GRADIENT PRESERVING QUANTIZATION

Mina Makar, Haricharan Lakshman, Vijay Chandrasekhar, and Bernd Girod

Information Systems Laboratory, Department of Electrical Engineering, Stanford University
{mamakar, harilaks, vijayc, bgirod}@stanford.edu

ABSTRACT

Local features are widely used for content-based image retrieval and object recognition. Most feature descriptors are calculated from the gradients of a canonical patch around repeatable keypoints in the image. In this paper, we propose a technique for designing quantization matrices that reduce the mean squared error distortion of the gradient derived from DCT-encoded canonical patches. Experimental results demonstrate that our proposed patch encoder greatly outperforms a JPEG encoder at the same encoding complexity. Moreover, our quantization matrices achieve lower gradient distortion and larger number of feature matches at the same bit-rate.

Index Terms— Image compression, image matching, gradient, quantization

1. INTRODUCTION

Many content-based image retrieval and augmented reality applications require the use of local image features. Examples of these robust local features include Scale-Invariant Feature Transform (SIFT) [1], Speeded Up Robust Features (SURF) [2] and Compressed Histogram of Gradients (CHoG) [3].

In an image matching framework, keypoints are detected from database and query images. These keypoints should be shift, scale and rotation invariant and also should be repeatable in different images of the same scene. After that, feature descriptors are calculated for every keypoint. Matching between two descriptors is evaluated using a suitable distance metric such as L_2 norm. All the above-mentioned feature descriptors [1, 2, 3] share the common framework that the descriptor consists of histograms of gradients in a patch located around the detected keypoint. For scale and rotation invariance, patches are oriented so that the maximum gradient is along the same direction and resized according to the scale of the detected feature. We refer to these oriented and scaled patches as canonical patches.

For applications where data are transmitted over a network for feature detection and image matching, it is desirable that the amount of data sent is as low as possible. The authors in [4] compare different solutions used to compress feature descriptors, especially SIFT. In our previous work [5], we present an initial study showing that one can alternatively transmit the compressed canonical patch and perform descriptor computation at the receiving end with only minor loss in performance. This has the advantage of allowing interoperability between systems using different feature descriptors. Also, moving descriptor computation to the receiving end is preferable when, e.g., a server in the cloud has more processing power than a mobile device acquiring the query image. A conventional image coding technique based on adaptive block-size discrete cosine transform (DCT) was used for patch encoding in [5].

In this paper, we propose a new patch encoder, designed for the specific requirements of an image matching system. Preserving a good visual quality is not important. However, it is important to preserve the feature descriptors calculated from the patch. We define a new distortion metric based on the patch gradients and optimize the rate-distortion performance according to this metric. This is achieved through designing a quantization matrix that preserves the patch gradients.

The remainder of the paper is organized as follows. Section 2 presents our proposed patch encoder and the modifications we apply to account for the requirements of an image matching system. In Section 3, we explain how to design a quantization matrix that preserves the patch gradients. Finally, in Section 4, we present experimental results showing the performance of the proposed quantization matrix in terms of lowering the distortion in the patch gradients and achieving more feature matches at the same bit-rate.

2. PATCH ENCODER

Since patches have different statistics and are used for different applications from natural images, We build our own *Patch ENCoder* (PENC) exploiting the fact that we only care to preserve the patch gradients since these gradients are used in the calculation of many local feature descriptors (SIFT, SURF, CHoG, etc.). The block diagram of PENC is shown in Fig. 1.

First, we extract canonical patches from the input image, by running keypoint detection algorithm based on detecting maxima in a difference of Gaussian (DoG) scale-space representation. The patches are extracted at a canonical scale and are oriented so that the maximum gradient is along the same direction. To ensure sending the patches at a reasonable bit-rate, patches are sampled using a 16×16 grid. However, we upsample the decoded patches to size 32×32 using bilinear interpolation before computing the corresponding descriptors to improve the matching performance.

To ensure low complexity and easy hardware implementation for PENC, we use an encoding pipeline similar to JPEG [6]. The first stage in PENC is a pre-processing stage which consists of Gaussian blurring of the patch followed by DC removal. Gaussian blurring adds more robustness to the patch against sub-pixel shifts between matching patches. DC can be discarded since only patch gradients are used for descriptor computation.

