

УДК 517.5

ФОРМИРОВАНИЕ ВЕЙВЛЕТ-ФУНКЦИЙ В ЗАДАЧЕ ИДЕНТИФИКАЦИИ МУЗЫКАЛЬНЫХ СИГНАЛОВ

А.С. Фадеев

Томский политехнический университет

E-mail: fas@aics.ru

Предложен подход, позволяющий формировать вейвлет-функции на основе периодических сигналов и фрагментов сигналов музыкальных инструментов. Рассмотрены необходимые и достаточные условия, предъявляемые к формируемым вейвлет-функциям. Описан эксперимент, позволяющий идентифицировать несколько гармонических составляющих сигнала, локализованных во времени. Показана возможность применения разработанного подхода в задачах идентификации сложных музыкальных сигналов.

Рост информационных технологий последних лет сделал возможным программные и аппаратные реализации многих математических задач, связанных с распознаванием образов. Благодаря развитию современных высокопроизводительных вычислительных систем решен ряд задач в области распознавания речи, графических объектов, обработки сейсмических данных, кардиограмм и ряда других. Менее изучены задачи в области распознавания музыкальной информации и возможность их применения в системах автоматизированной обработки информации [1]. Одной из таких задач является задача идентификации отдельных нот в одноголосных и многоголосных мелодиях, исполненных на определенных музыкальных инструментах.

Основу вычислительно эксперимента представляет математическая модель музыкального сигнала $F(t)$, полученного при звукозаписи мелодии определенного музыкального инструмента:

$$F(t) = A_1 n_1(t - \theta_1) + A_2 n_2(t - \theta_2) + \dots + A_N n_N(t - \theta_N) + h(t),$$

где $n_i(t)$ – амплитудно-временные характеристики отдельных нот мелодии; θ_i – временные смещения, определяющие начальные моменты звучания каждой ноты; A_i – громкость звучания отдельной ноты; $h(t)$ – сигнал помехи, вносимой звукозаписывающей аппаратурой; t – время.

Большинство музыкальных инструментов обладают свойством автомодельности, которое позволяет из временной функции одной ноты $n_0(t)$ определенного музыкального инструмента получить временную функцию любой другой ноты $n_i(t)$ этого музыкального инструмента масштабированием функции $n_0(t)$ вдоль оси времени:

$$n_i(t) = n_0\left(\frac{t}{m_i}\right),$$

где m_i – коэффициент масштабирования, i – положение ноты $n_i(t)$ по высоте относительно ноты $n_0(t)$. Для равномерно темперированного строя европейской музыки $m_i = 2^{-i/12}$, $i = 0, \pm 1, \pm 2, \dots$.

Например, нота «до» второй октавы расположена на 12 полутонов выше ноты «до» первой октавы и имеет частоту основного тона в два раза выше частоты основного тона ноты «до» первой октавы. Временная функция ноты «до» второй октавы $n_2(t)$ будет сжата в два раза относительно функции $n_1(t)$:

$$n_2(t) = n_1\left(\frac{t}{2}\right) = n_1(2t).$$

В качестве базовой ноты $n_0(t)$ может быть выбрана любая нота из диапазона определенного инструмента, вне зависимости от использования в конкретном музыкальном сигнале $F(t)$. Например, за $n_0(t)$ можно принять временную функцию ноты «ля» первой октавы с частотой основного тона $\nu = 440$ Гц.

Это свойство применено во всех современных музыкальных синтезаторах использующих метод волнового табличного синтеза [2]. Такие синтезаторы используют банк данных сигналов одной (базовой) ноты для каждого музыкального инструмента. При формировании мелодии одного инструмента сигнал базовой ноты каждый раз масштабируется по высоте основного тона на значение m_i . Каждая нота сдвигается во времени на величину θ_i и масштабируется по амплитуде на величину A_i согласно партии этого инструмента. Все сформированные функции для одного музыкального инструмента суммируются. В результате модель музыкального сигнала, формируемого синтезатором для конкретного музыкального инструмента, имеет вид:

$$F(t) = \sum_i A_i \cdot n_0\left(\frac{t - \theta_i}{m_i}\right).$$

Аналогично, для k -различных музыкальных инструментов, имитируемых синтезатором, музыкальный сигнал может быть представлен суммой сигналов всех нот, сыгранных в различные моменты времени с различной амплитудой:

