RESTORATION OF FAULTS CAUSED BY PATCHES ON OPTICAL AREA SOUND TRACKS

Dipl.-Inform. Andy KUIPER, Doctoral Degree Programme (1)
Dept. of Radio Electronics, FEEC, BUT
E-mail: xkuipe00@stud.feec.vutbr.cz

Ing. Lukáš DŽBÁNEK, Doctoral Degreee Programme (2)
Dept. of Radio Electronics, FEEC, BUT
E-mail: dzbanek@feec.vutbr.cz

Supervised by Dr. Milan Sigmund

ABSTRACT

Early motion picture recordings represent a historically valuable cultural inheritance which has to be preserved. This paper describes an attempt how applied patches on old motion pictures recordings can be restored using linear predictive coding.

1 INTRODUCTION OF THE PROBLEM AND GOAL OF THE WORK

At present more than 2 billion meters of films with substrates based on nitrocellulose are being stored in national and international film archives. Many of these films originate from the early period of sound film technology which started in the 1930s and represent a large amount of historical documents, which are being exposed to natural and potential destruction.

This aging process has lead to the fact that nowadays 90 % of all silent movies and half of all films produced before 1950 are irreparably destroyed.

Besides faults which arose due to aging, errors on sound tracks can arise from bad maintenance of the technical equipment (scratches) or careless storage (dust, fibers, dirt).

A third category which is subject of this paper is film patches. Patches are applied when two scenes are being joined together either caused by a scene cut or repairing ripped film material. A relatively large amount of sound track information is therefore irreparably lost.

2 OPTICAL SOUND TRACKS WITH FAULTS CAUSED BY PATCHES

Optical sound tracks are based on the principle that the encoding influences the intensity of passing light which is collected with a photo sensitive sensor during the reproduction. Therefore all optical soundtracks are compatible to each other and different types of encoding can be used with the same movie.

The Variable Density Code which can be seen in Fig. 1 (starting 1906) stores the sound information as a homogeneous gray level over the entire sound track width. Another variant called Variable Area Codes (starting 1911) [1] can be seen in Fig. 2 through 4. They use variable opaque and transparent areas and occur mostly in symmetric formations.

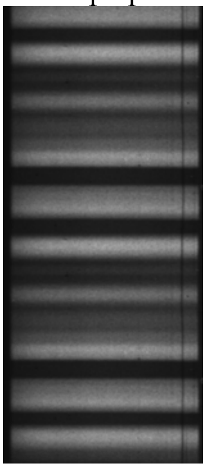

Fig. 1: *Variable density code*

Fig. 2: *Patched variable density code*

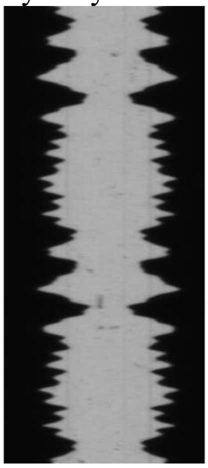

Fig. 3: *Double sided variable area code*

Fig. 4: *Patched double sided variable area code*

3 EFFECTS OF DISTURBANCES ON THE AUDIO SPECTRUM

Disturbances in a digital audio signal appear as a gap when frames are discarded and cannot be requested again due to the real time requirement. This gap is noticeable as clicking which is disturbing. Nearly the same affect appears on old optical sound track films when the film material breaks and has to repaired. These film-parts are re-joined using patches [1]. Such a patch can be seen in Fig. 5 with its corresponding audio signal in Fig. 6.

The effect of an applied sound patch on the sound signal of an optical sound track is quite drastic which the audio signal shows in Fig. 6. Due to the shape of the patch the sound is faded out and faded in again.

