Vector Permutation Encoding for the Uniform Sources

W. A. Finamore [1], S. V. B. Bruno, and D. Silva

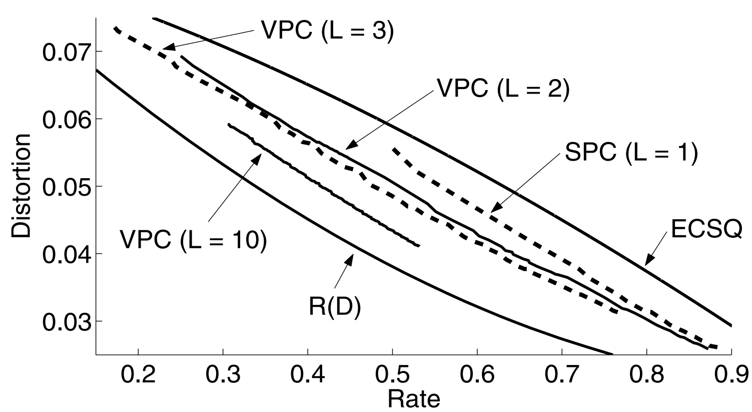
Pontifícia Universidade Católica do Rio de Janeiro — Brazil

[1] *e-mail: weiler@cetuc.puc-rio.br*

This paper extends the idea of a scalar permutation code (SPC) [1] to a higher dimensional case. A vector permutation code (VPC) is defined as a fixed-rate vector quantizer with a codebook $\mathcal{C}[\mathbf{y}_0] = \{\mathbf{y}_\ell\}_{\ell=0}^{M-1}$ consisting of all the distinct permutations of the columns of an $L \times n$ matrix $\mathbf{y}_0 = (\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_2, \ldots, \boldsymbol{\mu}_K, \ldots, \boldsymbol{\mu}_K)$, where each L-dimensional column-vector $\boldsymbol{\mu}_i$ appears n_i times. The size of the codebook is $M = \frac{n!}{\prod_{i=1}^{K} n_i!}$ and the rate of the code is $R = \frac{1}{nL} \log M$ bits/source-symbol. The optimal encoder reproduces a source output $\mathbf{x}$ by the permutation of $\mathbf{y}_0$ that best match $\mathbf{x}$, i.e., by the matrix $O(\mathbf{y}_0, \mathbf{x}) = \arg\min_{\mathbf{y} \in \mathcal{C}[\mathbf{y}_0]} d(\mathbf{x}, \mathbf{y})$, where $d(\mathbf{x}, \mathbf{y})$ is the distortion between $\mathbf{x}$ and $\mathbf{y}$. The average per-symbol distortion is thus $D = \frac{1}{nL} E[d(\mathbf{X}, O(\mathbf{y}_0, \mathbf{X})]$.

If the source can be characterized by a training set $\mathcal{T} = \{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(t)}, \ldots, \mathbf{x}^{(T)}\}$ formed by $L \times n$ matrixes, and the *composition vector* $\{n_i\}_{i=1}^{K}$ is fixed, a Lloyd-like algorithm can be used for (locally) optimum codebook design: (1) initialize $\mathbf{y}_0$ and set $S_i = \sum_{j=1}^{i} n_j$; (2) calculate $\overline{\mathbf{z}} = (\overline{\mathbf{z}}_1 \ \ldots \ \overline{\mathbf{z}}_n) = \frac{1}{T} \sum_{t=1}^{T} O(\mathbf{x}^{(t)}; \mathbf{y}_0)$, $\boldsymbol{\mu}_i = \frac{1}{n_i} \sum_{j=1+S_{i-1}}^{S_i} \overline{\mathbf{z}}_j$, and update $\mathbf{y}_0$; (3) repeat step 2 until $\mathbf{y}_0$ converges.

When $L \geq 2$, the matching operation for VPC has complexity $O(n^2 \sqrt{n} \log n)$ [2]; in spite of that, increasing the dimension provides an improvement of performance, as can be seen in the figure for a memoryless uniform source with a squared error criterion. In a related paper [3], we show that, as with SPC and ECSQ, the performance of VPC tends to that of ECVQ as blocklength $n \to \infty$. Based on this connection we also propose [3] an algorithm for *composition vector* design.



This work was partially supported by the Brazilian National Research Council (CNPq).

[1] T. Berger, F. Jelinek, and, J. K. Wolf, "Permutation Codes for Sources," *IEEE Trans. Inform. Theory*, vol.IT-18, pp.160-169, Jan. 1972.

[2] A. V. Goldberg, and R. Kennedy, "An Efficient Cost Scaling Algorithm for the assignment Problem," Tech Report CS Department, Stanford University, Stanford, CA, May 1995.

[3] W. A. Finamore, D. Silva, and S. V. B. Bruno, "Vector Permutation Code Design Algorithm", submitted to the *International Symposium on Information Theory*, Chicago, IL, 2004.

Proceedings of the Data Compression Conference (DCC'04)
1068-0314/04 $ 20.00 © 2004 IEEE