We then apply a 2-D 16×16 DCT on the pre-processed patches. Having the whole patch as a single transform block avoids blocking artifacts that may introduce false patch gradients. Each DCT coefficient is quantized using a uniform scalar quantizer where the quantization step is defined by the corresponding value in the quantization matrix Qmat. Qmat is multiplied by a quantization constant Q to adjust the overall bit-rate.

The final stage in PENC is entropy coding. PENC only encodes the AC coefficients as DC is already removed from the patch. Similar

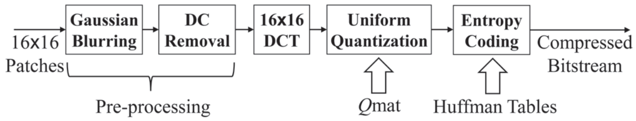

Fig. 1. Block diagram of the proposed patch encoder PENC. Patches are pre-processed through Gaussian blurring and DC removal. DCT coefficients are calculated and quantized using a quantization matrix Qmat. Finally, the quantized coefficients are entropy coded.

to JPEG standard, the quantized AC coefficients are zigzag scanned and then converted to symbols which concatenate a run of zeros and the category of the following magnitude. We refer to these symbols as (R, C) symbols. Refinement bits are added to obtain the exact magnitude of the AC coefficient. We fix the maximum run to be 16 and category to be 10. We also have two special symbols. The first symbol specifies the end of the patch. The other symbol indicates that we have 16 consecutive zeros and is used to chop a run longer than 16 into shorter runs. We train a Huffman table for these 162 symbols to obtain the variable length codewords for PENC.

We append a small header before PENC compressed bitstream. This header specifies Q and the number of patches and is used in the decoding process. For applications that require geometric consistency check of the image matching results, the final bitstream also includes the locations and orientations of the keypoints. The locations data are encoded using the method in [7] while the orientations are uniformly quantized and encoded using fixed length code.

3. QUANTIZATION MATRIX DESIGN

Based on the results of rate-distortion theory and reverse water-filling for optimal bit-rate allocation among transform coefficients, we know that using a constant quantization matrix (Qmat is the all ones matrix) results in the best rate-distortion performance assuming high bit-rate regions and using mean squared error (MSE) as the distortion measure.

We use $N \times N$ signals and refer to the original and quantized versions of the signal as f and f_q respectively. To account for the distortion in the gradients of the signal due to quantization, we define a cost function J_{grad} as the MSE in the signal gradients as follows

$$J_{grad} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} \left\| (\nabla f)_{ij} - (\nabla f_q)_{ij} \right\|^2 \tag{1}$$

Hence, the gradient SNR (GSNR) is defined as

$$\text{GSNR} = 10 \log_{10} \left(c^2 / J_{grad} \right) \tag{2}$$

where c is an arbitrary constant. We use $c = 255$ to match the convention used for regular PSNR of 8-bit images. If e is the error between f and f_q (i.e., $e = f - f_q$), the MSE distortion between f and f_q can be calculated from the DTFT domain as

$$J_{MSE} \propto \frac{1}{4\pi^2} \int_{\omega_x=-\pi}^{\omega_x=\pi} \int_{\omega_y=-\pi}^{\omega_y=\pi} \phi_{ee} (\omega_x, \omega_y) \, d\omega_x d\omega_y \tag{3}$$

where ϕ_{ee} is the power spectral density (PSD) of e. Using the properties of DTFT, we can write

$$J_{grad} \propto \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} \phi_{ee} (\omega_x, \omega_y) \, . \left(\omega_x^2 + \omega_y^2 \right) d\omega_x d\omega_y \tag{4}$$

Note that (4) can be viewed as filtering the PSD of e with a linear filter H such that

$$|H (\omega_x, \omega_y)| = \sqrt{\omega_x^2 + \omega_y^2} \tag{5}$$

Thus, if we were to encode the signal in the 2-D Fourier domain, we would expect a quantization matrix

$$Q\text{mat} (\omega_x, \omega_y) \propto \frac{1}{|H (\omega_x, \omega_y)|} = \frac{1}{\sqrt{\omega_x^2 + \omega_y^2}} \tag{6}$$

While (6) gives us a first insight and general guidance, additional issues must be addressed to come up with the best Qmat for a practical patch encoder. First, PENC uses a type-II DCT [8] rather than a Fourier transform. Second, most practical implementations for gradient calculation use the simplest finite gradient filter with impulse response $[-0.5\ 0\ 0.5]$ horizontally and vertically. The frequency response of this gradient filter is far from ideal. Third, the feature descriptors are calculated on an upsampled version of the decoded patch. Finally, we would like to take into account how filtering operations are performed near the borders of the patch.