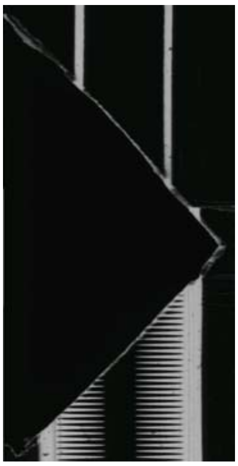

Fig. 5: *Sound track with patch*

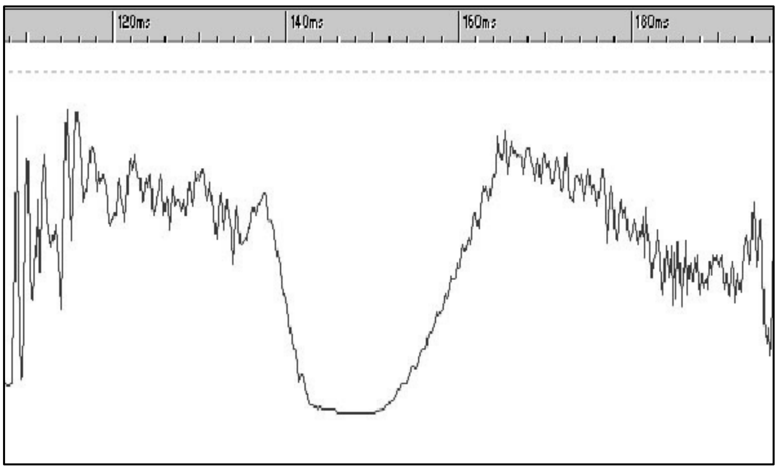

Fig. 6: *Corresponding audio signal*

4 LINEAR PREDICTIVE ANALYSIS

Linear predictive analysis is a powerful tool for speech processing. Its principle is to estimate which values could occur taking previous values into account. The brilliance of this method is that it is able to estimate speech parameters very precisely and is still very fast in computation. Linear prediction theory is well documented in the literature. Some of the landmark academic texts include [2], [3].

From a given signal $s(n)$ the linear predictive analysis tries to model the signal as a linear combination of M using its previous samples $s(n\text{-}1)$ to $s(n\text{-}M)$. Prediction of the speech sample is done in

$$\hat{s}(n) = \sum_{m=1}^{M} a_m \, s(n-m), \tag{1}$$

where a_m are the prediction coefficients and M denotes the number of previously used speech samples (order of the predictor).

In speech processing the autocorrelation method is used almost exclusively to calculate the predictor-coefficients from the autocorrelation coefficients $R(m)$. The autocorrelation method is advisable because of its computational efficiency and inherent stability. The most efficient method is Durbin's recursive procedure which is defined as follows:

Initialization:

$$e(0) = R(0) \tag{2}$$

$$a_{0,m} = 1 \tag{3}$$

Recursion:

$$a_{m,m} = \left[R(m) - \sum_{i=1}^{m-1} a_{i,m-1} \, R(m-i) \right] \Big/ e(m-1) \tag{4}$$

$$a_{i,m} = a_{i,m-1} - a_{m,m} \, a_{m-i,m-1} \qquad m = 1,\ldots, M \tag{5}$$

$$e(m) = \left(1 - a_{m,m}^2\right) e(m-1) \qquad i = 1,\ldots,m-1 \tag{6}$$

Equations (2) to (6) are solved recursively for m=1,2,...,M and the final solution is given as

$$a_{1,M}, a_{2,M}, \ldots, a_{M,M} \quad \Rightarrow \quad a_1, a_2, \ldots, a_M \tag{7}$$

An extended algorithm to calculate the predictor coefficients and their derivates in a noisy environment is shown in [4].

5 EXPERIMENTS AND RESULTS

The experiments using linear predictive analysis were done on a 1700 MHz Pentium IV with 382 Mbyte RAM as hardware using MathLab 6.5 from The MathWorks, Inc. as software environment. Main priority was set to the interpolation of faults from both sides which meant that preceding and subsequent data were used. The original speech signal which was used during the experimentation was sampled with 8000 Hz and had a resolution of 16 Bit which results therefore in a duration of about 1 second.

Fig. 7 shows the speech signal which was used for the linear predictive analysis. It features two gaps, each 17.25 ms in size. The gap denoted with A is in the transition area and the second gap B appears within the steady state area. The areas around A and B were analyzed separately, both with 30 and 300 LPC-coefficients.

