



Лектор: Лукин Алексей Сергеевич

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



- Автокорреляция: мера схожести 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



- 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

$$\begin{aligned} E[e^2(m)] &= E\left[\left(x(m) - \sum_{k=1}^P a_k x(m-k)\right)^2\right] = \\ &= E[x^2(m)] - 2 \sum_{k=1}^P a_k E[x(m)x(m-k)] + \sum_{k=1}^P a_k \sum_{j=1}^P a_j E[x(m-k)x(m-j)] = \\ &= r_{xx}(0) - 2r_{xx}^T a + a^T R_{xx} a \end{aligned}$$

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

Линейное предсказание



- Нахождение наилучших параметров регрессионной модели

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

- Приравниваем градиент нулю

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

- 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 \rightarrow меньше ошибка предсказания (но хуже стабильность вычислений)

Линейное предсказание



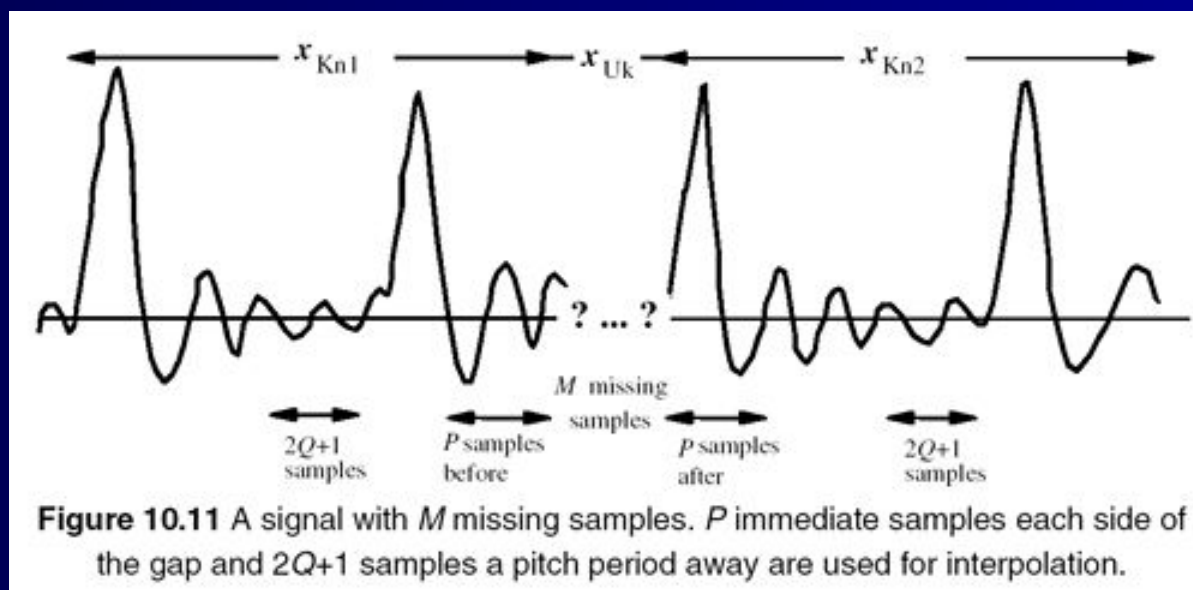
- Применения

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

LSAR-интерполяция



- Пусть неизвестный интервал окружен известными отсчетами:



LSAR-интерполяция



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

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

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

LSAR-интерполяция



- Перепишем, отделив неизвестные отсчеты:

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

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

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

LSAR-интерполяция



- Минимизируем ошибку предсказания:

$$\frac{\partial e^T e}{\partial x_{Uk}} = 2A_1^T A_1 x_{Uk} + 2A_1^T A_2 x_{Kn} = 0$$

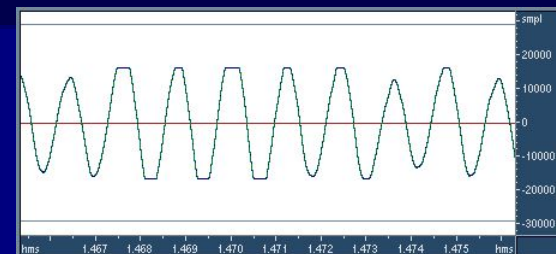
$$\hat{x}_{Uk}^{LSAR} = -(A_1^T A_1)^{-1} (A_1^T A_2) x_{Kn}$$

