# NOVEL VQ WITH CONSTRAINTS ON THE QUANTIZATION ERROR DISTRIBUTION

*Joachim Schenk and Frank Wallhoff and Gerhard Rigoll*

Institute for Human-Machine Communication
Department of Electrical Engeneering and Information Technology
Technische Universität München
Theresienstraße 90, 80333 München
{schenk,wallhoff,rigoll}@mmk.ei.tum.de

**ABSTRACT**

In this paper, we motivate and introduce a novel vector quantization (VQ) scheme for distributing the quantization error among the quantized features of a continuous feature vector in a predefined manner. This is done by defining ratios between the individual quantization errors of the features and shaping the Voronoi cells accordingly.

In a series of experiments we show that the novel approach is capable of either distributing the quantization error equally among the dimensions or realize an arbitrary distribution.

***Index Terms*—** vector quantization, quantization error, error shaping, Voronoi cells

## 1. INTRODUCTION

Vector quantization (VQ [1, 2]), i. e. the mapping of a sequence of continuous vectors to a discrete, one-dimensional sequence, is widely used in data compression [3], signal processing, and pattern recognition [4]. In speech coding, quantization error shaping in the spectral domain is performed [5] such that each spectral component of the signal contributes to the overall distortion equally [6].

The quantization error shaping presented here targets a different problem. In machine learning, Hidden Markov Models (HMMs [4]) are used for classification of dynamic sequences, distinguishing between continuous and discrete HMMs. While in continuous HMMs the distribution of the data is modeled by Gaussian mixtures, in discrete HMMs this distribution is modeled by the relative frequency of the vector quantized continuous data, outperforming continuous HMMs in terms of processing speed [4]. In [7], we showed for the features used in on-line handwriting recognition (HWR) of whiteboard notes (see [8]) that the quantization error is not distributed equally among the quantized features. Hence, the contribution of the features to the quantization process varies. In this paper, we present a novel VQ scheme which is capable of distributing the quantization error in a predefined manner (e. g. equally) among the dimensions.

The next section gives a brief description of the HWR system. Then, VQ is reviewed and summarized in Sec. 3. Section 4 introduces our novel VQ scheme, which is analyzed in the experimental section (Sec. 5). Conclusions and an outlook are given in Sec. 6.

## 2. HANDWRITING RECOGNITION SYSTEM

This section summarizes the on-line HWR system and its features used for the experiments in Sec. 5. Further details can be found in [7].

The handwritten data, which is recorded with the EBEAM-System and represented by sample vectors $\mathbf{s}(t)$, is heuristically segmented into lines [8]. Then, preprocessing and normalization is performed, and features are extracted from the sample vector and form a 24-dimensional feature vector $\mathbf{f}(t) = (f_1(t), \ldots, f_{24}(t))$. The features listed below can be divided into two groups: *on-line* and *off-line* features. While the continuous on-line features are derived from the pen's trajectory, the discrete off-line features evaluate a bitmap gained by binarization of the handwritten script. The on-line features are:

$f_1$ : indicating the pen "pressure", i. e. $f_1 = 1$ if the pen tip is placed on the whiteboard and $f_1 = 0$ otherwise

$f_2$ : velocity equivalent

$f_{3,4}$ : $x$-and $y$-coordinate (high pass filtered)

$f_{5,6}$ : angle $\alpha$ of spatially resampled and normalized strokes (coded as $\sin\ \alpha$ and $\cos\ \alpha$, "writing direction")

$f_{7,8}$ : difference of consecutive angles $\Delta\alpha = \alpha(t) - \alpha(t-1)$ (coded as $\sin\ \Delta\alpha$ and $\cos\ \Delta\alpha$, "curvature")

$f_9$ : logarithmized aspect $v$ of the trajectory between the points $\mathbf{s}(t-\tau)$ and $\mathbf{s}(t)$, whereby $\tau < t$ denotes the $\tau^{\text{th}}$ sample point before $\mathbf{s}(t)$: $f_9 = \text{sign}(v) \cdot \lg(1 + |v|)$, where $\lg(\cdot) = \log_{10}(\cdot)$