To account for all the previous practical considerations, we formulate the problem in discrete domain and represent filtering, interpolation and DCT transform using matrix multiplications. We convert f and f_q to column vectors $\mathbf{f}$ and $\mathbf{f_q}$ via column-wise scanning. Both $\mathbf{f}$ and $\mathbf{f_q}$ are of length N^2. The horizontal and vertical gradients can be calculated with a $[-0.5\ 0\ 0.5]$ filter through multiplication with matrices $\mathbf{G_x}$ and $\mathbf{G_y}$ shown in Fig. 2 for $N = 4$. Consider the discrete gradient cost function J_{dg}

$$\begin{aligned} J_{dg} &= \|\mathbf{G_x} (\mathbf{f} - \mathbf{f_q})\|^2 + \|\mathbf{G_y} (\mathbf{f} - \mathbf{f_q})\|^2 \\ &= (\mathbf{f} - \mathbf{f_q})^T \mathbf{G_x}^T \mathbf{G_x} (\mathbf{f} - \mathbf{f_q}) + (\mathbf{f} - \mathbf{f_q})^T \mathbf{G_y}^T \mathbf{G_y} (\mathbf{f} - \mathbf{f_q}) \\ &= (\mathbf{f} - \mathbf{f_q})^T \mathbf{W} (\mathbf{f} - \mathbf{f_q}), \end{aligned} \tag{7}$$

where $\mathbf{W} = \mathbf{G_x}^T \mathbf{G_x} + \mathbf{G_y}^T \mathbf{G_y}$

Since we are interested in preserving the gradients in the patch upsampled by two horizontally and vertically, we define the bilinear interpolation matrix $\mathbf{B}$ of dimensions $4N^2 \times N^2$. Fig. 2 shows an example matrix $\mathbf{B}$ for the case of $N = 4$. We are interested in minimizing the cost function J_{dgu} which depends on the error in the gradients between the upsampled original patch and the upsampled quantized patch. Note the change in dimensions of $\mathbf{G_x}$ and $\mathbf{G_y}$ to be $4N^2 \times 4N^2$.

$$\begin{aligned} J_{dgu} &= (\mathbf{Bf} - \mathbf{Bf_q})^T \mathbf{W} (\mathbf{Bf} - \mathbf{Bf_q}) \\ &= (\mathbf{f} - \mathbf{f_q})^T \mathbf{B}^T \mathbf{W} \mathbf{B} (\mathbf{f} - \mathbf{f_q}) \end{aligned} \tag{8}$$

Since quantization is performed in the DCT domain, we refer to the 2-D DCT transform matrix as $\mathbf{T}$ (see Fig. 2) and the DCT coefficients of $\mathbf{f}$ and $\mathbf{f_q}$ as $\mathbf{c}$ and $\mathbf{c_q}$ respectively.

$$\begin{aligned} J_{dgu} &= \left(\mathbf{T}^{-1}\mathbf{c} - \mathbf{T}^{-1}\mathbf{c_q} \right)^T \mathbf{B}^T \mathbf{W} \mathbf{B} \left(\mathbf{T}^{-1}\mathbf{c} - \mathbf{T}^{-1}\mathbf{c_q} \right) \\ &= (\mathbf{c} - \mathbf{c_q})^T \mathbf{T} \mathbf{B}^T \mathbf{W} \mathbf{B} \mathbf{T}^T (\mathbf{c} - \mathbf{c_q}) \end{aligned} \tag{9}$$

Diagonal elements in $\mathbf{TB}^T\mathbf{WBT}^T$ represent the relative importance of each DCT coefficient in the cost function. We define the matrix $\mathbf{S} = \mathbf{TB}^T\mathbf{WBT}^T$ at diagonal elements and $\mathbf{S} = 0$ elsewhere. Elements in $\mathbf{S}$ are normalized such that $\sum_{k=1}^{N^2} \mathbf{S}_{kk} = N^2$. Similar to (6), assuming that we are operating in the high-rate regime, the quantization matrix that minimizes J_{dgu} at a specified bit-rate is given by

$$Q\text{mat}_k = \frac{Q}{h_k}, \text{ where } h_k = \sqrt{\mathbf{S}_{kk}} \text{ and } k = 1, ..., N^2 \tag{10}$$

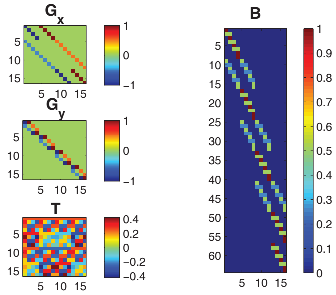

Fig. 2. Matrices used in Qmat design ($N = 4$): Gradients $\mathbf{G_x}$ and $\mathbf{G_y}$, DCT transform $\mathbf{T}$ and bilinear interpolation $\mathbf{B}$.

