

УДК 004.043

## СПОСОБ ИЗУЧЕНИЯ РАЗНООБРАЗИЯ ОРНИТОФАУНЫ МЕТОДОМ ВЕЙВЛЕТ-АНАЛИЗА

©2013 М.Р. Богданов, А.А. Думчиков

Башкирский государственный педагогический университет им. М. Акмуллы, г. Уфа

Поступила 25.06.2013

Предложен подход распознавания голосов птиц методом вейвлет-анализа. Обсуждаются математические аспекты разработки использованных ортогональных вейвлетов. Предложено программное обеспечение, распознающее голоса птиц в режиме постпроцессинга.

**Ключевые слова:** голоса птиц, вейвлет-анализ, цифровая обработка сигналов, экология.

Птицы играют важную роль в различных экосистемах. Информация об их видовом составе и численности представляет интерес для специалистов из различных областей. Учет, особенно ночных, птиц затруднен. Было установлено, что звуковые сигналы несут много ценной информации. По голосам птиц можно судить об их видовом составе, численности и миграциях [1]. Для распознавания сигналов со сложными спектральными характеристиками широко применяется метод вейвлет-анализа [2]. В связи с этим возникла идея проверить возможность распознавания голосов птиц методом вейвлет-анализа. Целью предлагаемой работы является математическая проработка процесса распознавания голосов птиц и разработка программного обеспечения, предназначенного для распознавания голосов птиц в режиме постпроцессинга.

### МАТЕРИАЛ И МЕТОДЫ

В исследовании использовалась выборка из 269 голосов птиц. Выборка включала в себя голоса 166 видов птиц России, представленных от 1 до 14 вариантами. Варианты являлись разными трелями птиц одного вида, например, пение или беспокойство. Исходные звуковые файлы предварительно были преобразованы в единый звуковой формат: частота дискретизации 5512 Гц, разрядность 8 бит, моно, импульсно-кодовая модуляция. Далее, проводилось создание вейвлетов. Вейвлет – это некая функция, заданная таблично (небольшой массив чисел), определенным образом характеризующая сигнал [2]. Среди этих функций важную роль играют ортогональные вейвлеты. У такого вейвлета коэффициент корреляции с целевым сигналом стремится к единице, в то время как по отношению к другим сигналам корреляция этого вейвлета стремится к нулю [3].

Для создания вейвлета необходимо избавиться от избыточности исходного сигнала и выявить характерную область, адекватно описывающую этот сигнал. Устранение избыточности сигнала произ-

водится с помощью процедуры детектирования. Степень разреженности исходного сигнала задается размером так называемой рамки детектирования. Выбранная область детектированного сигнала подвергается процедуре нормирования. Нормированный участок детектированного сигнала и является вейвлетом [4]. Нормирование позволяет устранить влияние амплитуды сигнала на процесс распознавания.

Важно определить оптимальную длину вейвлета. Длинные вейвлеты более информативны, но могут давать ошибочную корреляцию с нецелевыми сигналами. Создав хорошие (ортогональные) вейвлеты, мы можем использовать их в дальнейшем для распознавания неизвестных сигналов [5].

В ходе экспериментов было установлено, что оптимальной длиной вейвлета является 70 отсчетов, оптимальный размер рамки детектирования – 100 отсчетов.

Рассмотрим процедуру создания вейвлета на примере голоса выпи большой. Вначале необходимо считать исходный звуковой сигнал и записать его в массив (рис. 1).

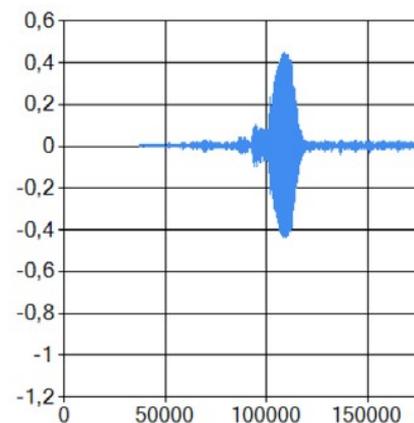


Рис. 1. Исходный массив голоса выпи большой

После этого требуется избавиться от избыточности сигнала. Для этого осуществляется процедура детектирования сигнала. Предположим, исходный сигнал записан в массив  $A_i$ ,  $i$  изменяется от 1 до  $N$ . Задается рамка считывания сигнала. Малый размер рамки повышает информативность, но обременяется большими накладными расходами в плане производительности. Как уже говорилось

Богданов Марат Робертович, к.б.н., доц. каф. прикладной информатики, e-mail: bogdanov\_marat@mail.ru; Думчиков Александр Александрович, студент, e-mail: redfox-ufa@gmail.com

выше, размер рамки был определен в 100 единиц. После этого весь диапазон массива  $A$  делится на  $N/100$  рамок. Внутри каждой рамки находили сумму квадратов  $sum_j = \sum_{i=1}^{100} a_i^2$ , где  $j$  – номер рамки, а  $i$  – индекс элемента массива  $a$  внутри  $j$ -й рамки. Далее, получали массив детектированного сигнала  $B_j$ , где элементы массива  $B$  вычисляются по фор-

муле:  $b_j = \sqrt{\frac{sum_j}{100}}$ . На рис. 2 показан детектированный сигнал голоса выпи большой. Видно, что размер его массива в сто раз меньше исходного.