$$f(t) = \sum_k \sum_i A_i \cdot n_0^k\left(\frac{t - \theta_i}{m_i}\right),$$

где $n_0^k(t)$ – временная функция сигнала базовой ноты k -го музыкального инструмента; θ_i – временной интервал смещения ноты $n_0^k(t)$; m_i – масштаб ноты $n_i^k(t)$ относительно базовой ноты $n_0^k(t)$, задающий частоту основного тона; A_i – амплитуда звучания ноты $n_i^k(t)$.

Функция $f(t)$ представляет собой идеализированную модель музыкального сигнала, полученного аналогично звукозаписи нескольких музыкальных инструментов (оркестра, ансамбля).

Соответственно модели музыкального сигнала $f(t)$, задачу распознавания музыкального сигнала

можно представить как задачу идентификации амплитуд нот с определенным масштабом m_i и смещением во времени θ_i для всех k -музыкальных инструментов, присутствующих в исследуемом сигнале $f(t)$. Частным случаем этой задачи может быть задача идентификации мелодии отдельно записанного музыкального инструмента или мелодии сгенерированной музыкальным синтезатором.

При выборе математического аппарата для исследования частотно-временных свойств сигнала $f(t)$, который является нестационарным и неперiodическим сигналом, было решено применить непрерывное вейвлет-преобразование (НВП) [3]:

$$Wf(\tau, s) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{s}} f(t) w\left(\frac{t-\tau}{s}\right) dt,$$

где $w(t)$ – материнский вейвлет; s – коэффициент масштабирования вейвлета; τ – коэффициент сдвига вейвлета.

Одной из особенностей НВП является формирование семейства вейвлетов $w_{s,\tau}(t)$ путем сдвигов τ и масштабирования s материнского вейвлета $w(t)$. Формирование семейства вейвлетов схоже с системой формирования семейств нот $n_i^k(t)$ из одной базовой $n_0^k(t)$ путем сдвигов θ_i и масштабирования m_i . Поэтому использование НВП в данной задаче весьма органично.

Процедура выбора функции материнского вейвлета остается эмпирической для каждой конкретной задачи и сводится к перебору функций материнских вейвлетов в НВП до достижения желаемого результата. Исследования свойств вейвлет-функций [4] показали, что наилучшие графические представления результатов НВП получены при схожести частотных спектров сигнала $f(t)$ и вейвлета $w(t)$.

Для каждого конкретного масштаба s функция $Wf_s(\tau)$ аналогична взаимной корреляционной функции сигналов $w_s(t)$ и $f(t)$ и описывает как степень сходства формы двух сигналов, так и их взаимное расположение друг относительно друга по оси времени:

$$Wf_s(\tau) = \int_{-\infty}^{\infty} f(t) w_s(t-\tau) dt, \quad w_s(t) = w\left(\frac{t}{s}\right).$$

Известно, что значения взаимной корреляционной функции максимальны при совпадении функций [6]. В данном случае значения $Wf_s(\tau)$ будут максимальными для таких τ , при которых функции $f(t)$ и $w_s(t-\tau)$ равны: $f(t) = w_s(t-\tau)$.

Очевидно, что кроме сдвига τ , для выполнения этого условия необходимо и равенство двух функций в каждой точке t .

Вейвлет-функция, используемая в НВП должна удовлетворять ряду необходимых условий [5]:

1. Ограниченности во времени:
 $w(t) \rightarrow 0$, при $t \rightarrow \infty$;
2. Кусочной непрерывности функции $w(t)$.
3. Интегрируемости с равенством нулю:

$$\int_{-\infty}^{\infty} w(t) dt = 0.$$

Примером материнской вейвлет-функции, удовлетворяющей всем приведенным условиям, может служить вейвлет Морле (рис. 1).

