representation is given in *Section 2*. The proposed image retrieval method and region merging technique are described in *Section 3*. In *Section 4* we present the experimental evaluation of the proposed method and comparison with other state of the art DCT-phase based image retrieval methods. Conclusions and Future work are discussed in *Section 5*.



(a) Original (b) Magnitude only (c) Phase only

Figure 1: (a) Monochrome test image, (b) IDCT applied to the magnitude array with constant phase, (c) Reconstruction when IDCT is applied over binary-valued phase arrays, when magnitude is set to one.

2. DCT-PHASE FOR IMAGE RETRIEVAL

A study on the significance of DCT-phase in images was reported in [4] where it is shown that the DCT-phase, in spite of its binary value {0,π}, conveys significant amount of information about its associated image. In this section we explain the DCT-phase of an image and its ternary-valued representation used for the new correlation metric proposed in *Section 3*.

2.1. DCT-phase of an image

Let $s(n,m)$ be a 2-dimmensional N x M sequence. The DCT of $s(n,m)$ is denoted by $S_{DCT}(j,k)$. It can be expressed in terms of its absolute value, $|S_{DCT}(j,k)|$, and its corresponding phase term, $S'_{DCT}(j,k)$ as:

$$S_{DCT}(j,k) = |S_{DCT}(j,k)|\, S'_{DCT}(j,k) \qquad (1)$$

Hereafter we refer to $S'_{DCT}(j,k)$ as the DCT-phase term, which can take two values, i.e. $S'_{DCT}(j,k) \in \{+1,-1\}$, as $S_{DCT}(j,k)$ is real valued.

The "DCT-phase only image" is the inverse DCT [10] of $S'_{DCT}(j,k)$ whereas the "DCT-magnitude only image" is the inverse DCT of $|S_{DCT}(j,k)|$. *Figure 1* shows an example of an image, its DCT-magnitude only image and its DCT-phase only image. As we can see from the example, high amount of information is conveyed in the DCT-phase.

Most digital images are compressed with JPEG compression standard, which uses DCT on 8x8 blocks of the image. In JPEG compressed images, the DCT coefficient signs of the 8x8 blocks can be obtained directly, with low complexity computation, by partial entropy decoding of the encoded bitstream [4]. We thus perform image retrieval in the compressed domain using the DCT-phase of the 8x8 blocks of the image.

2.2. DCT-phase representation

We represent the DCT-phase with values from the ternary set $\{+1,\ \alpha, -1\}$, where the '+1' symbol corresponds to $S'_{DCT}(j,k) = +1$ (0 phase), and the '−1' symbol corresponds to $S'_{DCT}(j,k) = -1$ ('π' phase). The symbol 'α' corresponds to zero magnitude of the DCT coefficients, $|S_{DCT}(j,k)| = 0$.

This ternary-valued representation provides an improvement over the binary-valued representation $\{+1, -1\}$, as it removes the ambiguity of phase for zero magnitude DCT coefficients, which occur quite frequently after the quantization step while compressing images in JPEG format.

3. RETRIEVAL METHOD FOR OCCLUDED IMAGES

Figure 2 shows the block diagram of the proposed image retrieval method for occluded images. This method first finds the best matching image, using a new correlation metric with ternary-valued DCT-phase which improves matching performance, especially in the case of occluded images (see *Section 4*). After the best matching image is found, intermediate values of the correlation metric calculation are used for region merging in order to retrieve the non-occluded portion of the image.

3.1. Correlation metric

As explained in *Section 2*, we first compute the ternary-valued DCT-phase of the 8x8 blocks of an image I_Q with horizontal and vertical pixel resolution of W and H respectively, obtaining a W x H matrix of ternary symbols. Hence forth, we express this matrix as θ^{hk}_{Query}, where *h* = 0,1,…,(H/8)-1 and *k* = 0,1,…,(W/8)-1. The indexes *h* and *k* identify the corresponding 8x8 block of the image I_Q.