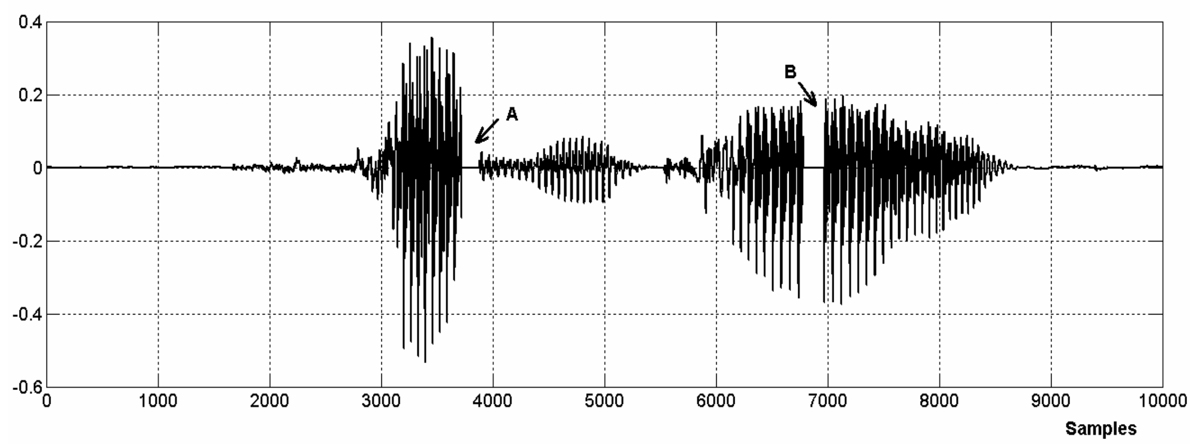

Fig. 7: *Speech signal with faults*

The first attempt of interpolating a missing area in an audio signal was done with the transition area. In Fig. 8 the difference between the prediction using 30 LPC-coefficients (drawn through line) and 300 LPC-coefficients (dashed line) can be seen. The same was done with the data after the gap which can be seen in Fig. 9 with 30 and 300 LPC-coefficients.

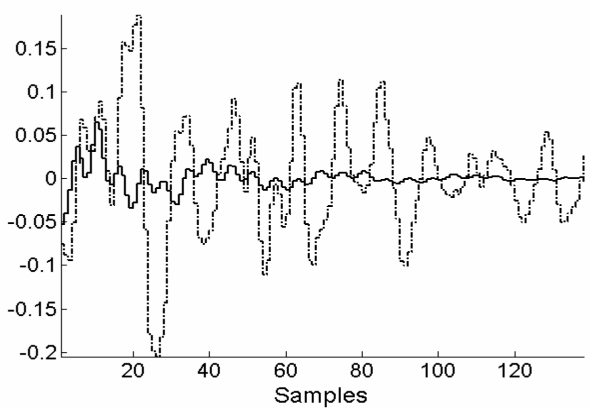

Fig. 8: *Predicted signal of the gap in the transition area using preceding data*

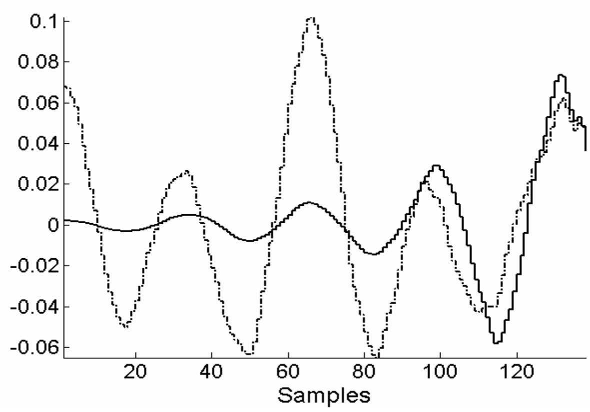

Fig. 9: *Predicted signal of the gap in the transition area using subsequent data*

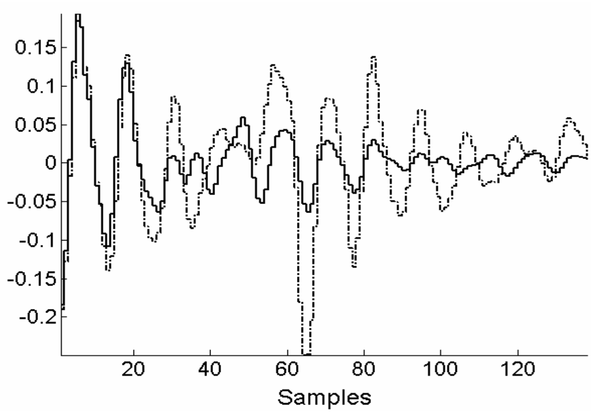

Fig. 10: *Predicted signal of the gap in the steady state area using preceding data*

