AN EFFICIENT IMAGE RETRIEVAL METHOD UNDER DITHERED-BASED QUANTIZATION SCHEME

A. Chaker[1]*, M. Kaaniche*[2]*, A. Benazza-Benyahia*[1]*, M. Antonini*[3]

[1] University of Carthage,
Higher School of Communication of Tunis,
COSIM Research Lab., Tunisia
{amani.chaker,benazza.amel}@supcom.tn

[2] Institut Galilée, L2TI,
Univ. Paris 13, Sorbonne Paris Cité, France
mounir.kaaniche@univ-paris13.fr

[3] I3S - UMR 7271 - CNRS
Univ. Nice Sophia Antipolis, France
am@i3s.unice.fr

ABSTRACT

Recent research works have shown the impact of the quantization techniques on the performance of standard image retrieval systems when datasets are compressed in a lossy mode. In this work, we propose to design an efficient retrieval method well adapted to wavelet-based compressed images. Our objective is to recover features of the original image (herein the moments of the unquantized subbands) directly from the quantized coefficients. To this end, we propose to apply a dithered quantization technique satisfying some specific conditions. Then, the estimated moments of the wavelet subbands are used in an appropriate way to construct the feature vectors of the database images. Experimental results show the interest of the proposed image retrieval method compared to the state-of-the-art ones.

***Index Terms*—** Content based image retrieval, compressed images, dithering quantization, feature extraction.

1. INTRODUCTION

Content-Based Image Retrieval (CBIR) systems allow an effective access to target images among huge databases by describing appropriately the visual cues [1]. At this level, it is worth to note that very often, images undergo a compression due to limitations of storage and bandwidth resources. To this end, it is used to resort to the popular image coding standards JPEG [2] and JPEG2000 [3] respectively based on Discrete Cosine Transform (DCT) and Wavelet Transform (WT). In this case, the most reported WT-based CBIR methods consist in directly extracting from the original wavelet coefficients of the query and database (DB) images the salient descriptors [4–7]. While most of these approaches are well adapted for losslessly encoded images (i.e *unquantized* transformed coefficients), a particular attention should be paid to the effect of the *quantization* in the context of lossy data compression. Indeed, some studies have shown that lossy compression adversely affects the image retrieval performance of several popular CBIR techniques especially when the query and database images (called also model images) are quantized at very different bitrates [8–11]. Therefore, designing more efficient indexing methods that account for the quantization step becomes a real challenge. However, to the best of our knowledge, only few works have addressed this problem [11–13]. More precisely, we have recently proposed to apply a recompression strategy before proceeding to the feature extraction step in order to impose on both query and DB images to have similar qualities [11]. To this end, the wavelet coefficients of the higher quality image are firstly reconstructed, and then requantized at the bitrate of the lower quality image. It should be noted that this method has been inspired from a similar recompression approach developed previously in the context of DCT-based CBIR systems [12, 14]. In addition, we have recently proposed a retrieval method that consists in estimating the statistical parameters of the original wavelet coefficients directly from the quantized ones according to the maximum likelihood criterion [15]. The Laplacian distribution was retained to model the wavelet coefficients distribution. Then, the estimated statistical parameters of the different subbands are selected as a feature.

The main objective of this paper is to improve the retrieval performances of the aforementioned approaches by designing salient features from the quantized subbands which are robust to compression effects. To this end, we propose to resort to a dithered quantization scheme that allows to reconstruct the second order moments of the original subbands from their quantized counterparts. Then, the estimated moments are used to build the descriptors for the indexing step.

The remainder of this paper is organized as follows. In Sec. 2, we first review the wavelet based CBIR systems. Then, in Sec. 3, we describe the proposed retrieval method. Finally, experimental results are given in Sec. 4 and conclusions are drawn in Sec. 5.

