Лектор: Лукин Алексей Сергеевич

MSU GRAPHICS & MEDIA LAB

Линейное предсказание, интерполяция аудио

План

- Корреляция и автокорреляция, их применения
- Линейное предсказание
  - Авторегрессионная модель сигнала
  - Нахождение коэффициентов регрессии
  - Применения
    - Сжатие
    - Интерполяция
- LSAR-интерполяция звука
- Подавление искажений перегрузки и щелчков

Корреляция

- Корреляция (кросс-корреляция): мера похожести двух сигналов при различных сдвигах *k* одного сигнала относительно другого

$$r_{xy}(k) = E[x(m)y(m-k)]$$

- Оценка корреляции $$r_{xy}(k) = \frac{1}{N}\sum_{m=0}^{N-1} x(m)y(m-k)$$
- Вычисление через FFT $$r_{xy}(m) = x(m) * y(-m)$$
- Применения корреляции: поиск похожих фрагментов сигналов, поиск сдвига кадра в видео

Автокорреляция

MSU GRAPHICS & MEDIA LAB

- Автокорреляция: мера похожести сигнала на собственные сдвинутые копии $r_{xx}(k) = E[x(m)x(m-k)]$

$$R_{xx} = E[x(m-m_1)x(m-m_2)]$$

- Оценка автокорреляции $r_{xx}(k) = \frac{1}{N}\sum_{m=0}^{N-1} x(m)x(m-k)$
- Вычисление через FFT $r_{xx}(m) = x(m) * x(-m)$
- Применение автокорреляции: оценка основного тона звукового сигнала, поиск периодичности

Линейное предсказание

MSU GRAPHICS & MEDIA LAB

- Линейное предсказание (LPC) $\bar{x}(m) = \sum_{k=1}^{P} a_k x(m-k)$
- Ошибка предсказания $e(m) = x(m) - \bar{x}(m) = x(m) - \sum_{k=1}^{P} a_k x(m-k)$
- Авторегрессионная модель сигнала

$$x(m) = \sum_{k=1}^{P} a_k x(m-k) + e(m)$$

Линейное предсказание

- Нахождение наилучших параметров регрессионной модели

$$E[e^2(m)] = E\left[\left(x(m) - \sum_{k=1}^{P} a_k x(m-k)\right)^2\right] =$$

$$= E[x^2(m)] - 2\sum_{k=1}^{P} a_k E[x(m)x(m-k)] + \sum_{k=1}^{P} a_k \sum_{j=1}^{P} a_j E[x(m-k)x(m-j)] =$$

$$= r_{xx}(0) - 2r_{xx}^T a + a^T R_{xx} a$$

$$\frac{\partial}{\partial a} E[e^2(m)] = -2r_{xx}^T + 2a^T R_{xx} \qquad \frac{\partial}{\partial a} = \left(\frac{\partial}{\partial a_1}, \frac{\partial}{\partial a_2}, \ldots, \frac{\partial}{\partial a_P}\right)$$

Линейное предсказание

MSU GRAPHICS & MEDIA LAB

- Нахождение наилучших параметров регрессионной модели

$$\frac{\partial}{\partial a} E[e^2(m)] = -2r_{xx}^T + 2a^T R_{xx} \qquad \frac{\partial}{\partial a} = \left(\frac{\partial}{\partial a_1}, \frac{\partial}{\partial a_2}, \ldots, \frac{\partial}{\partial a_P} \right)$$

- Приравниваем градиент нулю

$$R_{xx} a = r_{xx} \Rightarrow a = R_{xx}^{-1} r_{xx}$$

- Матрица R_{xx} – тёплицева, обращаем рекурсией Левинсона-Дурбина за P^2 операций

Линейное предсказание

- Составляющие ошибки
  1. Особенности сигнала, не описываемые моделью
  2. Неточность параметров модели
  3. Шум
- Как выбрать число параметров модели?
  - ► Модель порядка P может точно моделировать смесь P/2 синусоид с различными частотами и амплитудами
  - ► Выше порядок → меньше ошибка предсказания (но хуже стабильность вычислений)

