

The same procedure is applied to a question asked by user. Then some kind of intersection of two sets is obtained. The first set is previously obtained set for text documents and the second set is obtained by applying morphological analysis and hypotheses generation procedures to text representation of a question. The resulting answer set is obtained by extraction of words that are consistent with the question word from that intersection and sorting it in decreasing order of ranks.

For now proposed method is restricted to provide only one word answers without taking into account the context in which the words are used. In future those restrictions will be removed.

The proposed method of generating and ranking of hypotheses was implemented with a number of restrictions and tested on a database of news articles (over 3500 documents). Studies of the developed system showed positive results and proved its suitability to the task. Now the main goal of research is to develop algorithms of deductive, inductive and abductive logical inference. Those algorithms will help to answer such questions, the answers to which are not contained in the indexed documents in explicit form.



МНОГОМАСШТАБНЫЙ МЕТОД СПЕКТРАЛЬНОГО ВЫЧИТАНИЯ ДЛЯ ПОДАВЛЕНИЯ ШУМОВ В АУДИОСИГНАЛАХ

Ткаченко М.С., Лукин А.С.

МГУ им. М.В. Ломоносова, ф-т вычислительной математики и кибернетики,
лаборатория математических методов обработки изображений

1. Введение

Шум – это нежелательный сигнал, который возникает при передаче или измерении другого, чистого сигнала. Шумы различают по их спектральным свойствам: стационарные (не меняющиеся во времени) и нестационарные. Аддитивным называют шум, который суммируется с чистым сигналом $y[t]$ и не зависит от него: $x[t] = y[t] + noise[t]$, где t – время, $x[t]$ – наблюдаемый сигнал. Постоянное шипение микрофона, усилительной аппаратуры или гул электросети – примеры аддитивных стационарных шумов.

Одним из первых устройств для очистки аудио был широкополосный гейт. Предполагалось, что шум тише полезного сигнала. Как только мощность сигнала опускалась ниже некоторого порога, гейт переставал пропускать его [1]. Затем стали использоваться многополосные гейты, способные эффективно подавлять шум в нескольких частотных полосах.

Следующим шагом в этой области считается метод спектрального вычитания (МСВ). Результат его работы сильно зависит от выбора банка фильтров и точности оценки спектра подавляемого шума [2]. Один из недостатков МСВ – т.н. «музыкальный шум» – артефакт, возникающий вследствие того, что коэффициенты оконного преобразования Фурье (Short Time Fourier Transform: STFT) шума оказываются статистически случайными, что приводит к неравномерному подавлению шума по частоте и времени.

Спектрограмму сигнала можно рассматривать как изображение. Это позволяет адаптировать алгоритмы шумоподавления для изображений под задачи аудио. К примеру, использование метода нелокального усреднения (Non-Local Means) позволяет значительно снизить степень музыкального шума путем сглаживания спектрограммы.

В работе рассматриваются методы подавления аддитивных стационарных шумов с использованием многомасштабных преобразований (Multi-Resolution FFT: MR FFT), которые позволяют эффективнее использовать различные фильтры (в т.ч. фильтры удаления шума). Это достигается благодаря адаптивному подбору частотно-временного разрешения STFT с учетом особенностей сигнала. Такой подход основывается на учёте локальных свойств сигнала: оценки степени размытия его спектрограммы.

2. Метод спектрального вычитания (МСВ)

Метод спектрального вычитания устраняет аддитивные стационарные шумы [3]. Этапы работы МСВ:

1. Разложение сигнала с помощью оконного преобразования Фурье (STFT), компактно локализирующего энергию сигнала.

2. Составление вычитаемого «слепок шума» (noise footprint) – оценки спектра шума. Слепок шума получается усреднением по времени амплитуд спектра, взятых из заранее подготовленного участка шума, не содержащего полезного сигнала: $FOOTPRINT[f] = \frac{1}{K} \sum_{t=1}^K NOISE[f, t]$, где $NOISE[f, t]$ – амплитудный

спектр шума, f – индекс коэффициента Фурье, соответствующий частоте, t – номер текущего STFT окна, K – количество окон в участке с шумом.