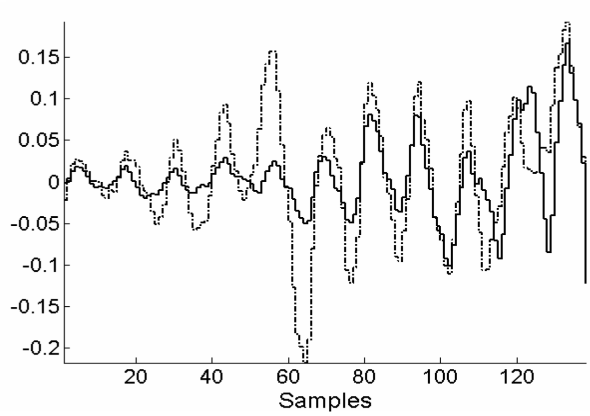

Fig. 11: *Predicted signal of the gap in the steady state area using subsequent data*

The second attempt used the same procedure as described previously using the steady state area instead of the transition area. The results of the prediction can be seen in Fig. 10 for the preceding and in Fig. 11 for the subsequent part. The dotted line denotes the linear prediction with 300 LPC-coefficients and the drawn through line the usage of 30 LPC-coefficients for prediction.

In order to combine the calculated signals before and after the fault a cross-fading function was applied which fades out slowly from the preceding side while fading in from the subsequent and the other way round. The resulting audio signals can be seen in Fig. 12 and 13 whereby the dotted line denotes the original signal, the dashed line the interpolated signal with 300 coefficients and the drawn through line the signal with 30 coefficients.

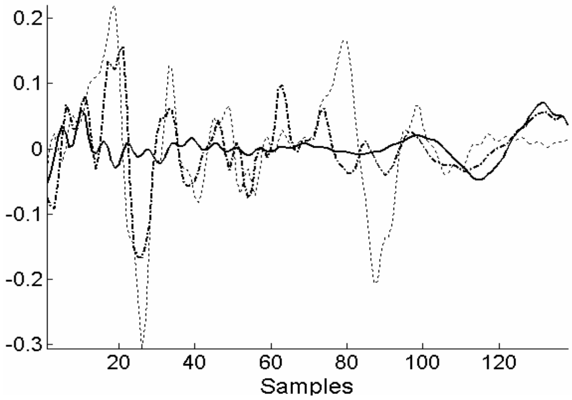

Fig. 12: *Result of the linear prediction analysis in the transition area*

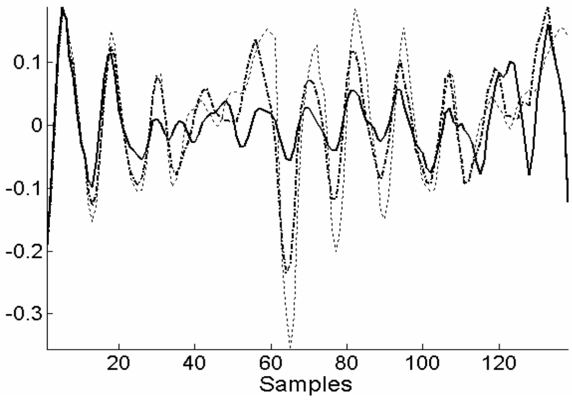

Fig. 13: *Result of the linear prediction analysis in the steady state area*

6 CONCLUSIONS AND FUTURE WORK

It can be seen that the used approach shows fair results with 30 LPC-coefficients and a gap of about 20 ms and good results using 300 LPC-coefficients. As expected this method proved to be more suitable for usage in the steady state part of a speech signal than in the transition area.

At the moment this attempt of reconstructing missing audio data works in 1-D domain only. As it is possible to convert 1-D audio data back to digital images future work could include the conversion of faulty 2-D audio data, e.g. an optical sound track to a 1-D signal for reconstruction of the gap and then converting it back to the 2-D audio signal.

REFERENCES

[1] Webers, J.: Handbuch der Film- und Videotechnik. Francis Verlag, 2000.

[2] Markel, J., Gray, A. H.: Linear Prediction of Speech. New York, Springer, 1976.

[3] Juang, B. H., Rabiner, L. R.: Fundamentals of Speech Recognition. Englewood Cliffs, NJ, Prentice Hall, 1993.

[4] Lee, L. M., Wang, H. C.: An Extended Levinson-Durbin Algorithm for the Analysis of Noisy Autoregressive Process. IEEE Signal Processing Letters, Vol. 3, 1996, No. 1, pp.13-15.