An improved image retrieval algorithm for JPEG2000 compressed images

Amani Chaker [1], Mounir Kaaniche [2], Amel Benazza-Benyahia [1]
[1] COSIM Lab., SUP'COM, Carthage University, Tunisia
[2] Université Paris-Est, LIGM and UMR-CNRS 8049, 77454 Marne-la-Vallée, France
amani.chaker@gmail.com, mounir.kaaniche@univ-mlv.fr, benazza.amel@supcom.rnu.tn

Abstract— *In this paper, we are interested in studying the impact of the compression on the performances of some wavelet-based retrieval systems. Firstly, we show that the quantization operation of JPEG2000 has a negative effect on the performance of these content-based image retrieval systems for a given feature extraction method. Moreover, in this work, we aim at designing a novel retrieval strategy in order to reduce the performance drop resulting from quantizing the query and database images at different bitrates. More precisely, instead of directly comparing these images, we propose to perform the comparison step at similar qualities by re-compressing the high quality image at the bitrate of the low quality one. Experimental results corroborate the gain achieved by the proposed strategy.*

Keywords— *Content-based image retrieval, image compression, JPEG2000 standard, retrieval performance.*

I. Introduction

With the improvements of image acquisition systems, there is a proliferation of large image databases (DB) often browsed by the public or experts in different applications. For instance, the volumes of remotely sensed data gathered by the U.S. agencies NASA, NOAA, and USGS have dramatically grown resulting in about 18 petabytes by 2010 [1]. Development of image compression techniques is an active research area which aims at reducing the memory requirements and/or transmission delays of the image database. To this end, Wavelet-Transforms (WT) based image coders have been successfully investigated since they provide very compact multiscale representations of the input images that are suitable for progressive decoding (e.g. telebrowsing applications). These are the reasons why image compression standard JPEG200 [2] employs lifting schemes (second generation of WT).

Furthermore, the DB management system should enable an easy image search from the underlying DB [3]. To this purpose, Content-Based Image Retrieval (CBIR) approach is considered as an appropriate solution thanks to an automatic indexing of the images only based on the visual content [4]. As a consequence, a challenging issue is to combine WT-based image coding with CBIR within a common framework. In other words, it is desirable that the DB management system allows also fast and accurate image retrieval from the stored wavelet coefficients. In this respect, several wavelet-based image retrieval methods were already proposed [5], [6], [7]. Most of the reported works have implicitly considered that the features are extracted from the uncompressed wavelet coefficients both for the query and the DB images. However, in practice, the query and the DB images (called hereafter model images) are quantized with compression ratios not necessarily the same. Depending on the retained features at the indexing step, the loss of information due to quantization at different compression ratios may lead to a decrease of the retrieval performance as it has been recently observed in the context of DCT-based CBIR systems [8], [9].

The objective of this paper is twofold. First, we show that JPEG2000 quantization has also an effect on the performance of WT-based CBIR systems for some feature extraction methods. Secondly, we propose a novel retrieval strategy to improve the robustness w.r.t. compression of such CBIR systems.

The remainder of this paper is organized as follows. Section II is dedicated to a brief description of the JPEG2000 standard and the most current feature extraction techniques operating in the WT domain. In Section III, we present a novel strategy for improving the retrieval performances of JPEG2000 compressed images. In Section IV, experimental results are given in order to evaluate the gain achieved by the proposed retrieval strategy. Finally, some conclusions are drawn in Section V.