To correlate a query image I_Q and a reference image I_{Ref}, we first multiply element-by-element their corresponding DCT-phase arrays, θ^{hk}_{Query} and $\theta^{hk}_{Reference}$:

$$\theta^{hk}_{Q,Ref}(i,j) = \theta^{hk}_{Query}(i,j) \cdot \theta^{hk}_{Reference}(i,j) \qquad (2)$$

Table 1 shows the possible outcomes for such a multiplication where $\theta^{hk}_{Query}(i,j)$ and $\theta^{hk}_{Reference}(i,j)$ can take values {+1, α, −1} and the outcome belongs to the quinary set $\{+1,\ \alpha, \alpha^2, -\alpha, -1\}$. An exact match is found when the outcome belongs to $\{+1,\ \alpha^2\}$, which forms the principal diagonal of the Matrix of possible outcomes given in *Table 1*.

		$\theta^{hk}_{Reference}(i,j)$		
		$+1$	α	-1
$\theta^{hk}_{Query}(i,j)$	$+1$	$\mathbf{+1}$	α	-1
	α	α	$\boldsymbol{\alpha^2}$	$-\alpha$
	-1	-1	$-\alpha$	$\mathbf{+1}$

Table 1: Possible outcomes for $\theta^{hk}_{Q,Ref}(i,j)$ from Eq. 2

After performing the multiplication between the query and the reference image, each element of the resultant arrays is assigned one of three possible "labels", Valid (V), Not-valid (N) and a

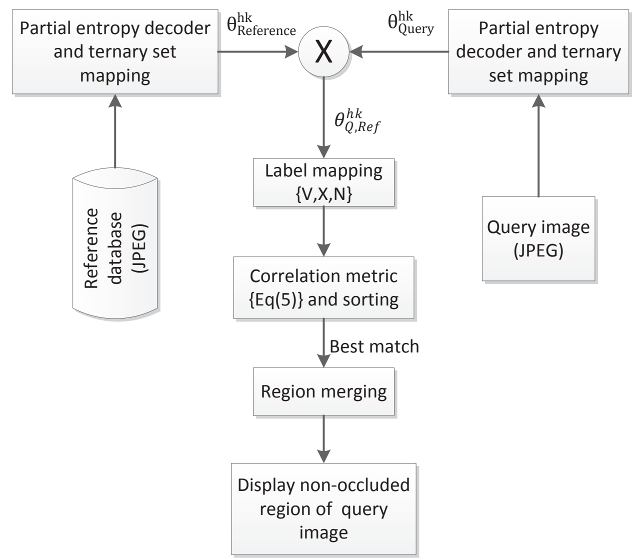


Figure 2: Block diagram of the proposed image retrieval method for occluded images

Don't care (X), which correspond to $\{+1\}$, $\{\alpha, -\alpha, -1\}$ and $\{\alpha^2\}$ respectively, as shown in *Table 2*.
The rationale behind labeling $\{\alpha^2\}$ as don't care (X) is that in JPEG compression the majority of DCT coefficient values are quantized to zero and these zero-valued DCT coefficients do not provide significant information for image retrieval using phase correlation.

		$\theta^{hk}_{Reference}(i,j)$		
		$+1$	α	-1
$\theta^{hk}_{Query}(i,j)$	$+1$	**V**	N	N
	α	N	**X**	N
	-1	N	N	**V**

V = Valid
N = Not-valid
X = Don't care

Table 2: Labels assigned to possible outcomes

The similarity between the query and the reference image is calculated using the following correlation metric:

$$d(I_Q, I_{Ref}) = \frac{\sum(V)}{\sum(V) + \sum(N)} \qquad (3)$$

Using Eq. 3, we calculate the correlation between the query image and each image within the searched database. The best matching image corresponds to the maximum correlation value in the searched database.

For clean non-occluded images, i.e. no distortion between the query image and its reference stored in the database, the similarity measure for the best match is one. If there is occlusion, the similarity measure gives the approximate percentage of non-occlude area.

3.2. Region merging

In the case of occluded images, once the best match is retrieved from the database using the metric in Eq. 3, we can find the region in the image which is not been occluded by performing region merging using the labels (Valid, Not-valid, Don't Care) assigned previously in the DCT-phase domain.