3. Вычитание (в обобщенном смысле) амплитудного спектра шума из амплитудного спектра сигнала:

$$Y[f, t] = \max \{ X[f, t] - k \cdot W[f, t], 0 \}$$

Здесь $X[f, t]$ и $W[f, t]$ – амплитудные спектры сигнала и шума соответственно, $Y[f, t]$ – амплитудный спектр очищенного сигнала, а k – коэффициент подавления. Фазовый спектр очищенного сигнала полагается равным фазовому спектра зашумленного сигнала.

4. Обратное преобразование STFT – синтез результирующего сигнала.

Этот метод работает с фиксированными размерами окон STFT. Большой размер окна позволяет добиться хорошего разрешения по частоте, а значит, с ним можно наиболее точно производить спектральное вычитание, не «задевая» по частоте полезный сигнал. Но чем больше размер окна, тем хуже разрешение по времени, и тем сильнее происходит размытие спектра транзиентов (резких всплесков энергии). Таким образом, при использовании больших размеров окон возникает эхо и размытие транзиентов во времени, а при использовании коротких – шумоподавление оказывается неэффективным.

Для решения этих проблем можно использовать многомасштабные преобразования.

3. Многомасштабные преобразования

При использовании банков фильтров с постоянным разрешением невозможно получить одновременно хорошее разрешение, как по частоте, так и по времени. Необходимо на этапе обработки локально выбирать наилучшее разрешение, подходящее для данного сигнала (рис. 1).

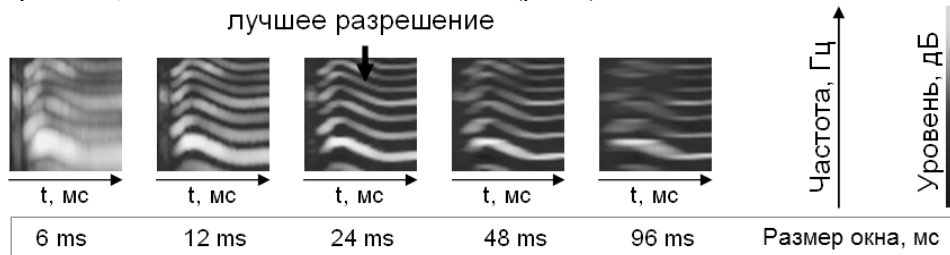


Рис. 1. Размытие спектра и выбор лучшего разрешения

Банки фильтров с адаптивным частотно-временным разрешением, в частности многомасштабные преобразования, помогают решить эту проблему [4, 5].

В основе MR FFT лежат:

- вычисление STFT, взятых от одного сигнала, с разными разрешениями (размерами окна),
- оценка степени разреженности соответствующих областей спектра,
- микширование спектрограмм с учетом их разреженности.

Вначале выполняется независимая друг от друга работа эффектов с фиксированными размерами STFT (к примеру, работа фильтров шумоподавления, основанных на МСВ).

Результирующие спектрограммы передаются в микшер и разбиваются по времени и частоте на области Ω_r . Анализатор вычисляет их степени размытия. Чем выше степень размытия, тем с меньшим весом Ω_r будет суммироваться в итоговый сигнал.

Степень разреженности для каждой области Ω_r вычисляется, как отношение среднего арифметического

модулей коэффициентов $a_i \in \Omega_r$ к их среднему квадратичному.

$$S_{\Omega_r} = \frac{\sum_{i=1}^N a_i}{\sqrt{N} \sqrt{\sum_{i=1}^N a_i^2}}$$

Таким образом, для длинных тональных сигналов будет подбираться STFT с наибольшим разрешением по частоте (большой размер окна), а для резких всплесков энергии, щелчков и событий с быстрой атакой – с наименьшим по частоте, но наилучшим по времени (маленький размер окна).

4. Система шумоподавления

На основе описанного алгоритма разработана система шумоподавления с широкими возможностями настройки параметров работы. С её помощью подобраны наиболее оптимальные настройки.

На практике установлено, что высокое качество результатов достигается при обработке и анализе как минимум трёх спектрограмм с разными размерами STFT. Размер каждого окна STFT выбирается в два раза больше предыдущего, начиная с 12 мс.