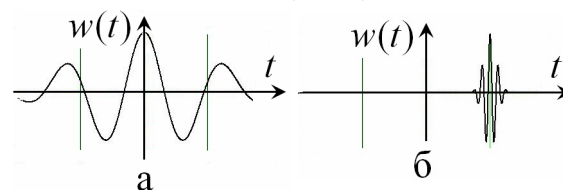


Рис. 1. Вейвлеты семейства Морле: а) материнский вейвлет Морле, полученный масштабированием $s=0,2$; б) вейвлет семейства Морле, полученный сдвигом на $\tau=5$ и масштабированием $s=2$

В силу условий временной локализации, наложенных на функцию материнского вейвлета, значение функции $Wf_s(\tau)$ тем больше, чем более точно вейвлет масштаба s $w_s(t-\tau)$ совпадет с локальным участком сигнала $f(t)$. Для точного совпадения на локальных временных отрезках функции вейвлета и сигнала должны быть равны.

В данной работе предлагается использовать функцию материнского вейвлета, сформированную из локального участка исследуемого музыкального сигнала. Для более узкой задачи – идентификации нот музыкального инструмента, обозначенной выше, материнский вейвлет можно сформировать из локального участка функции сигнала базовой ноты этого инструмента $n_0(t)$ (рис. 2).

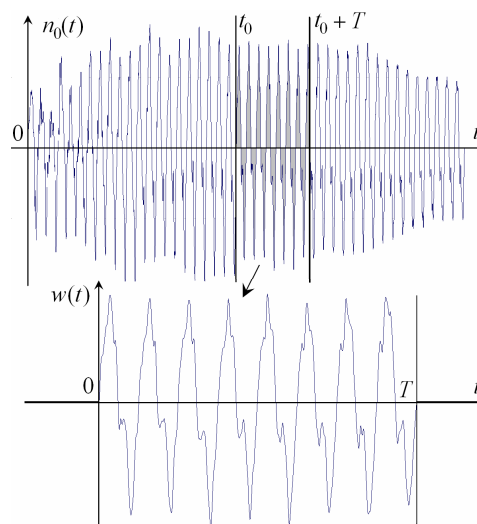


Рис. 2. Функция $n_0(t)$ и сформированный на основе ее фрагмента вейвлет $w(t)$

Для формируемого вейвлета должны выполняться изложенные условия 1–3.

1. Ограниченности во времени

$$w(t) = 0, \text{ при } t \notin [0, T];$$

$$w(t) = n_0(t+t_0), \text{ при } t \in [0, T].$$

Здесь t_0 характеризует момент времени, с которого значения вейвлет-функции $w(t)$ равны значе-

ниям функции сигнала $n_0(t)$, а значение T равно длительности фрагмента сигнала $n_0(t)$, совпадающего с вейвлетом $w(t)$. Значения t_0 и T выбираются таким образом, чтобы обеспечить остальные необходимые условия для вейвлет-функций [5].

2. Кусочной непрерывности

Функция базовой ноты $n_0(t)$ на всем интервале существования является непрерывной, так как описывает колебания физического тела с конечной массой во времени и по определению не может иметь разрывов.

Для обеспечения кусочной непрерывности функции материнского вейвлета необходимо выполнение условия равенства нулю начального и конечного значений функции $n_0(t)$ на интервале $[t_0, T+t_0]$:

$$n_0(t_0)=0 \text{ и } n_0(T+t_0)=0.$$

3. Интегрируемости с равенством нулю

Одним из свойств музыкальных инструментов является отсутствие гармонических составляющих с частотой меньше частоты основного тона ноты. В том числе, в сигналах музыкальных инструментов отсутствует нулевая гармоника [1]. Это свойство позволяет обеспечить нулевое среднее для $n_0(t)$ на интервале $[t_0, T+t_0]$ при наличии на данном интервале целого числа периодов в функции $n_0(t)$.

Таким образом, для формирования вейвлета, обладающего наибольшими избирательными способностями к сигналу $n_0(t)$, необходимо использовать периодический участок сигнала $n_0(t)$ с нулевыми начальным $n_0(t_0)$ и конечным $n_0(T+t_0)$ моментами, такой, чтобы $\int_{t_0}^{T+t_0} n_0(t)dt = 0$.

Идентификация нот музыкального произведения состоит в определении частот основных тонов самих нот и времени начала и длительности их звучания. Одним из свойств функции вейвлета, определяющего основную частотно-временную избирательную способность НВП, является локализация вейвлета в частотной и временной областях [5]. При НВП каждый вейвлет одного семейства (полученный из одного материнского вейвлета) образует частотно-временное окно конечного размера. Площадь окна для вейвлетов одного семейства на плоскости сдвиг-масштаб всегда остается постоянной [3] (рис. 3).