Линейное предсказание

MSU GRAPHICS & MEDIA LAB

- Применения
  - Реставрация сигнала (интерполяция/экстраполяция пропущенных отсчетов)
  - Компрессия сигнала (достаточно хранить коэффициенты модели и сигнал ошибки)

LSAR-интерполяция

- Пусть неизвестный интервал окружен известными отсчетами:

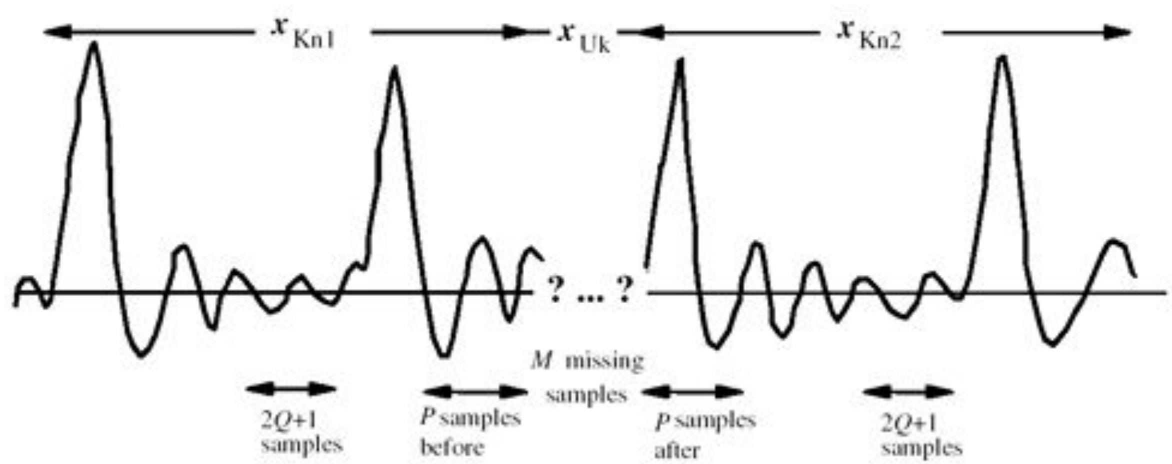


Figure 10.11 A signal with M missing samples. P immediate samples each side of the gap and $2Q+1$ samples a pitch period away are used for interpolation.

По материалам книги S. Vaseghi “Advanced Digital Signal Processing and Noise Reduction”

LSAR-интерполяция

- Запишем ошибку линейного предсказания (предполагая, что коэффициенты известны):

$$
\left(\begin{array}{c}
e(P) \\ e(P+1) \\ \vdots \\ e(k-1) \\ \hline
e(k) \\ e(k+1) \\ e(k+2) \\ \vdots \\ e(k+M+P-2) \\ e(k+M+P-1) \\ \hline
e(k+M+P) \\ e(k+M+P+1) \\ \vdots \\ e(N-1)
\end{array}\right)
=
\left(\begin{array}{c}
x(P) \\ x(P+1) \\ \vdots \\ x(k-1) \\ \hline
x_{Uk}(k) \\ x_{Uk}(k+1) \\ x_{Uk}(k+2) \\ \vdots \\ x(k+M+P-2) \\ x(k+M+P-1) \\ \hline
x(k+M+P) \\ x(k+M+P+1) \\ \vdots \\ x(N-1)
\end{array}\right)
-
\left(\begin{array}{cccc}
x(P-1) & x(P-2) & \cdots & x(0) \\
x(P) & x(P-1) & \cdots & x(1) \\
\vdots & \vdots & \ddots & \vdots \\
x(k-2) & x(k-3) & \cdots & x(k-P-1) \\ \hline
x(k-1) & x(k-2) & \cdots & x(k-P) \\
x_{Uk}(k) & x(k-1) & \cdots & x(k-P+1) \\
x_{Uk}(k+1) & x_{Uk}(k) & \cdots & x(k-P+2) \\
\vdots & \vdots & \ddots & \vdots \\
x(k+M+P-3) & x(k+M+P-2) & \cdots & x_{Uk}(k+M-2) \\
x(k+M+P-2) & x(k+M+P-1) & \cdots & x_{Uk}(k+M-1) \\ \hline
x(k+M+P-1) & x(k+M+P) & \cdots & x(k+M) \\
x(k+M+P) & x(k+M+P+1) & \cdots & x(k+M+1) \\
\cdots & \cdots & \ddots & \cdots \\
x(N-2) & x(N-3) & \cdots & x(N-P-1)
\end{array}\right)
\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_P \end{pmatrix}
$$

$$e(X_{Uk}, a) = x - Xa$$

По материалам книги S. Vaseghi “Advanced Digital Signal Processing and Noise Reduction”

