RESTORATION OF LOCAL DEGRADATIONS IN AUDIO SIGNALS

Milan BREJL
Faculty of Electrical Engineering and Computer Science
Technical University of Brno
Antonínská 1
662 09 Brno
Czech Republic
E-mail: xbrejl00@krel.fee.vutbr.cz

Abstract

The paper presents an algorithm for restoration of local degradations in audio signals. The theoretical foundations and basic suggestions of this algorithm were published in [1]. A complete description of restoration process and some improvements are presented here.

Keywords:

signal restoration, AR model, local degradation detection.

1. Introduction

The problem of signal restoration has often to be solved in some applications. The method depends on the character of signal degradation. Some types of degradation in audio signals have a local character - clicks, scratches, etc. This means that some parts of a signal are kept uncorrupted while the other parts are substantially degraded. To restore such signals the corrupted parts must be localized and uncorrupted parts must be used to obtain the best possible estimation of the signal character and apply it to the restoration of the corrupted parts.

2. A Posterior Probability Method for the Localisation of Corrupted Parts

The corrupted parts in audio signals can be regarded as additive bursts of noise with a zero-mean Guassian distribution. The variance of noise is marked σ_n^2.

The location of noise bursts can be modelled using a two-state Markov chain. One state—state 0—represents an uncorrupted sample and the second state—state 1—represents a corrupted sample. The probabilities of transmissions between these states are P_{01} and P_{10}. The probabilities of continuations in these states are P_{00} and P_{11} where

$$P_{00} = 1 - P_{01}, \tag{1}$$

$$P_{11} = 1 - P_{10}. \tag{2}$$

The probability P_{00} affect the density of noise bursts occurrence and P_{11} affect the length of these bursts.

An autoregressive model exist for the uncorrupted parts of signal. The order of AR model is p and the AR model coefficients $a_1, \ldots, a_p$ are computed in such a way that a zero-mean white noise process e_n, called excitation noise, has minimal variance σ_e^2.

$$\sum_{k=0}^{p} a_k \cdot s_{n-k} = e_n, \tag{3}$$

where s_n is the n-th sample of uncorrupted signal. The AR model coefficients can be arranged to a matrix $\mathbf{A}$ as follows

$$\mathbf{A} = \begin{bmatrix} a_p & a_{p-1} & \cdots & a_1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & a_p & a_{p-1} & \cdots & a_1 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \cdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_p & a_{p-1} & \cdots & a_1 & 1 & 0 \\ 0 & 0 & \cdots & 0 & a_p & a_{p-1} & \cdots & a_1 & 1 \end{bmatrix} \tag{4}$$

and then

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{e} \tag{5}$$

where $\mathbf{x}$ is a vector of a signal and $\mathbf{e}$ is a vector of a zero-mean white noise process.

Assuming these, formulas for assigning a possible location of noise bursts a probability value were derived in [1]. The locations of noise bursts are determined by a detection vector $\mathbf{i}$. When $\mathbf{x}_c$ is vector of corrupted signal, $P(\mathbf{i}|\mathbf{x}_c)$ is the posterior probability of certain noise bursts locations given by $\mathbf{i}$.

$$P(\mathbf{i}|\mathbf{x}_c) \approx \frac{P(\mathbf{i})\left(\frac{\sigma_e}{\sigma_n}\right)^l \exp\left(-\frac{1}{2\sigma_e} E_{\min}\right)}{|\Phi|^{1/2}}, \tag{6}$$

$$E_{\min} = E_0 - \Theta^T \mathbf{x}_{MAP}(\mathbf{i}), \tag{7}$$

$$\mathbf{x}_{MAP}(\mathbf{i}) = \Phi^{-1}\Theta, \tag{8}$$

$$\Phi = \mathbf{A}(\mathbf{i})^T \mathbf{A}(\mathbf{i}) + \frac{\sigma_e^2}{\sigma_n^2}\mathbf{I}, \tag{9}$$

$$\Theta = -\left(\mathbf{A}(\mathbf{i})^T \mathbf{A}(\tilde{\mathbf{i}})\mathbf{x}_c(\tilde{\mathbf{i}}) + \frac{\sigma_e^2}{\sigma_n^2}\mathbf{x}_c(\mathbf{i})\right), \tag{10}$$

$$E_0 = \mathbf{x}_c(\tilde{\mathbf{i}})^T \mathbf{A}(\tilde{\mathbf{i}})^T \mathbf{A}(\tilde{\mathbf{i}})\mathbf{x}_c(\tilde{\mathbf{i}}) + \frac{\sigma_e^2}{\sigma_n^2}\mathbf{x}_c(\mathbf{i})^T \mathbf{x}_c(\mathbf{i}). \tag{11}$$