$f_{10,11}$ : angle $\varphi$ between the line $[\mathbf{s}(t-\tau), \mathbf{s}(t)]$ and lower line (coded as $\sin\ \varphi$ and $\cos\ \varphi$, "vicinity slope")

$f_{12}$ : the length of the pen-trajectory between the sample points $\mathbf{s}(t-\tau)$ and $\mathbf{s}(t)$, normalized by the $\max(|\Delta x|; |\Delta y|)$ ("vicinity curliness")

$f_{13}$ : average square distance to each point and the line $[\mathbf{s}(t-\tau), \mathbf{s}(t)]$

The off-line features are:

$f_{14-22}$ : a $3 \times 3$ subsampled bitmap slid along the pen-trajectory ("context map") to incorporate a $30 \times 30$ partition of the currently written letter's actual image

$f_{23,24}$ : number of pixels above respectively beneath the current sample point $\mathbf{s}(t)$ (the "ascenders" and "descenders")

As the values of the features may vary in different ranges, each dimension $d$ of the feature vector is normalized to a mean of $\mu_d = 0$ and variance of $\text{var}_d = 1$. After feature extraction, the handwritten data is recognized by a discrete HMM-based recognizer [7]. Combined segmentation and classification is provided by the well-known Viterbi-Algorithm [9]

## 3. QUANTIZATION

As mentioned in the introduction, quantization denotes the mapping of a sequence $\mathbf{F} = (\mathbf{f}(1), \ldots, \mathbf{f}(T))$ of $T$ continuous, $D$-dimensional vectors $\mathbf{f}(t) \in \mathbb{R}^D$ to a discrete, one-dimensional sequence $\hat{\mathbf{f}} = (f(1), \ldots, f(T))$ of codebook indices $\hat{f}(t) \in \mathbb{N}$ in the codebook $\mathbf{C} = (\mathbf{c}(1), \ldots \mathbf{c}(N_{\text{cdb}}))$ containing $|\mathbf{C}| = N_{\text{cdb}}$ centroids $\mathbf{c}(i) \in \mathbb{R}^D$. For $D = 1$, this mapping is called *scalar*, in all other

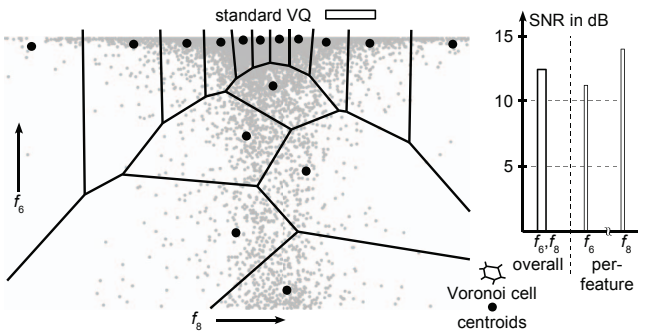


**Fig. 1**. Voronoi cells and centroids for joint VQ of the features $f_6$ and $f_9$ (left), overall and (unevenly distributed) per-feature SNR (right).