2. WAVELET-BASED CBIR SYSTEM

2.1. Conventional WT-based compression

WT has been attracting much attention in the most recent image compression algorithms. For instance, the recent JPEG2000 coding standard operates in the WT domain and

A. Chaker has benefited from the EU Erasmus Mundus Alyssa programme a mobility grant to conduct this work at I3S Lab.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

outperforms the JPEG standard especially at very low bitrate. Furthermore, it enables a scalable decoding [16]. A wavelet-based coding system is composed of three modules: wavelet transform, quantization and entropy coding. More precisely, a discrete WT decomposition can efficiently be applied according to a Lifting Scheme (LS) [17]. For an image, a separable 1D LS is often performed over the lines and columns. This procedure, repeated over J stages, results in an approximation image and $3J$ detail subbands oriented horizontally, vertically and diagonally.

In the following, X_j will denote the j^{th} subband with $j = \{1, \ldots, 3J+1\}$. Once the WT is performed, the wavelet coefficients are generally quantized by a uniform scalar quantizer with quantization step q_j and shift parameter $a_j \in [-\frac{1}{2}, \frac{1}{2})$. Thus, for the input X_j, the quantized output $\overline{X}_j$ is given by:

$$\begin{array}{rl} \overline{X}_j = & Q[X_j] = q_j(a_j + n + \frac{1}{2}) \\ & \text{if } q_j(a_j + n) \leq X_j \leq q_j(a_j + n + 1) \end{array} \tag{1}$$

where $n \in \mathbb{Z}$ is the set of integers and, q_j denotes the quantization step chosen at the j^{th} subband. The quantization steps $q_1, q_2, \ldots, q_{3J+1}$ are generally adjusted thanks to a rate-distortion allocation algorithm [18].

2.2. Feature extraction

In a WT-based CBIR system, the relevant features are extracted from the wavelet coefficients. The most popular and simplest one describes each subband by its moments especially the energy [4, 19]. Other systems resort to a statistical parametric modeling of the coefficient distribution, the parameters being considered as relevant signatures. Indeed, many univariate models such as the Generalized Gaussian (GG) distribution [5] and the generalized Gamma distribution [20] have been used to successfully capture the sparsity of the wavelet coefficients. Despite their effectiveness, statistical models are dedicated to the unquantized coefficients X_j which are considered as realizations of a continuous random variable. These models are no longer valid for quantized coefficients $\overline{X_j}$ which are samples of a discrete random variable.

3. PROPOSED IMAGE RETRIEVAL METHOD UNDER DITHERED QUANTIZATION SCHEME

3.1. Motivation

Motivated by the fact that popular WT-based indexing methods retain the first p^{th}-order moments μ_j^p of all detail subbands as salient features [4,19], it would be interesting to have a quantization scheme that allows to recover the p^{th}-order moments of the original wavelet subbands μ_j^p from those of the quantized ones $\bar{\mu}_j^p$ for $j = 1, \ldots, 3J$. To this end, based on a previous work developed in [21] where the relationship between μ_j^p and $\bar{\mu}_j^p$ is derived, we propose to resort to a dithering quantization technique in order to design efficient and robust retrieval method. It is worth noting that dithering is a multi-purpose method which has been mainly used to reduce the quantization artifacts. Indeed, it has been originally introduced in speech and video processing to reduce the perceptual distortion due to compression [22]. Moreover, it has been employed in sensor network application by operating at the sensor node before data transmission [23]. In what follows, we present the concept of the dithered quantization and then propose a retrieval approach adapted to such quantized images.

3.2. Dithered quantization

For a general uniform scalar quantization with an input X_j, it was shown in [21] that, for a given $p \in \mathbb{N}^*$, $\bar{\mu}_j^p$ is the sum of two terms A_p and B_p such as A_p is a linear combination of the moments μ_j^r with $r = 0, \ldots, p$ and, B_p is a bias term involving the characteristic function φ_{X_j} of X_j:

$$A_p = \frac{1}{p+1} \sum_{r=0}^{p} C_r^{p+1} \left(\frac{q_j}{2}\right)^{p-r} \mu_j^r [p \oplus r \oplus 1] \tag{2}$$