∞	112.4	56.7	38.4	29.5	24.3	21	18.9	17.4	16.5	15.9	15.5	15.4	15.4	15.5	15.5
112.4	79.7	50.8	36.5	28.7	23.9	20.8	18.7	17.3	16.4	15.8	15.5	15.4	15.4	15.4	15.5
56.7	50.8	40.5	32.2	26.6	22.7	20.1	18.3	17	16.1	15.6	15.3	15.2	15.2	15.3	15.4
38.4	36.5	32.2	27.8	24	21.2	19.1	17.6	16.5	15.8	15.3	15.1	15	15.1	15.1	15.2
29.5	28.7	26.6	24	21.7	19.6	18.1	16.9	16	15.4	15	14.8	14.8	14.9	15	15
24.3	23.9	22.7	21.2	19.6	18.2	17	16.1	15.4	15	14.7	14.6	14.6	14.6	14.8	14.8
21	20.8	20.1	19.1	18.1	17	16.2	15.5	14.9	14.6	14.4	14.3	14.4	14.5	14.6	14.7
18.9	18.7	18.3	17.6	16.9	16.1	15.5	14.9	14.5	14.3	14.1	14.1	14.2	14.3	14.4	14.5
17.4	17.3	17	16.5	16	15.4	14.9	14.5	14.2	14	14	14	14.1	14.2	14.4	14.5
16.5	16.4	16.1	15.8	15.4	15	14.6	14.3	14	13.9	13.9	13.9	14.1	14.2	14.4	14.5
15.9	15.8	15.6	15.3	15	14.7	14.4	14.1	14	13.9	13.9	14	14.1	14.3	14.4	14.6
15.5	15.5	15.3	15.1	14.8	14.6	14.3	14.1	14	13.9	14	14.1	14.2	14.4	14.6	14.7
15.4	15.4	15.2	15	14.8	14.6	14.4	14.2	14.1	14.1	14.1	14.2	14.4	14.6	14.8	14.9
15.4	15.4	15.2	15.1	14.9	14.6	14.5	14.3	14.2	14.2	14.3	14.4	14.6	14.8	15	15.1
15.5	15.4	15.3	15.1	15	14.8	14.6	14.4	14.4	14.4	14.4	14.6	14.8	15	15.2	15.3
15.5	15.5	15.4	15.2	15	14.8	14.7	14.5	14.5	14.5	14.6	14.7	14.9	15.1	15.3	15.5

Fig. 3. Proposed gradient-preserving Qmat for PENC. $N = 16$ and $Q = 16$.

The normalization of $\mathbf{S}$ results in the same total weighting for DCT coefficients as the identity matrix $\mathbf{I}$. This ensures fair comparison between Qmat and the constant quantization matrix when both matrices are scaled up by the same Q. Finally, Qmat is arranged as an $N \times N$ matrix. Fig. 3 presents Qmat generated using our proposed technique and used in PENC ($N = 16$ and $Q = 16$).

4. EXPERIMENTAL RESULTS

We evaluate our scheme in the context of the ongoing standardization efforts, Compact Descriptors for Visual Search (CDVS) [9] conducted by the Motion Picture Experts Group (MPEG). Experiments are performed using images from the CDVS dataset. These images are divided into five classes to cover different image matching applications. These classes are *text/graphics*, *museum paintings*, *video frames*, *buildings/landmarks*, and *common objects*.

4.1. Rate-Distortion Performance

We vary the quantization constant $Q \in \{16, 32, 64, 128, 256\}$ to cover the whole bit-rate range. A different Huffman table is trained for each Q in the cases of constant Qmat and the proposed Qmat using the (R, C) statistics generated from encoding the patches of 5000 images from the distractor images used in CDVS evaluation (not part from the CDVS dataset). This guarantees fair comparison in entropy coding between different schemes.

