Non-stationary Noise Estimation Based on Non-negative Matrix Factorization

Kwang Myung Jeon¹, Nam In Park¹, Hong Kook Kim¹, Kwang Il Hwang², and Myung Kyu Choi²

¹ School of Information and Communications
Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Korea
{kmjeon, naminpark, hongkook}@gist.ac.kr
² Digital Imaging Business, Samsung Electronics Co., Ltd., Gyeonggi-do 443-742, Korea
{core.choi, srkim}@samsung.com

Abstract. In this paper, we apply a non-negative matrix factorization (NMF) technique to propose a method of estimating noise occurring in non-stationary environments. In the proposed method, the basis matrix of the target noise is first obtained via NMF training. The noise basis is then applied to estimate an activation matrix of the target noise from the noisy signal. The proposed method is finally applied to reduce the auto-focus (AF) noise in a digital camera. It is shown from the experiment that the proposed method provides a better estimate of the AF noise than a conventional method based on signal-presence probability.

Keywords: Noise estimation, non-negative matrix factorization (NMF), non-stationary noise, auto-focus (AF) noise

1 Introduction

Most speech or audio enhancement algorithms estimate noise spectrum prior to denoising. For instance, a priori signal-to-noise ratio (SNR) and/or a posteriori SNR must be estimated for performing a denoising operation using the Wiener filtering [1] or minimum mean squared error (MMSE) algorithms [2]. Thus, an accurate estimate of SNR plays a crucial role in the performance of denoising methods. In general, noise estimation might work well under stationary noise conditions, but it might not be successful to estimate non-stationary noises that occur in most real environments. To cope with this limitation, several noise estimation methods based on signal-presence probability have been proposed [3][4]. However, the performance of such methods could be limited when the noise level was highly unpredictable, as in autofocus (AF) noise generated by a digital camera.

To mitigate this problem, a noise estimation method based on non-negative matrix factorization (NMF) [5] is proposed in this paper. In the proposed method, noise spectrum is estimated using prior knowledge of the target noise, while the conventional methods based on signal-presence probability estimate noise spectrum by tracking the
local minimum [4]. Thus, the proposed method is able to more accurately estimate non-stationary noise such as AF noise than the conventional ones.

Following this introduction, Section 2 proposes a non-stationary noise estimation method using NMF. Next, Section 3 evaluates the performance of the proposed noise estimation method. Finally, Section 4 concludes this paper.

2 Proposed Non-stationary Noise Estimation Using Non-negative Matrix Factorization

In an additive noise environment, a spectral magnitude of the input noisy signal, \( Y \), is represented as \( Y = S + D \), where \( S \) is a spectral magnitude of clean audio signal and \( D \) is that of noise signal. The goal of the proposed method is to estimate \( D \) from \( Y \) without prior knowledge of \( S \). This can be achieved by applying an NMF technique [5] because all the spectral magnitudes are non-negative. To this end, a modified discrete cosine transform (MDCT) is applied to noisy signal in order to obtain \( Y \) and several analysis frames are concatenated for the NMF technique. In other words, \( Y \) becomes an \( i \times \mu \) matrix, where \( i \) is the MDCT point and \( \mu \) is the number of concatenated analysis frames. Thus, \( S \) and \( D \) are also \( i \times \mu \) matrices. In the NMF framework, \( D \) is represented as \( D = B_D A_D \), where \( B_D \) and \( A_D \) are a basis matrix and activation matrix of \( D \), respectively. Thus, the proposed noise estimation method provides \( B_D \) in the training part while \( A_D \) is obtained in the estimation part.

First, the training part of the proposed method obtains \( B_D \) from noise signals by using NMF. The NMF training is performed iteratively by the following equations of [6]

\[
B_D^t = B_D^{t-1} \otimes \frac{D (A_D^{-1})^T}{B_D^{t-1} A_D^T A_D^{-1} I}, \quad (1)
\]

\[
A_D^t = A_D^{t-1} \otimes \frac{B_D^t D (B_D^t A_D^t)^T}{B_D^{t-1} A_D^T A_D^{-1} I}, \quad (2)
\]

where \( t \) is an iteration index, and \( I \) is a matrix with all elements equal to unity. Moreover, both multiplication \( \otimes \) and division indicate element-wise operators. In Eqs. (1) and (2), \( B_D \) and \( A_D \) are an \( i \times a \) basis matrix and an \( a \times \mu \) activation matrix of the target noise, respectively, where \( a \) is the number of bases and is a controllable parameter in NMF. The iterative procedure described above is terminated if the difference of an objective function according to the iteration is less than a pre-defined threshold. That is, the objective function is defined as [6]

\[
obj(t) = \sum_{i,p} D_{i,p} \otimes \log \frac{D_{i,p}}{(B_D^t A_D^t)_{i,p}} - D_{i,p} + (B_D^t A_D^t)_{i,p}, \quad (3)
\]
Thus, if $\theta > -1$ and $\theta < 0$, then NMF training is finished. In this paper, $\theta$ is manually set to 0.1 by trading off between the iteration number and the estimation accuracy of the NMF operations. After terminating the iteration, we set $B_0 = B_0$.