$P(\mathbf{i})$ is prior probability for detection vector $\mathbf{i}$ computed from Markov chain model. The notation $\mathbf{x}(\mathbf{i})$ marks the elements of vector $\mathbf{x}$ in positions corresponding to corrupted samples detected by $\mathbf{i}$, $\mathbf{A}(\mathbf{i})$ marks the columns of matrix $\mathbf{A}$ in such positions and $\mathbf{x}(\tilde{\mathbf{i}}), \mathbf{A}(\tilde{\mathbf{i}})$ marks elements or columns in positions corresponding to uncorrupted samples. Probabilities of various noise bursts locations determined by various $\mathbf{i}$ can be compared using these formulas and the most probable $\mathbf{i}$ can be determined as the right detection vector. $\mathbf{x}_{MAP}(\mathbf{i})$ is a vector of samples by which the corrupted samples can be substituted. So vector $\mathbf{x}_{MAP}$ represents the restored signal.

3. The Restoration Algorithm

For computation of probability value $P(\mathbf{i}|\mathbf{x}_c)$ two parameters of Markov model P_{00} and P_{11}, variances σ_n^2 and σ_o^2 and AR model coefficients $a_1, ..., a_p$ must be known. One possible way to get them is, as suggested in [1], to adapt their original estimations from restored signal during the sequential restoration algorithm.

The sequential algorithm works with a window of the signal. Before the window the signal is unrestored, in the window several possibilities of noise burst locations presented by vectors $\mathbf{i}_1, ..., \mathbf{i}_m$ are considered and after the window the signal is restored. One sample is added to the window in each step considering both possibilities—the sample is corrupted or uncorrupted—for each vector of $\mathbf{i}_1, ..., \mathbf{i}_m$. So the number of possible detections is doubled to $2m$: $\mathbf{i}_1, ..., \mathbf{i}_{2m}$. Now for each vector $\mathbf{i}$ its probability value $P(\mathbf{i}|\mathbf{x}_c)$ is computed. Vectors $\mathbf{i}$, whose logarithmic probability is distanced from the most probable vector $\mathbf{i}$ more than a certain threshold is, are not taken into consideration. If the number of remaining vectors $\mathbf{i}$ is greater than a certain maximum, the less probable vectors are not included too.

If the sample at the end of the signal window is detected as uncorrupted by all remaining vectors $\mathbf{i}$, the sample can be put out of the window and interpreted as restored. If next sample is also detected by all vectors $\mathbf{i}$ as uncorrupted, it can be put out too etc. until a sample that is detected by some vectors $\mathbf{i}$ as corrupted is at the end of the signal window. If the sample at the end of the signal window is detected by some vectors $\mathbf{i}$ as corrupted no sample is put out of the window and the window is one sample longer in the next step. The vectors $\mathbf{i}$ detecting this sample as corrupted may have small probabilities and may be left from consideration in next steps. If the window reaches its given maximum length, the sample at the end of the signal window is restored following the most probable vector $\mathbf{i}$. If the sample is detected as uncorrupted by this vector $\mathbf{i}$, no restoration is made; if detected as corrupted, correspondent sample from the vector $\mathbf{x}_{MAP}$ is supplied to the restored signal. Though the sample at the end of the signal window is detected as corrupted by all vectors $\mathbf{i}$, the restoration of this sample is not made until the window reaches its maximum length because the supplement by the vector $\mathbf{x}_{MAP}$ is better when a longer block of signal is used.

Every time a sample is put out of the signal window as a coming restored signal sample, the adaptation of original signal and noise bursts parameters is made. The algorithm of trimming the signal window when the samples at the end of the window are certainly uncorrupted enables to make more adaptations and keep the signal parameters more actual.

