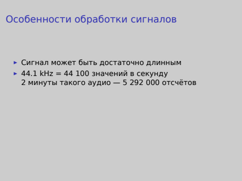
Оконное преобразование Уолша

Калабин Григорий Александрович, 523 группа
Научный руководитель: к. ф.-м. н., доц. С.М. Машарский
Рецензент: д. ф.-м. н., проф. В.Н. Малозёмов

Санкт-Петербургский Государственный Университет
Математико-Механический факультет
Кафедра Исследования Операций

Весна, 2013

Здравствуйте, меня зовут Григорий Калабин, мой научный руководитель — Сергей Михайлович Машарский, а рецензент — Василий Николаевич Малозёмов. Тема моей дипломной работы — Оконное преобразование Уолша. Я применяю данный подход для обработки цифровых дискретных аудио сигналов, поэтому для начала я расскажу про особенности их обработки



Первое, что нужно учитывать при обработке аудио, то, что объёмы данных могут быть достаточно большими, поэтому обрабатывать сигналы необходимо с помощью быстрых алгоритмов Теперь рассмотрим определения объектов, используемых в работе

Особенности обработки сигналов

- Сигнал может быть достаточно длинным
- 44.1 kHz = 44 100 значений в секунду
  2 минуты такого аудио — 5 292 000 отсчётов
- Изменения аудио во время прослушивания
  Пример: эквализация

Зачастую возникает необходимость изменять сигнал во время прослушивания. Подобная ситуация возникает при работе с эквалайзером, например, когда мы хотим устранить некоторые шумы. Теперь рассмотрим определения объектов, используемых в работе

Особенности обработки сигналов

- Сигнал может быть достаточно длинным
- 44.1 kHz = 44 100 значений в секунду
  2 минуты такого аудио — 5 292 000 отсчётов
- Изменения аудио во время прослушивания
  Пример: эквализация
- Обработка сигнала в реальном времени
  Пример: передача голоса

Ещё одна особенность, часто встречающаяся в наши дни — потоковая обработка аудиосигналов. Например, когда мы разговариваем по телефону, наш голос оцифровывается, проходит обработку (такую, как шумоподавление и эхоподавление) и передаётся собеседнику. Априори неизвестно сколько продлится разговор, и обрабатывать сигнал нужно в реальном времени. Теперь рассмотрим определения объектов, используемых в работе

Предварительные сведения

Сигналом длины $N \in \mathbb{N}$ будем называть функцию x:

$$x : \mathbb{Z} \to \mathbb{C}, \quad \mathrm{supp}(x) = 0 : N - 1 \tag{1}$$

Множество сигналов обозначим как $\mathbb{C}_N$.
Потоком данных будем называть бесконечный в обе стороны сигнал.

Сигналом длины N будем называть целочисленную функцию комплексного аргумента с носителем от нуля до N.

Пространство сигналов обозначим как $\mathbb{C}_N$. Поток данных — бесконечный в обе стороны сигнал

Предварительные сведения

Рассмотрим

$$N = 2^s,\ \Delta_\nu = 2^{\nu-1}.$$

Пусть

$$k, j \in 0 : \Delta_{\nu+1} - 1,$$

$$k = (k_{\nu-1}, k_{\nu-2}, \ldots, k_0)_2\,,\ j = (j_{\nu-1}, j_{\nu-2}, \ldots, j_0)_2.$$

Положим

$$\{k, j\}_\nu = \sum_{\alpha=0}^{\nu-1} k_\alpha j_\alpha\,, \quad k, j \in 0 : \Delta_{\nu+1} - 1.$$

Функции

$$v_k(j) = (-1)^{\{k,j\}_s}\,, \qquad k, j \in 0 : N - 1 \tag{2}$$

называются *дискретными функциями Уолша*.

Мы будем рассматривать сигналы, длина которых представляется степенью двойки. Дискретными функциями Уолша называются функции, задаваемые формулой два.

Предварительные сведения

Дискретные функции Уолша образуют базис в $\mathbb{C}_N$. Этот базис называется *базисом Уолша-Адамара*.
Пусть $j \in 0 : 2^\nu - 1$ и $j = (j_{\nu-1}, j_{\nu-2}, \ldots, j_0)_2$. Введём обозначение

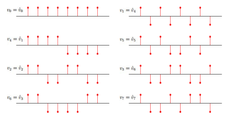
$$\mathrm{rev}_\nu(j) = (j_0, j_1, \ldots, j_{\nu-1})_2.$$

Частота функции v_k равна $\mathrm{rev}_s(k)$.
Обозначим $\hat{v}_k = v_{\mathrm{rev}_s(k)}$.
Функции $\hat{v}_k$ тоже образуют ортогональный базис в пространстве $\mathbb{C}_N$ который называется *базисом Уолша-Пэли*.

$$x = \frac{1}{N} \sum_{k=0}^{N-1} X(k) v_k = \frac{1}{N} \sum_{k=0}^{N-1} X(\mathrm{rev}_s(k)) \hat{v}_k. \qquad (3)$$

Известно, что функции Уолша образуют базис в $\mathbb{C}_N$, этот базис называется базисом Уолша-Адамара. Переупорядочим функции v_k по частоте, получим базис Уолша-Пэли, его и будем использовать.

Дискретные функции Уолша при $N = 2^3$



На рисунке представлены функции Уолша для $N = 8$

Будем использовать разложение сигнала по базису Уолша-Пэли, так как этот базис упорядочен по частоте. Поэтому необходимо уметь вычислять коэффициенты разложения сигнала по базису Уолша-Пэли.
Будем использовать быстрое преобразование Уолша, связанное с прореживанием по времени. Эта схема использует сложения в количестве $N \log_2 N$ операций и допускает обращение с той же скоростью.

