AN EFFICIENT STATISTICAL-BASED RETRIEVAL APPROACH FOR JPEG2000 COMPRESSED IMAGES

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

[1] COSIM Lab.,
Sup'Com, Univ. Carthage, 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

This paper deals with the problem of image retrieval when the database is presented in a compressed form, by using typically the JPEG2000 encoding scheme based on wavelet transform followed by an uniform scalar quantization. The state-of-the-art method aims at applying a preprocessing step before the feature extraction to reduce the difference in the compression qualities between the images. Our contribution consists in extracting robust features directly from the quantized coefficients. More precisely, assuming that the unquantized coefficients within a subband have a Laplacian distribution, we propose to estimate the distribution parameter from the quantized coefficients. Then, the estimated parameters of the whole subbands are used to build a salient feature for the indexing process. Experimental results show that the proposed retrieval approach significantly improves the state-of-the-art one.

Index Terms— Content based image retrieval, compressed images, wavelet domain, feature extraction, statistical model, retrieval performance.

1. INTRODUCTION

With the recent developments in data acquisition and display technologies, very huge image databases (DB) are continuously produced. As a result, the use of multimedia indexing systems is mandatory to enable fast access to the different images. Content Based Image Retrieval (CBIR) systems were found to be an efficient tool for managing large image databases based only on their visual contents [1]. In this context, the features are generally extracted from the spatial domain by using the color and texture information. However, to ensure fast transmission and compact storage of the involved data, images are often compressed. To this end, different compression methods have been developed such as JPEG and JPEG2000 standards respectively based on Discrete Cosine Transform (DCT) and Wavelet Transform (WT) [2]. Intuitively, the retrieval task of such compressed images can be simply performed by decoding them in order to recover their spatial versions to compute the classical features. To avoid such eventual decompression step, several retrieval techniques have been proposed so that the attributes are directly defined in the transform domains. In this respect, most of wavelet-based indexing techniques consist in computing the features from the original (i.e unquantized) wavelet coefficients of the query and DB images [3–5]. While such methods are appropriate for losslessly encoded images, their performance decrease in the case of lossy compression due to the loss of information resulting from the quantization step [6, 7]. The performance drop is more important when the query and DB images are quantized at very different bitrates. Thus, designing efficient indexing methods that account for the quantizer becomes a real challenge. However, only few works have addressed this issue [6, 8–10]. More precisely, to improve the retrieval performance, we have recently proposed to apply a recompression strategy before the indexing step [9] in order to impose on both query and DB images similar qualities. To this end, the wavelet coefficients of the higher quality image are firstly reconstructed, and then re-quantized at the bitrate of the lower quality image. A similar recompression approach has been considered in the context of DCT-based CBIR system [8, 11]. In addition, a retrieval method based on the moment preserving quantizer was presented in [10]. The idea consists in preserving some moments of the input signal and using them as salient features. While the latter work uses a specific quantization scheme, we propose in this paper to focus on the standard uniform scalar quantization. The objective is to alleviate the aforementioned drawbacks of the recompression approach by recovering the statistical parameters of the *original* wavelet coefficients directly from the *quantized* ones. To this end, we first consider the Laplacian distribution to model the distribution of the wavelet coefficient. Then, the estimated statistical parameter of this distribution will be used as relevant features in the retrieval procedure.

The remainder of this paper is organized as follows. In Sec. 2, we first review the typical wavelet-based coding schemes and the most popular feature extraction methods operating in the

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

WT domain. Then, we present the state-of-the art recompression approach for the retrieval of compressed images. 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. RETRIEVAL OF COMPRESSED IMAGES: STATE-OF-THE-ART