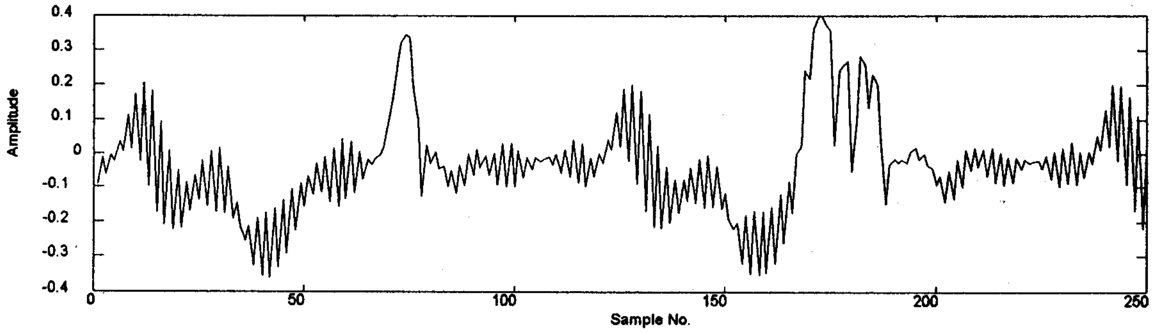

Fig. 1 Degraded signal

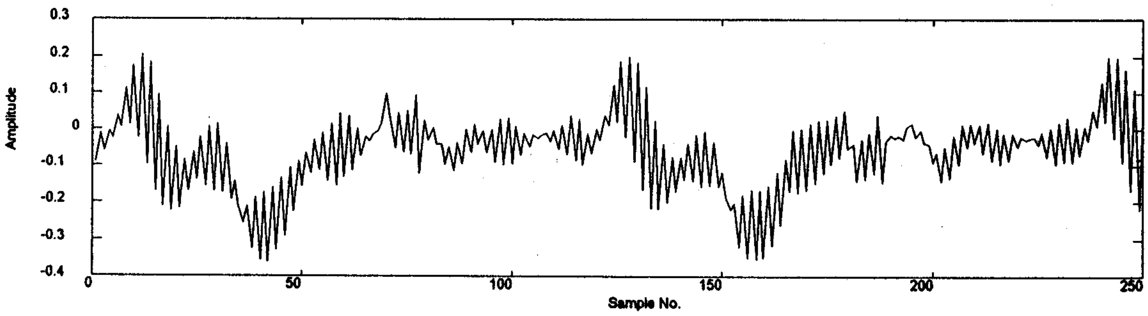

Fig. 2 Restored signal

Adaptation of σ_n^2 and σ_e^2 can be done using a exponential weighting of former values. If the coming restored sample has been uncorrupted originally, the adaptation of σ_e^2 is proceeded. If the sample has been corrupted, the adaptation of σ_n^2 is proceeded.

$$\sigma_e^2(n) = \lambda\sigma_e^2(n-1) + (1-\lambda)e_n^2, \qquad (12)$$

$$\sigma_n^2(n) = \lambda\sigma_n^2(n-1) + (1-\lambda)(x_{cn} - x_n)^2. \qquad (13)$$

where λ is a 'forgetting factor' whitch reflect's the estimated stacionarity of the original signal, e_n is sample of AR model error, x_{cn} is corrupted sample and x_n is the corresponding restored sample. σ_n^2 and σ_e^2 must be initialized by estimations.

Adaptation of AR model coefficients can be done by RLS algorithm [2]. RLS algorithm is method for adaptive setting of FIR filter coefficients to get the filter response as close as possible to the desired response. When sample x_n is fetched to the RLS algorithm as desired response and sample x_{n-1} as FIR filter input sample, the filter will be adapted to predicate one sample and the FIR filter coefficients will get function of AR coefficients. The formulas for RLS algorithm modified to this case are:

$$\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{x}(n)}{1+\lambda^{-1}\mathbf{x}^T(n)\mathbf{P}(n-1)\mathbf{x}(n)}, \qquad (14)$$

$$e_n = x_n - \mathbf{a}^T(n-1)\mathbf{x}(n), \qquad (15)$$

$$\mathbf{a}(n) = \mathbf{a}(n-1) + \mathbf{k}(n)e_n, \qquad (16)$$

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{x}^T(n)\mathbf{P}(n-1), \qquad (17)$$

where $\mathbf{x}(n)$ is vector of samples $x_{n-1}, x_{n-2}, ..., x_{n-P}$ and $\mathbf{a}(n)$ is vector of AR model coefficients $a_1, ..., a_p$ in the n-th step. The initialization of the algorithm is made by setting $\mathbf{a}(0) = \mathbf{0}$ and $\mathbf{P}(0) = \delta^1\mathbf{I}$ where δ is a small integer value like 0.1 or so.

Markov model parameters P_{00} and P_{11} can be adapted too but it is not necessary. Keeping the parameters constant at the initialization estimated values is sufficient. If the adaptation is processed anyway, a different 'forgetting factor' λ than in previous adaptations should be used, closer to 1, to consider a larger segment of signal.

4. Example

Result of restoration is shown on a part of a real string orchestra signal degraded by clicks. The parameters of restoration were set as follows: AR model order $p = 20$, maximum length of signal window 30, threshold of logarithmic probability distance 2, $\sigma_n^2 = 0.3$, $\sigma_e^2 = 0.003$, forgetting factor $\lambda = 0.999$, $P_{00} = 0.96$ and $P_{11} = 0.88$.

Comparing the restored and the degraded signal it is evident that the clicks were localized, removed and the supplied signal has a character corresponding to surrounding parts.

5. Conclusion

A synthesis of [1] and improvements presented here produces a powerful algorithm of restoration of local degraded signal. This algorithm can be implemented as a second process except noise removing in audio signal restorations or in other applications. A good result of restoration requires some experience with this algorithm for correct initial setting of restoration parameters. Then it is possible to remove most defects in signal authentically.

References

[1] GODSHILL, SIMON J. and RAYNER, PETER J. W.: A Bayesian Approach to the Restoration of Degraded Audio Signals, IEEE Transactions On Speech And Audio Processing, July 1995, Vol. 3, No. 4, pp. 267-278

[2] HAYKIN, SIMON: Adaptive Filter Theory, Prentice-Hall, 1991