After obtaining $B_0$ in the training part, the estimation part attempts to find $\hat{A}_0$ from $Y$. To obtain $\hat{A}_0$, an NMF iterative procedure is applied as

$$B_s^t = B_s^{t-1} \otimes \frac{Y}{B_s^{t-1}A_s^{t-1T}A_s^{T}}, \quad (4)$$

$$A'_v = A'^{t-1}_{v} \otimes \frac{B'_v^{T}Y}{B'_v^{T}B'_v A'^{t-1}_{v}}, \quad (5)$$

where $B_s$ and $A_s$ are an $i \times a$ basis matrix and an $a \times \mu$ activation matrix of $S$, respectively. Moreover, $B'_v = [B'_s; B_0]$ and $A'_v = [A'_s; \hat{A}_0]$. Similar to the termination condition of Eq. (3), the estimation procedure by NMF is also terminated by checking it out whether

$$\sum_{i,v} Y_{i,v} \otimes \log \frac{Y_{i,v}}{(B'_v A'_v)^{T})_{i,v}} - Y_{i,v} + (B'_v A'_v)^{T})_{i,v}$$

is going to be converged. Finally, the estimated noise spectrum $\hat{D}$ is obtained from $\hat{D} = B_0 \hat{A}_0$, where the procedure of Eqs. (4) and (5) is terminated at the $t$-th iteration.

### 3 Noise Estimation Experiment

The performance of the proposed noise estimation method was evaluated by measuring a mean squared error (MSE) between the true noise spectrum and the estimated noise spectrum by the proposed method. In the experiment, 10 input noisy signals of around 0 dB SNR were prepared by artificially adding AF noise recorded in a digital camera to clean audio signals from SQAM DB [8]. In this paper, $i$, $a$, and $\mu$ were set to 1024, 40, 350, respectively.

Table 1 compares the MSE of the proposed method and that of the conventional one based on signal-presence probability [4]. It was shown from the table that the proposed method reduced average MSE by 0.44, compared to the conventional method. This implies that the proposed method could estimate non-stationary noise with higher accuracy than the conventional method.

### 4 Conclusion

In this paper, we proposed a non-stationary noise estimation method based on NMF. To this end, NMF was iterative applied to get a basis matrix and an activation matrix.
of non-stationary noise in the training and the estimation part, respectively. It has been shown from the experiment that the proposed method estimated auto-focus noises with lower MSE than the conventional method did.

Acknowledgments. This work was supported in part by Samsung Electronics Co., and by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (MEST) (No.2011-0026201).

References


Table 1. Performance comparison of the noise estimation methods in MSE.

<table>
<thead>
<tr>
<th>Test signals</th>
<th>Conventional</th>
<th>Proposed</th>
<th>MSE difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1.41</td>
<td>0.99</td>
<td>0.42</td>
</tr>
<tr>
<td>B</td>
<td>1.30</td>
<td>0.97</td>
<td>0.33</td>
</tr>
<tr>
<td>C</td>
<td>1.34</td>
<td>1.05</td>
<td>0.29</td>
</tr>
<tr>
<td>D</td>
<td>1.53</td>
<td>1.03</td>
<td>0.50</td>
</tr>
<tr>
<td>E</td>
<td>1.50</td>
<td>0.92</td>
<td>0.58</td>
</tr>
<tr>
<td>F</td>
<td>1.54</td>
<td>1.02</td>
<td>0.52</td>
</tr>
<tr>
<td>G</td>
<td>1.30</td>
<td>0.95</td>
<td>0.37</td>
</tr>
<tr>
<td>H</td>
<td>1.29</td>
<td>0.91</td>
<td>0.37</td>
</tr>
<tr>
<td>I</td>
<td>1.28</td>
<td>1.07</td>
<td>0.20</td>
</tr>
<tr>
<td>J</td>
<td>1.53</td>
<td>0.93</td>
<td>0.60</td>
</tr>
<tr>
<td>Average</td>
<td>1.42</td>
<td>0.98</td>
<td>0.44</td>
</tr>
</tbody>
</table>