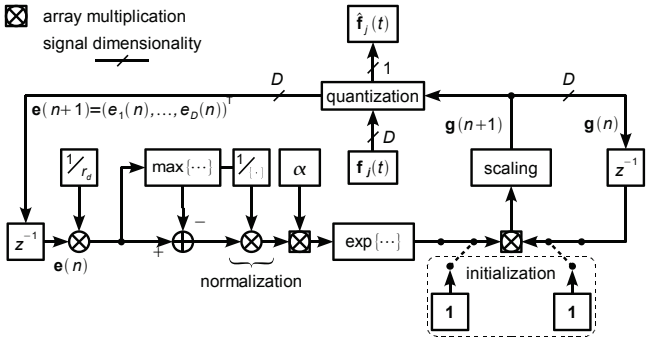


**Fig. 2**. VQ post processing depicted as control loop for achieving an equally distributed quantization error.

cases $D \geq 2$ *vector* quantization (VQ). The codebook $\mathbf{C}$ and its entries $\mathbf{c}(i)$ are derived from a training set $\mathcal{S}_\mathrm{t}$ with $|\mathcal{S}_\mathrm{t}| = N_\mathrm{t}$ training sequences $\mathbf{F}_j$, by partitioning the $D$-dimensional feature space defined by $\mathcal{S}_\mathrm{t}$ into $N_\mathrm{cdb}$ Voronoi cells $V_i$ represented by the centroids $\mathbf{c}(i)$ [1]. Herein, this is performed by the well-known $k$-Means algorithm as described e. g. in [2]. The $N_\mathrm{cdb} = 16$ Voronoi cells partitioning the space spanned by the two features $f_6$ and $f_9$ (see Sec. 2) and their corresponding centroids are shown in Fig. 1 (left).

Once a codebook $\mathbf{C}$ is generated, the assignment of the continuous sequence to the codebook entries is a minimum distance search

$$\hat{f}(t) = \underset{1 \leq k \leq N_\mathrm{cdb}}{\operatorname{argmin}} \; d(\mathbf{f}(t), \mathbf{c}(k)). \tag{1}$$

The quality of the VQ is measured by its distortion. In this paper, the signal-to-noise ratio (SNR) is used:

$$\mathrm{SNR} = 10\mathrm{lg}\frac{\bar{S}}{\bar{E}} = 10\mathrm{lg}\frac{\sum_{j=1}^{N_\mathrm{t}}\sum_{t=1}^{T_j}||\mathbf{f}_j(t)||^2}{\sum_{j=1}^{N_\mathrm{t}}\sum_{t=1}^{T_j}||\mathbf{f}_j(t)-\mathbf{c}(\hat{f}_j(t))||^2} \tag{2}$$

with $\bar{S}$ the average, square signal amplitude, $\bar{E}$ the average, square quantization error of all observations $\mathbf{F}_j$, and $\mathbf{f}_j(t)$ the $t^\mathrm{th}$ of $T_j$ feature vectors in the $j^\mathrm{th}$ of $N_\mathrm{t}$ sentences in the training set. Hence, the SNR is the average signal energy normalized by the distortion on a logarithmic scale [2]. As mentioned in Sec. 2, each dimension of the continuous feature vector, and therefore each feature, is normalized by its mean and variance value, yielding an average square signal amplitude of $\bar{s}_d = 1$ in each dimension. The SNR can then be expressed by the average square quantization error $\bar{e}_d$ of each feature:

$$\mathrm{SNR} = 10\mathrm{lg}\frac{\bar{S}}{\bar{E}} = 10 \cdot \left[\mathrm{lg}D - \mathrm{lg}\left(\sum_{d=1}^{D}\bar{e}_d\right)\right] \tag{3}$$

with $\bar{e}_d = \sum_{j=1}^{N_\mathrm{t}}\sum_{t=1}^{T_j}(f_{j,d}(t) - c_{\hat{f}_j,d}(t))^2$. The overall as well as the per-dimension SNR when quantizing the features $f_6$ and $f_9$ with the centroids $\mathbf{c}$ in Fig. 1 (left) is shown in Fig. 1 (right). As can be seen, although normalized, the per-feature SNR and hence, the quantization error are not equal. This has also been shown for different vector quantizers and a higher number of features in [7].

## 4. VORONOI CELL SHAPING

As pointed out in Sec. 3, the quantization error introduced by the quantization is not distributed equally among the dimensions. This section describes our approach for distributing the quantization error in a predefined manner among the dimensions, which consists of two stages: the centroid computation as described in Sec. 3 and a shaping of the Voronoi cells in order to achieve a distinct distribution of the quantization error.

