Restoration of variable density film soundtracks

Abdelâali Hassaïne, Etienne Decencière, Bernard Besserer

RESTORATION OF VARIABLE DENSITY FILM SOUNDTRACKS

A. Hassaïne[1], *E. Decencière*[1] *and B. Besserer*[2]

[1] Centre de Morphologie Mathématique
Ecole Nationale Supérieure des Mines de Paris
77305 Fontainebleau, France
{hassaine,decencie}@cmm.ensmp.fr

[2] Laboratoire Informatique, Image, Interaction
Université de La Rochelle
17042 La Rochelle, France
bernard.besserer@univ-lr.fr

ABSTRACT

The restoration of motion picture films has been an active research field for many years. The restoration of the soundtrack however has mainly been performed at the audio domain in spite of the fast that it is recorded as a continuous image on the film stock.

In this paper, we propose a new restoration method for variable density soundtracks. The method first detects and corrects accurately the azimuth deviation. A robust thresholding technique based on the minimization of the total variation is then performed to remove the remaining faults.

Restoration results are very promising and testify to the efficiency of our method.

1. INTRODUCTION

Most film stocks produced since the Thirties use an optical soundtrack to record the sound. According to the standard "Academy Optical Mono track" which was introduced by "The Academy of Motion Picture Arts and Sciences", the optical soundtrack is located in a space of 3 mm between the images of the film and the perforations. Therefore, duplicating a film makes it possible to duplicate its soundtrack at the same time. Two types of optical soundtracks exist: variable area soundtracks which contain a bright region located between two dark regions; the length of the bright region is directly proportional to the amplitude of the audio signal. And density variable soundtracks for which the brightness of the soundtrack, at each instant, is directly proportional to the signal audio amplitude (fig. 1).

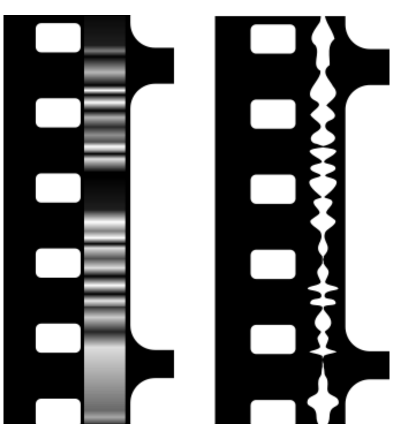


Figure 1: From left to right: variable density soundtrack, variable area soundtrack.

When the film is read, the soundtrack is lighted with a lamp through a narrow slit. The light intensity which goes through the soundtrack is measured and transformed to an electric signal by a photodetector. This signal is then converted to sound via a classical amplification chain (fig. 2). Note that this process is used for both variable area and variable density soundtracks.

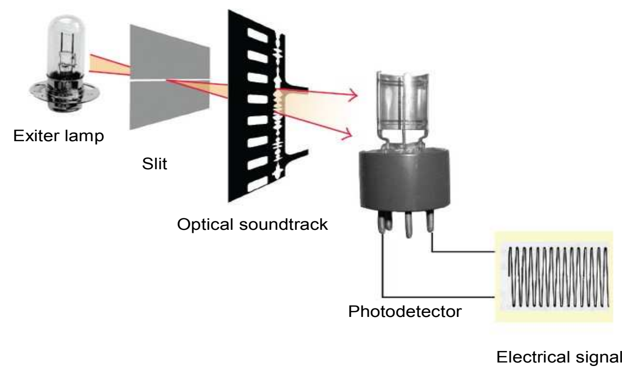


Figure 2: The process of reading a soundtrack.

Nowadays, optical recorders have a bandwidth of 20Hz to 14 kHz. The spatial resolution of the film stock used for optical soundtracks is about 100 lines per mm. Since a 35mm film travels at 456mm per second, the maximum "bandwidth" of the film itself as analog optical carrier does not exceed 22 kHz.

The defiler used in this work contains a linear camera able to scan 48000 lines of 512 pixel per second with a resolution of 3200 DPI and a gray level depth of 8 bits (fig. 3). This resolution is good enough for processing high frequencies (it meets the requirements of the "X"-curve [2]). For practical reasons, an image is saved each time 640 lines are captured. Therefore, we deal with images which are 512 pixels wide and 640 pixels high.

