VECTOR QUANTIZATION IN IMAGE SEQUENCE CODING

J. Huguet, L. Torres

Dept. of Signal Theory and Communications
E.T.S.I. Telecomunicació - UPC
Apdo 30.002 08080 BARCELONA, SPAIN

Abstract

A 3D Vector Quantization scheme and the subsequent results obtained in image sequence coding are presented as an efficient technique that exploits the correlation in space and time. A comparison between 2D and 3D vector quantizer performance in image sequence coding is given.

I. Introduction

The goal of a good coding technique is to reduce the information as much as possible, in order to be transmitted or stored efficiently. This is accomplished by removing the information redundancy contained in the original signal and sending or storing only the most significant parts of the image information. Among the great number of available techniques to extract the relevant part of the information signal, vector quantization (VQ) has proved to be very effective. Vector Quantization was used first with speech signal (1D-signal) and the good results observed led to extend its application to image signal (2D-signal). In a further stage it was used with image sequence signal (3D-signal). A brief revision of these advances in vector quantization and some comparative results are treated here.

II. Vector Quantization

In contrast to scalar quantization, vector quantization performs the quantization not treating each sample individually but previously grouping a number of consecutive samples forming a vector and then quantizing this vector. Let X be an input vector from the original signal and let Y represent a codeword or reproduction vector. The codebook contains all the available codewords and the quantizer replaces the input vector by the codeword that is more similar to X. Using the euclidean distance as a distortion measure, the quantizer searches for the codeword located at the minimum distance from the input vector. The quantization process is usually termed \( Y = Q(X) \). In a mathematical sense, the quantizers establishes a correspondence between a N-dimensional space \( V \) to a finite codebook \( W = \{ Y_1, \ldots, Y_M \} \).

This Paper has been supported by PRONTIC 105/88.
IV. Vector Quantization in Image Sequence Coding

The first intuitive way to use vector quantization in image sequence coding is to use a 2D vector quantizer, where each image of the sequence is quantized individually and independently of previous or future images of the sequences.

A closer look at the problem reveals that only the spatial correlation is exploited in a 2D vector quantizer. In case we want to exploit the spatial and temporal correlation contained in an image sequence, a 3D vector quantizer should be used. The image sequence, considered as a 3D signal, is divided into 3D cubes that will be the quantization units. Each cube contains the pixels within a small region in the image (spatially correlated) and the pixels of the next images of the sequence (time correlated). Thus it verifies that all pixels grouped within a cube are correlated.

Figure 2. Three dimensional cube from an image sequence in a 3D vector quantizer.

In order to use the LBG algorithm, we need to introduce a new measure of distortion between cubes. A simple square error has been chosen, calculated as follows for a MxNxK cube size

\[ d(X,Y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{k=0}^{K-1} (x_{mnk} - y_{mnk})^2 \]

where X is a cube of the image sequence and x_{mnk} a pixel within the cube. Similarly, Y represents a codeword cube of the codebook and y_{mnk} a pixel of that codeword. The distortion per pixel that will be used as a comparison measure between quantizers is computed as

\[ \text{distortion per pixel} = \frac{d(X,Y)}{(M \times N \times K)} \]

In the coding process, a full search is carried out for each cube implying the calculation of the distance to all the codewords. Full search is highly processor consuming when the codebook is large, but gives the best (optimum) results in the square error sense. In order to reduce the computational cost of the full search algorithm the fast codebook search algorithm described in [2] has been used, reducing the encoding complexity to less than the half when using a square distortion measure. This algorithm discards a bad codeword with very few operations without calculating the euclidean distance which requires N multiplications and 2N-1 additions / subtractions. A property of this algorithm is that it never discards a codeword whose distance is less than that of the best codeword tested so far. The final result is then the same as if all euclidean distances had been calculated and compared.

On the other hand, a tree search algorithm or a multistage structure [3] can be used to accelerate the encoding process at the cost of some quality degradation and additional complexity. In fact, many of the more sophisticated schemes used in 1D and 2D vector quantizer can also be applied to a 3D-VQ scheme because the underlying idea of the algorithms remains the same.