Fig. 4 shows sample images from the CDVS dataset. The left image is from the *text/graphics* class and the right image is from the *buildings/landmarks* class. We extract 200 patches with the highest Hessian score from each image. In Fig. 5, we show the first 30 original patches from each image. The original patches are encoded with two variants of PENC; the first uses constant Qmat and the second uses the proposed Qmat.

We plot the rate versus gradient SNR curves for the patches of the sample images. Gradient SNR (GSNR) is calculated based on (2) while using $\frac{1}{N^2} J_{dgu}$ (9) instead of J_{grad} to take all practical considerations into account. We observe a gain of 0.8 dB in GSNR at high rates due to the proposed Qmat. Conventional JPEG encoding of the patches (using MATLAB's implementation) is plotted for comparison. Both variants of PENC achieve a large improvement of more than 4 dB over JPEG. However, the gain due to the gradient-preserving quantization matrix is relatively modest.

4.2. Feature Matching Experiment

To study the effect of using the proposed Qmat on feature matching, we conduct the pairwise matching experiments described in the CDVS evaluation framework. This includes matching of 10319 pairs of images from all the five classes defined in the CDVS dataset. RANSAC [10] is used for geometric consistency check. Since, we are expecting improvements only for high bit-rate patches, we use a small quantization constant Q and vary the bit-rate by varying the number of encoded patches. We extract a maximum of 600 patches per image, arrange them in descending order of their Hessian score and at each bit-rate point we only encode a subset of these patches.

Fig. 6 shows the average number of feature matches per image averaged over all classes in the CDVS dataset. Query size includes encoded patches and geometric verification data. We use the smallest value of $Q = 16$ for comparing the performance of PENC with constant Qmat and with the proposed Qmat. The points on the curves correspond to sending 50, 100, 200, 300, 400 and 600 patches per image. The results show the improvement in feature matching due to the proposed Qmat. We achieve around 5% bit-rate reduction at the same average number of feature matches. Around $20 - 30\%$ of the matches are discarded during geometric verification; however, pre-RANSAC and post-RANSAC results show the same improvement.

On the image matching level, we declare that two images match if they have 6 or more post-RANSAC feature matches. Comparing the use of constant Qmat and the proposed Qmat in image matching, we find that both matrices result in very similar performance. Investigating image matching for different classes, the *video frames* class is the easiest to match with a true positive rate of 98% at the largest query size point, while the *buildings/landmarks* class is the most difficult with a true positive rate of only 77% at the largest query size point.

(a) (b)

Fig. 4. Sample images from the CDVS dataset. (a) *text/graphics* class and (b) *buildings/landmarks* class.

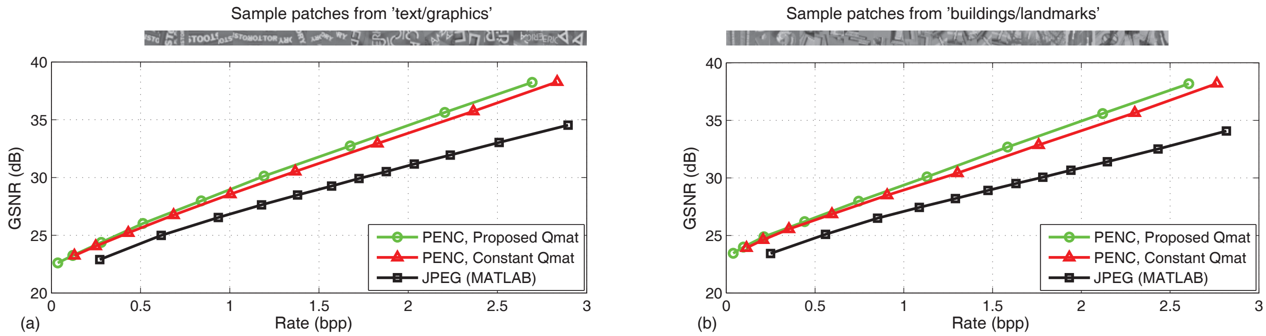

Fig. 5. Sample patches and rate-GSNR curves for the images in Fig. 4. (a) *text/graphics* class and (b) *buildings/landmarks* class