### 4.1. Preliminaries

The vector $\mathbf{r} = (r_1, \ldots, r_D)^\mathrm{T}$ is introduced, which contains coefficients $r_d$ corresponding to the features $f_d$. The goal of our VQ is to provide average per-dimension quantization errors $\bar{e}_d$ with

$$\bar{e}_1/r_1 = \bar{e}_2/r_2 = \ldots = \bar{e}_D/r_D \tag{4}$$

by shaping the Voronoi cells utilizing the distance measure

$$d(\mathbf{f}_j(t), \mathbf{c}(k)) = (\mathbf{f}_j(t) - \mathbf{c}(k))^\mathrm{T} \cdot \mathbf{G} \cdot (\mathbf{f}_j(t) - \mathbf{c}(k)) \tag{5}$$

with $\mathbf{G}$ a diagonal weight-matrix containing the weights $g_d$ of the features $f_d$, in Eq. 1. By selecting the weights $g_d$, the average quantization error $\bar{e}_d$ of the feature $f_d$ can be influenced. To show this property, the following relations between the weights $g_d$ are assumed:

$$g_1 = xg_2 = x^2g_3 = \ldots = x^{D-1}g_D = x^{D-1}g, \quad x, g > 1. \tag{6}$$

The feature $\mathbf{f}_j(t)$ is then assigned to $\mathbf{c}(k)$ instead of $\mathbf{c}(k')$ if

$$d(\mathbf{f}_j(t), \mathbf{c}(k)) < d(\mathbf{f}_j(t), \mathbf{c}(k')) \Rightarrow \tag{7}$$
$$\sum_{d=1}^{D}(f_{j,d}(t) - c_d(k))^2 x^{D-d} g < \sum_{d=1}^{D}(f_{j,d}(t) - c_d(k'))^2 x^{D-d} g.$$

Dividing Eq. 7 by $x^{D-1} \cdot g > 1$ yields

$$\begin{aligned}(f_{j,1}(t) - c_1(k))^2 + \ldots + \frac{(f_{j,D}(t) - c_D(k))^2}{x^{D-1}} < \\ (f_{j,1}(t) - c_1(k'))^2 + \ldots + \frac{(f_{j,D}(t) - c_D(k'))^2}{x^{D-1}}.\end{aligned} \tag{8}$$

As can be seen from Eq. 8, choosing $x = 1$ lets each feature contribute to the overall distortion $d(\mathbf{f}_j(t), \mathbf{c}(k))$ by its actual distance to the corresponding centroid dimension $c_d(k)$. However, when raising the value of $x$ the contribution of higher numbered features decays. Finally, when choosing $x \to \infty$ only the distance of the first feature contributes to the distortion and is therefore minimized.

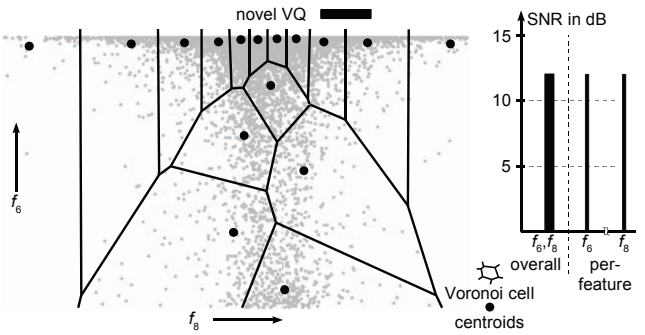


**Fig. 3**. Shaped Voronoi cells around the same centroids as in Fig. 1 (left), overall and (evenly distributed) per-feature SNR (right).

### 4.2. Weight Estimation

The analytic relation between the actual quantization error $\bar{e}_d$ of each feature $f_d$ and the corresponding weight $g_d$ is unknown. Hence, a control loop is used for recursively fitting the weights $g_d$, $1 \leq d \leq D$ to achieve an error distribution as defined by Eq. 4. The weight $g_{d_{\max}}$ of the feature $f_{\max}$ that differs most from the distribution defined by Eq. 4, $e_{d_{\max}}$ is not changed; all other weights are lowered by a factor depending on $\bar{e}_{d_{\max}} - \bar{e}_d > 0$. After an initialization with $g_d(0) = 1/D$ the weights $g_d(n)$ are recursively updated:

