low gate leakage current. The measured values of gate-drain breakdown voltage of over -90V, fT of 48 GHz and fmax of 108 GHz are the highest ever reported data for recessed-gate AlGaN/GaN HEMTs.

Acknowledgments: This work was supported at UIUC by ONR under contract no. N00014-01-1-1000, Air Force under contract no. AFOSR-40029, and Triquint Corporation.

© IEE 2001
Electronics Letters Online No: 20010999
DOI: 10.1049/el:20010999

V. Kumar, W. Lu, F.A. Khan, R. Schwindt and I. Adesida (Department of Electrical and Computer Engineering and Microelectronics Laboratory, University of Illinois at Urbana Champaign, Urbana, IL 61801, USA)
E. Piner (ATMI/Epitronics, Phoenix, AZ 85027, USA)

References

Threshold analysis in wavelet-based denoising
L. Zhang, P. Bao and Q. Pan

The hard threshold $t = \sigma \alpha$ is efficient in wavelet threshold-based nonlinear filtering. In general, the optimal constant $c$ would vary with the signal and the added noise. The nearly optimal choice of $c$ by minimizing $R(c)$, which is an equivalent function to the mean square error (MSE) of the recovered signal, is discussed. Experiment shows that $R(c)$ is consistent with the MSE.

Introduction: In wavelet-based signal denoising, an intuitive and efficient approach is to apply the preset threshold to the wavelet coefficients. Donoho [1] first gave a soft threshold $t = \sigma \alpha (2 \log M)$, where $\sigma$ is the standard deviation of noise and $M$ is the length of signal. While this threshold possesses some minimax properties, it is non-intuitive and varies with $M$. Pan and Zhang [2] used a hard threshold $t = \sigma \alpha$ to denoise signals effectively, where constant $c$ is chosen as $c \in (3, 4)$ based on the fact that most of the noise values will fall within 3–4 times of its standard deviation.

For a given $M$ or a fixed $c$, the two thresholds above will be invariant for all types of signals. In general, the optimal value of $t$ or $c$ would vary with the signals and noise. In this Letter, a function $F(c)$ that approximates an equivalent to the mean square error (MSE) of the recovered signal is constructed. The near optimal $c$ can be determined by minimizing $F(c)$.

Hard threshold-based denoising by wavelet transform: Suppose there is a sequence of observations $y_i = x_i + \eta_i$, $i = 1, 2, ..., M$, where $\eta_i = \mathcal{N}(0, \sigma^2)$ is white Gaussian noise. The goal is to estimate signal $X$ from $Y$. Donoho [1] first developed a wavelet shrinkage method by a soft threshold $t = \sigma \alpha (2 \log M)$. Some impressive results were reported based on hard thresholding [2]. This procedure can be described as follows. First the sequence $Y$ is transformed into wavelet coefficient $W_Y$, then hard threshold $t = \sigma \alpha$ is applied on $W_Y$:

$$W_Y(i) = \begin{cases} W_Y(i), & |W_Y(i)| \geq t \\ 0, & |W_Y(i)| < t \end{cases}$$

where $c \in (3, 4)$ is a constant; finally the estimation $\hat{Y}$ is reconstructed from $W_Y$.

In general, for different signals and noise, the optimal $c$ will be different. We now construct a function $R(c)$, nearly equivalent to the MSE of $\hat{Y}$, to determine $c$.

Determination of $c$: Orthogonal wavelet transform (OWT) is linear. Thus we have $W_Y = W_X + W_N$, where $W_Y$, $W_X$ and $W_N$ denote the OWT of observation $Y$, signal $X$ and noise $N$, respectively. Similarly, we have $W_Y = W_X + W_N$, where:

$$\begin{align*}
\hat{W}_X(i) &= W_X(i), |W_X(i)| \geq t \\
\hat{W}_X(i) &= 0, |W_X(i)| < t
\end{align*}$$

We also have $\hat{Y} = \hat{X} + \hat{N}$, where $\hat{Y}$, $\hat{X}$ and $\hat{N}$ are the inverse OWT of $W_Y$, $W_X$ and $W_N$.

Obviously, the optimal $c$, i.e., $t$, should minimize the MSE of $\hat{Y}$:

$$E[|\hat{Y} - X|^2] = \min$$

Because $E[|\hat{Y} - X|^2]$ is independent of $c$, it is equivalent to minimizing the following function:

$$\text{Error}(c) = E[|\hat{Y} - X|^2] = E[X^2]$$

Since OWT is orthonormal, from eqn. 2 and $E[\hat{W}_X W_N] = 0$, we get:

$$\text{Error}(c) = 2E[W_N^2]$$

Suppose $c$ points will be eliminated in $W_Y$, we have:

$$E[\hat{W}_N^2] = \frac{1}{M} \sum_{i=1}^{M} W_N^2(i) - \frac{1}{M} \sum_{j=1}^{M} W_N^2(i)$$

where $\hat{W}_N$ denotes the $c$ eliminated points in $W_N$.