LSAR-интерполяция

- Перепишем, отделив неизвестные отсчеты:

$$
\begin{pmatrix} e(k) \\ e(k+1) \\ e(k+2) \\ e(k+3) \\ e(k+4) \\ \vdots \\ e(k+P-1) \\ e(k+P) \\ e(k+P+1) \\ \vdots \\ e(k+M+P-2) \\ e(k+M+P-1) \end{pmatrix}
=
\begin{pmatrix}
1 & 0 & 0 & 0 & \cdots & 0 \\
-a_1 & 1 & 0 & 0 & \cdots & 0 \\
-a_2 & -a_1 & 1 & 0 & \cdots & 0 \\
-a_3 & -a_2 & -a_1 & 1 & \cdots & 0 \\
-a_4 & -a_3 & -a_2 & -a_1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
-a_P & -a_{P-1} & -a_{P-2} & -a_{P-3} & \cdots & 0 \\
0 & -a_P & -a_{P-1} & -a_{P-2} & \cdots & 0 \\
0 & 0 & -a_P & -a_{P-1} & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & -a_{P-1} \\
0 & 0 & 0 & 0 & \cdots & -a_P
\end{pmatrix}
\begin{pmatrix} x_{Uk}(k) \\ x_{Uk}(k+1) \\ x_{Uk}(k+2) \\ x_{Uk}(k+3) \\ \vdots \\ x_{Uk}(k+M-1) \end{pmatrix}
+
$$

$$
\begin{pmatrix}
-a_P & -a_{P-1} & -a_{P-2} & \cdots & -a_1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & -a_P & -a_{P-1} & \cdots & -a_2 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & -a_P & \cdots & -a_3 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -a_P & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & -a_1 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & -a_2 & -a_1 & 1 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & -a_3 & -a_2 & -a_1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0 \\
0 & 0 & 0 & \cdots & 0 & 0 & \cdots & -a_{P-1} & -a_{P-2} & -a_{P-3} & \cdots & -a_1
\end{pmatrix}
\begin{pmatrix} x(k-P) \\ x(k-P+1) \\ x(k-P+2) \\ \vdots \\ x(k-1) \\ 0 \\ \vdots \\ x(k+M) \\ x(k+M+1) \\ x(k+M+2) \\ \vdots \\ x(k+M+P-1) \end{pmatrix}
$$

$$e = A_1 x_{Uk} + A_2 x_{Kn}$$

По материалам книги S. Vaseghi “Advanced Digital Signal Processing and Noise Reduction”

LSAR-интерполяция

- Минимизируем ошибку предсказания:

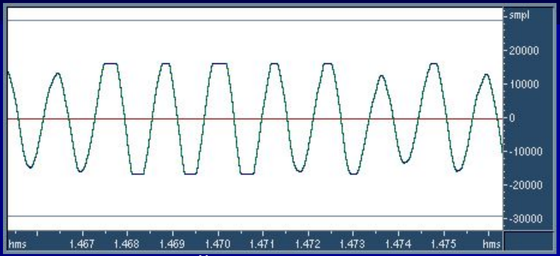
$$\frac{\partial \boldsymbol{e}^{\mathrm{T}}\boldsymbol{e}}{\partial \boldsymbol{x}_{\mathrm{Uk}}}=2\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{x}_{\mathbf{Uk}}+2\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{x}_{\mathrm{Kn}}=0$$

$$\hat{\boldsymbol{x}}_{\mathrm{Uk}}^{LSAR}=-\left(\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_1\right)^{-1}\left(\boldsymbol{A}_1^{\mathrm{T}}\boldsymbol{A}_2\right)\boldsymbol{x}_{\mathrm{Kn}}$$

- Проблема: коэффициенты LPC неизвестны
- Решение: вычислим их приблизительно, затем – оценим x_{Uk} и снова вычислим более точные коэффициенты…