2.1. Typical transform and quantizer based coding scheme

Due to the limitations of DCT at very low bitrates, we focus in this paper on the most commonly used WT based coding scheme retained in the JPEG2000 standard. More precisely, the wavelet coefficients are efficiently computed according to a Lifting Scheme (LS) [12]. A one dimensional LS consists of 3 modules: split, predict and update. The original 1D signal $a_j(k)$ is first divided into two polyphase components: its even $a_j(2k)$ and odd samples $a_j(2k+1)$. Then, each sample of one of the two resulting subsets (say the odd ones) is predicted from the neighboring even samples to generate the following detail coefficients $d_{j+1}(k)$:

$$d_{j+1}(k) = a_j(2k+1) - \mathbf{p}_j^\top \mathbf{a}_j(k) \tag{1}$$

where $\mathbf{p}_j$ is the vector of the prediction weights and $\mathbf{a}_j(k)$ is a reference vector containing some even samples used in the prediction step. Finally, the update step is applied to produce the approximation signal $a_{j+1}(k)$:

$$a_{j+1}(k) = a_j(2k) + \mathbf{u}_j^\top \mathbf{d}_{j+1}(k) \tag{2}$$

where $\mathbf{u}_j$ denotes the vector of update weights and $\mathbf{d}_{j+1}(k)$ is a reference vector containing the detail coefficients involved in the update step. In the case of an image, the 1D LS is separately applied along the rows then the columns (or inversely). Hence, an approximation subband and 3 detail ones oriented horizontally, vertically and diagonally are obtained. This decomposition is again repeated on the approximation sub-images over J resolution levels leading to $(3J+1)$ wavelet subbands. 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 a central deadzone. Thus, each coefficient $x_j(k,l)$ at position (k,l) is quantized to $\bar{x}_j(k,l)$ as follows:

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

where q_j denotes the quantization step retained in 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 [13, 14]. Finally, the quantized coefficients are entropy encoded.

2.2. Feature extraction in wavelet domain

The wavelet-based CBIR approaches aim at defining features from the wavelet coefficients. A conventional approach is based on the energy of the subband [15,16]. Another interesting alternative consists in considering a parametric model of the distribution of the WT coefficients [3, 17]. In this respect, the generalized Gaussian [4], the Gaussian mixture [18], and the generalized Gamma [19] distributions were found to be appropriate models of the coefficient distribution in a given subband. Powerful feature extraction approaches have also been developed based on multivariate statistical model to express the joint distribution of wavelet coefficients [5, 20]. Then, the parameters of the selected model for the $3J$ wavelet subbands constitute the feature vector that characterizes each image.
At the retrieval step, the feature vector of the query is finally compared with those of the DB images (called also model images) according to a given similarity measure.

2.3. Image retrieval after recompression

As aforementioned, the state-of-the-art retrieval approach appropriate for compressed data is based on the recompression strategy. It has been firstly applied in DCT domain in [8] and, then extended to the context of WT domain in our previous work [9]. Since the current work is devoted to wavelet-based quantized images, we describe in what follows the approach developed in [9]. Instead of comparing directly the query and DB images which may have different qualities, the key idea consists in obtaining two images with similar qualities before extracting the features. Based on the facts that the original wavelet coefficients cannot be perfectly reconstructed after quantization and the resulting quantization error increases at lower bitrates, the low quality image is kept unchanged whereas the high quality image is transformed to a low quality version. More precisely, for a quantization step q_j, the reconstructed wavelet coefficients $\tilde{x}_j(k,l)$ are firstly computed from the finely quantized ones $\bar{x}_j(k,l)$:

$$\tilde{x}_j(k,l) = \begin{cases} (\bar{x}_j(k,l) + \gamma)\, q_j & \text{if } \bar{x}_j(k,l) > 0 \\ (\bar{x}_j(k,l) - \gamma)\, q_j & \text{if } \bar{x}_j(k,l) < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

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 [21]. Then, the reconstructed wavelet coefficients $\tilde{x}_j(k,l)$ are re-quantized with the quantization steps of the image having lower quality. After that, the feature vectors are built by taking the energy of all the quantized (resp. re-quantized) wavelet subbands of the lower (resp. higher) quality image. Finally, features of query and model images are compared using the Normalized Euclidean Distance (NED) [16].

