Chapter 15
A Fast Image Encoding Algorithm Based on the Pyramid Structure of Codewords

Ahmed A. Radwan
Minia University, Egypt

Ahmed Swilem
Minia University, Egypt

Mamdouh M. Gomaa
Minia University, Egypt

ABSTRACT

This article presents a very simple and efficient algorithm for codeword search in the vector quantization encoding. This algorithm uses 2-pixel merging norm pyramid structure to speed up the closest codeword search process. The authors first derive a condition to eliminate unnecessary matching operations from the search procedure. Then, based on this elimination condition, a fast search algorithm is suggested. Simulation results show that, the proposed search algorithm reduces the encoding complexity while maintaining the same encoding quality as that of the full search algorithm. It is also found that the proposed algorithm outperforms the existing search algorithms.

INTRODUCTION

A standard vector quantization (VQ) is an efficient data compression technique that has been widely applied to image and speech coding (Gray, 1984; Linde, Buzo, & Gray, 1980). A vector quantizer of rate r bits/sample and dimension k is a mapping from a k-dimensional vector space R^h into some finite subset of it, $Y = \{y_i; i = 1, 2, \dots, N\}$ where, $N = 2^{kr}$. The subset Y is called a codebook and its elements are called codewords or reproducing vectors. In the conventional full search method to encode an input vector $x = (x_1, x_2, \dots, x_k)$, we have to find its distance from each of N codewords, and then compare these distances to find the best-matched codeword.

A complete description of a vector quantization process includes three phases, namely training,

DOI: 10.4018/978-1-60960-563-6.ch015

encoding and decoding. The original signal is first segmented into individual vectors. The training phase is a codebook design procedure which tries to find the best set of representatives from a large set of input vectors using clustering algorithms like the generalized lioyd algorithm (GLA) (Gray, 1984; Linde, Buzo, & Gray, 1980). The GLA consists of alternatively applying the nearest neighbor condition and centroid condition to generate a sequence of codebooks with decreasing average distortion. The algorithm operates as follows: It begins with an initial codebook of size N. It partitions the training set into N sets of vectors using the nearest neighbor rule which associates each training vector with one of the N codeword's in the codebook. This partitioning defines the vector encoder mapping $Q(.)$. Then, a new codebook is created with the new codeword's being the centroids of each of the N partition regions of input vectors. This new codebook defines the vector decoder mapping $\hat{Q}$ (.). The algorithm repeats these steps and calculates the average distortion for the new codebook over the entire training set at the end of each iteration. When the average distortion fails to decrease by a certain threshold amount, it is assumed that a local minimum in distortion has been reached and the algorithm terminates.

The encoding phase finds the best matched codeword for a test vector and uses the index of the codeword to represent it. A full codebook search could be used in an encoder of vector quantization to find the codeword which is the nearest neighbor to test vector. The decoding phase is simply a table look-up procedure which uses the received index to deduce the reproduction codeword. The particularly simple table look-up decoding procedure makes VQ an attractive method of data compression in practice. Both of the training and encoding phases are computation-intensive procedures. This limits the applicability of VQ.

To find the best-matched codevector, we need a matching criterion and a codebook search algorithm. The most popular matching criterion is the Euclidean distance. When the squared Euclidean distance is used as the distortion measure, the distance between x and $y_i = (y_{i1}, y_{i2}, \ldots, y_{ik})$ can be expressed as

$$d^2(x, y_i) = || x - y_i ||^2 = \sum_{j=1}^{k} (x_j - y_{ij})^2; i = 1, 2, ..., N, \tag{1}$$

where x is the current image block, y_i is the i^{th} codeword, j represents the j^{th} element of a vector, $k(= n \times n)$ is the vector dimension and N is the codebook size. Then a best-matched codeword with minimum distortion, which is called winner afterwards, can be determined straightforwardly by

$$d^2(x, y_w) = \min [d^2(x,y_i)]; i = 1, 2, ..., N \tag{2}$$

where y_w means winner and subscript "w" is the index of the winner. Once winner has been found, VQ only transmits the index of winner instead of the winner itself so as to reduce the amount of image data. Because the exact same codebook is also available at the receiving end, an image can be decoded by using the received winner index via inverse look-up table method easily and ultimately the winner itself is pasted to the corresponding position of the original image block to reconstruct it. Therefore, VQ features a very heavy encoding process due to a lot of distance computations and an extremely simple decoding process. VQ is asymmetric and the time-consuming encoding is a critical time bottleneck for its practical applications. Based on VQ property, it can be predicted that image compression in VQ method is very applicable to the broadcasting-type communication systems since they need a lot of low-cost decoding implementations at the receiving end.

From (1), each distance calculation requires k multiplications and $2k-1$ additions (and subtrac-