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

Исправление перегрузки

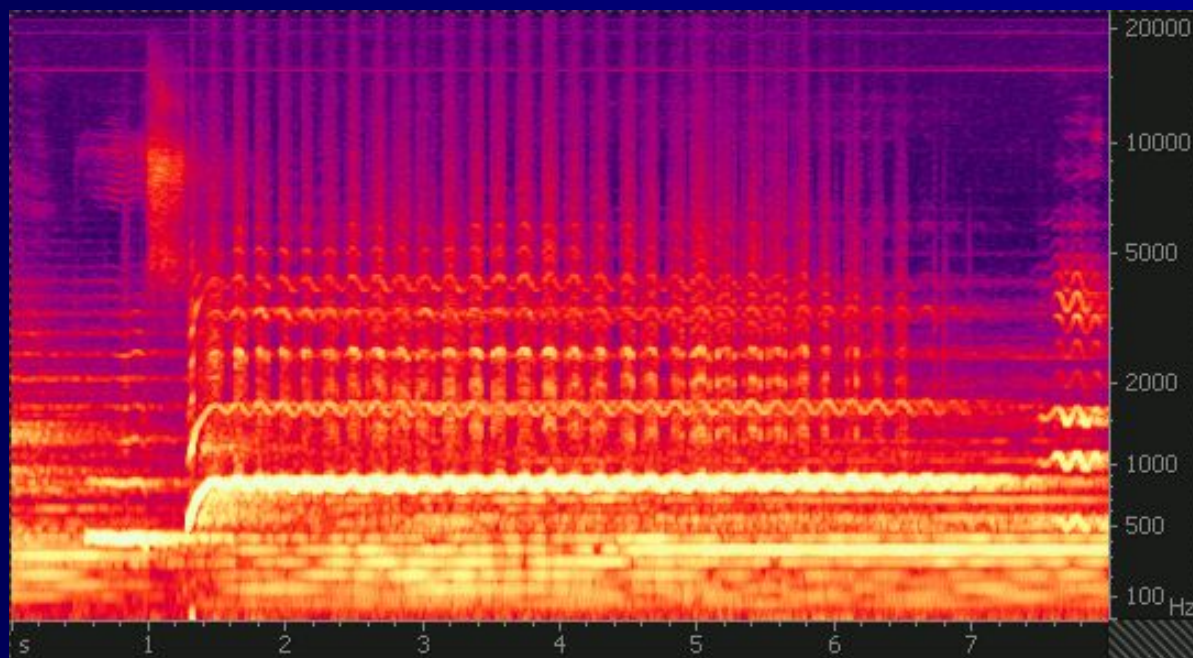


- Исходная запись с перегрузкой
(clipping)



короткий фрагмент волны

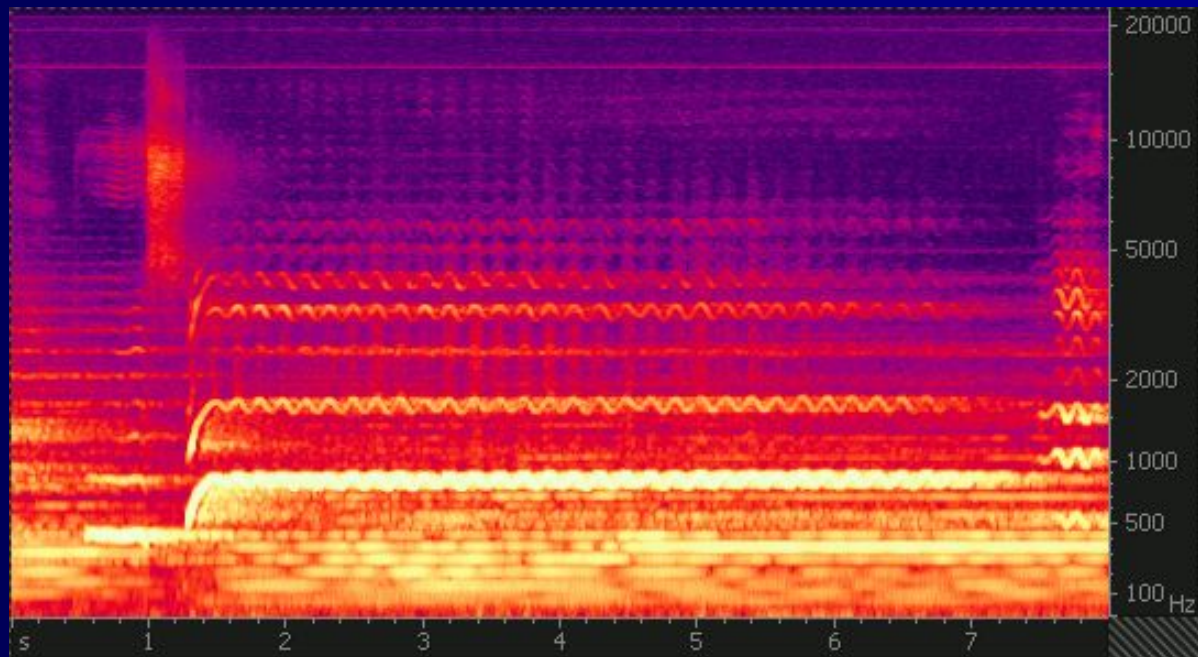
женский вокал с перегрузкой: спектрограмма



