CLASSIFIED CONTEXT QUANTIZATION OF VQ INDEXES FOR IMAGE COMPRESSION

Yun Gong

Department of Electrical Engineering
Widener University
Chester, PA 19013, USA
Email:yun.gong@widener.edu

ABSTRACT

A classified context quantization (CCQ) technique is proposed to code basic image VQ indexes in the setting of high order context models. The context model of an index is first classified into one of three classes according to the smoothness of the image area they represent. Then the index is coded with a context quantizer designed for that class. Experimental results show that CCQ achieves about three percent improvement over the previous best results of image VQ by conditional entropy coding of VQ indexes (CECOVI), and does so at a lower computational cost.

1. INTRODUCTION

Vector quantization (VQ) has been intensively researched for more than twenty years, and it has produced the best performance in many practical problems [1]. However, for image compression, direct application of VQ currently seems not to be the most efficient technique compared to many transform coding methods such as wavelet image coding. The reason for the inferior performance of image VQ is operational, i.e., practical VQ can't afford to completely exploit the high-order pixel correlations in images by using very large vector dimensions. Many research efforts have been made to improve the performance of image VQ. In a recent work [2], Wu et al. developed a frame work of conditional entropy coding of VQ indexes (CECOVI) and utilized a simple Bayesian-type method to estimate high order conditional probabilities of VQ indexes that explore inter-codevector correlations of basic VQ. Their experimental results show about 20 percent improvement over the previous best results of image VQ.

In this paper, we study the same image VQ indexes problem as in CECOVI. First, we follow the same arguments to generate VQ codevectors and their corresponding indexes. In general, VQ indexes are just labels of codevectors, and they do not have physical significance pertaining to the original signals. However, in case of image compression, if the codevectors are arranged in ascending order of their average energy and the corresponding indexes are interpreted as intensity values of pixels, it can be observed that the indexes preserve the basic structure of the original image. Therefore, there must exist some structures between the indexes which represent higher order pixel correlations. Second, we investigate the same context models of VQ indexes. Suppose a VQ index sequence are generated by an ordered codebook in raster scan order. A modeling context of four events $\{x_1, x_2, x_3, x_4\}$, which are the immediate neighboring VQ indexes to the current VQ index x, is studied:

$$\begin{array}{ccc} x_3 & x_2 & x_4 \\ x_1 & x & \cdots, \end{array}$$

where x and x_i, $i = 1, 2, 3, 4$ are drawn from alphabet $\mathcal{A} = \{1, 2, \cdots, n\}$, and n is the size of the codebook.

In order to minimize the coding rate of VQ indexes, the number of bits used to code the index sequence should be a function of $p(x|x_1, x_2, x_3, x_4)$. But $p(x|x_1, x_2, x_3, x_4)$ is generally unknown in practice, and has to be estimated on the fly based on past observations in the coding process. It is well known that directly estimating $p(x|x_1, x_2, x_3, x_4)$ has the "model cost" [2, 3] problem of high order context modeling. In CECOVI, Wu et al. used Bayes' theorem to tackle the problem. The major downside of CECOVI is its complexity.

A recent work [3] by Chen provides us another perspective to look at the above VQ index problem. Chen introduced a general context quantization concept. In this paper, we apply the context quantization technique to code image VQ indexes. We propose a new method, classified context quantization (CCQ) to deal with a variety of neighboring indexes that represent blocks with distinct features of images such as edge or

This research was supported by the NSF-PFI project under Grant EEC-0332490.

nonedge areas. In CCQ, a classifier determines the class for the modeling context of four neighboring indexes. Then the current index is coded with a context quantizer designed specially for that class. As demonstrated later, our CCQ method achieves better performance than CECOVI with less complexity. The remainder of this paper is organized as follows. To introduce notation and to be self-contained, we briefly describe the context quantization technique in the context of image VQ indexes in Section 2. We then discuss the classification of context models and the design of classified context quantizers in detail in Section 3. The experimental results are presented in Section 4 followed by conclusion in Section 5.

