Image and Vision Computing 28 (2010) 1637–1644

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

A fast vector quantization encoding algorithm based on projection pyramid with Hadamard transformation

Ahmed Swilem *

Computer Science Department, Faculty of Science, Minia University, 61111, Egypt

ARTICLE INFO

Article history:
Received 6 March 2009
Received in revised form 2 February 2010
Accepted 11 May 2010

Keywords:
Hadamard transform
Projection pyramid
Vector quantization

ABSTRACT

Vector quantization (VQ) for image compression requires expensive time to find the closest codevector in the encoding process. In this paper, a fast search algorithm is proposed for projection pyramid vector quantization using a lighter modified distortion with Hadamard transform of the vector. The algorithm uses projection pyramids of the vectors and codevectors after applying Hadamard transform and one elimination criterion based on deviation characteristic values in the Hadamard transform domain to eliminate unlikely codevectors. Experimental results are presented on image block data. These results confirm the effectiveness of the proposed algorithm with the same quality of the image as the full search algorithm.

1. Introduction

Recent developments in multimedia and computer networks have resulted in widespread use of electronic file transmissions to replace traditional postal mail. However, due to the large size of multimedia data files and the bandwidth restrictions of computer networks, data transmission is inefficiency. Data can be compressed to reduce its size, improving the efficiency of its size, improving the efficiency of its transmission across computer networks. So far, vector quantization (VQ) [8,9,16] has long been a well-celebrated lossy compression technique that guarantees the achievement of a satisfactory balance between image quality and compression ratio [5]. At the same time, because of its simple and easy implementation, VQ has been very popular in a variety of research fields such as speech recognition and face detection [7]. Even in real-time video-based events detection [15] and the anomaly intrusion detection systems [34], VQ has been exploited recently to learn and collect some representative patterns and then to identify similar feature vectors or detect unusual activities.

VQ can be defined as a mapping Q from a k-dimensional Euclidean space R^k to a finite set $Y=\{y_1,y_2,...,y_N\}$ of vectors in R^k called the codebook. Each representative vector y_i in the codebook is called a codevector. Traditionally, VQ can be divided into three procedures: codebook design, encoding and decoding. The codebook design procedure is executed before the other two procedures for VQ. The goal of the codebook design is to construct a codebook Y from a set of training vectors using clustering algorithms like the generalized Lloyd algorithm (GLA) [16]. This codebook is used in both the image encoding/decoding procedures. In the encoding procedure, for each training vector x, the index i of the closest codevector to the vector x is found. The codevector y_i must give minimum distortion and satisfy $d^2(x,y_i)<d^2(x,y_j)$, $\{j=1,2,...,N;i\neq j\}$, where

$$d^2(x,y_i) = \sum_{j=1}^{k}(x_j-y_{ij})^2, \qquad (1)$$

so the codevector y_i now represents the vector x. The decoding procedure is simply a table look-up procedure that uses the received index i to deduce the reproduction codevector y_i, and then uses y_i to represent the input vector x.

Using the squared Euclidean distance criteria in Eq. (1), the computational cost of finding the best suitable codevector in encoding and codebook design imposes practical limits on the codebook size N and the vector dimension k. When N and/or k become larger, the computation complexity problem occurs for full codebook search. Researchers have proposed numerous fast search approaches to speed up codebook matching process, including standard VQ, tree structure VQ (TSVQ) [3,23–25], and lattice VQ (LVQ) [2,19,22,32]. The standard VQ algorithms can be classified into three classes. The first class uses one or more constraint inequality in the spatial domain, the second class exploits the topological structure of vectors and the last one utilizes a transformed constrain inequality.

There are many algorithms follow the first class as in the following: An algorithm for fast nearest neighbor search presented by Orchard [20] precomputes and stores the distance between each pair of codevectors. Given an input vector x, the current best codevector y_b, and a candidate codevector y_j, if $d(x,y_j)\leq d(x,y_b)$, then $d(y_j,y_b)\leq 2d(x,y_b)$. Graphically, this constrains the search area within a

* Tel.: +20862328790.
E-mail address: ahmed200eg@yahoo.com.

0262-8856/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2010.05.002

sphere centered on the current best codevector, with a radius of twice the smallest distortion calculated so far. The equal-average nearest neighbor search (ENNS) algorithm [10] uses the mean as a constraint value to reject impossible codevectors. The equal-average equal-variance nearest neighbor search (EENNS) [13] uses the mean and the variance as two individual inequalities to reduce the search area and reject the codevectors that are not contained in this area. The improved algorithm termed (IEENNS) [1] uses the mean and the variance in one inequality to reduce the search area. Wu and Lin [33] presented a new kick out condition based on the norms of codevectors. Two lossy design methods were described in [27,28] using a hyperplane partitioning rule. Lu and Sun [18] presented the equal-average equal-variance equal-norm nearest neighbor search (EEENNS) algorithm, which uses three significant features to reject many impossible codevectors. A fast codebook design algorithm for entropy constrained vector quantization was introduced in [29] using a new constraint called the angular constraint.