Исправление перегрузки



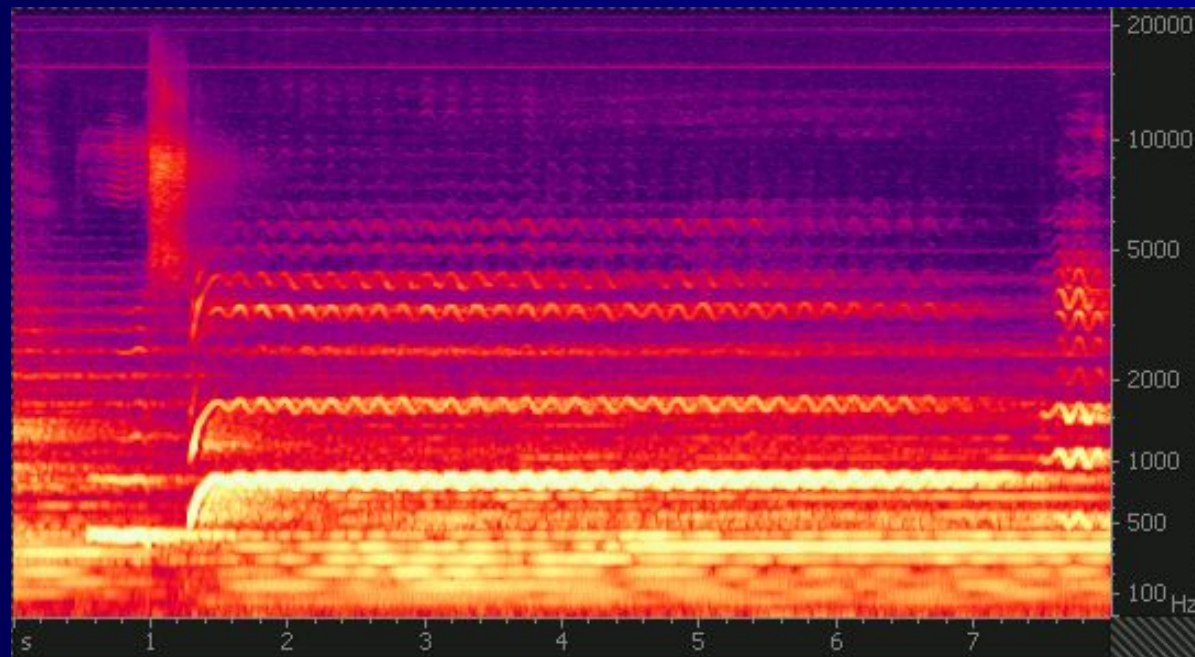
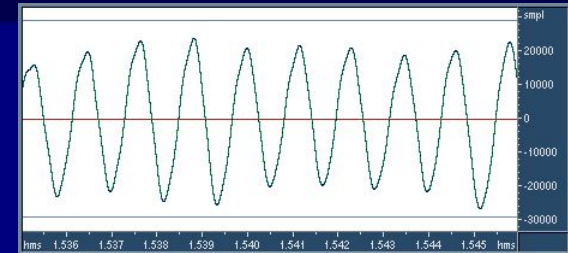
- После одной итерации LSAR-интерполяции