II. JPEG2000-based CBIR

A. A brief review of JPEG2000

After its adoption as an ISO standard for still image coding, JPEG has presented the problem of blocking artifacts which is a prohibitive shortcoming for the development of very low bit rate applications. To alleviate such drawback, the recent image compression standard JPEG2000 has been adopted. It has the advantage of offering a broad range of features for emerging applications [2]. The JPEG2000 encoder is composed of several operations: wavelet transform, quantization, entropy coding and bit-stream organization. More precisely, this new compression scheme consists firstly in applying a discrete wavelet decomposition to the original data to be encoded. To this end, a Lifting Scheme (LS)-based implementation is employed [10]. More precisely, a prediction and update steps are performed on the input signal $s_j(m,n)$ yielding the two following sub-signals: the detail one $d_{j+1}(m,n)$ and the approximation signal $s_{j+1}(m,n)$:

$$d_{j+1}(m,n) = s_j(m,2n+1) - \mathbf{p}_j^\top \mathbf{s}_j(m,n) \qquad (1)$$

$$s_{j+1}(m,n) = s_j(m,2n) + \mathbf{u}_j^\top \mathbf{d}_{j+1}(m,n) \qquad (2)$$

where $\mathbf{p}_j$ (resp. $\mathbf{u}_j$) is the prediction (resp. update) vector, and $\mathbf{s}_j(m,n)$ (resp. $\mathbf{d}_{j+1}(m,n)$) is a reference vector containing the even image samples (resp. detail coefficients) used in the prediction (resp. update) step. Note that this 1D processing is applied along the lines then the columns (or inversely) to generate an approximation subband and three detail subbands oriented horizontally, vertically and diagonally. A multiresolution representation of the input image over J resolution levels is obtained by recursively repeating these steps to the resulting approximation coefficients, yielding $(3J+1)$ subbands. In the following, x_j will denote the j^{th} subband whose height and width dimensions are designated by M_j and N_j with $j \in \{1, \ldots, 3J+1\}$. Once this multiscale transform is performed, the coefficients of each subband x_j are quantized by using a uniform quantizer with a central deadzone. Thus, for each coefficient located at position (m, n), the output $\bar{x}_j$ of the quantizer is given by:

$$\bar{x}_j(m,n) = \text{sign}(x_j(m,n)) \left\lfloor \frac{|x_j(m,n)|}{q_j} \right\rfloor \tag{3}$$

where q_j denotes the quantization step. It is worth noting that small (resp. high) q_j values correspond to high (resp. low) bitrates r_j and, result in a high (resp. low) reconstructed subband quality.

A bit allocation among the different subbands is carried out in order to compute the quantization steps q_1, q_2, $\ldots$, q_{3J+1} (and, hence the related bitrates r_1, $\ldots$, r_{3J+1}) according to a rate-distortion criterion. Indeed, the average distortion in the wavelet domain is minimized subject to a constraint on the total available bitrate R:

$$R = \frac{\sum_{j=1}^{3J+1} M_j N_j r_j}{\sum_{j=1}^{3J+1} M_j N_j}. \tag{4}$$

This constrained minimization problem can be solved using the Lagrangian optimization approach [11]. An entropy coding of the quantized coefficients is carried out after the quantization. It employs a context modeling to cluster the bits of the quantized wavelet coefficients into groups with similar statistics to improve the efficiency of the arithmetic coder. Finally, the output of the arithmetic coder is collected into packets and a bitstream is generated according to a predefined syntax. Obviously, once this bitstream is received, it remains to decode it in order to recover the quantized coefficients and, then to apply an inverse LS to restitute the lossy version of the input image.

B. Feature extraction in the WT domain

A wavelet-based CBIR consists of computing relevant features directly from the resulting coefficients. The most popular and fast technique aims at retaining the first moments of the subband x_j as a salient feature [5], especially the energy E_j of the subband [12], [13]:

$$E_j = \frac{1}{M_j N_j} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} (x_j(m,n))^2. \tag{5}$$

The absolute mean A_j of x_j has also been considered as a measure of the coefficient sparsity [13], [14] :

$$A_j = \frac{1}{M_j N_j} \sum_{m=1}^{M_j} \sum_{n=1}^{N_j} |x_j(m,n)|. \tag{6}$$

Several texture features such as inertia, entropy and local homogeneity can also be employed [15]. These statistical features are generally computed for all high-pass subbands to form the co-occurrence signature of an image [13], [14]. Another alternative is to exploit the sparsity of the WT representation by resorting to a parametric modeling of the (monomodal) distribution of each wavelet subband x_j [5], [13]. Very often, the texture information is shaped by means of parametric models of the marginal densities of the wavelet coefficients in every subband such as the generalized Gaussian distribution [6], the gamma distribution [16] and the Gaussian mixture model [17]. The related feature vector of the image is built by gathering the features of all the subbands. At this level, it is worth noting that the so far proposed statistical models concern the uncompressed coefficients x_j considered as realizations of a continuous random variable. These models are no longer valid for quantized coefficients $\bar{x}_j$ which are samples of a discrete random variable.