2. CONTEXT QUANTIZATION

$p(x|x_1, x_2, x_3, x_4)$ involves a space of n^5 points, where $x, x_i \in \mathcal{A} = \{1, 2, \cdots, n\}$. Define the conditional set $\mathcal{C} = \{(x_1, x_2, x_3, x_4), x_i \in \mathcal{A}\}$, and the size of the set $\mathcal{C}$ is $N = n^4$. The basic idea of context quantization is to find a partition of $\mathcal{C}$, which we can denote as $\mathcal{M} = \{m_k, k = 1, 2, \cdots, K\}$, such that $K \ll N$, and all $p(x|x_1, x_2, x_3, x_4)$, where $(x_1, x_2, x_3, x_4) \in m_k$, can be represented by one conditional probability $p(x|m_k)$. The first step in the development of quantization techniques is to define a distortion measure. Since the relative entropy $D(p||q)$ is a measure of the inefficiency of assuming that the distribution is q when the true distribution is p [5], the relative entropy between $p(x|x_1, x_2, x_3, x_4)$ and its quantization value $p(x|m_k)$

$$D(p(x|x_1, x_2, x_3, x_4)||p(x|m_k)) = \sum_{x \in \mathcal{A}} p(x|x_1, x_2, x_3, x_4) \log_2 \frac{p(x|x_1, x_2, x_3, x_4)}{p(x|m_k)} \quad (1)$$

was chosen in [3] as the distortion measure. The overall quantization distortion can then be expressed by

$$D = \sum_{(x_1,x_2,x_3,x_4) \in \mathcal{C}} D(p(x|x_1, x_2, x_3, x_4)||p(x|m_k)) \times p(x_1, x_2, x_3, x_4) \quad (2)$$

The principal objective of context quantizer design is: for a given number of quantization levels K, find an optimal partition $\mathcal{M} = \{m_k, k = 1, 2, \cdots, K\}$ for $\mathcal{C}$ so that the distortion defined by (2) is minimized. A Lloyd style iterative algorithm was given in [3] to find optimal context quantizers after the nearest neighbor condition and the centroid condition for the design problem were proved to be satisfied. From the centroid condition, we can calculate the quantization value $P(x|m_k)$ for each partition m_k as follows

$$P(x|m_k) = \frac{\sum_{m_k} P(x|x_1, x_2, x_3, x_4) P(x_1, x_2, x_3, x_4)}{\sum_{m_k} P(x_1, x_2, x_3, x_4)} \quad (3)$$

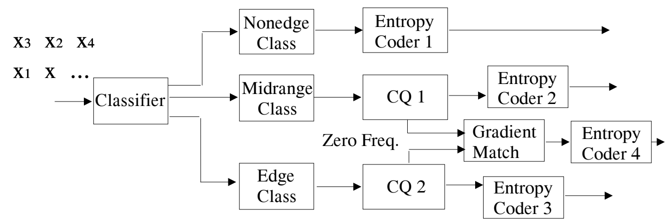

Fig. 1. Classified Context Quantization

The Lloyd style iterative algorithm provides a general guideline for finding optimal quantizers. However, it is well known that the step of finding the optimal partition for a given quantizer is very computational intensive when too many training samples are involved, which is the case of quantizing $p(x|x_1, x_2, x_3, x_4), x, x_i \in \mathcal{A}$. Furthermore, the neighboring indexes could represent blocks of an image with very different features. So, image VQ indexes should be considered as samples from a composite source. Those motivate us to develop the following classified context quantization (CCQ) technique. CCQ can be viewed as the counterpart of classified vector quantization [4] in the setting of context quantization.

3. CLASSIFIED CONTEXT QUANTIZATION OF VQ INDEXES

3.1. Context Quantization Classifier

We first classify image VQ indexes into two groups: nonedge and edge, according to the characteristics of their corresponding codevectors. We define the mean of each codevector $Y = [y_1, y_2, \cdots, y_L]$ by $m = \frac{1}{L} \sum_{i=1}^{L} y_i$, the variance $\sigma^2 = \frac{1}{L} \sum_{i=1}^{L} (y_i - m)^2$, and the relative variance $\delta = \sigma/m$. An index associated with a codevector with $\delta < \epsilon$ is classified as a nonedge index, otherwise it is called as an edge index. Here, ϵ is a small number. Denote $\mathcal{E}$ as the set with all edge indexes, and $\mathcal{S}$ as the set with all nonedge indexes. Recall that $\mathcal{A}$ is the complete index set of size n. Then, $\mathcal{A} = \mathcal{E} + \mathcal{S}$. If the size of $\mathcal{S}$ is r, the size of $\mathcal{S}$ is $n-r$. Normally, we choose a small ϵ and make r small compared to $n - r$.