For the array $\theta_{Q,Ref}^{hk}$ of a given 8x8 block to which labels (V,N,X) have been assigned , we compute the mean of the number of labels which are either valid or don't cares. That is:

$$m^{hk} = \frac{1}{64}\sum_{i=0}^{7}\sum_{j=0}^{7} 1_{\{V,X\}}\left(\theta_{\text{Query}}^{\text{hk}}(\text{i},\text{j})\right) \quad (4)$$

If this mean is above a given threshold, m_{thres}, the 8x8 block is classified as a matching block. After all the matching blocks are found, those which are connected to at least one adjacent matching block are retained, and the others are discarded. The non-occluded portion of the retrieved image is formed by all the retained blocks.

4. EXPERIMENTAL EVALUATION

In the next sub-sections we explain the databases used for experimental evaluation, the testing procedure and the obtained results.

4.1. Experimental Database

The reference database used for the tests contains 1000 images from the COREL photograph data set used in [11] which are natural color JPEG compressed images with QVGA resolution (320x240).

The reference database was randomly divided into two datasets known as 'datasetA' and 'datasetB' each containing 500 images. A sample image from 'datasetA' is shown in *Figure 3(a)*.

An "occluded image database" was generated by occluding 37.5% of each image in 'datasetA' with part of an image in the 'datasetB': the lower-left triangular portion of the image was replaced by the same region from the occluding image. A sample image from this dataset is shown in *Figure 3(b)*.

To generate a database of fused images, called 'datasetFused', each image in the 'datasetA' was merged with an image from 'datasetB' using wavelet decomposition by taking the mean for both approximations and details coefficients (wfusimg function in Matlab). A sample image from this dataset is shown in *Figure 3(c)*.

Images in 'datasetA' were JPEG compressed with two different compression ratios (CR): 60% and 30%; creating two compressed databases 'dataset30CR' and 'dataset60CR'. A sample image from dataset30CR is shown in *Figure 3(d)*.

4.2. Testing Procedure and results

The proposed image retrieval method was implemented in MATLAB. The method was tested under the different conditions of occlusion, fusion and compression, corresponding to the datasets explained in *Section 4.1*. When testing a dataset, each image in a dataset was successively used as a query, searching it on the reference database. The best match always corresponded to the query image.

The ratio of the correlation of Best Match (BM) to the correlation of the Second Best Match (SBM) is used to measure the performance of the retrieval method. When testing a dataset, we calculate the mean and the standard deviation of this ratio. A high mean with low standard deviation is an indication of good performance and robustness.

Table 3 gives the summary of the resulting mean (m) and standard deviation (σ) of the BM/SBM correlation ratio, for the different datasets tested. *Figure 3(e)* to *3(h)* show the histograms of the BM/SBM correlation ratio for each dataset.