Once the features of the model images are extracted, the retrieval procedure can be applied. The query image is indexed using the same kind of features that have been considered for the database images. Then, the objective is to find in the DB the candidate model images whose feature vectors are closer to that of the query one, according to a given similarity measure. The most widely used one is the Normalized Euclidean Distance (NED) [13].

III. Proposed image retrieval under compression

A. Basic strategy

In this section, we assume that the original versions of the DB and the query are not available as they have been replaced by compressed versions according to a JPEG2000 coder presented in the previous section. In other words, the related feature vectors are computed from these lossy versions the only ones to be available. A straightforward solution to design a CBIR consists of *directly* comparing the query and model images as it is generally considered in the case of uncompressed data. However, it is clear that a lossy compressed image differs from its original and therefore their corresponding features may also be different. This suggests that the quantization will affect the performance of the retrieval procedure. To illustrate this point, we propose to perform some retrieval experiments by considering different compression qualities. More precisely, we select 40 textures images of size 512×512 obtained from the VisTex database [18]. Each one is then divided into 16 non-overlapping images of size 128×128, resulting in a DB of 640 images. The ground truth is obtained as follows: the 16 sub-images generated from a single original one are assumed to be similar. At the retrieval step, all DB images are used as query ones. The sixteen images belonging to the same family of a query image are considered as the relevant images for this query. The DB images are compressed at bitrates R_{M} whose values are in the set $\{1.5, 1, 0.8, 0.5, 0.25, 0.1\}$ bpp. The retrieval performances are evaluated in terms of precision $PR = \frac{N^{\mathrm{r}}}{N}$ and recall $RC = \frac{N^{\mathrm{r}}}{N^{t}}$, where N^{r} is the number of output images considered as relevant, N^{t} represents the total number of relevant images in the database and N is the number of all returned images. We retain also the whole set of the energy $\left(E_j\right)_{j\in\{1,\ldots,3J+1\}}$ as a feature vector to compare the different images.

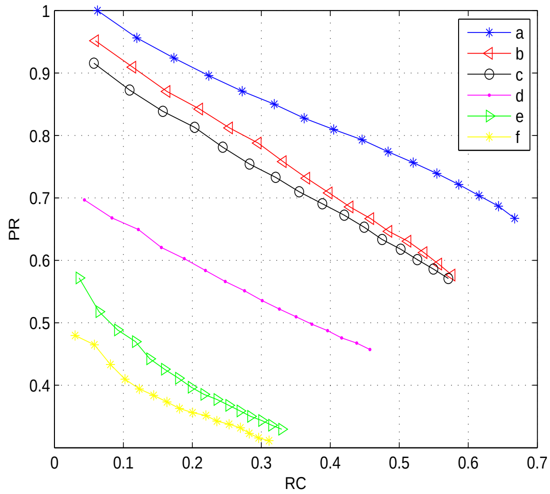


Fig. 1. Precision versus recall obtained for a query uncompressed images, when model images are : (a) uncompressed, compressed at R_{M} = (b) 1.5 bpp, (c) 1 bpp, (d) 0.8 bpp, (e) 0.5 bpp, (f) 0.25 bpp.

Fig. 1 provides the $PR - RC$ plots for different bitrates R_{M}. While the blue curve corresponds to the image retrieval algorithm performed on the uncompressed data (i.e on the original wavelet coefficients of the query and model images), the remaining ones are generated after compressing (i.e quantizing) the model images at the different bitrates R_M (the query one is kept as uncompressed). It can be noticed that a significant drop in the performances of the basic CBIR system occurs when model images are compressed at very low bitrates.