In the second class, the algorithms exploit the topological structure of codevector to avoid unnecessary codevector matching procedure. Lee and Chen [14] proposed a fast search algorithm based on the mean pyramid search (MPS) for codebook design using the squared Euclidean distance as distortion measure. This algorithm uses the mean pyramids of codevectors to reject many unmatched codevectors, thus drastically speeds up the search process in encoding and codebook design. Pan et al. [21] improved the encoding search process by adopting the variance pyramid in addition to the mean pyramid using the squared Euclidean distance as the distortion measure. The method uses a virtual distance between the input vector and the tested codevector at any level of the pyramid structure. This distance consists of the squared mean distance and the variance distance. Song and Ra [26] provided another technique using L_2-norm pyramid of codevectors. They used a modified distortion measure in the multilevel L_2-norm pyramid which is heavy from many multiplication operations. In [30], a high-speed closest search algorithm for VQ using the projection pyramid of the vectors was established. In this algorithm, a multilevel inequality for a simple and much lighter modified distortion measure was derived based on the pyramids of codevectors. By employing this inequality the procedures of codevector search for encoding or codebook design are speeded up. Another fast algorithm for vector quantization was introduced in [31] based on the mean pyramid and Hadamard transform using the squared Euclidean distance as distortion measure.

In the third class, there are also some transform domain-based algorithms, for example, wavelet transform based partial distortion search (WTPDS) algorithm [11] which uses wavelet transform and Hadamard transform based partial distortion search (HTPDS) [17]. Also, Jiang et al. [12] introduced another algorithm which uses Hadamard transform based on norm-ordered search (NOS). Another technique based on Hadamard transform was introduced in [6] which is efficient in the case of high dimensional.

In this paper, a high-speed closest codevector search algorithm for VQ is presented. The proposed algorithm is developed by combining the idea of the projection pyramid with the Hadamard transform to minimize the dimension of the vectors in the Hadamard projection pyramid levels. This minimization will reduce the computation cost required to construct the Hadamard projection pyramid. A new multilevel inequality is derived for a new lighter modified distortion based on the pyramid structure of the codevectors. By employing this inequality, many codevectors will be rejected. Hence, many distortion computations will be saved and the procedures of encoding can be speeded up.

The paper is organized as follows. In Section 2, some basic definitions and properties as a guide to understand our algorithm are introduced. In Section 3, some related previous work are presented and analyzed. In Section 4, the proposed algorithm is discussed in detail. The experimental results are given in Section 5. Section 6 concludes the paper.

2. Basic definitions and properties

In this section, we introduce some basic definitions and properties about Hadamard transform domain which is the basis of improved Hadamard transform based fast codevector search algorithm [6] and the pyramid data structure which is the basis of all pyramid search algorithms [4].

2.1. Hadamard transform

Hadamard transform (also known as the Walsh–Hadamard transform) H_m is a squared matrix of size $2^m \times 2^m$ and its elements is 1 or -1, with $H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ and $H_m = \begin{bmatrix} H_{m-1} & H_{m-1} \\ H_{m-1} & -H_{m-1} \end{bmatrix}$. For any vector x with k-dimension, where $k = 2^m$, $(m \geq 0)$, the Hadamard transformed vector $\hat{x}$, is given by:

$$\hat{x} = H_m x. \tag{2}$$

And the Hadamard transformed deviation of the vector $\hat{x}$ can be defined as:

$$V_{\hat{x}} = \sqrt{\sum_{i=2}^{k} \hat{x}_i^2}. \tag{3}$$

Note that Eq. (3) does not take the first component of the vector $\hat{x}$ into account as in [6].

From the definition of Hadamard transform, it is clear that the first component of the transformed vector $\hat{x}$ is equal to the sum of all components in the original vector x, because all values in the first row of Hadamard matrix are ones.

2.2. Pyramid data structure

Image pyramid data structure was originally developed for image coding by Burt and Adelson [4]. In this data structure, an image is represented hierarchically with each level corresponding to a reduced-resolution approximation. Given an image Z_n of size $2^n \times 2^n$, its pyramid can be defined as a sequence of matrices $\{Z_0, Z_1,..., Z_{r-1}, Z_r, Z_{r+1},..., Z_n\}$, where an image Z_{r-1} in level $r-1$ has a size of $2^{r-1} \times 2^{r-1}$ and is a half reduced-resolution version in both directions of Z_r. Note that Z_0 has only one pixel. A pyramid data structure can be formed by successively performing appropriate operations over 2×2 neighboring pixels in the next lower level. Therefore, the value of a pixel $z_{r-1}(i,j)$ in level $r-1$ is obtained from the values of the corresponding 2×2 neighboring pixels $z_r(2i-1, 2j-1)$, $z_r(2i, 2j-1)$, $z_r(2i-1, 2j)$ and $z_r(2i, 2j)$ in level r.

There are many types of image pyramids, for example, the mean pyramid [14] or the L_2-norm pyramid [26]. A general description of the pyramid data structure employing the idea of the projection is given by Eq. (4). The vector composed of corresponding values of 2×2 neighboring pixels is projected to a reference vector $\tilde{u}_1 = (1,1,1,1)$ in the 2×2-dimensional space. The pyramid pixel value $z_{r-1}(i,j)$ in level $r-1$ is:

$$\begin{aligned} z_{r-1}(i,j) &= \left[\frac{1}{\alpha}\sum_{g=0}^{1}\sum_{h=0}^{1}\left(z_r(2i-g, 2j-h)\right)^{\beta}\right]^{\frac{1}{\beta}} \\ &= \left[\frac{1}{\alpha}\,\tilde{u}_1\left(z_r^{\beta}(2i-1,2j-1), z_r^{\beta}(2i,2j-1), z_r^{\beta}(2i-1,2j), z_r^{\beta}(2i,2j)\right)^T\right]^{\frac{1}{\beta}}, \end{aligned} \tag{4}$$

where $(\cdot)^T$ denotes the vector transpose and α, β are two constants to define the kind of the projection and the pyramid structure. The mean pyramid structure [14] and the L_2-norm pyramid structure [26] are special cases of this pyramid structure. When $\alpha = 4$ and $\beta = 1$, the