Спектрограммы разбиваются по времени и частоте на несколько прямоугольных областей Ω_r . Важный этап – это сопоставление Ω_r и их размеров на разных спектрограммах. Каждая из них имеет размер по частоте в 1/16 от количества частотных полос для конкретной спектрограммы и по времени w :

$$v = (\text{Size}-1) \cdot \frac{\text{MixFFT}}{\text{CurrentFFT}} + \left\lceil \frac{\text{MixFFT}-\text{CurrentFFT}+1}{\text{CurrentFFT} \cdot \text{InvOverlap}} \right\rceil,$$

$$w = \begin{cases} v, & v > 1 \\ 1, & v \leq 1 \end{cases}$$

где $MixFFT$ – размер STFT для спектрограммы, в которую суммируется результат, $CurrentFFT$ – размер STFT исследуемой спектрограммы, $InvOverlap$ – обратная величина от степени перекрытия окон STFT (1/4, по умолчанию). Все величины задаются в виде отчетов (samples). $Size$ – размер области Ω_r для спектрограммы с минимальным размером STFT. Параметр $Size$ равный 64 оказывается достаточным для корректной работы анализатора.

Отметим, что суммирование производится в спектрограмму с наименьшим размером STFT, иначе переключение между областями будет происходить не достаточно быстро.

Благодаря ходу с перекрытием на 1/4 по частоте и времени всех Ω_r , производится плавное переключение при микшировании спектрограмм.

5. Результаты

Достигнута высокая производительность алгоритма, возможна обработка аудио в реальном времени. Помимо нескольких модификаций МСВ, в систему интегрирован метод нелокального усреднения из области шумоподавления для изображений [6].

Из таблицы 1 видно, что использование многомасштабных преобразований весьма эффективно, и их работа даже в совокупности с простым методом спектрального вычитания оказывается более чем удовлетворительной: в обработанных фонограммах значительно повышается степень удаления шума и не нарушается структура сигнала.

На основе данной системы разработан интернет-сервис – общедоступная система шумоподавления, оснащенная автоматическим поиском слежка шума [7].

Метод	Подавление шума	Комментарий
МСВ 11 мс	13.10 дБ	Размер окна – 11 мс. Недостаточная степень устранения шума.
МСВ 93 мс	13.15 дБ	Размер окна – 93 мс. Появление эхо.
МСВ на основе многомасштабных преобразований	16.10 дБ	Остается музыкальный шум. Запись кажется приглушенной.
Метод нелокального среднего на основе многомасштабных преобразований	16.70 дБ	Минимальное количество артефактов.

Таблица 1. Результаты тестирования системы шумоподавления

6. Заключение

В работе рассмотрен метод многомасштабных преобразований, с помощью которого производится адаптивный выбор размера STFT в зависимости от локальных свойств сигнала. Это повышает эффективность фильтров шумоподавления. В результате тестирования алгоритма установлены оптимальные размеры STFT анализируемых спектрограмм, размеры суммируемых областей и иные ключевые параметры.

В качестве альтернативы МСВ задействован более эффективный метод нелокального усреднения, адаптированный из области сглаживания изображений. В результате сильно снизился уровень музыкального шума.

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009–2013 годы.

Литература

1. David Gibson. The Art of Mixing: A Visual Guide to Recording, Engineering, and Production. REV, 2005. ISBN13: 9781931140454. pp. 52-53, 162 p.
2. Gilbert Strang, Truong Nguyen. Wavelets and Filter Banks. Wellesley-Cambridge Press, 1996. ISBN: 0961408871, 9780961408879. pp. 103-105, 490 p.
3. Saeed V. Vaseghi. Advanced Digital Signal Processing and Noise Reduction. John Wiley & Sons Ltd, 2000. ISBN: 0471626929. pp. 5-8, 466 p.
4. Alexey Lukin, Jeremy Todd. Adaptive Time-Frequency Resolution for Analysis and Processing of Audio. AES'06, October 2006, 10 p.
5. Karin Dressler. Sinusoidal Extraction Using an Efficient Implementation of a Multi-Resolution FFT. DAFx'06, September 2006, 6 p.
6. Antoni Buades, Bartomeu Coll, Jean-Michel Morel. A non-local algorithm for image denoising. CVPR'05, Vol. 2, June 2005, pp 60-65.
7. Демонстрационная версия онлайн-денойзера: <http://makseq.com/makseq>

A MULTIREOLUTION SPECTRAL SUBTRACTION ALGORITHM FOR NOISE SUPPRESSION IN AUDIO SIGNALS

Tkachenko M., Lukin A.