B. Image retrieval after recompression

To better understand the quantization effect on the image retrieval performance, we have considered another context of experiments in which the query and model images are both quantized. These results are illustrated in Fig. 2 where the query image is compressed at $R_{\mathrm{Q}} = 0.5$ bpp and the model images are all compressed at R_{M} belonging to the set $\{1.5, 1, 0.8, 0.5, 0.25, 0.1\}$ bpp.
It can be observed that the best improvements are achieved

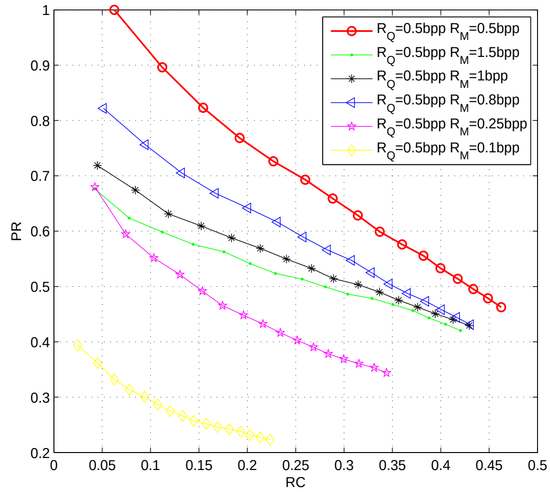


Fig. 2. Precision versus recall obtained under compression for a query compressed at 0.5 bpp.

when the query and model images are quantized at the same bitrate (i.e having similar qualities).
Based on this observation, we propose to improve the retrieval performance of a compressed dataset by forcing the query and model images to have similar qualities. While the effect of compression on image retrieval performance has been studied in the literature by considering JPEG compressed datasets [19], we present hereafter a more general framework well adapted to the JPEG2000 standard. In other words, the retrieval strategy recently proposed by Edmunson and Schaefer for DCT-based CBIR systems [19] is extended in this work to the context of JPEG2000-based CBIR ones. This strategy could be considered as a preprocessing step performed before extracting features when images with different qualities are involved. More precisely, when one image has a lowest quality than the other one, the key idea of our approach is to produce a higher quality image from the lowest one or inversely. Since the original wavelet coefficients cannot be recovered from the quantized ones and, the quantization error (after the reconstruction procedure) becomes much important at lower bitrates, we propose to perform the comparison at the low quality level. In this respect, we will consider the two following cases:

1) Case A: query image is more compressed than the model ones ($R_{\mathrm{Q}} < R_{\mathrm{M}}$);
2) Case B: model images are more compressed then the query one ($R_{\mathrm{Q}} > R_{\mathrm{M}}$).

Concerning case B, we propose to keep all the model images at the lower quality and generate a low quality version from the high quality query image. More precisely, from the finely quantized coefficients $\bar{x}_j(m,n)$, the related reconstructed wavelet coefficients $\tilde{x}_j(m,n)$ are firstly computed as follows:

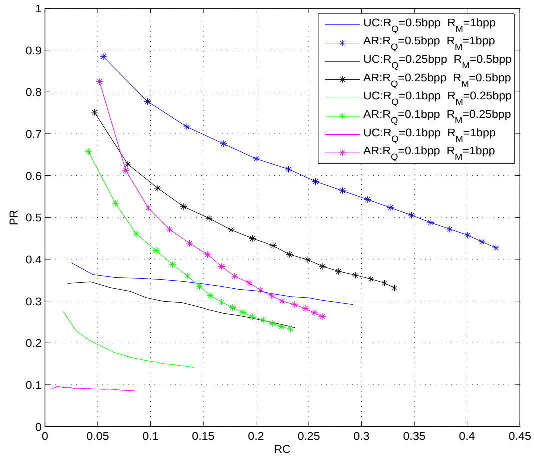
$$\tilde{x}_j(m,n) = \begin{cases} (\bar{x}_j(m,n)+\gamma)\,q_j & \text{if } \bar{x}_j(m,n) > 0 \\ (\bar{x}_j(m,n)-\gamma)\,q_j & \text{if } \bar{x}_j(m,n) < 0 \\ 0 & \text{otherwise,} \end{cases}$$