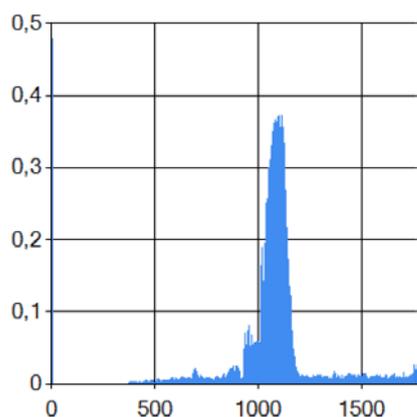


Рис. 2. Детектированный сигнал голоса выпи большой

Произведя детектирование сигнала, выбираем область, характерную для конкретного сигнала. Мы остановились на длине вейвлета в 70 отсчетов.

Выбрав характерную область детектированного сигнала  $B_j$ , где  $j$  изменяется от  $B_{min}$  до  $B_{max}$  ( $m = B_{max} - B_{min}$ ), осуществляем нормирование детектированного сигнала  $B_j$  на отрезке от  $B_{min}$  до  $B_{max}$ . Нормирование проводили следующим образом. Вначале находим среднее арифметическое  $r = \frac{1}{m} \sum_{B_{min}}^{B_{max}} B_j$ , сумму квадратов отклонений

$$sum_{sqr} = \sum_{B_{min}}^{B_{max}} (b_j - r)^2 \text{ и дисперсию } \sigma = \sqrt{\frac{sum_{sqr}}{m}}$$

Зная  $r$  и  $\sigma$ , можно найти вейвлет, являющийся массивом  $V_j = \frac{B_j - r}{\sigma}$  (рис. 3).

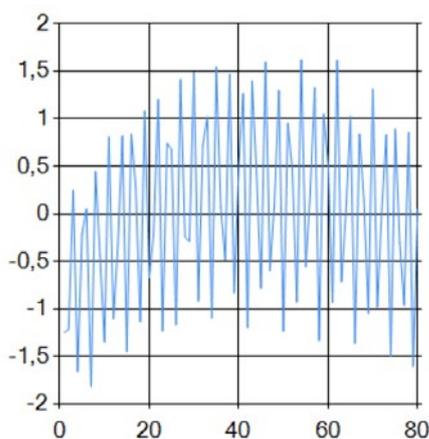


Рис. 3. Вейвлет голоса выпи большой

Процесс распознавания проходил следующим образом. На вход системы подается неизвестный сигнал  $U_i$ . Далее, производится детектирование сигнала с такой же рамкой (в нашем случае 100 отсчетов), что и при создании вейвлета. Получаем массив детектированного сигнала  $C_j$  (в сто раз меньше отсчетов, чем в исходном сигнале). Далее выбирается шаг распознавания  $myStep$ , равный размеру вейвлета  $m$  (в нашем случае  $m=70$ ). Внутри шага распознавания определяется коэффициент корреляции участка исследуемого сигнала  $C_j$ , равного  $myStep$ , со сравнимым вейвлетом. Определение коэффициента корреляции  $COR_j$  проводили следующим образом. На исследуемом участке находим среднее арифметическое  $r$  и дисперсию  $\sigma$  способом, используемым при создании вейвлета. На основании  $r$  и  $\sigma$  нормировали выбранный участок исследуемого детектированного сигнала, получая массив  $D_j$ . Имея нормированный детектированный сигнал  $D_j$  и вейвлет  $V_j$ , находим коэффициент корреляции  $COR_j$ .

$$COR_j = \sqrt{\frac{\sum_j^{j+myStep} D_j \times V_j}{myStep}}$$

Найдя коэффициент корреляции на выбранном шаге, сдвигаем рамку распознавания на один элемент и возобновляем процесс (рис. 4).

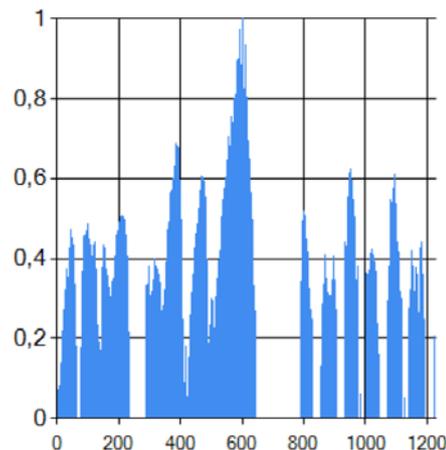


Рис. 4. Корреляционная функция между голосом выпи большой и вейвлетом выпи большой

Если коэффициент корреляции неизвестного сигнала и известного вейвлета близок к единице, программа принимает решение об успехе распознавания.

### РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Испытания программного обеспечения, разработанного в соответствии с вышеприведенными рекомендациями выявили некоторые проблемы. Так, все вейвлеты давали коэффициент корреляции, равный единице к целевому сигналу. Однако, чем больше использовалось сигналов, тем труднее было разработать вейвлеты, дающие низкие коэффициенты корреляции по отношению ко всем остальным

сигналам (менее 0,65). Эта работа очень усложнялась, когда сигналов становилось больше десяти. Была построена корреляционная матрица, включающая в себя коэффициенты корреляции между 269 сигналами и 269 вейвлетами. К сожалению, результаты оказались неоднозначными – часть вейв-

летов обладала высокой степенью ортогональности. Например, вейвлет сороки дал низкий коэффициент корреляции для 72 сигналов, вейвлет лугового луня – дал низкую корреляцию для 37-ми голосов птиц (табл.).

**Таблица.** Степень ортогональности различных вейвлетов

Вейвлет	Количество случаев относительно низкого коэффициента корреляции к звуковым сигналам ( $r < 0,65$ )
Сорока_1	72
Луговой лунь	37
Камышовка-барсучок	35
Вертишейка	30
Пеночка-трещотка	19
Белый аист	17
Чика	17
Канарейка	9
Большой пестрый дятел	7
Глухарь_3	5
Горлица обыкновенная	3
Ворон_1	2
Кенарь	2
Соловей_2	2
Большая синица_2	1
Дятел_2	1
Козодой_2	1
Крапивник_1	1
Пустельга_2	1
Садовая камышовка	1
Сова полярная	1
Сова ястребиная	1
Соловей_5	1
Соловей_13	1
Сплюшка	1
Широконоска	1

Стало очевидно, что надежному распознаванию всех голосов птиц мешают вейвлеты с высокой степенью ортогональности. Для решения этой проблемы было решено перейти к частичной ортогональности вейвлетов, используя при распознавании голосов птиц, эвристики, полученные при анализе корреляционной матрицы. Предлагается следующая стратегия.

Пример эвристики: так как коэффициент корреляции неизвестного сигнала с вейвлетом клуши превышает 0,99 и равен  $0,62 \pm 0,05$  для вейвлета большой синицы\_1, то исследуемый сигнал является голосом клуши.

1. Сравнивается первый звуковой сигнал со всеми вейвлетами, находится вейвлет с минимальным коэффициентом корреляции, результат заносится в таблицу. Например, для голоса балабана

минимальную корреляцию из всех вейвлетов дал вейвлет лесного конька ( $r=0,42$ ).

2. Удаляется ранее использованный вейвлет (лесного конька) из числа вейвлетов, участвующих в распознавании (в нашем случае, после этого остается 268 вейвлетов).

3. Берется новый звуковой файл (Бекас) и снова строится вейвлет из числа имеющихся 268 экземпляров, дающих минимальный к нему коэффициент корреляции (в нашем случае это – вейвлет вертишейки с  $r=0,59$ ).

4. Удаляется вейвлет вертишейки из числа вейвлетов, участвующих в распознавании.

5. Повторяются этапы с первого по четвертый до последнего звукового сигнала.

Важным моментом здесь является то, что классифицирующим множеством являются сами вейв-

леты. Так как мы постепенно отбрасываем наиболее ортогональные вейвлеты, постепенно накапливаются вейвлеты с высокими коэффициентами корреляции. Для решения этой проблемы можно расширить классифицирующее множество. Предположим, мы хотим создать систему распознавания голосов ста видов птиц России. Для повышения надежности распознавания можно классифицирующее множество расширить за счет вейвлетов, например тропических птиц или даже вейвлетов звуков автомобилей. То есть, сто звуков российских птиц будут сравниваться с двумястами-тремястами вейвлетами.

Разработанное программное обеспечение позволяет классифицировать различные трели птиц од-

ного вида. Распознаются 166 видов России с надежностью 0,87. Удастся распознать различные трели птиц одного вида.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Богданов М.Р., Богданова А.В., Габидуллин Ю.З.* Беспроводные сенсорные сети в экологии // Информационные технологии. 2010. № 7. С.39-43.
2. *Mallat S.G.* A Wavelet Tour of Signal Processing. Academic Press, 1999. 637 p.
3. *Hubbard B.B.* World According to Wavelets: The Story of a Mathematical Technique in Making / 2-nd Edition. University Press (India) Pvt. Limited. 1988. 357 p.
4. *Nievergelt Y.* Wavelets Made Easy. Springer, 1999. 297 p.
5. *Aboufadel E., Schlicker S.* Discovering wavelets. Wiley, 1999. 125 p.

### THE APPROACH FOR DISCOVERING OF BIRD'S POPULATION WITH WAVELET ANALYSIS

©2013 M.R. Bogdanov, A.A. Dumchikov

Bashkir State Pedagogical University named after M. Akmullah, Ufa

The approach for recognition of bird's voices with wavelet analysis is proposing. Mathematical aspects of orthogonal wavelets designing for recognition are discovering. The software for bird's voices recognition in post processing mode is discussing.

**Keywords:** *bird's voices, wavelet analysis, digital signal processing, ecology.*