Moscow State University, Faculty of Computational Mathematics and Cybernetics,
Laboratory of Mathematical Methods of Image Processing

Noise is an undesirable signal appearing during transmission or measurement of another clean signal. There are two categories of noise, by spectral properties: stationary (the one that does not change in time) and non-stationary.

Additive noise is summed with the clean signal $y[t]$ and does not depend on it: $x[t] = y[t] + noise[t]$, where t is time and $x[t]$ is the observed signal. Constant hiss from a microphone or an amplifier, electric hum are the examples of additive stationary noise.

In this paper, we propose a method for suppression of additive stationary noise using a multiresolution STFT (short-time Fourier transform). This approach is able to improve quality of many audio processing algorithms (such as noise reduction) by adaptively varying STFT time-frequency resolution based on local properties of the signal: estimates of spectrogram sparsity.

The simple method of spectral subtraction (SS) is widely used for reduction of additive stationary noises. It works with a fixed STFT window size. Long windows provide good frequency resolution and achieve accurate separation of closely spaced noise and signal harmonics. However long windows sizes also lead to poor time resolution and increase smearing of transients (sharp attacks or fast changes in the signal). On the other hand, short-window STFT processing is inefficient in terms of frequency resolution and possible depth of noise suppression.

A filter bank with a fixed window size cannot provide good frequency and time resolution simultaneously. It is required to select best resolution for each local part of the signal during the processing. We propose using a multiresolution STFT to solve this problem.

The algorithm of MR STFT consists of three parts:

1. Calculation of several copies of the signal processed with different STFT window sizes;
2. Estimation of smearing (or sparsity) for local spectrogram areas for each resolution;
3. Mixing of the resulting spectrograms based on spectrogram sparsity estimates.

The effect of this adaptive algorithm is selection of high frequency resolution for tonal signals and selection of high time resolution for transients.

A noise reduction system based on MR STFT has been created. Several modifications of spectral subtraction rules have been implemented, including a highly effective method of Non-Local Means for smoothing of a “musical noise” artifact. During algorithm testing we have found optimal range of STFT sizes for the MR STFT frameworks, sizes of spectrogram mixing areas and other key parameters.

We have compared the performance of a simple spectral subtraction (SS) with several STFT window sizes, SS based on MR STFT (SS MR STFT), and SS with a Non-Local Means based on MR STFT. It has been shown that SS MR STFT produces results significantly different with a simple SS: after processing with a simple SS noise power decreased by 13.00 dB, while with SS MR STFT – by 16.10 dB. The application of a Non-Local Means smoothing has removed musical noise and further improved overall processing quality.



ОБ ОДНОМ СПОСОБЕ ЧАСТОТНОГО АНАЛИЗА ПЕРИОДИЧЕСКИХ СИГНАЛОВ

Колоколов А.С.

Институт проблем управления им. В.А.Трапезникова РАН, Москва

В настоящее время цифровой спектральный анализ является важным инструментом в практике анализа гармонических, периодических и квазипериодических сигналов. Примерами таких сигналов являются музыкальные звуки, вокализованные фрагменты речи, а также разного рода сигналы в задачах акустической диагностики.

Нахождению спектра сигнала $s(t)$ цифровыми методами предшествует его дискретизация с частотой f_s , в результате чего непрерывный сигнал представляется последовательностью дискретных отсчётов $s(n\Delta t)$, где n – номер отсчёта, принимающий значения $n = \dots - 2, -1, 0, 1, 2, \dots$, $\Delta t = 1/f_s$ – интервал квантования. В этом случае комплексный спектр отрезка дискретного сигнала $s(n)$, заданного на интервале

$$n = 0, 1, \dots, N - 1, \text{ определяется выражением} \quad \tilde{S}(m) = \sum_{n=0}^{n=N-1} s(n)w(n)e^{-2\pi jmn\Delta f\Delta t} = a(m) + jb(m),$$