where $0 \leq \gamma < 1$ is a reconstruction parameter chosen by the decoder. Note that choosing $\gamma = 0.5$ corresponds to a mid-point reconstruction as used in many encoding strategies [20]. Then, the reconstructed wavelet coefficients are *re-quantized* (i.e *re-compressed*) at the same bitrate of the image having a lower quality. In the case A, a similar procedure is applied by reversing the quantization of the model images and re-compressing them at the bitrate of the query image.

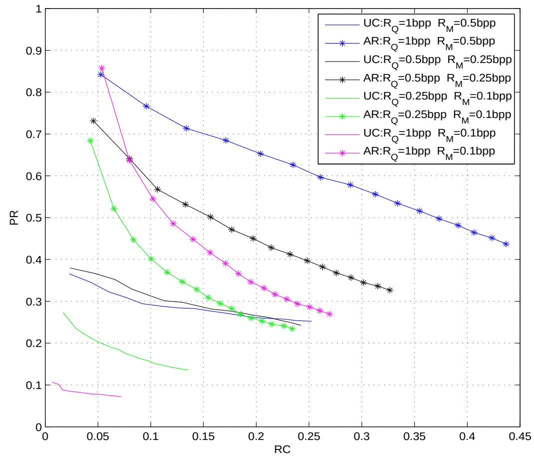
IV. Experimental results

WT-based CBIR are recommended for texture images. This is the reason why we have employed two different texture datasets: the first one is the classical selection of 40 textures from the Vision texture database previously described whereas the second one is the Outex texture collection [21], in particular the Outex_TR_00000 set. The latter dataset includes 319 different textures with 20 samples for each texture. Both datasets have been compressed by applying the JPEG2000 at different bitrates belonging to the set $\{1, 0.5, 0.25, 0.1\}$ bpp. Performances of the CBIR were evaluated in terms of precision PR and recall RC standard criterion. In order to show the benefits of the proposed compressed image retrieval algorithm, we will consider both cases A and B.

Concerning the indexing step, we have considered the two following features $\left(E_j\right)_{j\in\{1,\ldots,3J+1\}}$ and $\left(A_j\right)_{j\in\{1,\ldots,3J+1\}}$. Moreover, the NED is used as a similarity measure. Our experiments were divided into 2 rounds: in the first one, the query images belong to the DB whereas in the second round, the query images are not in the DB. The retrieval results corresponding to the first round are presented in Figures 3, 4 and 5. We first illustrate the retrieval performance when the query and the model images are compressed at different bitrates. In both cases A and B, we can note that recompression yields a significant improvement in the retrieval performance. The gain in performance becomes much more important when the difference of the bitrates $|R_{\text{M}} - R_{\text{Q}}|$ is large. For example, for case A, it can be noted from Fig. 5 that recompressing model images from $R_{\text{M}} = 1$ bpp to $R_{\text{Q}} = 0.1$ bpp results in higher gain than recompressing model images from $R_{\text{M}} = 0.25$ bpp to $R_{\text{Q}} = 0.1$ bpp. This can be explained by the quantization error which is more important at low bitrates. Indeed, since coarse (resp. fine) quantization may yield crucial (small) information loss, the reconstructed wavelet coefficients may be very different from (resp. close to) the original ones, and therefore the quality of the recompressed image will be much (less) affected and the retrieval gain after recompression becomes small (high). The second pass of experiments consists in employing query images not in the DB in order to get more