$$\tilde{g}_d(n+1) = g_d(n) \cdot \exp\left[\alpha \cdot \frac{\bar{e}_d(n)/r_d - \max_{1\leq\delta\leq D} \bar{e}_\delta(n)/r_\delta}{\max_{1\leq\delta\leq D} \bar{e}_\delta(n)/r_\delta}\right], \quad (9)$$

with $1 \leq d \leq D$ and an experimentally chosen step size $\alpha$. The normalization by $\max(\cdot)$ in Eq. 9 is necessary, due to the variation in the absolute value of the quantization errors. In order to prevent an infinite growth of the updated weights $\tilde{g}_d(n+1)$ their values are normalized such that they meet $\sum_{d=1}^{D} g_d(n+1) = 1$:

$$g_d(n+1) = \frac{\tilde{g}_d(n+1)}{\sum_{\delta=1}^{D} \tilde{g}_\delta(n+1)}. \quad (10)$$

The weight adaptation is continued until the change in the individual quantization errors $\bar{e}_d$ falls below a threshold. Fig. 2 shows the control loop in order to find the desired weighting values $g_d$. The result of applying the proposed VQ scheme on the centroid and feature distribution as depicted in Fig. 1 is shown in Fig. 3: after shaping the Voronoi cells, as is shown on the left hand side, the SNR of both features is the same. However, this is accompanied by a slight drop in the overall SNR.

## 5. EXPERIMENTS

The experiments presented in this section are conducted on the IAM-OnDB database, containing handwritten, heuristically line-segmented whiteboard notes [10]. The training set defined in the IAM-onDB-t1 benchmark (see [11]), contains $N_{\text{t}} = 5364$ text lines written on a whiteboard. On these text lines, preprocessing, normalization, and feature extraction is performed as explained in Sec. 2. The derived feature vectors $\mathbf{f}(t)$, containing the features $f_1, \ldots, f_{24}$ are vector quantized with and without our error shaping method using a varying number of codebook entries ($N_{\text{cdb}} = 10, 100, 1000$) in the following experiments.

Experiment 1 (*Exp. 1*): In the first experiment, all 24 features are first vector quantized with a standard $k$-Means VQ. Then, our novel VQ scheme using $r_1 = \ldots = r_{24} = 1$ is applied. Hence, the quantization errors and, due to the normalization, the SNR of each feature is required to be the same. The result is shown in Fig. 4. While regardless of the number of codebook entries the SNR is distributed unevenly among the dimensions in case of standard VQ, our novel approach yields an approximately even distribution except for the feature $f_1$. This can be explained with the binary nature of the first feature. In [7], we therefore suggest a special treatment when quantizing this feature. Our novel VQ scheme lowers the SNR of some features, while raising the SNR of others at the same time, yielding an equal contribution of each feature to the quantization process, which is accompanied by a slight drop in overall quantization performance.

Experiment 2 (*Exp. 2*): In the second and third experiment, the findings of *Exp. 1* are further investigated. The feature set is therefore narrowed to $\mathcal{X} = \{f_4, f_6, f_7, f_{14}, f_{17}, f_{18}, f_{21}\}$, containing three continuous on-line features and four discrete off-line features. The result, when applying both the standard VQ and our novel VQ approach with $r_4 = \ldots = r_{21} = 1$, is shown in Fig. 5 (top). Again, it can be observed that the SNR is distributed evenly among the dimension after proper Voronoi cell shaping. All features therefore contribute equally to the quantization process.

Experiment 3 (*Exp. 3*): In the last experiment, it is shown that our novel VQ approach is also capable of distributing the quantization error among the dimensions in an arbitrary manner: in two quantization scenarios $r_4 = 1; r_6 = \ldots = f_{21} = 4$ and $f_4 = f_6 = f_8 = \ldots = f_{21} = 4; f_7 = 1$ is chosen respectively. Following Eq. 2, this translates to an increase of the SNR value of $\Delta = 6\,\text{dB}$ for the features $f_4$ and $f_6$, respectively, when compared to the other features, which all share the same SNR value. This assumption is confirmed by the results shown in Fig. 5 (middle, bottom). As can be seen, the SNR-value of the selected feature is increased given the remaining features. Again, a slight drop in overall quantization performance can be observed.