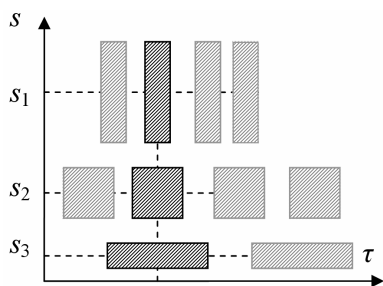


Рис. 3. Окно частотно-временной локализации вейвлета при различных значениях параметров сдвига τ и масштаба s

Однако в ходе проведения ряда экспериментов выяснилось, что при изменении самого вейвлета,

геометрические характеристики окна изменяются. Так, с увеличением количества периодов в материнском вейвлете (а, следовательно, и во всех вейвлетах семейства), окно вытягивается вдоль оси времени, сужаясь относительно оси масштабов, и, наоборот, с уменьшением числа периодов в материнском вейвлете, окно вытягивается вдоль оси масштабов s , сужаясь вдоль оси времени (рис. 4).

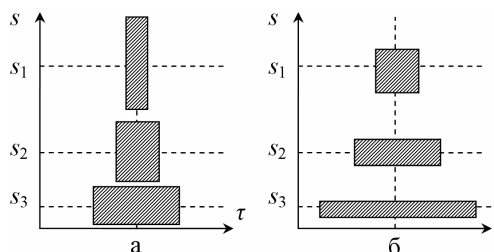


Рис. 4. Окно частотно-временной локализации вейвлетов с числом периодов а) $N=2$ и б) $N=8$

Для задач выявления элементарных составляющих (нот) в музыкальном сигнале, необходимо формирование семейства вейвлетов удовлетворяющих двум условиям:

1. разрешение по частоте (ширина окна вдоль оси масштабов Δs) должна обеспечивать такое разрешение, чтобы частоты основных тонов двух соседних ноты были различимы;
2. разрешение по времени (ширина окна вдоль оси времени $\Delta \tau$) должно позволять выполнять идентификацию всех нот минимально возможной длительности.

1. Разрешение по частоте

С целью оценки разрешающей способности искусственно создаваемых материнских вейвлетов был проведен эксперимент. Целью эксперимента являлось определение количества периодов в вейвлете, позволяющего идентифицировать частотные масштабы m_i всех нот, присутствующих в сигнале одновременно (при НВП эквивалентом m_i будут масштабы вейвлета s , относительно материнского).

В эксперименте использовано трезвучие «домажор» первой октавы. Частоты основных тонов нот этого трезвучия соответствуют гармоническим сигналам с частотами 261,6, 329,6 и 392 Гц. Продолжительность сигнала условно выбрана равной 0,2 с: $f(t) = \sin(2\pi 261,6t) + \sin(2\pi 329,6t) + \sin(2\pi 392t)$.

Для исследования тестового сигнала были созданы материнские вейвлеты $w_i(t)$. Количество периодов гармонического сигнала в вейвлетах составляло 1, 2, 4, 8 и 16 периодов соответственно (рис. 5).

Для тестового сигнала $f(t)$ было произведено НВП со всеми семействами вейвлетов $w_i(t)$. Для каждого преобразования были построены графические интерпретации результатов в виде трехмерных моделей и карт проекций изолиний. На рис. 6 приведены карты проекций изолиний результатов НВП для семейств вейвлетов $w_i(t)$. Ось ординат каждой карты проекций изолиний представляет собой ось масштабов вейвлета s , Ось абсцисс — ось временных

сдвигов τ . τ_0, τ_1 – время начала и окончания сигнала $f(t)$. Оттенками серого отображается амплитуда результатов НВП: более темные участки соответствуют большей амплитуде результатов НВП (рис. 6).