Case A



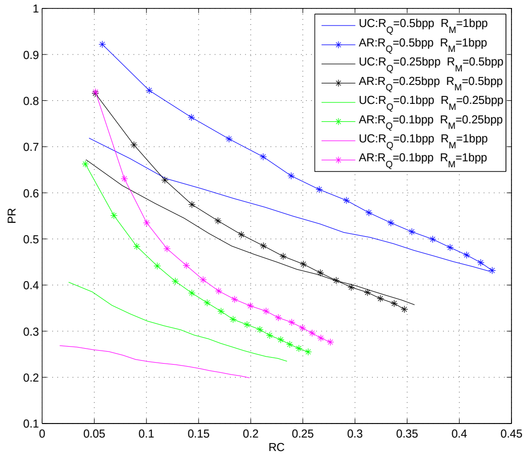
Case B

Fig. 3. Retrieval performance of JPEG2000 compressed images (taken from the VisTex database) using the absolute mean feature: under compression (UC) and after recompression (AR).

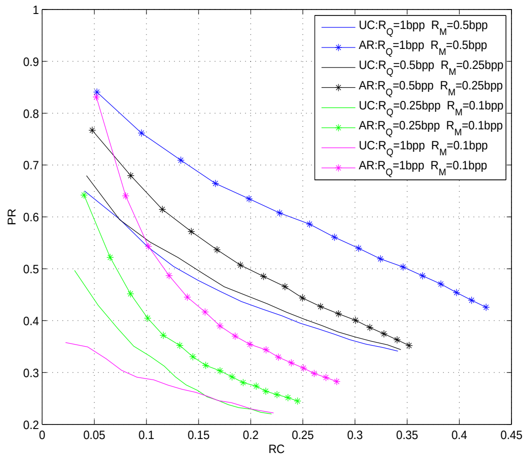
reliable performances. To this purpose, one image of each class has been excluded from the DB to be considered as a query image. Figures 6 and 7 provide retrieval results for respectively the absolute mean and energy features. The same conclusions as in the first part of experiments are still valid.

V. Conclusion

In this paper, we have investigated the influence of JPEG2000 image compression on retrieval performance. Experiments show that JPEG2000 quantization has a negative impact on retrieval performance especially when either query or model images are compressed at low bitrates. To improve the performances, we have proposed an efficient method based on the comparison of images at similar qualities by re-compressing images with higher quality at the bitrate of the

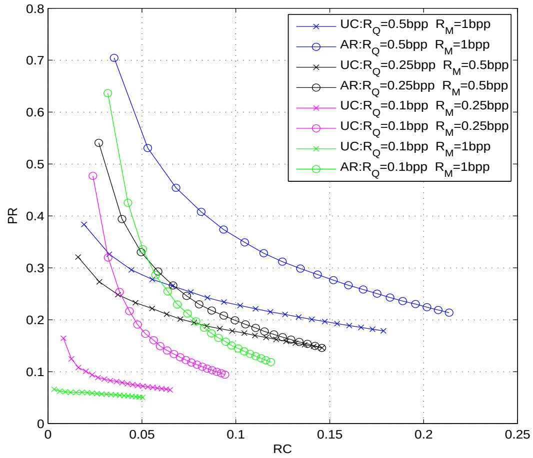


Case A

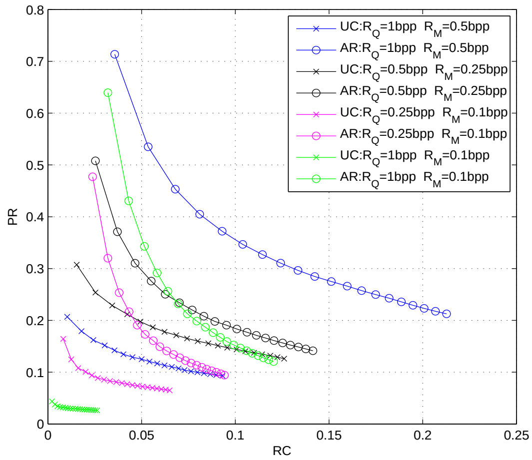


Case B

Fig. 4. Retrieval performance of JPEG2000 compressed images (taken from the VisTex database) using the energy feature: under compression (UC) and after recompression (AR)



Case A



Case B

Fig. 5. Retrieval performance of JPEG2000 compressed images (taken from the Outex_TR_00000 database) using the energy feature: under compression (UC) and after recompression (AR).

low quality image. Experiments prove the benefits drawn from our approach when two basic features are employed (energy and absolute mean of wavelet coefficients). Future work should aim at extending this approach to more sophisticated salient features in the WT transform domain.