Unfortunately, the optical soundtrack undergoes the same type of degradations as the images of the film (dust, scratches, ...etc). Moreover, since the soundtrack is close to the extremity of the film, it is sometimes degraded by abrasion or altered over a large surface due to moisture. The soundtrack might also be badly exposed. This problem is due to the diffusion of light during different duplication processes.

In most cases, the restoration of the soundtrack is performed in the sound domain, using signal processing methods, in spite of the fact that it is recorded as a continuous

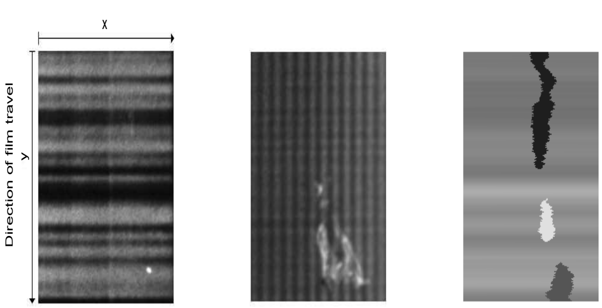
Figure 3: The scanner used in our work.

image between the film images and the perforations. Working with the image representation of the soundtrack has several advantages. First, the defects are visible at the image level. Most importantly, an image-based restoration makes it possible to preserve the authentic sound as it was originally recorded, whereas an audio-based restoration might cause the loss of this authenticity (by filtering a noise which was originally present in the sound for instance).

This work is undertaken within the "Resonances" project which was launched in 2005 and aims to propose an automatic restoration method for optical soundtracks. We proposed in [3] a restoration method for variable area soundtracks at the image level which gives generally good results. In this paper, we propose a new method for the restoration of variable density soundtracks using image processing tools.

2. EXPERIMENTAL DATABASES

Two variable density soundtracks have been used in this study. Both are taken from old French movies. The first one contains music (fig. 4(a)), whereas the second one contains voice (fig. 4(b)). We have also numerically generated a soundtrack corresponding to a restaurant conversation, on which we added simulated faults by putting random shapes at random places (fig. 4(c)).



(a) (b) (c)

Figure 4: (a) Cut-off of an image of the first soundtrack. (b) Cut-off of an image of the second soundtrack. (c) Cut-off of the simulated soundtrack.

Notice that the second track is vertically ridged. The reason for this remains unknown, sound experts from "GTC-Eclair Group" suppose however it is due to a grid that has been put to control the exposure.

Finally, in an ideal variable density soundtrack, each line at right angle to the direction of film travel has a constant gray value. When these lines of supposed constant gray values are not exactly at right angles to the direction of film travel, we talk about "azimuth deviation". This problem is due to the misalignment of either the line scan camera or the film recorder's optical slit. We will see further in this paper that the azimuth deviation in the digitized images is not nil, especially in the second database.

3. PREVIOUS WORK

Extremely few papers have been published on the restoration of variable density soundtrack at the image level. Streule proposed in [7] an image-based restoration system of optical soundtracks. For the restoration of density variable soundtracks, Streule only proposes an algorithm for azimuth correction. This algorithm does not detect the amount of azimuth deviation automatically and it only works in some special cases, as we will see further in this paper.

The method proposed by Richter et al. in [6] starts by correcting the fixed pattern noise which is due to sensors imperfection. These corrections are hardware-implemented in cameras nowadays and are not useful anymore [1]. The authors also proposed a statistical method for the restoration of each scanned line: First, by analyzing the gray values histogram of a specific line, the most recurring gray value is determined. After that, a lower and a upper threshold are set around this value. The pixels which have a gray value within these two thresholds remain unchanged, and the pixels which have a gray value outside these thresholds are replaced with the average gray value of all the pixels within the thresholds. Unfortunately, in case of very contrasted defects, the most recurring gray value does not correspond to the "correct gray value" of the processed line, but to the gray value of the defect in question. Moreover, we tested this method on our databases, and it generates a background noise due to sudden changes between lines.

J. Valenzuela patented in [9] several operations related to the digitization and processing of variable density soundtracks. The azimuth correction is claimed in this patent, but it is done by aligning the scan camera and not numerically. The patent claims also the operation of eliminating pixel values that deviate above a user defined threshold, but no method is proposed to determine this threshold automatically.