Nonedge regions of an image usually appear continuous. if the neighboring indexes are nonedge indexes, it is highly possible that the current VQ index is a nonedge index too. On the other hand, neighboring edge indexes make the current index unpredictable. Therefore, with different neighboring indexes, $p(x|x_1,$

$x_2, x_3, x_4)$ should be estimated and quantized in different ways. We classify the four-event context model of the neighboring indexes into three different classes as shown in Figure 1. If $x_1 = x_2 \in S$, the context model is defined as a nonedge class. If $x_1, x_2 \in S$, but $x_1 \neq x_2$, the context model is said to be a midrange class. The context model belongs to an edge class for all other cases.

The nonedge class is the most straightforward case. First, we estimate $p(x|s)$ directly by counting the number of training samples that are classified into this class since $p(x|s)$, where $s = x_1 = x_2 \in \mathcal{S}$ and $x \in \mathcal{A}$, has a manageable size of $n \times r$ $(< n^2)$. Then we use an arithmetic coder for $p(x|s)$. The continuity of nonedge regions makes it likely that $\max_{x \in A} p(x|s) = p(s|s)$, and $p(x|s)$ is a highly skewed probability for each given s. We expect that the arithmetic coder would work well on this class. For the midrange class and the edge class, we design two context quantizers, CQ1 and CQ2 respectively; then apply entropy coding. Details is given in the next subsection. Context quantization does not solve the zero frequency problem caused by sample sparsity when estimating high-order conditional probabilities. Zeros or very small numbers are not considered in the quantizer design process. For this case, we use the gradient match method [7] to predict the current index from its available neighbors, then perform entroy coding.

3.2. Context Quantization of Midrange and Edge Classes

Since a good estimation of $p(x|x_1, x_2, x_3, x_4)$ requires too large number of samples, we consider only two conditioning events, i.e., we estimate $p(x|x_1, x_2)$ for both midrange and edge classes. Context quantization for both classes is to find an optimal quantizer $p(x|m_k), k = 1, 2, \cdots, n$, to represent all $p(x|x_1, x_2)$, for $(x_1, x_2) \in m_k$. The quantization reduces the size of context models from $n \times r^2$ to n^2 for the midrange class, and from n^3 to n^2 for the edge class. The Lloyd style iterative algorithm given in [3] is used to design the context quantizer. Recall that the iterative algorithm converges to a local optimal value. It is very important to choose an appropriate initial quantizer. In order to make arithmetic coding efficient for probability $p(x|m_k)$, we choose an initial quantizer with $p(x|m_k)$, $k = 1, 2, \cdots, n$, highly skewed and centered at $x = k$. Such initial quantizer can be obtained by the following steps.

- Step 1: Use a sufficiently large set of images to get a good estimate of $p(x|x_1, x_2)$.
- Step 2: Partition the conditional set $\mathcal{C} = \{(x_1, x_2), x_i \in \mathcal{A}\}$ of size $n \times r^2$ or n^2 into $\mathcal{M} = \{m_k, k = 1, 2, \cdots, n\}$ of size n such that for each $(x_1, x_2) \in m_k$, $\max_{x \in \mathcal{A}} p(x|x_1, x_2) = p(k|x_1, x_2)$.
- Step 3: Calculate the quantization value $p(x|m_k)$, $k = 1, 2, \cdots, n$, using equation (3).

Table 1. Results for LENA and PEPPER

		LENA	PEPPER
Nonedge	size	4256	2813
	rate	0.099	0.108
Midrange	size	4559	4092
	rate	0.144	0.145
Edge	size	4287	5512
	rate	0.251	0.197
Zero freq.	size	2900	3585
	rate	0.355	0.318
Overall	size	16384	16384
	rate	0.202	0.198

3.3. Dealing with Zero Probabilities

Even with only two conditioning events, estimation of $p(x|x_1, x_2)$ through counting requires a very large number of samples. Plus, the training set may not be sufficiently representative to cover all the cases of $p(x|x_1, x_2)$, $x, x_1, x_2 \in \mathcal{A}$. It is possible that some of $p(x|x_1, x_2)$ have zero or very small frequency. Zero or very small probabilities are not considered when the context quantizers are designed. In other words, there is no quantization value $p(x|m_k)$ to be assigned to any zero or very small $p(x|x_1, x_2)$. We need to find another way to encode indexes with such context model of zero or small probability. We re-examine the four-event modeling context from a prediction point of view, i.e., we try to predict the current index x from its neighboring indexes (x_1, x_2, x_3, x_4) available for both encoder and decoder by using the gradient match method [7]. Most cases of zero or small probabilities come from the edge class which represents a complex area of an image. A prediction approach exploring the gradient continuity property is expected to have better performance than one using the spatial continuity property [6]. The total gradient match distortion (gmd) is the sum of distortions from vertical, horizontal and diagonal directions. See [7] for details. The codevector with the smallest *gmd* is chosen to be the reconstructed block whose index is the prediction of x, denoted as $\hat{x}$. $p(x|\hat{x})$ will be used for arithmetic coding in the next step of the encoding process.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed CCQ, 30 monochrome images (couple crowd barbara tiffany man zelda lake mandrill boat, plus their 90^o, 180^o rotations) are used as the training set. The LBG algorithm is applied to build the codebook of size 256. The codevectors are then arranged in ascending order of their norm. $\delta = 0.02$ is used to classify the codevectors and corresponding indexes. LENA and PEPPER are chosen to be two test images in order to compare with results in [2]. In Table 1, the number of VQ indexes in three classes plus the zero frequency case and the corresponding rates are presented. The total number of indexes also includes the first row, the first column and the last column of indexes that need to transmit first using an basic probability model without considering the neighboring indexes. We compare our CCQ method with CECOVI [2] in Table 2. Two to three percent of decrease on coding rate is observed from the proposed CCQ method.