It is almost true that if $|W_N(i)| < t$ then $|\hat{W}_N(i)| < t$. This statement is validated by the tests on two typical signals in Fig. 1. Let $t = \sigma \alpha$ and $c$ increase from 1 to 4 with step-length 0.3. Denote $K$ the number of points satisfying $|\hat{W}_N(i)| < t$ and $K_N$ the number of points satisfying both $|W_N(i)| < t$ and $|\hat{W}_N(i)| < t$. The averaged results of $K/N$ generated by the Monte Carlo experiment in signal-to-noise ratio (SNR) are listed in Table 1. In fact, when $t \geq 0$, for noise in any scale, $K/N$ is nearly equal to 1, which implies that when $|W_N(i)| < t$, $|\hat{W}_N(i)| < t$ holds with high probability. The wavelets used in the experiments are Haar (for Blocks) and four taps orthogonal wavelet (for Bumps).

Fig. 1 Two typical signals and their noisy versions
- a Blocks
- b Bumps
- c Noisy Blocks
- d Noisy Bumps
Table 1: Averaged value of $K_\gamma/K$, where $K$ is the number of points satisfying $|W_\gamma(i)| < t$ and $K_\gamma$ the number of points satisfying both $|W_\gamma(i)| < t$ and $|W_\gamma(i)| > t$. $W_\gamma$ and $w_\gamma$ are the OWT of observed signal and noise, respectively. Threshold $t = \alpha\sigma$.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>1.0</th>
<th>1.5</th>
<th>1.6</th>
<th>1.9</th>
<th>2.2</th>
<th>2.5</th>
<th>2.8</th>
<th>3.1</th>
<th>3.4</th>
<th>3.7</th>
<th>4.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.9</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>15.9</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>14.9</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>13.9</td>
<td>0.998</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>12.9</td>
<td>0.997</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>11.9</td>
<td>0.996</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>10.9</td>
<td>0.996</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>9.9</td>
<td>0.995</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>

For the points that $|W_\gamma(i)| > t$, there exist some points satisfying $|W_\gamma(i)| < t$. These points can be considered as the random samplings of all the points satisfying $|W_\gamma(i)| < t$. Thus

$$
\sum_{i=1}^{k} W_\gamma^2(i) \approx k \cdot E[|W_\gamma|^2] \approx k \int_{-\infty}^{\infty} x^2 e^{-x^2/(2\sigma^2)} \frac{e^{2\sigma f(c)}}{\sqrt{2\pi\sigma f(c)}} dx
$$

where $\sigma f(c) = [1/(2\pi\sigma)]^2 e^{2\sigma^2 f(c)}$. From eqns. 3 - 5, we define the following function:

$$
F(c) = 2 \left( \sigma^2 - \frac{k}{2\sigma f(c)} \int_{-\infty}^{\infty} x^2 e^{-x^2/(2\sigma^2)} \frac{e^{2\sigma f(c)}}{\sqrt{2\pi\sigma f(c)}} dx \right) - E[|W_\gamma|^2]
$$

$F(c)$ is approximately equal to $Error(c)$, which is equivalent to the MSE of recovered signal $Y$. The minimum of $F(c)$ can be computed to determine the nearly optimal value of $c$.

Fig. 2 $F(c)$ and $Error(c)$ of Blocks

--- --- $F(c)$
--- --- $Error(c)$

Experimental results: Two typical signals, Blocks and Bumps, with their noisy versions are shown in Fig. 1. To obtain the constant $c$ for the hard threshold-based denoising method, the functions $F(c)$ and $Error(c)$ of the noisy signals in Fig. 1 are calculated with $c \in (1, 6)$ and shown in Fig. 2 and Fig. 3. It is observed that the shapes of $F(c)$ and $Error(c)$ are identical. Through $F(c)$ a small interval of $c$ could be obtained, which could be considered approximately as the minimal position of $Error(c)$, hence the MSE of the recovered signals. In Fig. 2 and Fig. 3, the intervals are $c = 3.5 - 4.0$ for Blocks and $c = 3.0 - 3.3$ for Bumps.

Fig. 3 $F(c)$ and $Error(c)$ of Bumps

--- --- $F(c)$
--- --- $Error(c)$

Conclusion: The scheme presented in this Letter is adaptive to signal and noise and then the near optimal $c$ is signal and noise dependent.

© IEE 2001
Electronics Letters Online No. 20010971
DOI: 10.1049/el:20010971
L. Zhang and P. Bao (Department of Computing, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong)
E-mail: cbsao@comp.polyu.edu.hk
Q. Pan (Department of Automatic Control, Northwestern Polytechnical University, Xi'an, 710072, People's Republic of China)

References


GMM based on local PCA for speaker identification

Changwoo Seo, Ki Yong Lee and Joohun Lee

An efficient Gaussian mixture modelling (GMM) method based on local principal component analysis (PCA) with vector quantisation (VQ) for speaker identification is proposed. The proposed method firmly partitions the data space into several disjoint regions by VQ and then performs PCA in each region. Finally, the GMM for the speaker is obtained from the transformed feature vectors in each region. Compared to the conventional GMM method with diagonal covariance matrix, under the same performance, the proposed method requires less storage and shows faster results.

Introduction: The Gaussian mixture modelling (GMM) method with diagonal covariance matrix is increasingly being used for both speaker identification and verification [1]. In a speaker recognition system, a larger feature set is preferable to enhance speaker recognition. Also, since the elements of feature vectors extracted from a speech signal are correlated, a larger number of mixtures is necessary in order to provide a good approximation [2]. However, the increase in the number of the feature vectors and mixtures causes other problems. For example, the recognizer using a higher dimension of feature set requires more parameters to characterise...