Исправление перегрузки



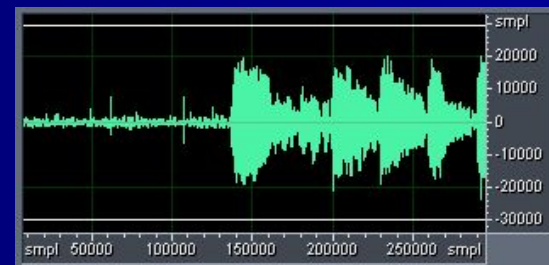
- После трех итераций



Исправление щелчков



- Исходная запись со щелчками
(clicks)



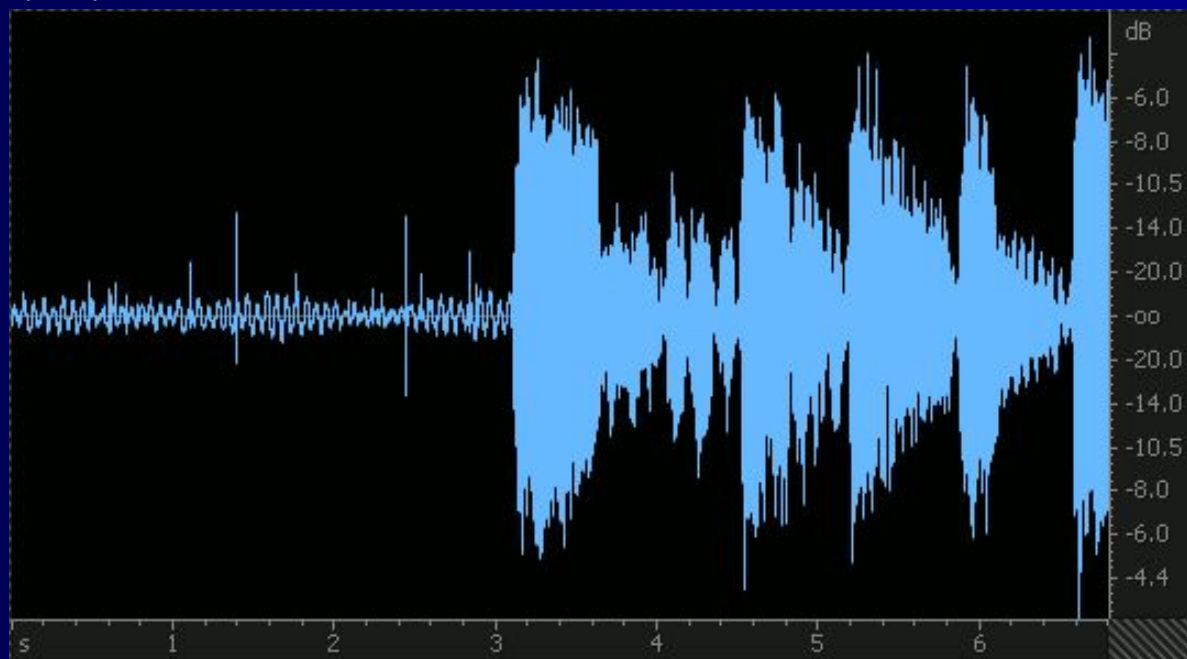
- Детектирование щелчков
 - ▶ Анализ разности между соседними отсчетами
 - ▶ Анализ ошибки LPC (в т.ч. – многополосный)
 - ▶ Анализ спектрограммы
- Интерполяция щелчков

Исправление щелчков



- Исходная запись со щелчками
(clicks)