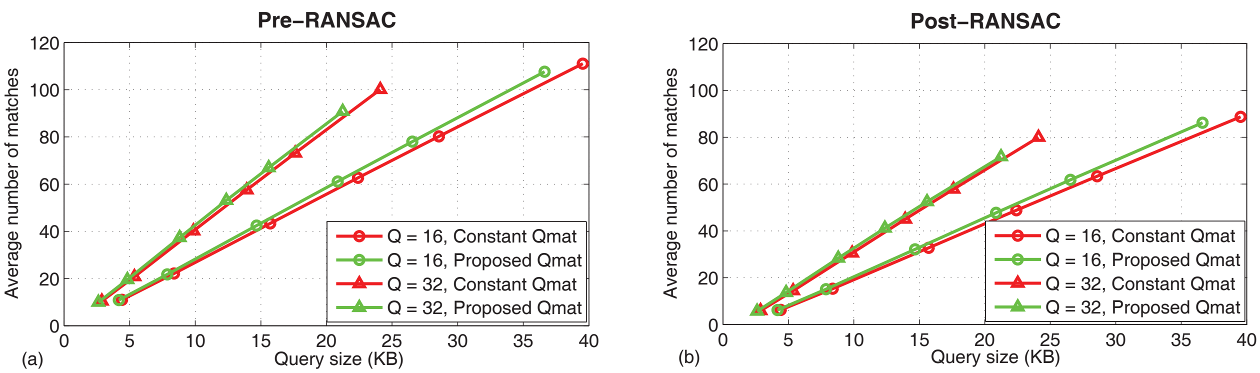

Fig. 6. Pairwise matching experiment for high bit-rate patches. (a) pre-RANSAC and (b) post-RANSAC.

Performing the same experiment at $Q = 32$, we observe that the gain due to the proposed Qmat is decreasing and the performance of constant Qmat and the proposed Qmat are nearly the same. This indicates that in feature matching, our proposed Qmat is only useful for high bit-rate patches which is the important region for applications where the matching performance has higher priority than the patches bit-rate. As we increase Q (low bit-rate patches), we observe that constant Qmat has better performance than the proposed Qmat in feature matching. This is because at large Qs, the difference between the DCT coefficients error variance is large and our assumptions in Section 3 are no longer valid.

5. CONCLUSIONS

We present an efficient method for encoding canonical image patches intended for the calculation of local feature descriptors. The proposed patch encoder uses a 16×16 DCT and a gradient-preserving quantization matrix.

Experimental results show that the proposed patch encoder outperforms a JPEG encoder in terms of gradient SNR by more than 4 dB at the same bit-rate. Using gradient-preserving quantization matrices can improve the feature matching performance at high bit-rates. Our results show a 5% bit-rate reduction for the same average number of feature matches.

6. REFERENCES

[1] D. Lowe, "Distinctive Image Features From Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, Nov. 2004.

[2] H. Bay, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features," *Proc. European Conference on Computer Vision (ECCV)*, Graz, Austria, May 2006.

[3] V. Chandrasekhar, G. Takacs, D. Chen, S. Tsai, R. Grzeszczuk, and B. Girod, "CHoG: Compressed Histogram of Gradients," *Proc. IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, June 2009.

[4] V. Chandrasekhar *et al.*, "Survey of SIFT Compression Schemes," *Proc. Second International Workshop on Mobile Multimedia Processing (WMMP)*, Istanbul, Turkey, Aug. 2010.

[5] M. Makar, C.-L. Chang, D. Chen, S. Tsai, and B. Girod, "Compression of Image Patches for Local Feature Extraction," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009.

[6] ITU-T and ISO/IEC JTC1, "Digital Compression and Coding of Continuous-Tone Still Images," *in ISO/IEC 10918-1 - ITU-T Recommendation T.81*, Sept. 1992.

[7] S. Tsai, D. Chen, G. Takacs, V. Chandrasekhar, J. P. Singh, and B. Girod, "Location Coding for Mobile Image Retrieval," *Proc. International Mobile Multimedia Communications Conference (Mobimedia)*, London, U.K., Sept. 2009.

[8] E. Feig, and S. Winograd, "Fast Algorithms for the Discrete Cosine Transform," *IEEE Transactions on Signal Processing*, vol. 40, pp. 2174-2193, Sept. 1992.

[9] Y. Reznik, G. Cordara, and M. Bober, "Evaluation Framework for Compact Descriptors for Visual Search," *in ISO/IEC JTC1/SC29/WG11/N12202*, July 2011.

[10] M. Fischler, and R. C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting," *Communications of ACM*, vol. 24, no. 6, pp. 381-384, June 1981.