References

[1] H. K. Ramapriyan, *Satellite imagery in Earth science applications*. Image Databases: Search and Retrieval of Digital Imagery, Edited by V. Castelli, L. D. Bergman, John Wiley & Sons, 2002.

[2] D. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Norwell, MA, USA: Kluwer Academic Publishers, 2001.

[3] A. D. Bimbo, *Visual information retrieval*. Morgan Kaufmann, 1999.

[4] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *Journal of Visual Communication and Image Representation*, vol. 12, pp. 1349–1380, 2000.

[5] M. K. Mandal, T. Aboulnasr, and S. Panchanathan, "Image indexing using moments and wavelets," *IEEE Consumer Electr.*, vol. 42, no. 3, pp. 557–565, August 1996.

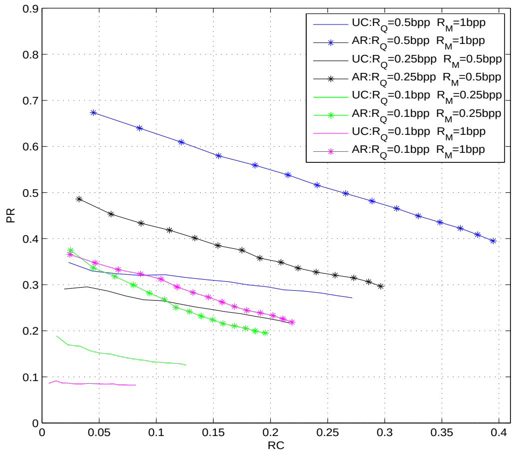
[6] M. N. Do and M. Vetterli, "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance," *IEEE Transactions on Image Processing*, vol. 11, no. 2, pp. 146–158, February 2002.

[7] S. Sakji-Nsibi and A. Benazza-Benyahia, "Indexing of multichannel images in the wavelet transform domain," in *IEEE International Conference on Communication Technologies: Theory & Practice*, Damascus, Syria, April 2008, pp. 1–6.

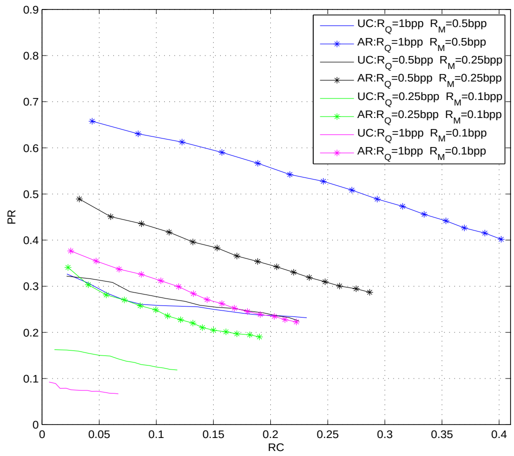
[8] E. Guldogan, K. C. O. Guldogan, S. Kiranyaz, and M. Gabbouj, "Compression effects on color and texture based multimedia indexing and retrieval," in *IEEE International Conference on Image Processing*, barcelona, Spain, September 2003, pp. 9–12.

[9] G. Schaefer, "Does compression affect image retrieval performance?" *International Journal of Imaging Systems and Technology*, vol. 18, no. 2-3, pp. 101–112, 2008.