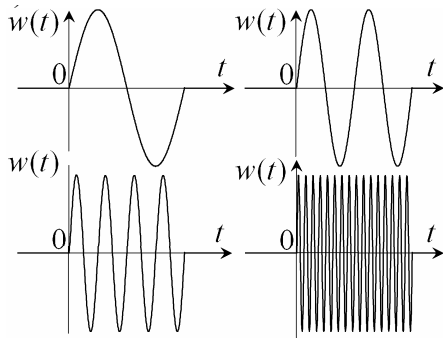


Рис. 5. Материнские вейвлеты с числом периодов гармонического сигнала 1, 2, 4 и 16

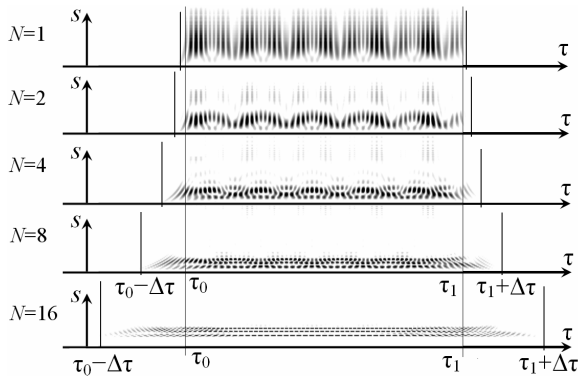


Рис. 6. Графические интерпретации результатов НВП для вейвлетов с различным числом периодов N гармонического сигнала

Как видно из рис. 6, при использовании вейвлета с одним периодом зона нечетких результатов в области временных сдвигов τ , продолжительностью $\Delta\tau$ достаточна мала, следовательно мала погрешность временных оценок сигнала. При этом разрешение по частоте настолько низкое, что для большей части масштабов s , результаты НВП имеют большие значения, означающие относительно равномерное присутствие частотных компонент на всем частотном диапазоне исследования. Хотя, в тестовом сигнале $f(t)$ присутствуют только три гармоники.

Из рис. 6 видно, что для вейвлета с 16 периодами на протяжении сигнала $f(t)$, в интервале от τ_0 до τ_1 , большие значения имеются только у результатов НВП для трех масштабов s , эквивалентных частотам 261,6, 329,6 и 392 сигнала $f(t)$. Для остальных значений s результаты НВП практически равны нулю. Низкое разрешение по времени формирует нечеткие результаты в моменты возникновения и затухания сигнала с промежуточными значениями амплитуды на интервалах продолжительностью $\Delta\tau$, не позволяя на этом интервале судить об изменениях амплитуды тестового сигнала.

Таким образом, семейство вейвлетов с одним периодом синусоидального сигнала при НВП дает высокое разрешение по времени, но крайне низкое по

частоте. Однако семейство вейвлетов с шестнадцатью периодами дает высокое разрешение по частоте (все гармоники, присутствовавшие в сигнале, однозначно идентифицируются), но низкое по времени.

2. Разрешение по времени

Нотная запись предполагает использование символов нот и пауз для обозначения элементов мелодии и промежутков мелодии, на которых ноты не звучат. Длительность нот, так же как и длительность пауз, кратна длительности звучания t_1 «целой» ноты – ноты максимально возможной длительности. Система нотной записи, состоящая из чередующихся нот и пауз, накладывает жесткие ограничения на моменты начала звучания нот и начала пауз. Моменты начала звучания нот дискретизованы с периодом дискретизации t_d , равном длительности самой короткой ноты. Как в классических, так и в современных музыкальных произведениях самой короткой по продолжительности считается нота с длительностью $t_{64}=1/64 \cdot t_1$. На практике, в силу технической сложности исполнения, ноты с длительностью t_{64} встречаются крайне редко. Практически самой короткой нотой можно считать длительность $t_{32}=1/32 \cdot t_1$. На рис. 7 приведен фрагмент двухголосной мелодии. В каждый момент времени звучит не более двух нот одновременно. Самой короткой является нота 2. Ноты 1, 3, 4 равны по длительности и в два раза длиннее ноты 2 и паузы. Если $t_d=t_{32}$, то нота 2 и пауза будут длительностью t_{32} , а ноты 1, 3, 4 – t_{16} .