3. PROPOSED IMAGE RETRIEVAL METHOD

3.1. Motivation

It is worth noting that the previous recompression approach for the retrieval of compressed images presents some limitations. First, when the images are quantized at very low bitrates, the related quantization error increases and thus, the recompression technique will negatively affect the feature relevance. Moreover, it implies an additional computational cost due to the reconstruction and requantization operations performed on all the database images. To overcome these drawbacks, it would be interesting to extract salient features *directly* from the quantized wavelet coefficients. To this end, by adopting the Laplacian model for the distribution of the wavelet coefficients, we propose to estimate the original distribution parameter from the quantized coefficients, and use it to compare the images during the indexing step. In what follows, the parameter estimation method from the quantized coefficients as well as the retrieval strategy will be described.

3.2. Estimation of the original parameter distribution

Although it is known that the Generalized Gaussian (GG) law gives the most accurate model to approximate the distribution of the wavelet coefficients [4], we consider here the Laplacian distribution which has also been used for subband image coding [22, 23]. Indeed, the Laplacian distribution allows an optimal trade-off between the simplicity and tractability of the model and modeling accuracy. The related probability density function f_j is given by:

$$\forall \xi \in \mathbb{R}, \quad f_j(\xi) = \frac{\lambda_j}{2} \exp^{-\lambda_j |\xi|}. \tag{5}$$

where $\lambda_j > 0$ denotes the distribution parameter. It can be estimated as the average of the original coefficients according to the Maximum Likelihood (ML) criterion. However, since we assume in this work that the original versions of the query and DB images are not available as they are compressed, we propose to estimate λ_j from the quantized coefficients $\bar{x}_j$ by considering again the ML criterion. It has been found that the estimated parameter $\hat{\lambda}_j$ is given by [23]:

$$\hat{\lambda}_j = -\frac{2}{q_j} \log(t_j) \tag{6}$$

where

$$t_j = \frac{-N_j^0 q_j + \sqrt{(N_j^0 q_j)^2 - 4(N_j q_j + 2S_j)(N_j^1 q_j - 2S_j)}}{2N_j q_j + 4S_j} \tag{7}$$

with $S_j = \sum_{\bar{x}_j(k,l)\neq 0} \mid \bar{x}_j(k,l) \mid$ and, N_j^0, N_j^1 are respectively the number of wavelet coefficients quantized to zero and non-zero values and, $N_j = N_j^0 + N_j^1$.

It should be noted that, at very low bitrate where the wavelet coefficients are coarsely quantized to zero, they could not carry any useful information: the value of t_j will tend to zero, and so $\hat{\lambda}_j$ will tend to infinity. This problem has also already been reported in [24].

3.3. Proposed retrieval strategy

Due to the estimation problem of the distribution parameter at very low bitrate, the feature vector should be appropriately defined according to the bitrate of the different wavelet subbands. More precisely, before building this vector, it is necessary to check if $\hat{\lambda}_j$ is a finite number. Based on many experiments, it can be observed that the coarse-scale (resp. fine-scale) subbands represent large (resp. small) percentage of the total bitrate. Indeed, the cases where $\hat{\lambda}_j$ is equal to infinity occur especially at low bitrates in the higher frequency subbands, whereas at middle and high bitrates, $\hat{\lambda}_j$ is often a finite number. Thus, the feature vector is built by taking the $\hat{\lambda}_j$ of the different wavelet subbands while omitting those whose values are equal to infinity. As a result, the resulting feature vectors of the query and DB images may have different sizes when they are compressed at different bitrates. For this reason, during the indexing step, 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 estimated parameter of the high frequency subbands. Once the feature vectors are generated, the retrieval procedure can be applied. The objective is to search the top candidate database images whose feature vectors $(\hat{\lambda}_j^{\text{db}})_{1\leq j\leq L}$ (where L is the length of the feature vector) are closer to that of the query one $(\hat{\lambda}_j^{\text{q}})_{1\leq j\leq L}$, according to a given similarity measure. More precisely, the database image I_j^{db} and the query one I_j^{q} are compared by computing the following measure:

