Vector Quantization Using the L_∞ Distortion Measure

V. John Mathews, *Senior Member, IEEE*, and Peter J. Hahn, *Student Member, IEEE*

***Abstract*—This paper considers vector quantization of signals using the L_∞ distortion measure. The key contribution is a result that allows one to characterize the centroid of a set of vectors for the L_∞ distortion measure. A method similar to the Linde–Buzo–Gray (LBG) algorithm for designing codebooks has been developed and tested. The paper also discusses the design of vector quantizers employing the L_∞ distortion measure in an application in which the occurrences of quantization errors with larger magnitudes than a preselected threshold must be minimized.**

I. Introduction

In many applications involving image compression with vector quantization, it is necessary to limit the maximum distortion introduced in the image during the coding process. An example involves image compression using perceptual threshold functions [1], [2]. Perceptual threshold functions for images define the amount of distortion that can be introduced into an image without being detected by human observers. An image compression system that constrains the quantization errors to be below the perceptual threshold values is *perceptually lossless*. One way in which we can constrain the maximum value of the quantization errors in a vector quantizer system is to employ the L_∞ distortion measure in which the distance between two K-dimensional vectors X and Y is defined as

$$\|X-Y\|_\infty = \max_{i\in[1,K]} |x_i - y_i| \tag{1}$$

where x_i and y_i denote the ith samples of X and Y, respectively. In addition to being useful in perceptual coding of images, vector quantization using L_∞ norm has the advantage of computational simplicity. Note that the selection of the nearest neighbor using the L_∞ distortion measure does not require any multiplications.

The objective of this paper is to discuss an approach for developing codebooks for vector quantizers employing the L_∞ distortion measure. The key contribution of this paper is a result that allows one to characterize the centroid of a set of vectors for the L_∞ distortion measure. A method similar to the Linde–Buzo–Gray (LBG) algorithm [3] for designing codebooks for this distortion measure is developed using our result. The technique is then extended to the situation in which occurrences of quantization errors larger than some preselected threshold should be minimized. This idea has been employed for designing vector quantizers that limit all the distortions to below a pre-selected threshold in [4]. The results can be extended to applications involving perceptual threshold functions, but such an extension is not considered in this paper.

Manuscript received August 7, 1996. This work was supported in part by NASA under Grant NAG 5-2200, by the NSF under Grant MIP 9016331, and by an IBM Departmental Grant. A version of this paper was presented at the IEEE International Conference on Image Processing, Washington, DC, October 23–26, 1995. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. R. Ansari.

The authors are with the Department of Electrical Engineering, University of Utah, Salt Lake City, UT 84112 USA (e-mail: mathews@ee.utah.edu).

Publisher Item Identifier S 1070-9908(97)01766-5.

II. Centroid Calculation

The following theorem characterizes a property of the centroid that minimizes the average L_∞ distortion for a set of vectors.

Theorem 1: Let $X_1, X_2, \cdots, X_N$ be a set of K-dimensional vectors. Let C be the centroid of the vectors that minimizes the average L_∞ distortion between C and the vectors in the above set. Assume that the centroid satisfies the property that the L_∞ distortion between C and any vector in the set is uniquely determined by one of the K indices. Then the ith element of C satisfies the property that it is the median of the ith samples of all the vectors in the set such that

$$\|X_l - C\|_\infty = |x_{l,i} - c_i|$$

where $x_{l,i}$ and c_i are the ith elements of X_l and C, respectively.

Proof: Recall that the L_∞ distortion between two vectors can be defined as the limiting case of the L_p distortion when $p \to \infty$, *i.e.*,

$$\|X-Y\|_\infty = \lim_{p\to\infty} \left\{ \sum_{i=1}^{K} |x_i - y_i|^p \right\}^{1/p}. \tag{2}$$

Our objective is to select C such that

$$\begin{aligned} J_\infty &= \sum_{l=1}^{N} \|X_l - C\|_\infty \\ &= \sum_{l=1}^{N} \lim_{p\to\infty} \left\{ \sum_{i=1}^{K} |x_{l,i} - c_i|^p \right\}^{1/p} \end{aligned} \tag{3}$$

is minimized over all choices of C. The approach we pursue is to differentiate

$$J_p = \sum_{l=1}^{N} \left\{ \sum_{i=1}^{K} |x_{l,i} - c_i|^p \right\}^{1/p} \tag{4}$$

with respect to each c_m, find the limit of the derivatives as $p \to \infty$, and then set the resulting equations to zero to solve for the centroid vector. The differentiation of J_p gives

$$\frac{\partial}{\partial c_m} J_p = -\sum_{l=1}^{N}\sum_{i=1}^{K}\left(\frac{|(x_{l,i}-c_i)|^p}{|(x_{l,m}-c_m)|^p}\right)^{1-p/p} \cdot \mathrm{sign}\{x_{l,m}-c_m\} \tag{5}$$