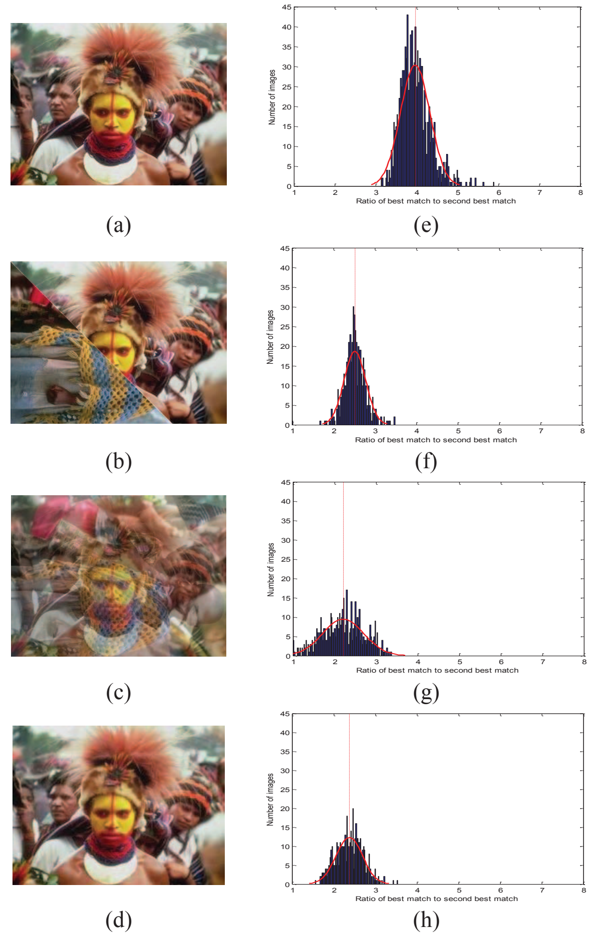


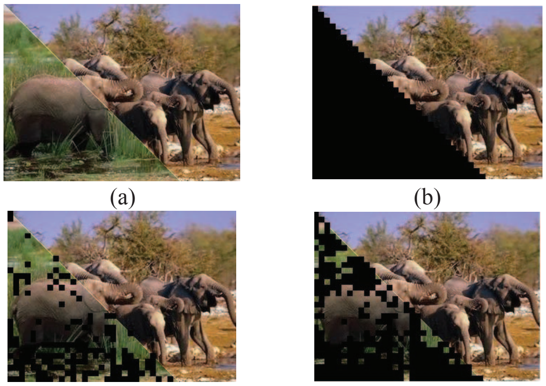
Figure 3: (a)-(d) Sample images from reference database, datasetOcclude, datasetFused and dataset30CR. Histogram of the BM/SBM correlation ratio for (e) reference dataset (f) datasetOcclude (g) datasetFused (h) dataset30CR.

It can be seen that the proposed image retrieval method provides a good distance between BM and SBM not only for the occluded set but also for reference, fused, and JPEG compressed images.

The proposed method was compared with other state of the art DCT-phase based methods for image retrieval [4], [5]. The method proposed in [5] uses a binary value for phase i.e., $\{+1,-1\}$ whereas [4] uses ternary-valued $\{+1,0,-1\}$ phase, but with a correlation metric that does not treat differently the "α^2" case, i.e. the case in which both DCT-magnitudes are zero (see *Subsection 3.1*).

	Proposed method		Method in [5]		Method in [4]	
Dataset	**m**	**σ**	**m**	**σ**	**m**	**σ**
Reference database	3.97	0.36	1.41	0.18	1.26	0.09
datasetOcclude	2.52	0.26	1.19	0.09	1.16	0.06
datasetFused	2.21	0.49	1.31	0.14	1.19	0.08
dataset30CR	2.37	0.32	1.04	0.03	1.13	0.05
dataset60CR	3.73	0.39	1.39	0.18	1.26	0.09

Table 3: Summary of test results for the different datasets



(a) (b)

(c) (d)

Figure 4: (a) Occluded image (b) Region Merging(RM) using our proposed method (c) RM on [5] (d) RM on [4]

Our region merging method successfully retrieved the non-occluded regions of the images. *Figure 4* shows an example of our proposed region merging method (b) and of region merging using the DCT-phase representation of [5] and [4] in (c) and (d)

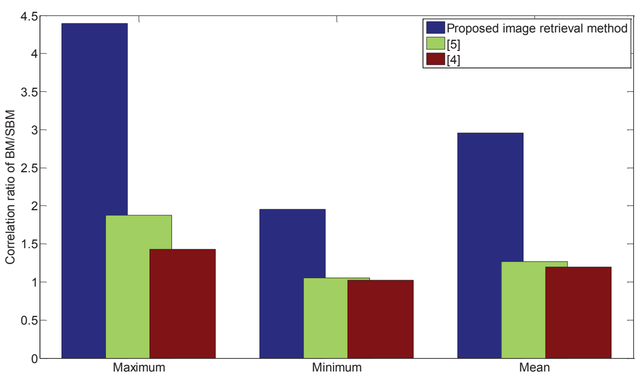


Figure 5: Comparisons of maximum, minimum and mean of the correlation ratio of BM/SBM

respectively. A threshold (m_{thres}) of 0.75 was used. It is seen that the DCT-phase representation of [5] and [4] is not suitable for accurately retrieving the non-occluded portion of the image.