## 6. CONCLUSION AND OUTLOOK

In this paper, we introduced a novel VQ scheme realizing constraints on the distribution of the quantization error. The constraints are defined by describing the ratio between the per-feature average quantization error after quantization. The implementation consists of two steps: first the centroids are found; then, their corresponding Voronoi cells are shaped accordingly. In a series of experiments we applied our novel VQ scheme on the features used in on-line HWR of whiteboard notes. The general purpose of our VQ scheme could be shown, however revealing a draw back: while the constraints are satisfactorily met, the overall quantization performance drops.

In future work, we therefore plan to combine the two stage approach presented here into an one stage VQ, i. e. the centroids are found and the Voronoi cells are shaped together. We also plan to utilize the here presented VQ to perform feature selection in combination with VQ: it has to be assured that all features contribute equally to the quantization process, as the significance of the feature can either be influenced by (im-)proper quantization or by the expressiveness for the given recognition task. This can be realized with our novel VQ approach in two ways: the expressiveness of certain features can be measured either by combining them in a feature subset and distributing quantization error distributed equally among the features or these features are quantized with a higher SNR than the remaining features.

Although the features on which the novel VQ scheme has been applied are native to HWR, the VQ scheme is not — therefore it can also be useful for the signal processing and speech community.

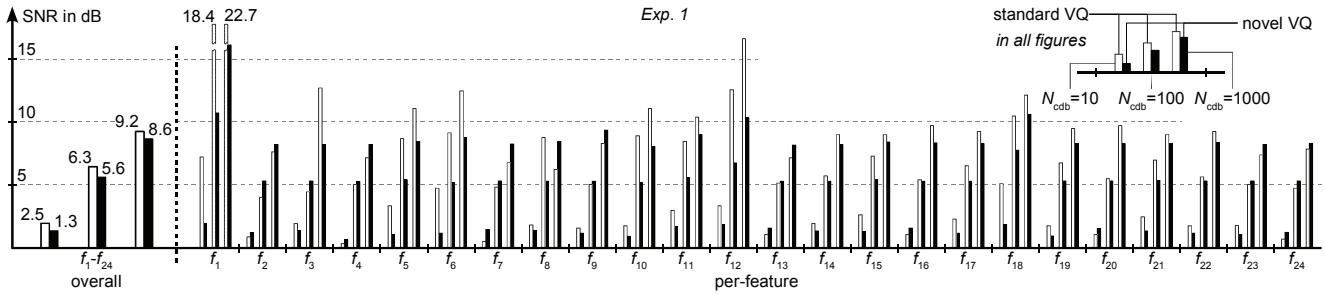


**Fig. 4**. Results for *Exp. 1*: Overall and per-dimension SNR derived on all features $f_1 - f_{24}$ used in on-line HWR of whiteboard notes for codebook sizes of $N_{\text{cdb}} = 10, 100, 1000$ applying standard VQ (each on the left) and our novel VQ (each on the right) in order to achieve an equal distribution of the quantization error and hence the same SNR in each dimension (i. e. $r_1 = \ldots = r_{24} = 1$).