Для вычисления компонент разложения по базису Уолша-Пели будем использовать быстрое преобразование Уолша со скоростью $N \log_2 N$

Оконная функция

Оконной функцией, или *окном*, называется весовая функция, которая равна нулю вне заданного интервала. Например, *прямоугольное окно*:

$$w_{rect}(k) = \begin{cases} 1, & \text{если } L \le k \le R \\ 0, & \text{иначе} \end{cases}$$

где $[L;R]$ — заданный интервал.

Оконной функцией будем называть весовую функцию, равную нулю вне заданного интервала.

Идея оконного преобразования

Имеется исходный поток данных, который необходимо обработать.

1. Зафиксируем $N \in \mathbb{N}$.
2. Определим сигнал x, как очередные N отсчётов исходного потока данных.
3. Произведём обработку x и получим $x' \in \mathbb{C}_N$.
4. Запишем x' в некоторый буфер.
5. Будем повторять шаги 2–4 пока не обработаем весь поток данных. Буфер будет содержать результат обработки исходного потока.

Длиной окна будем называть натуральное число N. Возьмём очередные N отсчётов исходного потока данных, обработаем их и запишем в некий буфер. Будем повторять процедуру, пока не закончим обработку, в итоге буфер будет содержать результат обработки.

Формальные определения прямого и обратного преобразования Уолша даны на следующих двух слайдах

Определение оконного преобразования Уолша

Пусть у нас есть некоторый поток данных x и прямоугольная оконная функция w с носителем $0 : N-1$. Оконным преобразованием Уолша потока данных x называется матрица X, задаваемая следующим равенством:

$$\mathbf{STWT}\{x\}(d,k) \equiv X(d,k) = \frac{1}{N}\sum_{j=0}^{N-1} x[j+d]w[j]\hat{v}_k(j), \tag{4}$$

где $k \in 0 : N-1,\ d \in (-\infty; +\infty)$. Индекс d матрицы соответствует точке во времени, а индекс k — частоте. Таким образом, данная матрица позволяет по времени и частоте узнать магнитуду исходного сигнала.

прямое... (раз-два-три)

Обратное оконное преобразование Уолша

Для каждого оконного спектра $X(d,k)$ вычислим обратное преобразование Уолша и получим часть исходного потока данных:

$$x(d,j) = \sum_{k=0}^{N-1} X(d,k)\hat{v}_k(j), \quad j \in 0 : N-1,\ d \in (-\infty;+\infty).$$

Поскольку в данной работе рассматривается оконное преобразование с прямоугольным окном, то можно перейти от матрицы $x(d,j)$ к исходному потоку данных $x(t)$ следующим образом:

$$x(t) = \frac{1}{N}\sum_{j=0}^{N-1} x(t-j,j), \quad t \in (-\infty;+\infty). \tag{5}$$

и обратное... (раз-два-три)

Результаты

В рамках данной работы было реализовано оконное преобразование Уолша на языке java и были проведены различные эксперименты.

Название	Инструмент	Длит.	Размер, bytes (samples)	Примечание
vivaldi.wav	Скрипка	15 сек	2 921 548 (1 460 752)	Антонио Вивальди, «Весна», часть 1 Аллегро
chopin.wav	Фортепиано	13 сек	1 178 548 (589 203)	Фридерик Шопен, «ноктюрн №2»
metallica.wav	Гитара	24 сек	2 190 902 (1 095 429)	Группа «Metallica», «Nothing else matters»
voice.wav	Речь	3 сек	702 644 (351 300)	Женская речь

Было установлено, что при помощь оконного преобразования Уолша возможна достаточно эффективная обработка сигналов.

В рамках данной работы были написаны два приложения для демонстрации того, что с помощью оконного преобразования Уолша возможна эффективная обработка сигналов.

Были проведены различные эксперименты на различных аудиозаписях, информация о которых представлена в таблице.

Первым приложением написаным в рамках данной работы является эквалайзер. С его помощью можно изменять мощность восьми диапазонов частот.

Выяснилось, что при обнулении $15/16$ верхних частот результат вполне различим. Этот факт говорит о том, что возможно эффективно сжимать аудио с помощью оконного преобразования Уолша, например, для передачи информации по узким каналам связи.

Фильтрация частот

Был проведён эксперимент в котором из спектра вырезались верхние частоты, то есть к исходному сигналу применялся простейший низкочастотный фильтр.
При проведении эксперимента использовалось прямоугольное окно без наложений с длиной равной 4096 отсчётам.
В результате было получено, что при вырезании $15/16$ верхних частот результат фильтрации вполне узнаваем.
Можно сделать вывод, что верхние частоты при разложении сигнала по базису Уолша воспринимаются человеком значительно меньше, чем низкие.

2016-03-18

Сжатие звука

Для сжатия данных использовался алгоритм DEFLATE. Компрессия применялась не к каждому отдельному окну, а ко всему сигналу целиком.

Название файла	Без фильтрации	1/2 спектра	1/4 спектра	1/8 спектра
vivaldi.wav	1.25	2.32	4.36	8.21
metallica.wav	1.22	2.28	4.32	8.15
chopin.wav	1.21	2.27	4.31	8.11
voice.wav	1.21	2.26	4.27	8.10

Второе приложение, написанное в рамках данной работы позволяет сжимать аудио на основе оконного преобразования Уолша. В таблице представлены коэффициенты сжатия при указанной ненулевой части спектра.