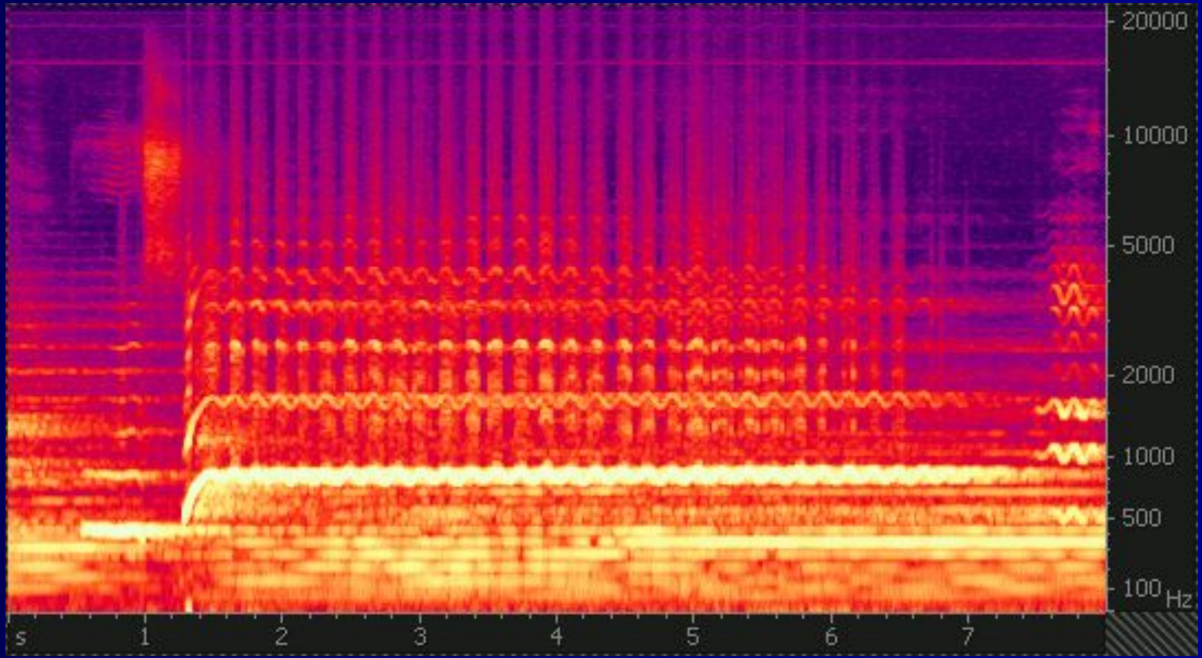
Исправление перегрузки

- Исходная запись с перегрузкой *(clipping)*



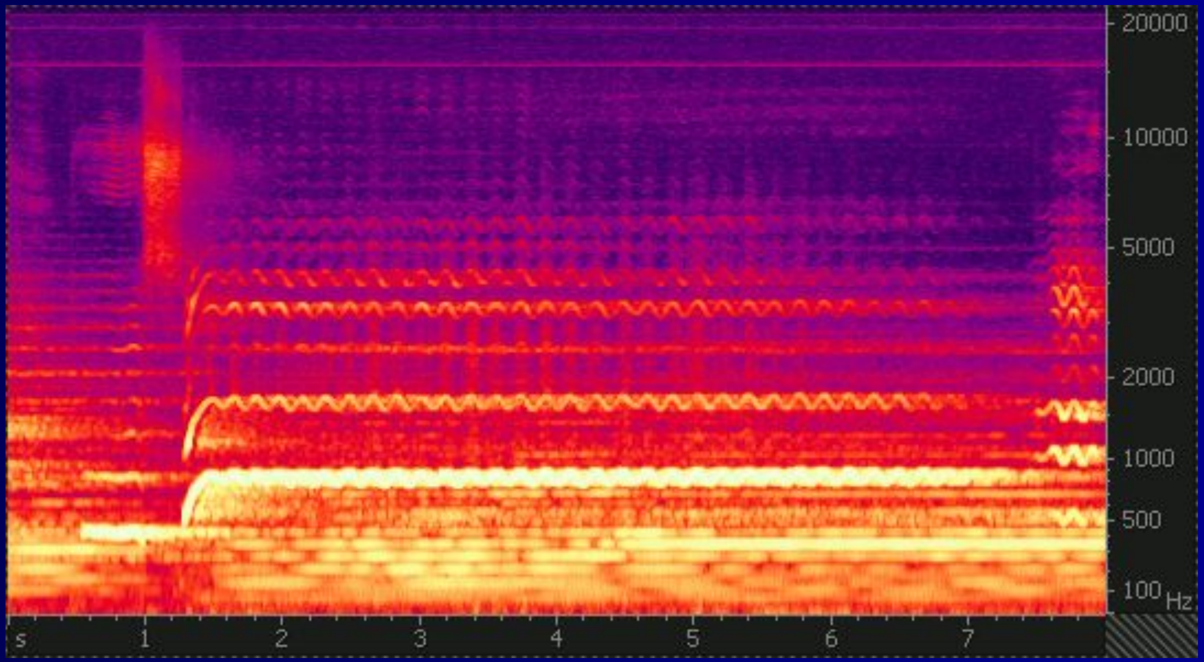
короткий фрагмент волны

женский вокал с перегрузкой: спектрограмма



Исправление перегрузки

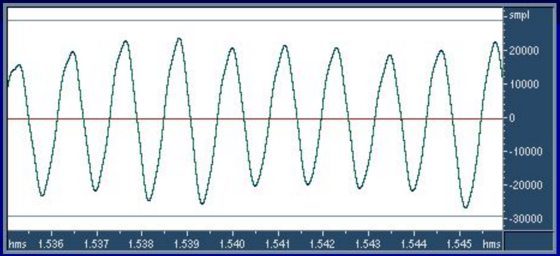
- После одной итерации LSAR-интерполяции