As a conclusion, the methods we found in the literature propose some simple solutions that can only deal with small defects. For the treatment of the whole range of defects we could find in our database, a new method must be developed.

4. PROPOSITION

The proposed approach consists of two steps. First, we correct the azimuth deviation of each scanned line. Then, we restore the gray level profile of the resulting lines.

4.1 Azimuth correction

In variable density soundtracks, an azimuth deviation results in the loss or reduction of high frequencies [4]. For this reason, even in the late 50's, an azimuth deviation of more than 0.000381 cm for a slit length of 0.24384 cm was considered inacceptable ([7] and [8]). The width of the scanned soundtracks being, in our case, 512 pixels, the azimuth deviation must not exceed 0.8 pixels. It is therefore important to consider subpixel accuracy to perform a good correction of azimuth.

As shown in figure 5(a), for angles up to 45 degrees, each pixel of a "skewed line" can be composed of the intensities of either two or three original pixels. The algorithm proposed by Streule [7] considers that each pixel is always composed of exactly two original pixels, the results obtained with his method are therefore not accurate enough.

We assume that the intensity of a new pixel is given by a weighted sum of all the pixels contributing to its intensity. The weights are done by the contributing area of each original pixel. We consider separately the case of two contributing pixels (fig. 5(b) and three contributing pixels (fig. 5(c)). After performing some geometrical calculations, we obtain Algorithm 1 that computes the intensities of a line skewed by δy in the image I.

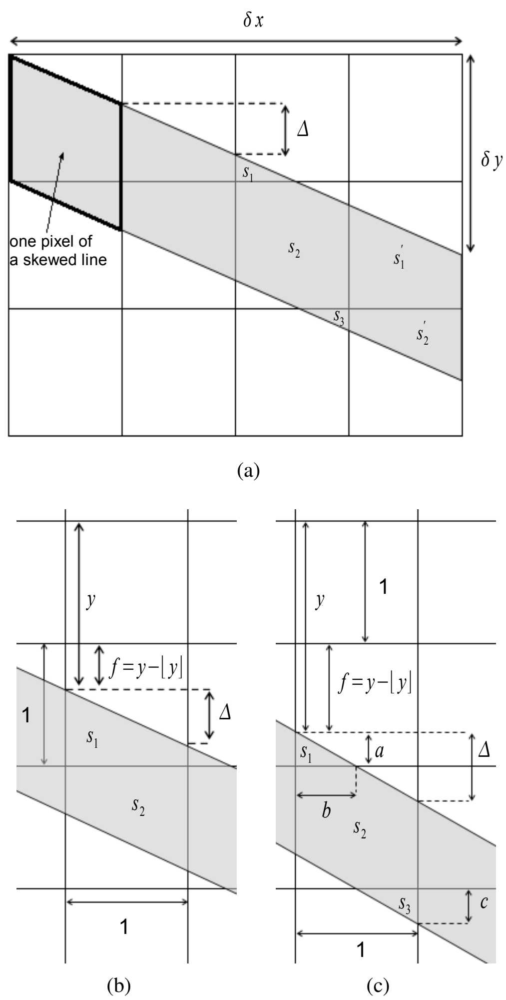


Figure 5: (a) Illustration of a skewed line. (b) Pixel composed of the intensities of two original pixels. (c) Pixel composed of the intensities of three original pixels.

To compute a skewed line with a negative δy, we use a similar algorithm with ceiling functions ($\lceil\ \rceil$) instead of floor functions ($\lfloor\ \rfloor$).

Algorithm 1 Calculation of the intensities of a skewed line

$\Delta \leftarrow |\delta y|/width$
$y \leftarrow$ *initial value of y*
for $x = 1$ to width **do**
  if $\lfloor y \rfloor = \lfloor y+\Delta \rfloor$ **then**
    $f \leftarrow y - \lfloor y \rfloor$
    $s_1 \leftarrow 1 - \Delta/2 - f$
    $s_2 \leftarrow 1 - s_1$
    $current\_intensity \leftarrow s_1 \cdot I[x, \lfloor y \rfloor] + s_2 \cdot I[x, \lfloor y \rfloor + 1]$
  else
    $f \leftarrow y - \lfloor y \rfloor$
    $a \leftarrow 1 - f$
    $b \leftarrow a/\Delta$
    $s_1 \leftarrow a \cdot b/2$
    $c \leftarrow \Delta \cdot (1-b)$
    $s_3 \leftarrow (1-b) \cdot c/2$
    $s_2 \leftarrow 1 - s_1 - s_3$
    $current\_intensity \leftarrow s_1 \cdot I[x, \lfloor y \rfloor] + s_2 \cdot I[x, \lfloor y \rfloor + 1] +$
    $s_3 \cdot I[x, \lfloor y \rfloor + 2]$
  end if
  $y \leftarrow y + \Delta$
end for

