



Цифровая обработка сигналов

Лабораторная работа № 5

Эффект роботизации

Содержание

1 Теоретические сведения.....	3
1.1 Обратное КВПФ.....	3
1.2 Эффект роботизации голоса.....	4
2 Задание на лабораторную работу.....	5
3 Формат сдачи.....	6
4 Контрольные вопросы.....	6

Организация:	Самарский университет
Подразделение:	Кафедра геоинформатики и информационной безопасности
Версия:	2023.09.29

Вечно хмурые лица
В вечно грязных автобусах
И сутулые спины
От обиды и гордости,

Вечно мутные души,
Вечно грубые фразы.
Не привыкшие слушать.
Не меняются фазы.

Ольга Фокина,
[«Девочка с апельсинами»](#)

1 Теоретические сведения

1.1 Обратное КВПФ

Данный подраздел основан на материале [Invertibility of overlap-add processing](#).

Большой класс алгоритмов в обработке звука построен по классической схеме, состоящей из трёх шагов:

- 1) вычислить КВПФ;
- 2) провести какую-то операцию над спектром;
- 3) вычислить обратное КВПФ.

При этом на первом шаге сегменты обычно берутся с перекрытием и с умножением на весовую функцию. Возникает вопрос: как из избыточного и, возможно модифицированного, представления КВПФ вычислить обратное преобразование?

Пусть входной сигнал $x[n]$ разбивается на сегменты длительностью N с шагом H (от английского hop – «скачок»). При этом $H \leq N$ (в случае строгого неравенства сегменты перекрываются). Используемое весовое окно обозначим как $w = \{w[n]\}_{n=0}^{N-1}$. При этом подразумевается, что $w[n] = 0$ при $n < 0$ или $n \geq N$.

Определим взвешенный сегмент сигнала как

$$x_t[n] = x[n] \cdot w[n - tH],$$

где $t \geq 0$ — целое число, определяющее порядковый номер сегмента. Отметим, что для сегмента используется абсолютная адресация оригинального сигнала $x[n]$, то есть ненулевые отсчёты сегмента принадлежат диапазону $[tH, tH + N - 1]$.

Тогда ДПФ сегмента:

$$X_t[m] = \sum_{n=tH}^{tH+N-1} x_t[n] \cdot e^{-j\frac{2\pi}{N}m(n-tH)}, \quad 0 \leq m < N.$$

Обратное ДПФ спектра сегмента:

$$y_t[n] = \begin{cases} \frac{1}{N} \sum_{m=0}^{N-1} X_t[m] \cdot e^{j\frac{2\pi}{N}m(n-tH)}, & tH \leq n < tH + N; \\ 0, & \text{иначе.} \end{cases}$$

Предположим, что спектр не подвергается модификации, тогда $y_t = x_t$.

Введём произвольный параметр $a > 0$ и рассмотрим сумму по всем сегментам:

$$\begin{aligned} \sum_t y_t[n] w^a[n - tH] &= \sum_t \left(\frac{1}{N} \sum_{m=0}^{N-1} X_t[m] \cdot e^{j\frac{2\pi}{N}m(n-tH)} \right) w^a[n - tH] = \\ &= \sum_t \left(\frac{1}{N} \sum_{m=0}^{N-1} \left(\sum_{k=tH}^{tH+N-1} x_t[k] \cdot e^{-j\frac{2\pi}{N}m(k-tH)} \right) e^{j\frac{2\pi}{N}m(n-tH)} \right) w^a[n - tH] = \end{aligned}$$

$$\begin{aligned}
&= \sum_t \left(\frac{1}{N} \sum_{m=0}^{N-1} \sum_{k=tH}^{tH+N-1} x_t[k] \cdot e^{-j \frac{2\pi}{N} m(k-n)} \right) w^a[n - tH] = \\
&= \sum_t \sum_{k=tH}^{tH+N-1} x_t[k] \underbrace{\left(\frac{1}{N} \sum_{m=0}^{N-1} e^{-j \frac{2\pi}{N} m(k-n)} \right)}_{\delta[k-n]} w^a[n - tH] = \\
&= \sum_t \sum_{k=tH}^{tH+N-1} x_t[k] \cdot \delta[k - n] \cdot w^a[n - tH] = \sum_t x_t[n] \cdot w^a[n - tH] = \\
&= \sum_t x[n] \cdot w[n - tH] \cdot w^a[n - tH] = x[n] \sum_t w^{a+1}[n - tH].
\end{aligned}$$