V Results

We have used both a 3D vector quantizer to code an image sequence named "ma" (that contains 24 images) for three different cube sizes and a fixed number of codewords chosen to be 1024, using the full search algorithm. The images were 256x256 pixels and each pixel ranged from 0 to 255 (gray level). The results obtained with a 3D and 2D vector quantizer are presented in tables I and II, while Fig. 3 contains a quantized image of each example (we just show one image of the sequence as it is impossible to include the whole sequence, and even that would not help to completely figure out the quality of image sequence seen in real time). Thus for a 4x4x2 pixel cube size and 4096 codewords gives a 0.375 bits/pixel with a mean distortion per pixel of 9. The same bit rate in a 2D vector quantizer for a 4x4 pixels block size (the same spatial size) limits the codebook size to 64 codewords and yields a distortion of 30. Notice that the first case of table I and II are in fact the same and so are their results. The codebook was generated in all cases from the same long sequence.

The (S/N) of tables I and II is taken as the peak signal to noise ratio, computed as

\[ (S/N) = 10 \log (255^2 / \text{dist. per pixel}) \]
The images of Fig. 3 clearly reveal that the quality of the image sequences quantized with the 3D vector quantizer are significantly better than those quantized with the 2D vector quantizer for the same bit rate. In the 2D case, the effect of reducing the bit rate drastically reduces the codebook size, with a great influence on the final quality of the images.

VI Conclusions

These results confirm that a 3D-VQ scheme gives better quality images than a 2D-VQ, specially when low bit rates are desired. In other words, additional redundancy is removed when grouping pixels in three dimensional blocks. It is obvious that a 2D vector quantizer is a particular case of a 3D vector quantizer (the first case of table I shows that). On the other hand, as the codebook is larger the computational cost required in the full search encoding process is also greater. That is why a great number of fast encoding algorithms for vector quantization have grown up in the last years.

References

"An algorithm for vector quantizer Design"  
IEEE Transactions on Communication.  

"Fast codebook search algorithm in vector quantization". Submitted for publication to Vision  
Communications and Image Processing'90. Lausanne. 2-4 October 1990

"Vector Quantization"  

"Image sequence using vector quantization".  
<table>
<thead>
<tr>
<th>Cube Size</th>
<th>Number of codewords</th>
<th>Bit Rate (bits/pixel)</th>
<th>Distortion per pixel</th>
<th>(S/N) dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>4x4x1</td>
<td>4096</td>
<td>0.75</td>
<td>8</td>
<td>39.10</td>
</tr>
<tr>
<td>4x4x2</td>
<td>4096</td>
<td>0.375</td>
<td>9</td>
<td>38.58</td>
</tr>
<tr>
<td>4x4x4</td>
<td>4096</td>
<td>0.18</td>
<td>10</td>
<td>38.13</td>
</tr>
</tbody>
</table>

**Table 1.** Results obtained with a 3D Vector Quantizer with the image sequence "ma".

<table>
<thead>
<tr>
<th>Cube Size</th>
<th>Number of codewords</th>
<th>Bit Rate (bits/pixel)</th>
<th>Distortion per pixel</th>
<th>(S/N) dB</th>
</tr>
</thead>
<tbody>
<tr>
<td>4x4</td>
<td>4096</td>
<td>0.75</td>
<td>8</td>
<td>39.10</td>
</tr>
<tr>
<td>4x4</td>
<td>64</td>
<td>0.375</td>
<td>30</td>
<td>33.35</td>
</tr>
<tr>
<td>4x4</td>
<td>8</td>
<td>0.18</td>
<td>67</td>
<td>29.87</td>
</tr>
</tbody>
</table>

**Table 1.** Results obtained with a 2D Vector Quantizer with the image sequence "ma" for the same bit-rate of table 1 (the number of codewords is modified accordingly).

**Figure 3.** Single image of the sequence quantized with different block (or cube) size when using a 2D (or 3D) vector quantizer scheme. (a) Original image of the sequence. (b) Cube of 4x4x1 pixels. (c) Cube of 4x4x2 pixels. (d) Cube of 4x4x4 pixels. (e) Block of 4x4 pixels and 64 codewords. (f) Block of 4x4 and 8 codewords.