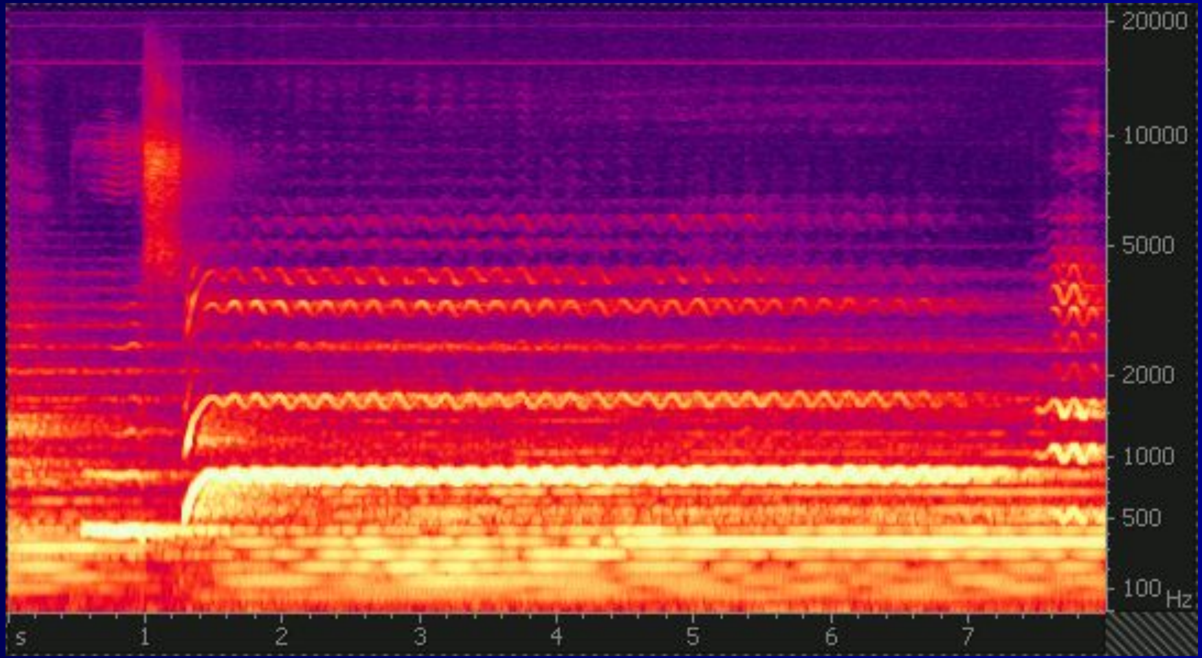


Исправление перегрузки

MSU GRAPHICS & MEDIA LAB

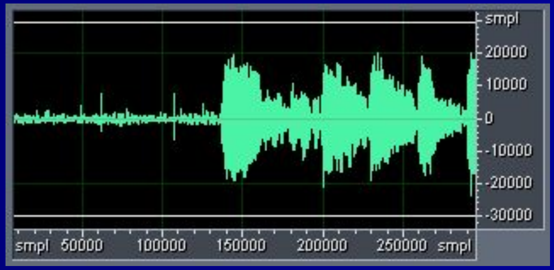
- После трех итераций





Исправление щелчков

- Исходная запись со щелчками
  (clicks)

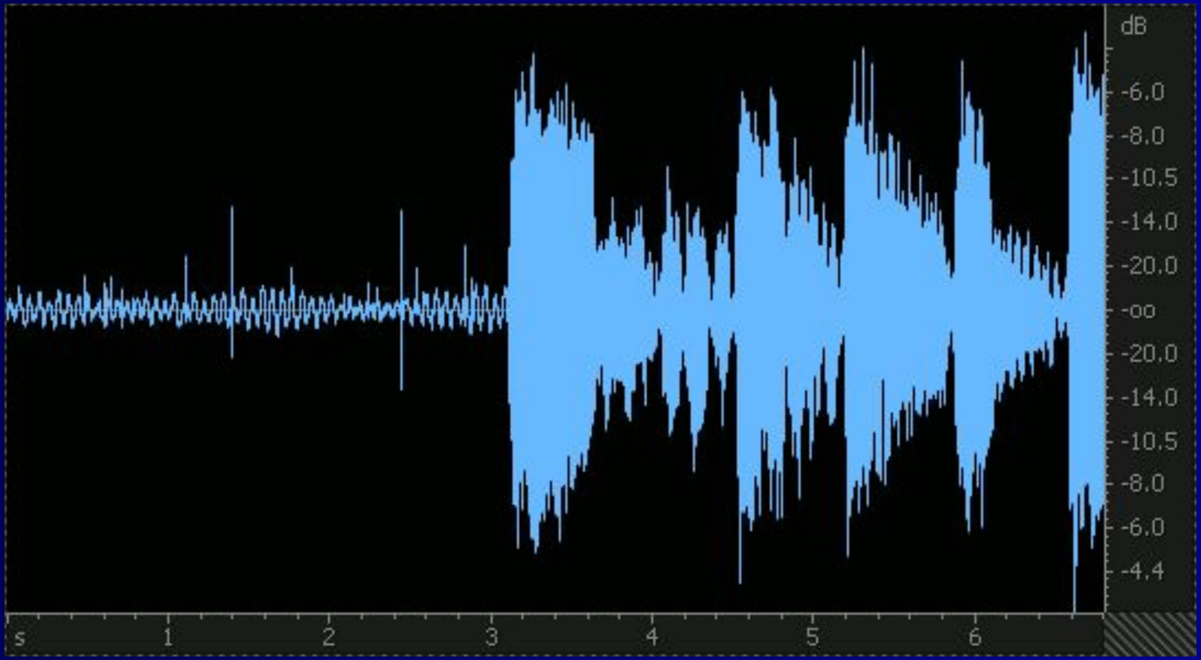


- Детектирование щелчков
  - ► Анализ разности между соседними отсчетами
  - ► Анализ ошибки LPC (в т.ч. – многополосный)
  - ► Анализ спектрограммы
- Интерполяция щелчков