Since all pixels of a line at right angle to the direction of film travel are supposed to be constant in a variable density soundtrack, we are interested in the value of δy for which all the obtained intensities are equal, or at least as close to each other as possible. This value is the one which minimizes the variance of these intensities. Unfortunately, because of local faults, this measure might give wrong results and has to be averaged over long sections of the soundtrack.

In the present work, we computed for each value of δy from -5 to 5 with a step of 0.1, the sum of variances of the obtained intensities over 6400 lines. Figure 6 shows the results for the first 6400 lines of the second soundtrack. The value $\delta y = -2.8$ minimizes the sum of variances and is the one to be used to correct the azimuth deviation. By processing in a similar way for the first soundtrack, the obtained result is $\delta y = -0.5$.

After correcting the azimuth deviation using the obtained value of δy, we noticed indeed a gain in high frequencies. Fig. 7 shows the spectrogram corresponding to a part of the second soundtrack before and after azimuth correction. Notice that for high frequencies, the gray levels in the spectrogram after azimuth correction are darker. This confirms that high frequencies are more significant after azimuth correction.

4.2 Restoration of the resulting lines

The obtained results might contain some local faults. To remove them, we decided to proceed as follows: For each line at right angle to the direction of film travel, we compute the average value v. Let t be a threshold such as, all the pixels that have a gray value larger than $v+t$ will be substituted by $v+t$ and the pixels that have a gray value smaller than $v-t$ will be substituted by $v-t$. The resulting gray values will be within $[v-t, v+t]$. If t is set to a small value, the dynamic of the gray values of each line will be very small, and the average gray value of intensities might change suddenly when

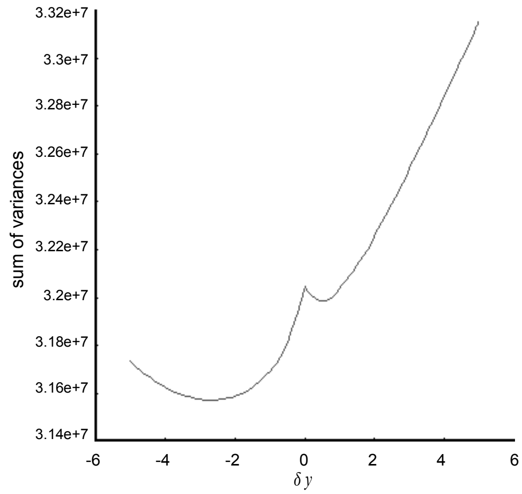


Figure 6: Graph showing the sum of variances VS. the amount of deviation.

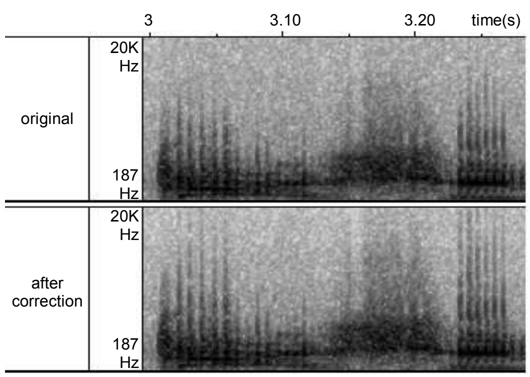


Figure 7: Audio spectrogram before and after azimuth correction.

moving from one line to the next. If in contrast, t is set to a large value, only the pixels that have a gray value very different from the average will be substituted and the final result will not be of a significant improvement. We have chosen to set t to the value which minimizes the noise, and minimizes therefore, the average gray value when moving from one line to the next.

In practice, to set t automatically we proceed, for each value of t, as follows:

- Let $L_{y,t}$ be a line "thresholded" at t (by the method previously defined).
- We compute for each line $L_{y,t}$ the sum of the gray values $sum_{y,t}$.
- Then, we compute the total variation [5] of the obtained sums: $total\ variation = \sum_{y=1}^{N} \left| sum_{y,t} - sum_{y-1,t} \right|$ (N being the number of lines, empirically set to 6400 lines).
- The chosen value of t is the one that minimizes the total variation.