$$D(I_j^{\text{db}}, I_j^{\text{q}}) = \sum_{j=1}^{L} \widetilde{D}(\hat{\lambda}_j^{\text{db}} \parallel \hat{\lambda}_j^{\text{q}}) \tag{8}$$

where $\widetilde{D}(\hat{\lambda}_j^{\text{db}} \parallel \hat{\lambda}_j^{\text{q}})$ represents the Kullback-Leibler Divergence (KLD) that measures the dissimilarity between the two Laplacian distributions of parameters $\hat{\lambda}_j^{\text{db}}$ and $\hat{\lambda}_j^{\text{q}}$. It is given by [4]:

$$D(\hat{\lambda}_j^{\text{db}} \parallel \hat{\lambda}_j^{\text{q}}) = \log\left(\frac{\hat{\lambda}_j^{\text{db}}}{\hat{\lambda}_j^{\text{q}}}\right) + \frac{\hat{\lambda}_j^{\text{q}}}{\hat{\lambda}_j^{\text{db}}} - 1. \tag{9}$$

4. EXPERIMENTAL RESULTS

Experiments were carried out on two different texture datasets. The first one is a selection of images from the popular MIT Vision Texture (VisTex) database [25], which consists of 40 textures that have been widely used for texture image retrieval purpose. Each 512×512 pixel image 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 [26] which is a large collection of 476 images of different textures. The images, of size 512×512, are split 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 are 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 study the compression effect on the retrieval performance, all the images are compressed at different levels that range from low to high bitrate by applying the 9/7 wavelet transform followed by an uniform scalar quantization with a deadzone of size q_j. The retrieval performances are evaluated in terms of precision $PR = \frac{R^{\mathrm{r}}}{R}$ and recall $RC = \frac{R^{\mathrm{r}}}{R^{\mathrm{t}}}$, where R^{t} is the total number of relevant images in the database, R^{r} is the number of output images selected as relevant and R represents the number of returned images. Note that R is equal to 16 in our experiments.

First, Fig. 1 shows the plots of the PR versus RC for the quantized VisTex database by taking different quantization steps, and so different bitrates R_Q and R_M for both the query and model images, respectively. For each image, the energy of all the quantized wavelet subbands are taken as salient features. Then, the resulting feature vectors of query and model images are compared using the normalized Euclidean distance [16]. The intuitive method which consists in comparing directly the images under different compression qualities is denoted by UC. The state-of-the-art strategy [8, 9], which aims at comparing images after recompression step, is designated by AR. It can be seen from Fig. 1 that recompressing the images at *similar* qualities improves the retrieval performance compared to the case where features are taken under compression. Moreover, the recompression strategy yields a significant gain especially where the query and DB images have very different qualities (i.e the difference of the bitrates R_Q and R_M is large).

Now, the proposed retrieval approach based on the distribution parameter estimation, which is denoted in what follows by DPE, is also compared to the conventional method. To this end, different combinations of bitrates assigned to the query and model images have been considered. Thus, as shown in Figures 2 and 3, our proposed approach outperforms the state of the art retrieval method based on the recompression technique.

5. CONCLUSION

In this paper, we have presented a novel image retrieval approach to deal with a database available in the compressed domain. More specifically, based on the Laplacian statistical model, we have proposed to estimate directly from the quantized wavelet coefficients the distribution parameter of the original wavelet coefficients. The estimated parameter is then chosen as salient feature. Experimental results have shown the benefits which can be drawn from the proposed approach compared to the state-of-the-art one based on the recompression technique. Future works aim to provide a general framework by considering the well-known statistical Generalized Gaussian model.