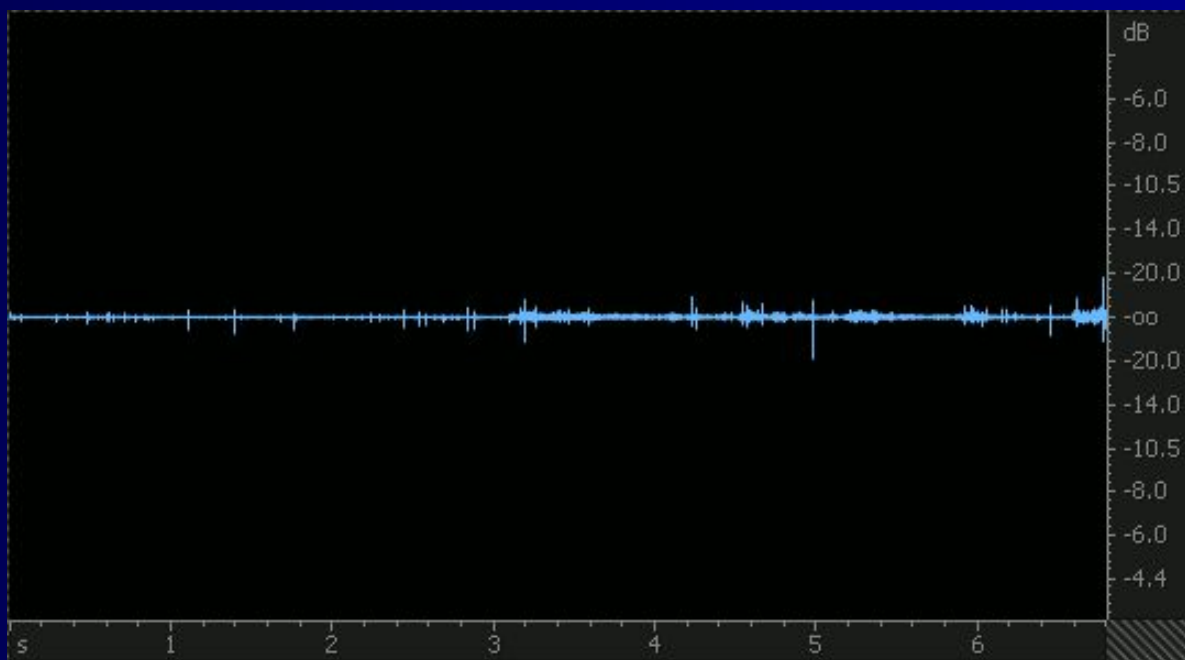
оркестр, записанный с виниловой пластинки: волна



Исправление щелчков



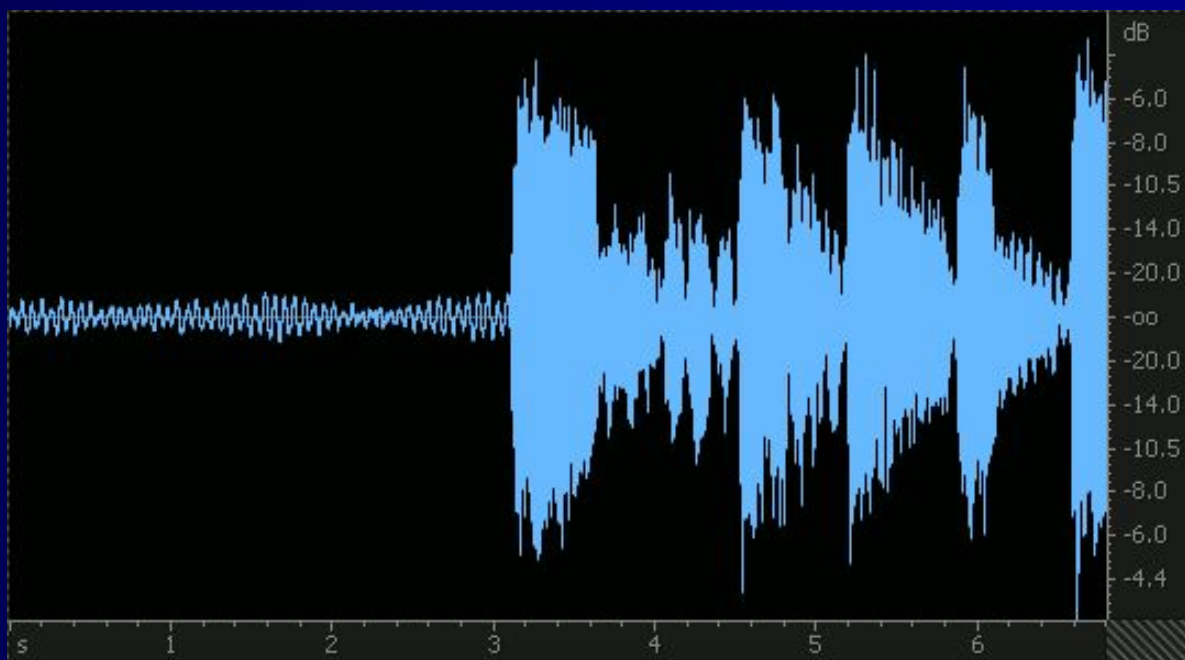
- Остаточный сигнал LPC
 - ▶ Порядок предсказания – 100 (но можно было и меньше)