Исправление щелчков

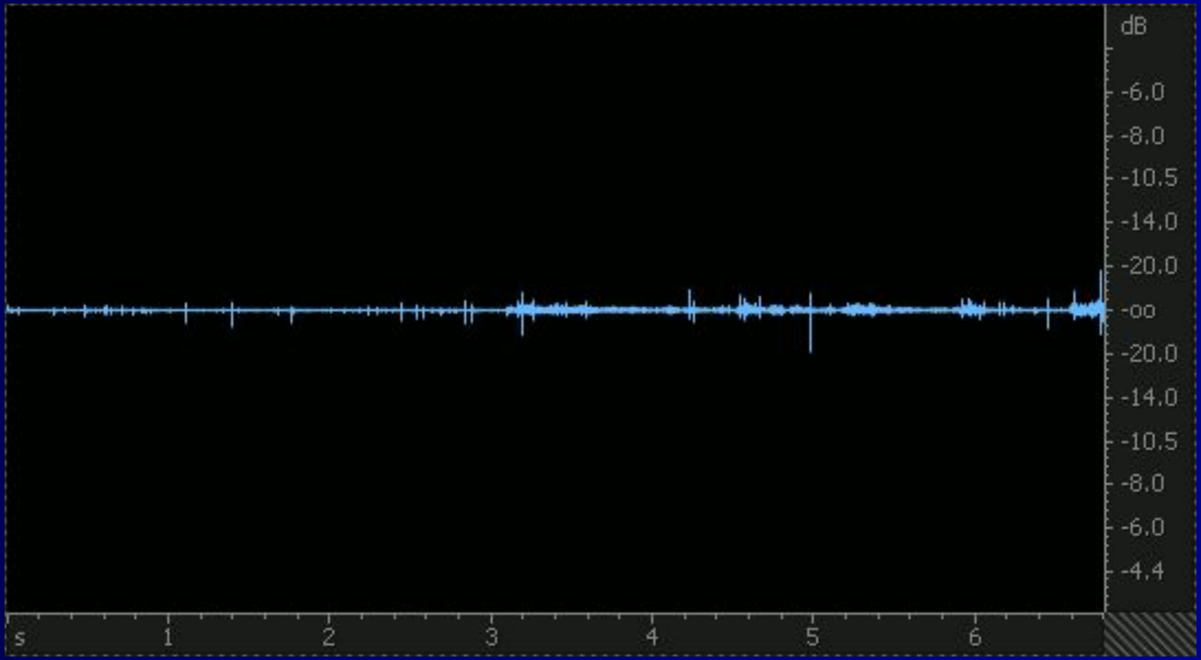
- Исходная запись со щелчками
  (clicks)