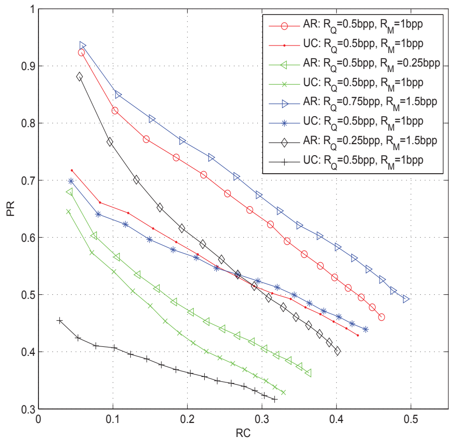

Fig. 1. Precision versus recall of wavelet-based compressed images (taken from the VisTex database): under compression (UC) and after recompression (AR).

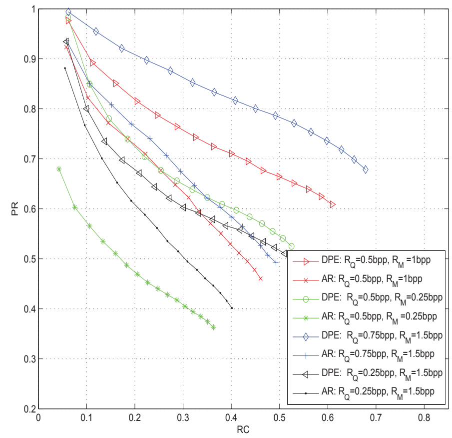

Fig. 2. Precision versus recall of wavelet-based compressed images (taken from the VisTex database) of the proposed approach (DPE) and the recompression one (AR).

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

[3] 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.

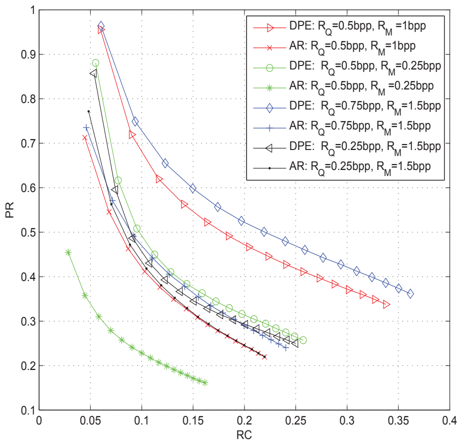

Fig. 3. Precision versus recall of wavelet-based compressed images (taken from the STex database) of the proposed approach (DPE) and the recompression one (AR).

[4] 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.

[5] 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.

[6] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] T. André, M. Cagnazzo, M. Antonini, and M. Barlaud, "JPEG2000-compatible scalable scheme for wavelet-based video coding," *EURASIP Journal on Image and Video Processing*, vol. 2007, 2007.

[15] 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.

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

[17] M. S. Allili, "Wavelet modeling using finite mixtures of generalized Gaussian distributions: application to texture discrimination and retrieval," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1452–1464, 2012.

[18] 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.

[19] 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.

[20] R. Kwitt, P. Meerwald, and A. Uhl, "Efficient texture image retrieval using copulas in a Bayesian framework," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2063–2077, 2011.

[21] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard," *Signal Processing: Image Communication*, vol. 17, no. 1, pp. 3–48, 2002.

[22] K. A. Birney and T. R. Fischer, "On the modeling of DCT and subband image data for compression," *IEEE Transactions on Image Processing*, vol. 4, no. 2, pp. 186–193, 1995.

[23] J. R. Price and M. Rabbani, "Biased reconstruction for JPEG decoding," *IEEE Signal Processing Letters*, vol. 6, no. 12, pp. 297–299, 1999.

[24] A. Ichigaya, M. Kurozumi, N. Hara, Y. Nishida, and E. Nakasu, "A method of estimating coding PSNR using quantized DCT coefficients," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 251–259, 2006.

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

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