[10] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis*,

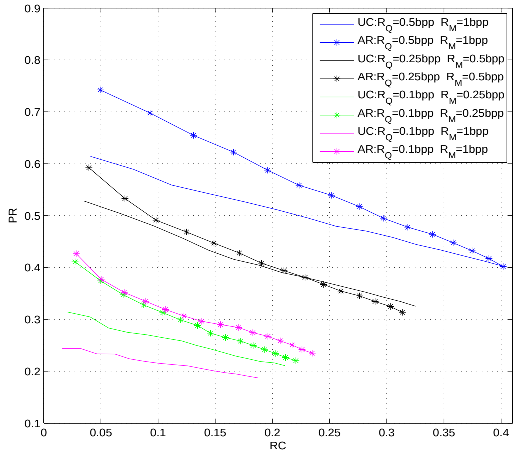


Case A

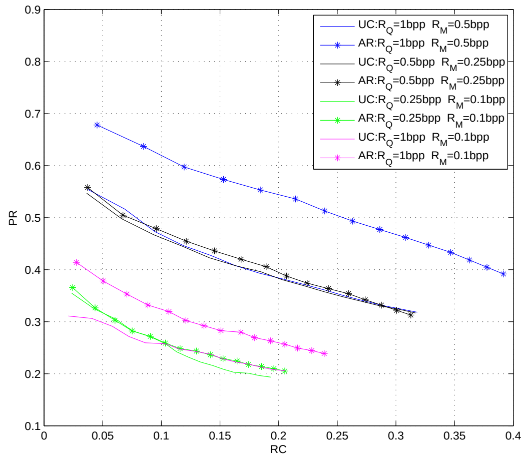


Case B

Fig. 6. Retrieval performance of JPEG2000 compressed images (taken from the VisTex database) using the absolute mean feature: under compression (UC) and after recompression (AR) when query images are not in the database.



Case A



Case B

Fig. 7. Retrieval performance of JPEG2000 compressed images (taken from the VisTex database) using the energy feature: under compression (UC) and after recompression (AR) when query images are not in the database.

vol. 3, no. 2, pp. 186–200, April 1996.

[11] Y. Shoham and A. Gersho, “Efficient bit allocation for an arbitrary set of quantizers [speech coding],” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, no. 9, pp. 1445–1453, 1988.

[12] J. R. Smith and S. F. Chang, “Transform features for texture classification and discrimination in large image databases,” in *IEEE International Conference on Image Processing*, vol. 3, Austin, TX, USA, November 1994, pp. 407– 411.

[13] G. V. Wouwer, P. Scheunders, and D. V. Dyck, “Statistical texture characterization from discrete wavelet representations,” *IEEE ransactions on Image Processing*, vol. 8, no. 4, pp. 592–598, April 1999.

[14] M. Pi, C. Tong, S. Choy, and H. Zhang, “A fast and effective model for wavelet subband histograms and its application in texture image retrieval,” *Image Processing, IEEE Transactions on*, vol. 15, no. 10, pp. 3078–3088, 2006.

[15] R. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 3, no. 6, pp. 610–621, 1973.

[16] R. J. Mathiassen, A. Skavhaug, and K. Bo, “Texture similarity measure using Kullback-Leibler divergence between Gamma distributions,” in *European Conference on Computer Vision*, Copenhagen, Denmark, May-June 2002.

[17] H. Yuan, X. P. Zhang, and L. Guan, “Content-based image retrieval using a gaussian mixture model in the wavelet domain,” in *SPIE, Visual Communications and Image Processing*, Lugano, Switzerland, July 2003.

[18] “MIT vision and modeling group. vision texture,” http://vismod.www.media.mit.edu, [Online].

[19] D. Edmundson and G. Schaefer, “Recompressing images to improve image retrieval performance,” in *International Conference on Acoustics, Speech and Signal Processing*, vol. 4, Kyoto, Japan, March 2012, pp. 1541–1544.

[20] M. Rabbani and R. R. Joshi, “An overview of the JPEG 2000 still image compression standard,” *Signal processing: Image communication*, vol. 17, no. 1, pp. 3–48, 2002.

[21] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen, and S. Huovinen, “Outex-new framework for empirical evaluation of texture analysis algorithms,” in *International Conference on Pattern Recognition*, vol. 1, 2002, pp. 701–706.