From *Table 3* it can be seen, by comparing the mean of the BM/SBM correlation ratios, that the proposed image retrieval method outperformed other methods [4],[5] by an average factor of 2.46 and 2.32 respectively over different datasets. We can see that the method proposed in [4] which also used a ternary-valued DCT-phase representation, showed a lower mean than [5] which used a binary-valued DCT-phase. However, method [4] had the least standard deviation. *Figure 5*, compares the maximum, the minimum and the mean of the BM/SBM correlation ratio for the proposed image retrieval method and the methods in [4] and [5], for all datasets tests combined. We can see from the figure that the proposed method outperforms [4] and [5].

5. CONCLUSION AND FUTURE WORK

In this paper we proposed an efficient image retrieval method for occluded images based on DCT-phase. The proposed method uses a novel correlation metric for ternary-valued DCT-phase. A new region merging method is introduced, which is used to extract the non-occluded regions from the retrieved image. The proposed image retrieval method showed an average BM/SBM correlation ratio of 2.96, when tested with different datasets containing reference images, occluded images, fused images and images compressed with different JPEG compression ratios. Experimental evaluation also showed that the proposed image retrieval method performs better than the current state of the art DCT-phase based image retrieval methods by at least a factor of 2.

Future work includes performing tests with camera acquired images for handheld real time image retrieval system and improving the detection of occlusion with a combination of thresholding the peak value of the correlation metric and further analysis on the reconstructed image from region merging.

6. ACKNOWLEDGEMENT

This work was partly supported by the Swiss Federal Office for Professional Education and Technology (OPET) through the Innovation Promotion Agency (CTI) under the Grant CTI 8811.2 PFNM-NM ("PictoBar" project).

7. REFERENCES

[1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 22, pp. 1349-1380, 2000.

[2] M. Shneier and M. Abdel-Mottaleb, "Exploiting the JPEG compression scheme for image retrieval," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 18, pp. 849-853, 1996.

[3] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An efficient iris recognition algorithm using phase-based image matching," in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, 2005, pp. II-49-52.

[4] J. Bracamonte, M. Ansorge, F. Pellandini, P.-A. Farine, W.-K. Leow, M. Lew, T.-S. Chua, W.-Y. Ma, L. Chaisorn, and E. Bakker, "Efficient Compressed Domain Target Image Search and Retrieval." vol. 3568, ed: Springer Berlin / Heidelberg, 2005, pp. 594-594.

[5] I. Ito and H. Kiya, "DCT Sign-Only Correlation with Application to Image Matching and the Relationship with Phase-Only Correlation," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, 2007, pp. I-1237-I-1240.

[6] F. Arnia, I. Iizuka, M. Fujiyoshi, and H. Kiya, "Fast Method for Joint Retrieval and Identification of JPEG Coded Images Based on DCT Sign," in *Image Processing, 2007. ICIP 2007. IEEE International Conference on*, 2007, pp. II - 229-II - 232.

[7] I. Ito and H. Kiya, "A new class of image registration for guaranteeing secure data management," in *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on*, 2008, pp. 269-272.

[8] P. Ayyalasomayajula, S. Grassi, N. Deurin, P. Farine, and T. Gueguen, "Implementation of an image recognition algorithm on the DM6446 DaVinci processor," in *Education and Research Conference (EDERC), 2010 4th European*, 2010, pp. 175-179.

[9] L. J. Chipman, T. M. Orr, and L. N. Graham, "Wavelets and image fusion," in *Image Processing, 1995. Proceedings., International Conference on*, 1995, pp. 248-251 vol.3.

[10] K. R. Rao and P. Yip, *Discrete cosine transform: algorithms, advantages, applications*: Academic Press Professional, Inc., 1990.

[11] J. Z. Wang, L. Jia, and G. Wiederhold, "SIMPLIcity: semantics-sensitive integrated matching for picture libraries," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 23, pp. 947-963, 2001.