The curve in figure 8 shows the relationship between the total variation and the threshold value for the 6400 first lines of the first soundtrack. The threshold that minimizes the total variation is 15. The same curve shows a similar behavior for both the second and the simulated soundtrack but achieves its minimum at 58 in the second soundtrack, because of the relatively large dynamic of gray values in this soundtrack, and achieves its minimum at 13 in the simulated soundtrack.

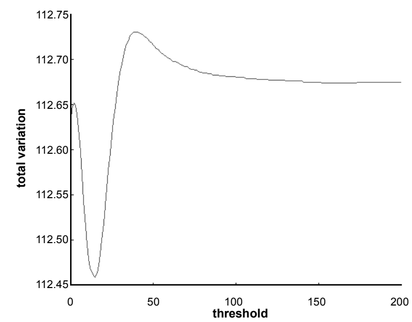
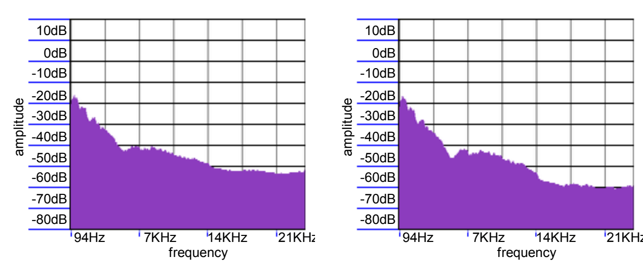


Figure 8: Total variation VS. the threshold value.

Figure 9 illustrates restoration examples taken from our three soundtracks. Notice that the faults are generally removed or at least reduced. The ridges in the second soundtrack are not removed, and we believe, without an a priori knowledge on whether the dark or the bright ridges should be favored, this task cannot be achieved.

In the digitized soundtracks, the defects are only located in some regions and not present all along the soundtrack, their effect is therefore insignificant on the audio domain. The restoration method preserves the quality of the audio signal and it does not generate parasitic frequencies (in contract to Richter's method). In the simulated soundtrack however, we noticed a significant reduction of the noise caused by the added defects. The frequency spectrums of this soundtrack before and after the restoration shows a reduction of high frequencies corresponding to noise (fig. 10).



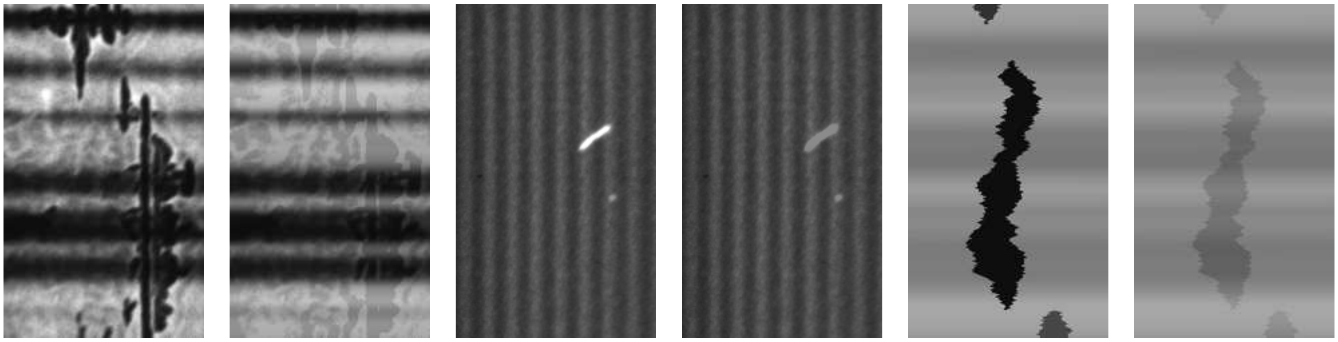
(a) (b)

Figure 10: Frequency spectrum of the simulated soundtrack (a) before restoration. (b) after restoration.

A final step aiming to replace all the gray values of the same line by their 16 bits average value can be considered to produce better looking images, but this step will not change the resulting audio signal.