оркестр, записанный с виниловой пластинки: волна



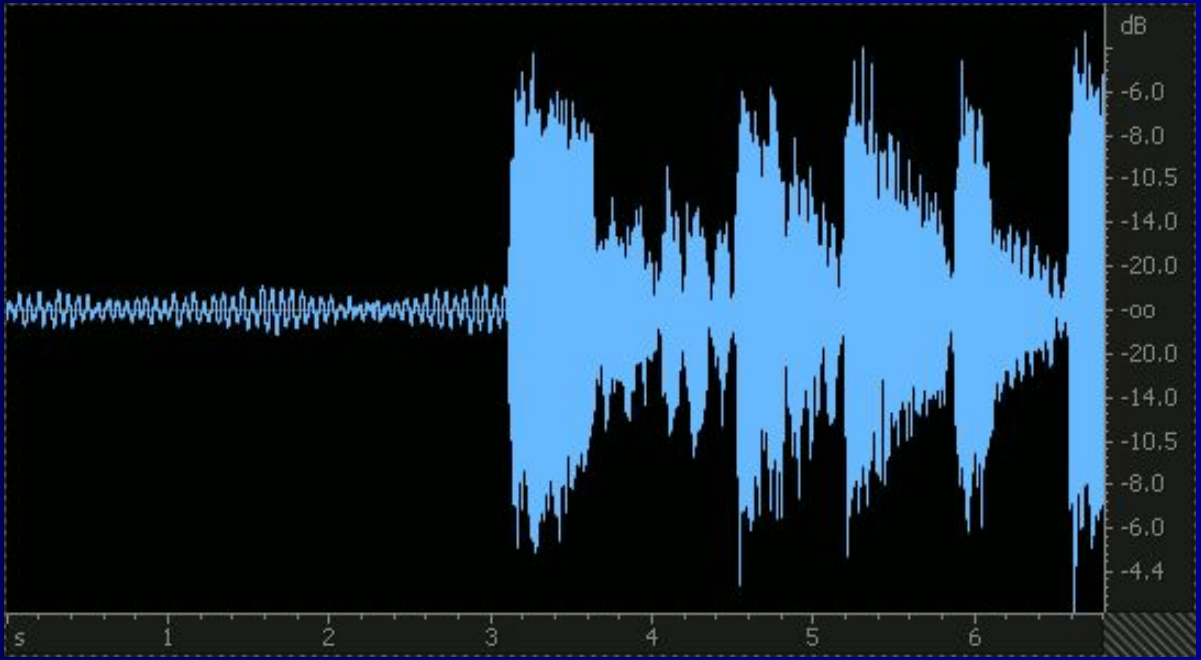
Исправление щелчков

- Остаточный сигнал LPC
  - ► Порядок предсказания – 100 (но можно было и меньше)



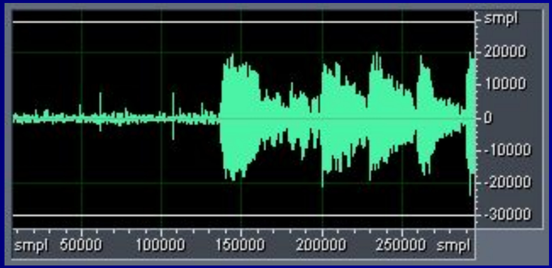
Исправление щелчков