Выражая отсюда $x[n]$ через $y_t[n]$, получаем следующую формулу сложения с перекрытием для обратного КВПФ:

$$x[n] = \frac{\sum_t y_t[n] w^a[n - tH]}{\sum_t w^{a+1}[n - tH]}.$$

Полученное выражение корректно только при отличном от нуля знаменателе. Это значит, что на используемое окно $w = \{w[n]\}_{n=0}^{N-1}$ нужно наложить некоторые ограничения. Перечислим возможные варианты, в порядке возрастания строгости:

- 1) **NOLA** (Nonzero OverLap Add): $\sum_t w^{a+1}[n - tH] \neq 0 \quad \forall n$;
- 2) $w[n] > 0, \quad 0 \leq n < N$;
- 3) **COLA** (Constant OverLap Add): $\sum_t w^{a+1}[n - tH] = C, \quad C \neq 0 \quad \forall n$;
- 4) **Strong COLA**: $\sum_t w^{a+1}[n - tH] = 1 \quad \forall n$.

В случае, если спектр $X_t[m]$ подвергается модификации ($y_t \neq x_t$), то может просто-напросто не существовать сигнала во временной области, соответствующего изменённому спектру. При $a = 1$ КВПФ восстановленного сигнала будет максимально близко к изменённому КВПФ в среднеквадратичном смысле. Доказательство данного факта выходит за рамки данной работы (интересующиеся могут обратиться к работе [Signal Estimation for Modified STFT](#)). Таким образом, в повсеместных реализациях обратного КВПФ (например, [istft](#) из пакета `scipy`) вышеприведённая формула используется при $a = 1$.

1.2 Эффект роботизации голоса

Один из простых алгоритмов для получения эффекта роботизации голоса состоит из трёх шагов:

1. Рассчитать КВПФ.
2. В спектре КВПФ обнулить все фазы.
3. Рассчитать обратное КВПФ.

2 Задание на лабораторную работу

Часть 1. Из репозитория курса скачайте заготовку `lab5.py`. В скачанном файле имплементируйте функцию `idft` для вычисления обратного ДПФ. Реализация данной функции должна быть своя (запрещается использовать готовые функции `dft/fft` из каких-либо библиотек).

В файле уже предоставлен готовый тест для проверки вашей реализации. Для проверки запустите файл — вы должны увидеть строку `OK (skipped=2)`.

Часть 2. В том же файле имплементируйте функцию `real_istft` для вычисления обратного КВПФ (см. формулу в подразделе [1.1](#)). Обратите внимание на следующие моменты.

1. Функция должна работать аналогично функции `istft` из библиотеки `scipy` (ознакомьтесь с документацией!). См. аргументы вызова этой функции в проверяющем методе-тесте `_test_istft`, чтобы понять, какое именно поведение ожидается от вашей реализации.

2. При реализации используйте функцию `idft` из первой части задания.

3. Реализация данной функции должна быть своя — запрещается использовать готовые функции `dft/fft/stft/istft` из каких-либо библиотек.

В файле уже предоставлены два готовых теста (`test_istft_unmodified` и `test_istft_modified`) для проверки вашей реализации, но они помечены аннотацией `@unittest.skip` (чтобы эти тесты не мешались в первой части). Для проверки уберите вышеупомянутые аннотации и запустите файл — вы должны увидеть строку `OK`.

Часть 3. В том же файле `lab4.py` в функции `main` реализуйте эффект роботизации (см. подраздел [1.2](#)).

С помощью микрофона запишите короткий отрывок речи. Можно использовать тот же самый входной файл, что и для прошлой лабораторной работы. В качестве начального значения размер сегмента взять равным 20 мс. Далее подобрать параметры по своему вкусу, изменяя следующие параметры `stft`: размер сегмента `nperseg`; величину перекрытия сегментов `noverlap`; вид оконной функции `window`.

Примечание. Допускается наложение эффекта роботизации одновременно с другими эффектами (как реализованными на прошлых лабораторных работах, так и

других, самостоятельно реализованных). При наложении других эффектов никаких ограничений на использование готовых пакетов нет, но для получения дополнительного балла математика любого реализуемого эффекта должна быть полностью и детально вами разобрана.

3 Формат сдачи

Предоставить входной wav-файл, выходной wav-файл с наложенным эффектом роботизации (и, возможно, другими эффектами) и скрипт с выполненным заданием.

4 Контрольные вопросы

1. Вычислите КВПФ для сигнала $x[n] = \{1, 2, 3, 4\}$ с размером сегмента в 2 отсчёта и перекрытием в 1 отсчёт.
2. Формула сложения с перекрытием для обратного КВПФ.