This method fails to restore images for which the fault's area is larger than the rest of the image (fig. 11). We believe that this type of images can be restored by performing an interpolation of the audio signal in the frequency domain.

When using a 2.13 GHz Pentium 4 processor with 2GB RAM, the determination of the amount of azimuth deviation takes about 30 seconds and can be optimized by computing the variance periodically instead of computing it for all the



(a) (b) (c) (d) (e) (f)

Figure 9: Cut-off of the first soundtrack (a) before restoration. (b) after restoration. Cut-off of the second soundtrack (c) before restoration. (d) after restoration. Cut-off of the simulated soundtrack (e) before restoration. (f) after restoration.

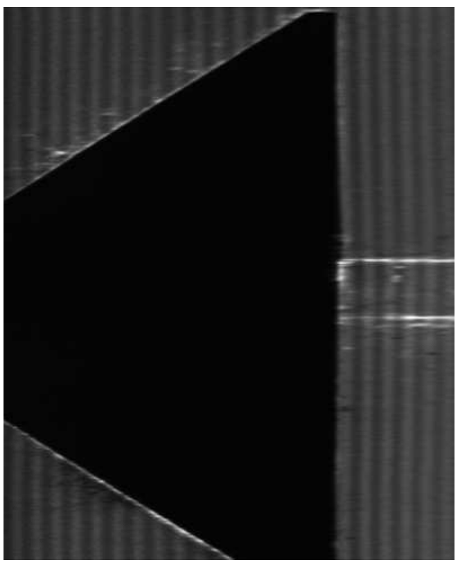


Figure 11: A defect present over a large area.

lines. Azimuth correction takes 60 seconds for processing 64000 lines. Determining the threshold value takes about 300 seconds (if we perform the operations for all the 256 possible thresholds). And the final correction using the obtained threshold takes 30 seconds for processing 64000 lines.

5. CONCLUSION

We proposed a new method for the restoration of variable density soundtracks. This method provides an accurate azimuth correction algorithm and a thresholding technique based on the minimization of total variation when moving from one line of the soundtrack to the next, thus allowing an automatic parametrization of the method. We are currently working on the development of restoration techniques that take into account the spatial distribution of defects instead of processing them on each line separately. The effect of bad exposure on variable density soundtracks is also of interest. Looking at the frequency domain to perform a restoration of defects located over a large area needs to be studied as well. Finally, all these developments have to be integrated into a complete restoration system that handles all the processes from the digitization of the soundtrack to the generation of the restored audio file.

6. ACKNOWLEDGMENTS

This work has been undertaken within the Resonances Project (http://www.riam-resonances.org/) and was made possible thanks to the financial help of the French Agence Nationale de la Recherche, through its RIAM program. The film material, as well as the expertise on motion picture optical soundtracks, were provided by N. Ricordel from the "Centre National de Cinématographie-Archives Françaises du Film" and by C. Comte from "GTC-Eclair Group".

REFERENCES

[1] ATMEL Application Note, *Flat Field Correction, How Flat Field Correction and Contrast Expansion are implemented in AViiVA cameras*. 2005.

[2] ISO2969: B-chain electro-acoustic response of motion-picture control rooms and indoor theatres – specifications and measurements (1987) International standard.

[3] E. Brun, A. Hassaïne, B. Besserer and E. Decencière, "Restoration of variable area soundtrack," in *Proc. of 2007 International Conference on Image Processing*, San Antonio, USA. september 2007.

[4] J. G. Frayne, H. Wolfe, *Elements of sound recording*. John Wiley & Sons, New York, 1949.

[5] S. Osher and L. Rudin, " Feature-oriented image enhancement using shock filters," *IEEE transactions on image processing* vol. 27, pp. 919–940, 1990.

[6] D. Richter, I. Kurreck and D. Poetsch, "Restoration of optical variable density sound track on motion picture films by digital image processing," in *Proc. of the 7th International Conference on Optimization of Electrical and Electronic Equipments*, University of Brasov, Romania, 2002.

[7] P. Streule, *Digital image based restoration of optical movie soundtracks*. Master Thesis. ETH Zurich. 1999.

[8] H. M. Tremaine, *Audio Cyclopedia*. Howard W. Sams & co., Indianapolis, First Printing, 1959.

[9] J. Valenzuela, *Digital reproduction of variable density film soundtracks*. United states patent 20060232745, 2006.