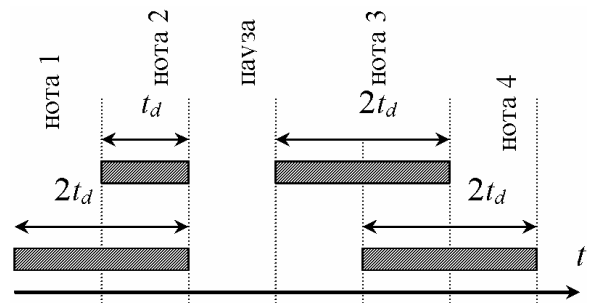


Рис. 7. Временная дискретизация начала звучания и продолжительности нот и пауз в многоголосной мелодии

При продолжительности вейвлета $T=t_d$ огибающая взаимно-корреляционной функции $Wf_s(t)$ вейвлета и совпадающего с ним по форме сигнала одной ноты (для конкретных значений s), вырождаясь в автокорреляционную функцию, будет иметь форму равностороннего треугольника с максимумом в центре звучания ноты [6] и шириной $2t_d$ (рис. 8, а). Если отбросить все значения $Wf_s(t)$, меньшие $Wf_{\text{om}}(t)$, то время идентификации ноты составит t_i . При $Wf_{\text{om}}(t)=0,5Wf_{\text{max}}(t)$, $t_i=0,5t_d$. Это значит, что время идентификации ноты равно половине ее длины.

При $T=0,5t_d$, огибающая корреляционной функции $Wf_s(\tau)$ вейвлета и совпадающего с ним по форме сигнала одной ноты будет иметь форму равнобокой трапеции, ширина верхней грани которой равна $0,5t_d$ (рис. 8, б). Если отбросить все значения $Wf_s(t)$, меньшие $Wf_{\text{om}}(t)$, то время идентификации ноты со-

ставит $t_{и}$. При $Wf_{ош}(t)=0,5Wf_{макс}(t)$, $t_{и}=0,5t_d$. Это значит, что время идентификации ноты равно ее длине.

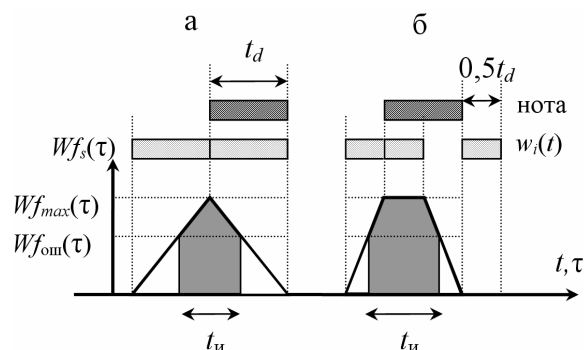


Рис. 8. Огибающие корреляционной функции ноты и вейвлета $w_i(t)$, соответствующего частоте основного тона ноты масштаба

При дальнейшем уменьшении ширины вейвлета T , и условия $Wf_{ош}(t)=0,5Wf_{макс}(t)$, время идентификации ноты $t_{и}$ останется постоянным, равным t_d .

Вывод: для временной идентификации ноты наименьшей длины t_d , необходимо использовать вейвлет длиной T , не большей t_d .

Запишем условия на продолжительность T вейвлета во времени:

1. $T \geq 16/\nu$, где ν – частота основного тона идентифицируемой ноты;
2. $T \leq t_d$.

Вычислим граничные значения частоты основного тона ноты ν , достоверная идентификация которой возможна и по времени, и по частоте: $T=16/\nu$, $T=t_d$, следовательно $t_d=16/\nu$ или $\nu=16/t_d$. Зависимость $\nu(t_d)$ определяет минимальную (граничную) частоту ν основного тона ноты продолжительностью t_d , которая будет однозначно идентифицирована вейвлетом с 16 периодами.