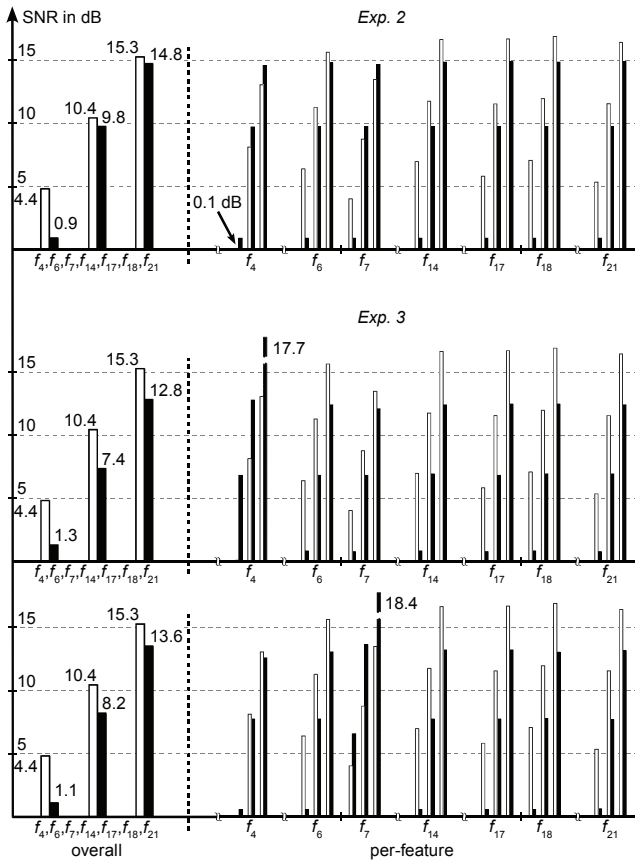


**Fig. 5**. Results of *Exp. 2* and *Exp. 3* (see labels in the figure): Overall and per-dimension SNR calculated on the feature subset $\mathcal{X} = \{f_4, f_6, f_7, f_{14}, f_{17}, f_{18}, f_{21}\}$ for codebook sizes of $N_{\text{cdb}} = 10, 100, 1000$ applying standard VQ (each on the left) and our novel VQ (each on the right). Equal distribution (i. e. $r_7 = \ldots = r_{21} = 1$, top), increase of $4^{\text{th}}$ feature's SNR (i. e. $f_4 = 1$; $f_6 = 4, \ldots = f_{21}$, middle), and increase of $7^{\text{th}}$ feature's SNR (i. e. $f_4 = f_6 = f_8 = \ldots = f_{21} = 4$; $f_7 = 1$, bottom).

# Acknowledgments

The authors sincerely thank H. Schenk for his vital comments.

# References

[1] J. Makhoul, S. Roucos, and H. Gish, "Vector Quantization in Speech Coding," *Proceedings of IEEE*, vol. 73, no. 11, pp. 1551 – 1588, 1985.

[2] R.M. Gray, "Vector Quantization," *IEEE ASSP Magazine*, pp. 4 – 29, 1984.

[3] T. Lookabaugh, E.A. Riskin, P.A. Chou, and R.M. Gray, "Variable Rate Vector Quantization for Speech Image and Video Compression," *IEEE Transactions on Communications*, vol. 41, no. 1, pp. 165 – 199, 1993.

[4] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proceedings of IEEE*, vol. 77, no. 2, pp. 257 – 285, 1989.

[5] H.-W. Nein and C.-T. Lin, "Incorporating Error Shaping Technique into LSF Vector Quantization," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 2, pp. 73 – 86, 2001.

[6] R.P. Cohn and J.S. Collura, "Incorporating Perception into LSF Quantization — Some Experiments," *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, pp. 1347 – 1350, 1997.

[7] J. Schenk and G. Rigoll, "Neural Net Vector Quantizers for discrete HMM-Based On-Line Handwritten Whiteboard-Note Recognition," *Proceedings of the International Conference on Pattern Recognition*, p. in press, 2008.

[8] M. Liwicki and H. Bunke, "HMM-Based On-Line Recognition of Handwritten Whiteboard Notes," *Proceedings of the International Workshop on Frontiers in Handwriting Recognition*, pp. 595 – 599, 2006.

[9] A. Viterbi, "Error Bounds for Convolutional Codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Transactions on Information Theory*, vol. 13, pp. 260 – 267, 1967.

[10] M. Liwicki and H. Bunke, "IAM-OnDB - an On-Line English Sentence Database Acquired from Handwritten Text on a Whiteboard," *Proceedings of the International Conference on Document Analysis and Recognition*, vol. 2, pp. 956 – 961, 2005.

[11] M. Liwicki and H. Bunke, "Combining On-Line and Off-Line Systems for Handwriting Recognition," *Proceedings of the International Conference on Document Analysis and Recognition*, pp. 372 – 376, 2007.