where $\oplus$ stands for modulo-2 operation and,

$$\begin{array}{rl} B_p = & \sum\limits_{n \neq 0} e^{-2i\pi n a_j} \sum\limits_{r=0}^{p-1} \left(\frac{q_j}{2}\right)^{p-r} i^{-r} \\ & \times \left(\sum\limits_{\lambda=0}^{(p-1)-r} \frac{p! i^{\lambda-1} (p \oplus r \oplus \lambda)}{r!(p-r-\lambda)!(n\pi)^{\lambda+1}} \varphi_{X_j}^{(r)} \left(\frac{2n\pi}{q_j}\right) \right) \end{array} \tag{3}$$

where $i^2 = -1$. It is obvious that if the characteristic function φ_{X_j} of the input X_j fullfills the following conditions:

$$\forall n \in \mathbb{Z}^*, \forall r = 0, \ldots, p-1, \quad \varphi_{X_j}^{(r)} \left(\frac{2n\pi}{q_j}\right) = 0, \tag{4}$$

then, the bias term B_p vanishes and therefore, the computation of the $\bar{\mu}_j^p$ becomes a linear combination of the original moments μ_j^r with $r = 0, \ldots, p$. This can be guaranteed by modifying the input of the uniform quantizer and adding a random variable R_j, called dither, to the original subband X_j before the quantization step. Such technique is known as dithered quantization. Thus, the quantized output $\widetilde{X}_j$ becomes:

$$\widetilde{X}_j = Q[X_j + R_j]. \tag{5}$$

Note that the dither R_j is independent of the input X_j and could be judiciously selected in order to cancel the B_p terms. Indeed, recall that the characteristic function $\varphi_{X_j+R_j}$ of the sum of the independent variables X_j and R_j is given by:

$$\varphi_{X_j+R_j} = \varphi_{X_j} \times \varphi_{R_j}. \tag{6}$$

Hence, it is enough to choose a dither variable whose characteristic function φ_{R_j} satisfies condition (4) in order to guarantee that $\varphi_{X_j+R_j}$ meets also this condition. A dithering characteristic function that satisfies this condition up to moment p is given by:

$$\forall t \in \mathbb{R}, \quad \varphi_{R_j}(t) = \text{sinc}^p(\frac{q_j t}{2}). \tag{7}$$

More precisely, in the case of $p = 1$, it is easy to show that a simple uniformly distributed dither over the range $[-q_j/2, q_j/2]$ leads automatically to the cancellation of B_1. Moreover, if $p = 2$, it is possible to select a dither as a sum of two independent variables uniformly distributed on $[-q_j/2, q_j/2]$ in order to cancel both B_1 and B_2. Such dither is the triangular distribution law on $[-q_j, q_j]$.

3.3. Moment reconstruction

By cancelling the B_p terms, the p^{th}-order moments of the dithered quantized coefficients $\widetilde{X}_j$, which will be denoted by $\tilde{\mu}_j^p$, becomes a linear combination of $\mu_j^0, \ldots, \mu_j^p$:

$$\tilde{\mu}_j^p = \sum_{r=0}^{p} c_r \mu_j^r \tag{8}$$

where

$$c_r \triangleq \sum_{t=0}^{p-r} \frac{p!}{(p-r-t+1)r!t!} \left(\frac{q}{2}\right)^{p-r-t} \mathsf{E}[R_j^t][p \oplus r \oplus t \oplus 1] \tag{9}$$

where $\mathsf{E}[.]$ denotes the mathematical expectation. For instance, the first and second order moments μ_j^1 and μ_j^2 of X_j from those of the dithered quantized subband $\widetilde{X}_j$ can be recovered as follows:

$$\mu_j^1 = \tilde{\mu}_j^1 - M_j^1, \tag{10}$$

$$\mu_j^2 = \tilde{\mu}_j^2 - 2M_j^1\tilde{\mu}_j^1 + 2(M_j^1)^2 - M_j^2 - \frac{q_j^2}{12} \tag{11}$$

where M_j^p denotes the p^{th}-order moment of the dither R_j.