where sign$(\cdot)$ denotes the sign of $(\cdot)$. Let us consider the quantity

$$\sum_{i=1}^{K}\left(\frac{|(x_{l,i}-c_i)|^p}{|(x_{l,m}-c_m)|^p}\right)^{1-p/p} \mathrm{sign}\{x_{l,m}-c_m\} \tag{6}$$

as $p \rightarrow \infty$. One can show that as p approaches ∞, the above term approaches one if $|x_{l,m}-c_m|$ is larger than the magnitudes of all other entries of the difference vector $X_l - C$. When $|x_{l,m}-c_m|$ is smaller than the magnitudes of at least one of the other entries of $X_l - C$, (6) approaches zero as p approaches ∞. Since the theorem assumes that there is always a single index that results in the maximum magnitude of $|x_{l,i}-c_i|$, we only need to consider the two cases discussed above. On the basis of the above discussion, we can express the partial derivative of J_∞ as

$$\frac{\partial}{\partial c_m} J_\infty = -\sum_{A_m} \mathrm{sign}\{x_{l,m}-c_m\} \tag{7}$$

where A_m denotes the set of values of 1 such that $|x_{l,m}-c_m|$ is equal to $\|X_l - C\|_\infty$. The value of c_m that makes the above expression zero is, indeed, the median of the set $\{x_{l,m}\colon l \in A_m\}$. This completes the proof of the theorem.

The theorem does not provide us with a direct method of calculating the centroid of a set of vectors, but it can be used to iteratively find the centroid as follows:

Initialize the centroid estimate with an arbitrary vector $\hat{C}_0$. Let $\hat{C}_i$ be the estimate of the centroid after the ith iteration. At the ith iteration, compute the L_∞ distortion between each vector X_l in the input set and $\hat{C}_{i-1}$. Partition the input vector set into K nonoverlapping subsets so that all input vectors for which the mth element determined the L_∞ distortion between $\hat{C}_{i-1}$ and the vectors are clustered into the mth subset. Let $\hat{c}_{i-1,m}$ denote the mth element of $\hat{C}_{i-1}$ and let $w_{i,m}$ represent the median of the mth elements of the vectors in the mth cluster. Then $\hat{c}_{i-1,m}$ is updated during the ith iteration as

$$\hat{c}_{i,m} = \hat{c}_{i-1,m} + \mu(w_{i,m} - \hat{c}_{i-1,m}) \tag{8}$$

where μ is a small positive constant. Experiments with images represented using 8-b integer pixels have shown that values of μ in the range 0.01 to 0.05 result in good convergence characteristics. The iterations can be terminated using an appropriate stopping criterion.

III. An Extension

We now consider a situation in which we desire to minimize the occurrences of quantization errors with magnitudes larger than a certain threshold. There are several applications in which such a criterion is useful. Image coding systems that employ perceptual threshold models require that the distortion introduced by the quantization process at any location in the image is below the predicted value of the perceptual threshold function at that location. Another application involves residual vector quantizers in which the residual vectors are coded only when the distortion in the vector exceeds a certain magnitude. By limiting the number of coded vectors with distortions above this threshold value at any stage, one can reduce the number of vectors that must be coded at subsequent stages.

We consider only the case of a constant threshold, even though a spatially varying threshold function is desirable in many applications. We modify the objective function for designing the codebook as follows to achieve the desired result. Given a training sequence $\{X_1, X_2, \cdots, X_L\}$, we desire to choose M codevectors $C_1, C_2, \cdots, C_M$ such that

$$J_\infty = \sum_{l=1}^{L} d_\infty(X_l, \hat{X}_l) \tag{9}$$

is minimized, where $\hat{X}_l$ is the code vector closest to X_l and $d_\infty(X,Y)$ represents a measure of the distortion between the vectors X and Y defined as

$$d_\infty(X,Y) = \begin{cases} 0; & \max_{i\in[1,K]} |x_i - y_i| < \tau \\ \max_{i\in[1,K]} |x_i - y_i| - \tau; & \text{otherwise.} \end{cases} \tag{10}$$

In the above equation, τ is a preselected threshold. A similar modification of the Euclidean cost function was used in [5]. Ideally, we would like to constrain all the distortions to be below this threshold. In situations where this is not possible, the system would attempt to design the codebook such that the occurrences and magnitudes of the distortions larger than τ are kept at the minimum possible level.

The following theorem characterizes the centroid of a set of vectors for the distortion measure in (10).

Theorem 2: Let $X_1, X_2, \cdots, X_N$ and C be as in Theorem 1 with the exception that the distortion measure is as defined in (10). Then the ith element of C satisfies the property that it is the median of the ith samples of all the vectors such that $\|X_l - C\|_\infty > \tau$ and $\|X_l - C\|_\infty = |x_{l,i} - c_i|$.

The proof is a straightforward extension of the proof of Theorem 1, and is therefore, omitted here. The codebook design algorithm of the previous section can be easily modified to suit the situation considered here.