Исправление щелчков



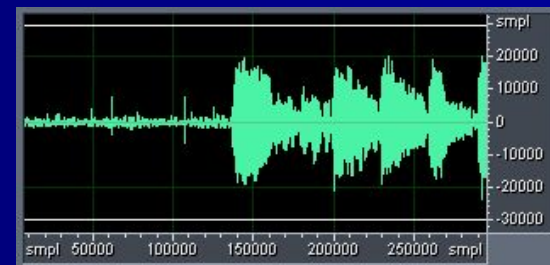
- После детектирования и интерполяции щелчков методом LSAR (3 итерации)



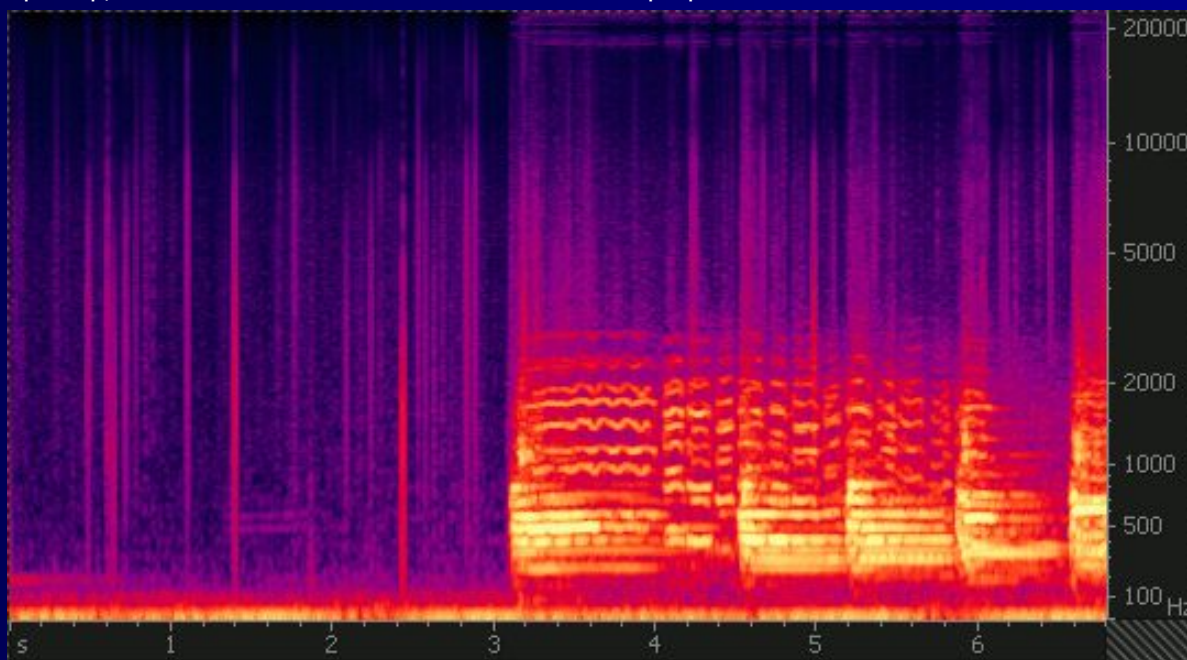
Исправление щелчков



- Исходная запись со щелчками
(clicks)



оркестр, записанный с виниловой пластинки: спектрограмма



Исправление щелчков



- После автоматического обнаружения и интерполяции щелчков