3.4. Appropriate feature extraction strategy

In this work, we propose to select the second order moment (i.e $p = 2$) of the wavelet subbands as a feature vector for each DB image. However, based on many experiments on the Vistex texture images [24], we have noticed that the reconstruction of the second order moments may be inexact when wavelet subbands are quantized at low bitrates in the higher frequency subbands. As a result, at low bitrate (resp. middle and high bitrates), we propose to build the feature vector by taking the estimate of the second order moment of only the low frequency subbands (resp. of all the subbands). Moreover, it is worth noting that the resulting feature vectors of the query and DB images may have different sizes when we confront two images with very different qualities, especially when query image is compressed at low bitrate and model images are compressed at high bitrate, or inversely. For this reason, we propose to adjust the descriptor vector dimension of images compressed at high bitrate to the size of that obtained for images compressed at low bitrate by omitting the reconstructed second order moment of the high frequency subbands.

4. EXPERIMENTAL RESULTS

Experiments are performed on two well known texture datasets: Vistex and Stex [24, 25]. The first one is the popular MIT Vision Texture database, which consists of 40 textures widely used for texture image retrieval purpose. Each image of size 512×512 is divided into 16 non-overlapping images, and thus, a database of 640 images of size 128×128 is obtained. The second one is the Salzburg Textures (Stex) database which is a large collection of 476 images of different textures. The images, of size 512×512, are divided into 16 non-overlapping subimages of size 128×128, which results in a database composed of 7,616 images. We assume that the 16 sub-images, generated from a single original one, are similar and considered as relevant images for each query belonging to these sixteen images. Note that all DB images are used as query ones. Moreover, in order to evaluate the proposed retrieval scheme in the compressed domain, the 9/7 lifting scheme is applied to all the database images over 3 resolution levels. Then, we consider the standard uniform quantization scheme (UQ) as well as the dithered one (DQ). The retrieval performances are evaluated in terms of precision $PR = \frac{R^{\text{r}}}{R}$ and recall $RC = \frac{R^{\text{r}}}{R^{\text{t}}}$, where R^{t} is the total number of relevant images in the database, R^{r} is the number of returned images that are relevant and R represents the number of returned images.

The first round of our experiments has been performed on the Vistex database. To this end, we will first compare the proposed DQ-based retrieval approach to the state-of-the-art method developed in [11, 12]. Let us recall that this method is appropriate for UQ-based quantized images and, involves a preprocessing step that aims at requantizing the high quality image to the low one before features extraction when the query and model images have different qualities. This method will be denoted by UQ-AR, since features are extracted after the requantization stage. Fig. 1 provides the plots of the precision versus recall when images are quantized at different bitrates R_Q and R_M, for both the query and model images, respectively. Note that these bitrates, used to encode the query and model images, are chosen in a way that allow us to obtain different image qualities that range from low to high. To this end, R_Q and R_M are set to $\{1.5, 1, 0.75, 0.5, 0.25, 0.1\}$ bpp. The retrieval performance of losslessly encoded images is also presented. For each image, the second order moment of wavelet subbands are taken as salient features. Then, the resulting feature vectors of query and DB images are compared using the normalized Euclidean distance (NED). It can be seen from Fig. 1 that the DQ-based retrieval approach leads to a significant improvement compared to the UQ-AR retrieval method.

Now, we aim to compare our proposed approach to the recent approach based on estimating the distribution parameter directly from quantized wavelet coefficients using maximum likelihood criterion [13]. This method will be denoted in what follows by DPE. Recall that the DPE approach assumes that the wavelet coefficients are modeled by a Laplacian law and, aims at estimating the original distribution parameter $\tilde{\lambda}_j$ from the quantized subbands. Then, the estimated parameters are chosen as a salient features. The Kullback-Leibler divergence (KLD) is used as a similarity measure. For fair comparison, the same features and similarity criterion will be retained during the indexing step. To this end, we will consider two experiments. In the first one, the moments of wavelet subbands will be used as salient features. The NED will be employed as a similarity measure. Thus, for the case of the DPE approach, the second order moments of wavelet subbands $\tilde{\mu}_j^2$ are computed from the estimated parameter of the Laplacian distribution $\tilde{\lambda}_j$ as follows:

$$\tilde{\mu}_j^2 = \Gamma(3)\tilde{\lambda}_j^2. \tag{12}$$

For the second experiment, we keep the DPE approach unchanged and adapt our proposed DQ-based approach by computing the distribution parameter of the Laplacian distribution $\tilde{\lambda}_j$ from the estimated second order moment by using also Eq. (12). Then, the KLD between the estimated parameters of the query and model images is computed. Thus, as shown in Figures 2 and 3, our proposed DQ-based retrieval approach outperforms the statistical based approach (DPE) for the different considered features and similarity measures. Finally, while previous results are obtained with the Vistex database, we should note that similar behavior has been also observed with the Stex database as shown in Fig. 4. All these results confirm the efficiency of the proposed method.