Both CCQ and CECOVI require $O(n^2)$ memory to store estimated probabilities. For CCQ, extra $n^2 + r^2$ memory is required to store context quantization information for the midrange and edge classes. The computational complexity of CCQ comes from the zero frequency case where $O(n)$ additions and comparisons are needed to find the smallest gradient match distortion for each index. As $O(n)$ operations are required by CECOVI for every index in an image, the zero frequency case in CCQ only constitutes about 20 percent of total indexes for the image. Furthermore, the final results of CECOVI are generated by a weighted sum of two order-one conditional probabilities and one order-two conditional probability for each index. Three weighting functions need to be sought heuristically for every index in the image. Finally, in CECOVI, the arithmetic coder for high-order conditional probabilities such as $p(x|x_1, x_2, x_4)$ of size n^4 needs to be implemented, and CCQ only requires the arithmetic coders for order-one conditional probabilities such as $p(x|m_k)$ of size n^2. In summary, it is fair to say that CCQ has less computational complexity with comparable memory requirement compared to CECOVI.

Table 2. Comparison of CCQ and CECOVI

	block size	CCQ PSNR (dB)	CCQ rate	CECOVI PSNR (dB)	CECOVI rate
LENA	4x4	30.38 (-0.6%)	0.202 (-3.4%)	30.55	0.208
	2x2	36.01 (+0.3%)	0.778 (-3.1%)	35.92	0.803
PEPPER	4x4	29.77 (-0.8%)	0.198 (-2.0%)	30.01	0.202
	2x2	34.74 (-0.1%)	0.772 (-3.4%)	34.79	0.799

5. CONCLUSION

We propose a classified context quantization (CCQ) method to code VQ indexes based on the context model introduced in [2]. The context models of VQ indexes are first classified into three classes according to the smoothness of the image area that their neighboring indexes represent. For the nonedge class, a simple order-one conditional probability is estimated for entropy coding. For the midrange and edge classes, context quantizers are designed by using the Lloyd style algorithm proposed in [3]. To deal with zero frequency cases, a gradient match method is utilized. The experimental results show that the proposed CCQ method reduces the bit rate of CECOVI by about two to three percent for the similar distortion, with lower computational complexity and the same order memory requirement.

6. REFERENCES

[1] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, New York: Kluwer Academic, 1992.

[2] X. Wu, J. Wen, and W. H. Wong, Conditional Entropy Coding of VQ Indexes for Image Compression, IEEE Transactions on Image Processing, Vol.8, No.8, pp.1005-1013, August 1999.

[3] J. Chen, Context Modeling Based on Context Quantization With Application in Wavelet Image Coding, IEEE Transactions on Image Processing, Vol.13, No.1, pp.26-32, January 2004.

[4] B. Ramamurthi and A. Gersho, Classified Vector Quantization of Images, IEEE Transactions on Communications, vol.COM-34, No.11, pp.1105-1115, November 1986.

[5] T.M.Cover and J.A.Thomas, Elements of Information Theory, New York: Wiley, 1991.

[6] S.C. Juan and C.Y. Lee, Entropy-Constrained Gradient-Match Vector Quantization for Image Coding, Proc. of ICASSP 1998, pp.2665-2668.

[7] Y. Gong, M. Fan and C.-M. Huang, "Lossless Coding on VQ Indexes for Image Compression," Proceedings of SCI'2001, Volume XIII, Image, Acoustic, Speech and Signal Processing: Part II, pp. 382-386, Orlando, July 2001.