Performing Vector Quantization Using Reduced Data Representation

Erickson Miranda, Guoqiang Shan, and Vasileios Megalooikonomou
Temple University

We propose an improvement to VQ by applying data and dimensionality reduction techniques on the original dataset and using different resolutions generated by the reduction techniques as input to the GLA with splitting at different stages of codebook generation. The different resolutions are pre-computed (a one time cost), and are used instead of the original dataset in each iteration of the GLA. We propose three variants of the reduced data representation VQ (RDR-VQ) based on the three reduced data representation techniques we use:

1. Reduced precision VQ (RP-VQ), where we use a simple rounding off of precision of the original dataset;
2. Wavelet RDR-VQ, where we use a discrete wavelet transform of the original dataset
3. PCA RDR-VQ, where we use a linear transform of the original dataset using PCA

The difference in our work with existing algorithms that use VQ with wavelets [1,2] is while these algorithms use multiple quantizers on the wavelet subbands, our method uses the different subbands as input to different iterations of a single quantizer.

Different resolutions of the dataset are used as input for the different Lloyd iterations. An important consideration is choosing an appropriate resolution level for each of these iterations. We call the different resolution levels used by the algorithm to compute the final codebook a resolution sequence.

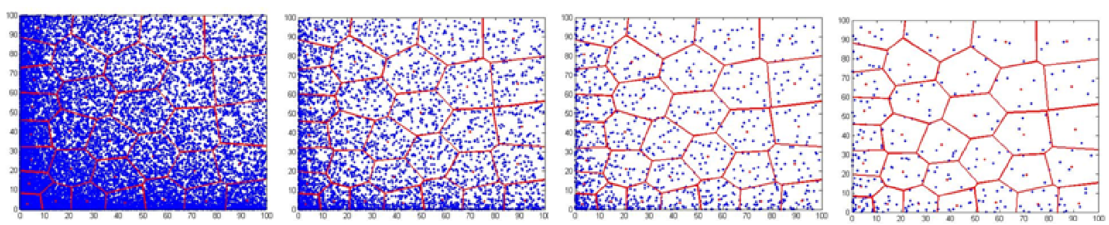

Figure 1. A 2D dataset with different resolutions (wavelet decomposition levels) shown with the code words.

Instead of applying the GLA directly, our algorithm first preprocesses the dataset, generates lower resolution representations of the dataset and chooses a resolution sequence. Then the GLA with splitting is applied on the different resolution levels of the data (given in the order provided by the resolution sequence) used as input instead of the original dataset.

Experimental results on both simulated datasets using three different random distributions and real data in the form of a combination of 20 color and grayscale images show that the proposed technique outperforms ordinary vector quantization in terms of mean squared error or running time. Moreover, other techniques for improving VQ can be combined with our proposed algorithm to further improve performance. For future work, the contribution of the resolution sequence in the performance of the quantizer needs to be further investigated.

This project was supported by National Science Foundation (NSF) research grants IIS-0237921 and IIS-0705215. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

[1] M. Effros, D. Dugatkin, "Multiresolution Vector Quantization", *IEEE Trans. on Information Theory* 50 (12), 2004, pp. 3130-3145.

[2] S. Yang, S. Mitra, "Wavelet Based Adaptive Vector Quantization for Encoding Large Color Images", *11th IEEE Symposium on Computer-Based Medical Systems* 208 (1998).