5. CONCLUSION

In this paper, we have presented a new image retrieval approach to deal with wavelet-based compressed images. More precisely, we have proposed to reconstruct the moments of the original wavelet subbands from those of the quantized ones. It was shown that dithered quantization can considerably simplify the estimation of the moments and therefore, allows us to design salient features robust to the compression effect. Experimental results have shown the benefits which can be drawn from the proposed approach compared to the state-of-the-art ones.

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," *Journal of Visual Communication and Image Representation*, vol. 12, pp. 1349–1380, 2000.

[2] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30–44, 1991.

[3] M. Rabbani and R. Joshi, "An overview of the JPEG2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, no. 1, pp. 3–48, 2002.

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

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

[6] S. Sakji-Nsibi and A. Benazza-Benyahia, "Copula-based statistical models for multicomponent image retrieval in the wavelet tranform domain," in *IEEE International Conference onImage Processing*, Cairo, Egypt, November 2009, pp. 253–256.

[7] C. Li, G. Duan, and F. Zhong, "Rotation invariant texture retrieval considering the scale dependence of gabor wavelet," *IEEE Transactions on Image Processing*, vol. 24, no. 8, pp. 2344–2354, 2015.

[8] E. Guldogan, O. Guldogan, S. Kiranyaz, K. Caglar, 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] D. Edmundson and G. Schaefer, "Fast JPEG image retrieval using optimised huffman tables," in *International Conference on Pattern Recognition (ICPR)*, 2012, pp. 3188–3191.

[11] A. Chaker, M. Kaaniche, and A. Benazza-Benyahia, "An improved image retrieval algorithm for JPEG 2000 compressed images," in *IEEE International Symposium on Signal Processing and Information Technology*, Ho Chi Minh City, Vietnam, December 2012, pp. 1–6.

[12] D. Edmundson and G. Schaefer, "Recompressing images to improve image retrieval performance," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Kyoto, Japan, March 2012, p. 4 pages.

[13] A. Chaker, M. Kaaniche, and A. Benazza-Benyahia, "An efficient retrieval strategy for wavelet-based quantized images," in *IEEE International Conference on Acoustics Speech and Signal Processing*, Vancouver, BC, Canada, May 2013, pp. 1493–1497.

[14] D. Edmundson, G. Schaefer, and M. E. Celebi, "Robust texture retrieval of compressed images," in *IEEE International Conference on Image Processing*, Orlando, USA, October 2012, pp. 2421–2424.

[15] A. Chaker, M. Kaaniche, A. Benazza-Benyahia, and M. Antonini, "An efficient statistical-based retrieval approach for JPEG 2000 compressed images," in *European Signal Processing Conference*, Nice, France, September 2015, pp. 1870–1874.

[16] H. Wendt, S. G. Roux, S. Jaffard, and P. Abry, "Wavelet leaders and bootstrap for multifractal analysis of images," *Signal Processing*, vol. 89, no. 6, pp. 1100–1114, 2009.

[17] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and computational harmonic analysis*, vol. 3, no. 2, pp. 186–200, 1996.

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

[19] 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*, Austin, TX, USA, November 1994, vol. 3, pp. 407– 411.

[20] S. K. Choy and C. S. Tong, "Statistical wavelet subband characterization based on generalized Gamma density and its application in texture retrieval," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 281–289, 2010.

[21] L. Cheded and P. A. Payne, "The exact impact of amplitude quantization on multi-dimensional, high-order moments estimation," *Signal Processing*, vol. 39, pp. 293–315, Sept. 1994.