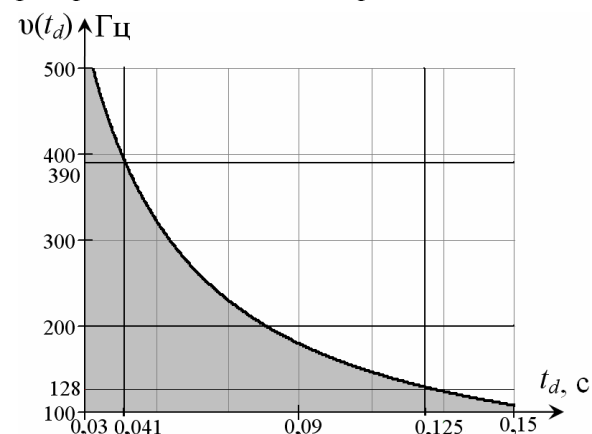


Рис. 9. Зависимость граничной частоты идентификации основного тона ноты от ее продолжительности

Темп исполнения большинства произведений, как правило, колеблется в диапазоне 60–180 четвертных долей в минуту, что соответствует времени звучания одной целой ноты $t_1=1,3...4,0$ с. Соответственно, звучание ноты длительностью t_{32} при быстром темпе

воспроизведения (180 ударов в минуту) составит $t_{32}=1,3/32=0,041$ с. Соответственно, длина вейвлета, способного идентифицировать временной интервал ноты наименьшей длительности при высоком темпе должна составлять не больше $t_d=t_{32}=0,041$ с. Значение граничной частоты идентификации основного тона ноты $\nu(t_d)=16/0,041=390$ Гц. Это означает, что точная идентификация времени начала и окончания звучания ноты возможна только для нот с частотой основного тона выше 390 Гц (рис. 9) (начиная с ноты «соль» первой октавы, частота основного тона которой составляет 392 Гц).

Для более медленного темпа исполнения произведений или для задач выявления нот большей длительности, чем t_{32} , граничная частота будет уменьшаться. Так, например, для современных клубных танцевальных композиций темп воспроизведения колеблется около значения 120 уд./мин, а самой короткой нотой является $t_d=t_{16}=0,125$ с. Значение граничной частоты идентификации основного тона ноты $\nu(t_d)=16/0,125=128$ Гц. Точная идентификация времени начала и окончания звучания ноты возможна только для нот выше «до» большой октавы с частотой основного тона 130,8 Гц. Большинство музыкальных произведений используют ноты из диапазона малой, первой и второй октав, что выше диапазона большой октавы.

Выводы

Предложенный подход позволяет создавать семейства материнских вейвлетов с заранее определенными избирательными способностями. Вейвлет, созданный из фрагмента определенного базового сигнала, позволяет выявлять не отдельные частотно-временные характеристики исследуемого сигнала, а частотно-временное поведение базового сигнала в исследуемом.

Экспериментально получены материнские вейвлеты для инструментов различных оркестровых групп: фортепиано, органа, скрипки, колокольчика и трубы. Все вейвлеты содержат по 16 периодов сигнала соответствующего музыкального инструмента с частотой основного тона 55 Гц. При вычислении НВП материнский вейвлет масштабируется так, чтобы очередной вейвлет $w_i(t)$ частотой основного тона совпадал с частотой основного тона ноты $n_i(t)$ музыкального инструмента.

Исследования семейств материнских вейвлетов, созданных на базе нот музыкальных инструментов, показали возможность выявлять частотные и временные параметры нот определенного инструмента в одноголосных и многоголосных мелодиях. Кроме того, в ряде экспериментов удалось идентифицировать мелодию определенного музыкального инструмента на фоне звучания другого [7].

Предполагается дальнейшее исследование применения данной методики в задачах, требующих выявления в сигнале фрагментов определенного семейства конечной длины на фоне других сигналов или помех.