- После детектирования и интерполяции щелчков методом LSAR (3 итерации)

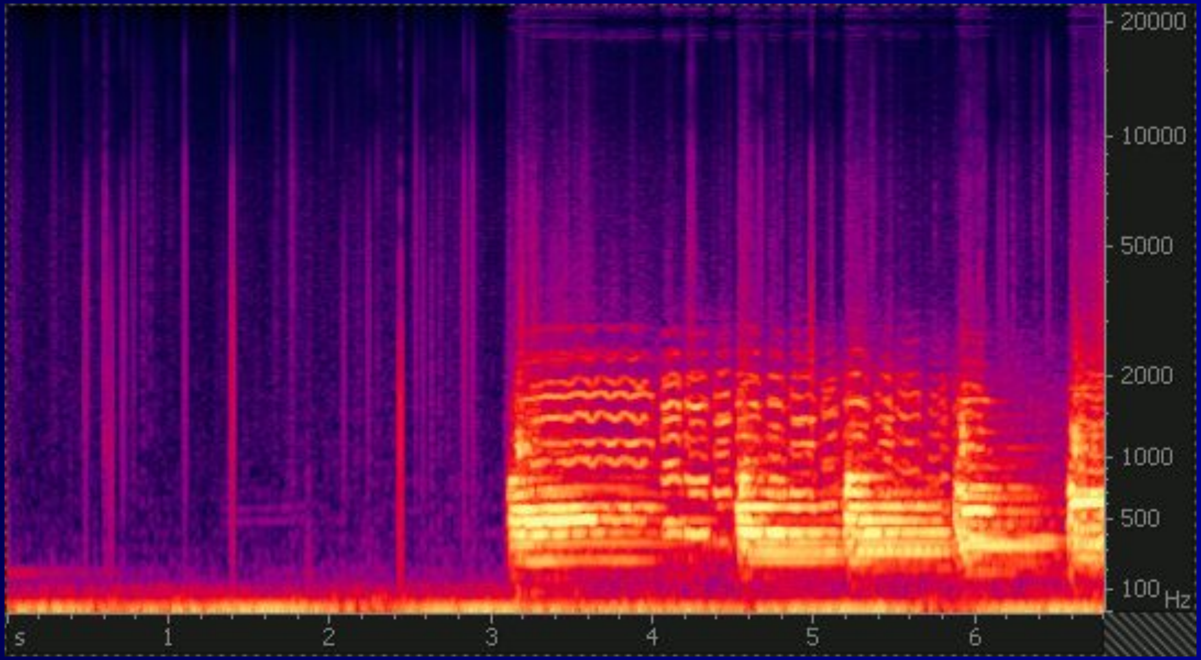


Исправление щелчков

- Исходная запись со щелчками
  (clicks)



оркестр, записанный с виниловой пластинки: спектрограмма



Исправление щелчков

- После автоматического обнаружения и интерполяции щелчков