[22] R. A. Wannamaker, "Psychoacoustically optimal noise shaping," *Journal of the Audio Engineering Society*, vol. 40, no. 7-8, pp. 611–620, 1992.

[23] F. Gustafsson and R. Karlsson, "Statistical results for system identification based on quantized observations," *Automatica*, vol. 45, no. 12, pp. 2794–2801, 2009.

[24] "MIT vision and modeling group. vision texture," `http://vismod.www.media.mit.edu`.

[25] R. Kwitt and P. Meerwald, "Salzburg texture image database," `http://www.wavelab.at/sources`.

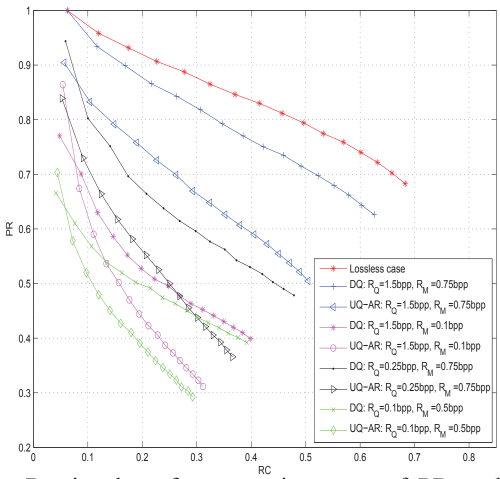

Fig. 1. Retrieval performance in terms of PR and RC of wavelet-based compressed images of the UQ-AR and the proposed DQ-based retrieval approaches.

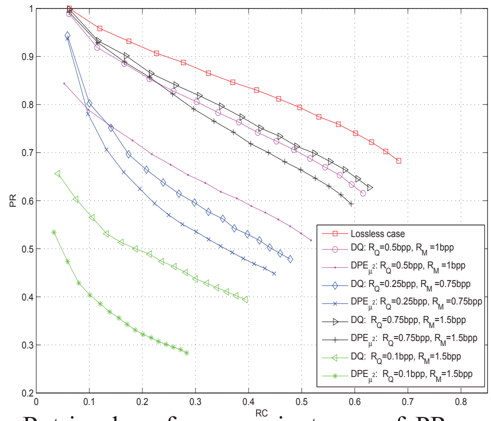

Fig. 2. Retrieval performance in terms of PR and RC of wavelet-based compressed images by using the proposed DQ-based approach and the modified version of the DPE one.

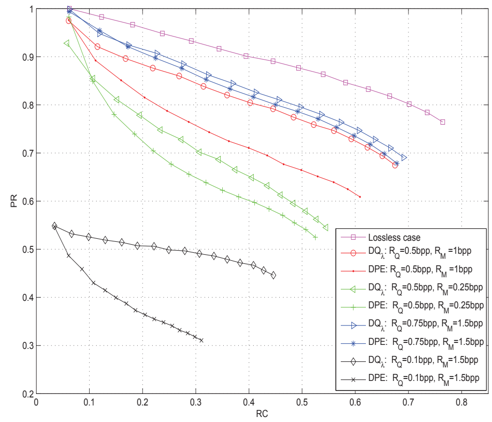

Fig. 3. Retrieval performance in terms of PR and RC of wavelet-based compressed images by using the DQ-based approach and the DPE one.

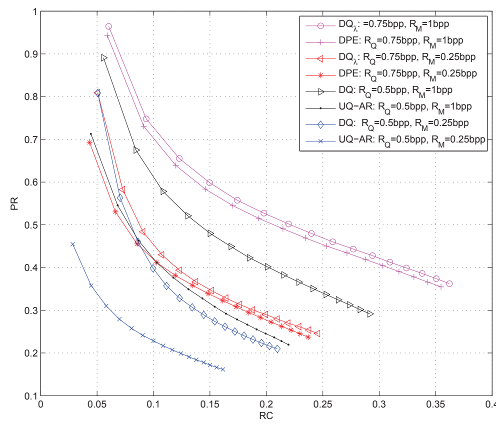

Fig. 4. Retrieval performance of the different approaches in terms of PR and RC for the Stex database.