СПИСОК ЛИТЕРАТУРЫ

1. Кудинов А.А. Использование распознавания образов для обработки и восстановления музыкальных сигналов: Дис. ... канд. техн. наук. – М., 2003. – 179 с.
2. Симаненков И.О. Цифровые синтезаторы музыкальных звуков // Компьютерра. – 1998. – № 32. – С. 21–24.
3. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения // Успехи физических наук. – 1996. – Т. 166. – № 11. – С. 1145–1170.
4. Фадеев А.С., Кочегурова Е.А. Выбор вейвлет-функций для анализа музыкальной информации // Современные техника и технологии: Труды XII Междунар. научно-практ. конф. студентов и молодых ученых. – Томск: ТПУ, 2006 – Т. 2. – С. 194–196.
5. Новиков Л.В. Основы вейвлет-анализа сигналов. – СПб.: ООО МОДУС+, 1999. – 152 с.
6. Курс лекций по сигналам и системам обработки сигналов. [Электронный ресурс] / авт. курса, А.В. Давыдов. – Электрон. текстовые дан. (32 файла) – [Екатеринбург]. – Режим доступа: <http://prodav.narod.ru/signals/index.html>, свободный. – Загл. с экрана. – Яз. рус., англ.
7. Кочегурова Е.А., Фадеев А.С. Вейвлет анализ в задаче идентификации музыкальной информации // Молодежь и современные информационные технологии: Сб. трудов IV Всеросс. научно-практ. конф. студентов, аспирантов и молодых ученых. – Томск: Изд-во ТПУ, 2006. – С. 149–151.
8. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MatLab. – М.: ДМК Пресс, 2005. – 304 с.
9. Фадеев А.С., Кочегурова Е.А. Подготовка результатов непрерывного вейвлет-преобразования к автоматизированной обработке // Известия Томского политехнического университета. – 2006. – Т. 309. – № 7. – С. 32–35.
10. Fujinaga I. Machine recognition of Timbre Using Steady-State tone of Acoustic musical Instruments // International Computer Music Conference. – Singapore, San Francisco, CA: ICMA, 2003. – P. 89–96.
11. Hebert S., Peretz I. Recognition of music in long-term memory: Are melodic and temporal patterns equal partners? // Memory & Cognition. – 1997. – V. 25. – № 4. – P. 518–533.

Поступила 19.10.2007 г.

УДК 004.93

ПРОСТОЙ АЛГОРИТМ ВЕЙВЛЕТ-СЖАТИЯ ПОЛУТОНОВЫХ И ЦВЕТНЫХ ИЗОБРАЖЕНИЙ

Д.В. Сидоров, А.Н. Осокин

Томский политехнический университет
E-mail: rauco@mail.ru

Разработан простой алгоритм вейвлет-сжатия, позволяющий реализовать недорогие аппаратно-программные платформы для систем видеонаблюдения и систем обработки телевидения высокой четкости в реальном времени.

Введение

Стандарт сжатия цифровых изображений JPEG2000 на основе вейвлет-преобразования считается одним из наиболее перспективных. Он состоит из двух обширных частей (JPEG2000 Part I, JPEG2000 Part II). В пределах стандарта существует множество программных реализаций (ACDSee, LeadTools, Mjp2000, Jasper, Lurawave и т. д.), значительно отличающихся степенью и скоростью сжатия изображений [1, 2]. Реализация JPEG2000 от ACDSee согласно [1] обладает наилучшими характеристиками: скоростью сжатия, степенью сжатия и качеством восстановленного изображения (далее при упоминании стандарта JPEG2000 будет рассматриваться реализация от ACDSee). JPEG2000 имеет явный выигрыш в 20...25 % в качестве восстановленного изображения по метрике PSNR [2, 3] (при равном коэффициенте сжатия) относительно обычного JPEG. Главной причиной ограниченного распространения JPEG2000 является сложность программной реализации и, как следствие, низкая скорость сжатия. Длительность процесса сжатия вызвана в большей степени сложностью этапов квантования, блочного кодирования и организации выходного

потока, занимающих порядка 80 % общего времени сжатия [2]. Поэтому, на основе первой и второй частей стандарта JPEG2000 (Part I, Part II) нами был разработан более простой и быстрый алгоритм вейвлет-сжатия QWC (Quick Wavelet Compress, название авторов), обладающий практически равным JPEG2000 качеством восстановленного изображения (при одинаковой степени сжатия).

Алгоритм сжатия QWC

Алгоритм QWC имеет аналогичную стандарту JPEG2000 схему сжатия (рис. 1).

Этапы предварительной обработки и вейвлет-преобразования стандарта JPEG2000 и алгоритма QWC совпадают. Согласно стандарту JPEG2000 предварительная обработка входного изображения состоит в выравнивании значений цветовых компонент RGB пикселей относительно нуля. Предварительная обработка изображения происходит только в том случае, если значения цветовых компонент RGB пикселей лежат в интервале $[0, 2^N - 1]$ (где N – разрядность бит цветовых компонент RGB или яркости пикселей изображения), тогда значение каждой цветовой компоненты RGB пикселей уменьшается на 2^{N-1} :