IV. Experimental Results

We present the results of one experiment that compared the performances of a predictive vector quantizer (PVQ) system [6] that uses the L_∞ distortion measure with that of another PVQ system that attempts to keep the quantization errors below some threshold. The PVQ system uses the same predictor structure as described in [7]. All the coefficients of the predictor had integer-power-of-two values, and therefore, the complete system was realized using a multiplication-free architecture.

Our system employed vectors of size 16 pixels formed from blocks of size 4×4 pixels each. A codebook containing 32 vectors were designed using a training sequence obtained by generating the prediction error sequences for three different

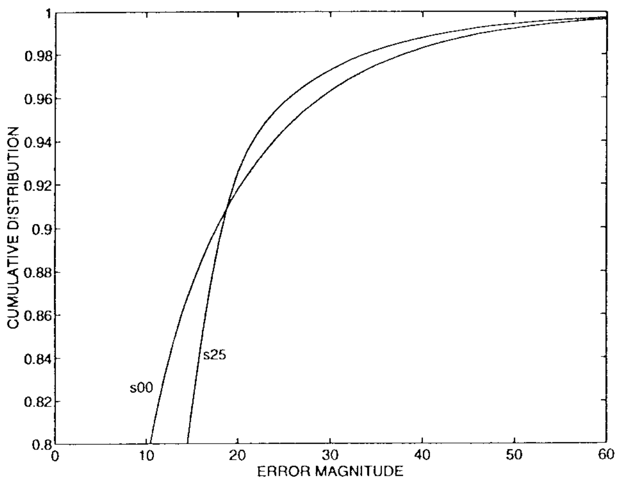

Fig. 1. CDF of the quantization errors.

monochrome images of size 512 × 512 pixels with 8-b per pixel resolution. The centroid calculations used a step size $\mu = 0.05$. We also created a different codebook of the same size for a vector quantizer system that attempts to minimize the occurrences of quantization errors with magnitudes larger than 25.

Fig. 1 displays the cumulative distribution functions (CDF) of the quantization errors estimated by computing the histogram of the error sequences when a fourth image that did not belong to the training set was quantized using the two codebooks. The curve marked "s00" corresponds to the system with no constraints imposed on the error sequences and the curve marked "s25" corresponds to the case for which the system attempted to reduce the occurrences of errors larger than 25 in magnitude. It can be clearly seen that the latter system succeeded in reducing the errors larger than 25 in this experiment. In fact, the number of pixels coded with error magnitudes larger than 25 was approximately 25% fewer for this case than the system that attempted to minimize the L_∞ distortion directly. All the experiments that we conducted produced similar results.

V. Concluding Remarks

This paper presented a method for designing codebooks for vector quantizers that employs the L_∞ distortion measure. The ideas were extended to an application in which it was desired to minimize the occurrences and magnitudes of the quantization errors larger than some preselected threshold. A large number of experiments have been conducted, and the results indicate that the technique that attempts to constrain the distortions was successful in reducing the occurrences of error magnitudes larger than the threshold. A considerable advantage the methods discussed in this paper have is their computational simplicity. We do not claim that the L_∞ distortion measure by itself provides better perceptual coding than L_2 or other similar distortion measures. Our experience with direct application of L_1, L_2, and L_∞ distortion measures is that the perceptual quality provided by all three measures is approximately the same. However, when combined with an appropriate perceptual threshold function, the L_∞ distortion measure and its variation presented in this paper should result in better perceptual quality. Some preliminary results in the design of distortion-limited vector quantizers using the distortion measure defined in (10) has been presented in [4]. The authors of this paper are at present studying these and other applications of vector quantizers that employ the L_∞ distortion measure.

References

[1] N. S. Jayant, J. D. Johnston, and R. J. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, pp. 1385–1424, Oct. 1993.

[2] K. S. Prashant, V. J. Mathews, and P. J. Hahn, "A new model of perceptual threshold functions for application in image compression systems," in *Proc. Data Compression Conf.*, Snowbird, UT, Mar. 1995, pp. 371–380.

[3] R. M. Gray, "Vector quantization," *IEEE Acoust., Speech, Signal Processing* vol. 32, pp. 4–29, Apr. 1984.

[4] P. J. Hahn and V. J. Mathews, "Distortion-limited vector quantization," in *Proc. Data Compression Conf.*, Snowbird, UT, Apr. 1996, pp. 340–348.

[5] R. Baseri and V. J. Mathews, "Vector quantization using visual masking functions," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, San Francisco, CA, Mar. 1993, pp. III-365–III-368

[6] V. Cuperman and A. Gersho, "Vector predictive coding of speech at 16 kbits/s," *IEEE Trans. Commun.*, vol. COM-33, pp. 685–696, July 1985.

[7] V. J. Mathews, "Multiplication free vector quantization using L_1 distortion measure and its variants," *IEEE Trans. Image Processing*, vol. 1, pp. 11–17, Jan. 1992.