Парламентское собрание Союза Беларуси и России

Постоянный Комитет Союзного государства

Оперативно-аналитический центр
при Президенте Республики Беларусь

Государственное предприятие «НИИ ТЗИ»

Полоцкий государственный университет

КОМПЛЕКСНАЯ ЗАЩИТА ИНФОРМАЦИИ

Материалы XXII научно-практической конференции

(Полоцк, 16–19 мая 2017 г.)

Новополоцк
2017

УДК 004(470+476)(061.3)
ББК 32.81(4Беи+2)
К63

К63 **Комплексная защита информации :** материалы XXII науч.-практ. конф., Полоцк, 16–19 мая 2017 г. / Полоц. гос. ун-т ; отв. за вып. С. Н. Касанин. – Новополоцк : Полоц. гос. ун-т, 2017. – 282 с.
ISBN 978-985-531-564-4.

В сборнике представлены доклады ученых, специалистов, представителей ггосударственных органов и практических работников в области обеспечения информационной безопасности Союзного государства по широкому спектру научных напрвлений.

Адресуется исследователям, практическим работникам и широкому кругу читателей.

Тексты тезисов докладов, вошедших в настоящий сборник, представлены в авторской редакции.

УДК 004(470+476)(061.3)
ББК 32.81(4Беи+2)

ISBN 978-985-531-564-4

Сверточная нейронная сеть является многослойной. Используются слои двух типов: сверточные и подвыборочные (многослойный персептрон). Сверточные и подвыборочные слои чередуются друг с другом. В свою очередь, каждый из этих слоев состоит из набора плоскостей, причем нейроны одной плоскости имеют одинаковые веса (так называемые общие веса), ведущие ко всем локальным участкам предыдущего слоя (как в зрительной коре человека).

Были проведены исследования зависимости качества прогнозов отказов турбогенератора от параметров алгоритма обучения и структуры нейронной сети. Результаты исследований показали, что качество прогнозирования зависит от разбиения отсчетов ряда спектрограмм на три множества – обучающее, тестирующее и контрольное. Наилучшее качество прогноза достигается при соотношении объемов выборок 250:150:100.

Эффективное решение задачи прогнозирования возможно только в том случае, если нейронная сеть обучается на большом объеме данных. В случае малоразмерной или некачественной обучающей выборки разработанный алгоритм не дает удовлетворительного результата.

Список литературы

1. Kim, Y. Convolutional neural networks for sentence classification / Y. Kim // arXiv preprint arXiv:1408.5882. – 2014.

2. Lakkaraju, H. Aspect Specific Sentiment Analysis using Hierarchical Deep Learning / H. Lakkaraju, R. Socher, C. Manning // NIPS Workshop on Deep Learning and Representation Learning. – 2014.

3. Natural language processing (almost) from scratch / R. Collobert [et al.] // The Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2493–2537.

4. Recurrent neural network based language model / T. Mikolov [et al.] // INTERSPEECH 2010: 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan (September 26-30, 2010). – 2010. – P. 1045–1048.

5. Ивченко В.Д. Диагностика и отказоустойчивость технических систем. Методы обработки информации и принятия решений. – М.: Машиностроение - 1, 2006. – 305 с.

ПРЕИМУЩЕСТВО ВЕЙВЛЕТ-ПРЕОБРАЗОВАНИЯ В ОБРАБОТКЕ РЕЧЕВЫХ СИГНАЛОВ

К.А. НАРОВ, В.К. ЖЕЛЕЗНЯК, И.Б. БУРАЧЕНОК

Полоцкий государственный университет

Введение. Сегодня идентификация и верификация личности по голосу является одной из важных задач. Основными направлениями развития технологии обработки речевого сигнала является распознавание диктора по голосу из заданного, ограниченного списка людей (идентификация личности) и подтверждение личности говорящего (верификация личности). Известные методы анализа звуков речи основываются на спектральной модели стационарного сигнала, однако в речевом сигнале самым информативным являются его частотно-временные характеристики. Использование вейвлетов в задачах обработки и распознавания речи продиктовано особенностями речевого (акустического) сигнала.

Целью работы является обоснование преимущества вейвлет-преобразования при анализе речевых сигналов для решения задачи идентификация и верификация личности по голосу. Классическим подходом для цифровой обработки акустического сигнала является преобразование амплитудно-временной зависимости в частотный вид. В большинстве современных алгоритмов обработки сигналов применяется прямое преобразование Фурье. Однако, несмотря на популярность преобразования Фурье для частотного представления сигнала, форматы этого представления и представления вейвлет имеют значительные отличия. Существуют также некоторые фундаментальные ограничения Фурье-анализа для представления нестационарных сигналов с быстрыми перепадами уровня сигнала. Эти серьезные ограничения были преодолены за счет применения специального аппарата представления сигналов и функции на основе нового математического базиса – вейвлетов. Использование вейвлетов является новым научным направлением, возникшим на стыке математики, информатики и техники связи. Благодаря работам Мейера, Добеши, Койфмана, Малла и других учёных [1,2] теория вейвлетов, за счёт реализации возможностей сжатия и фильтрации данных, анализа в базисе вейвлет-функций, позволяет решать следующие задачи: идентификации, моделирования, аппроксимации стационарных и нестационарных процессов, исследовать вопросы наличия разрывов в производных, осуществлять поиск точек склеивания данных, удалять в данных тренд, отыскивать признаки фрактальности информации и т.п. Сегодня вейвлеты нашли своё применение и при анализе тонкой структуры сигналов [3, 4]. В основу оценки тонкой структуры речевого сигнала положена теория разложения сигналов по специальным базисам (вейвлет-функциям), которые являются функциями двух аргументов: масштаба a и сдвига b.

Непрерывное вейвлет-преобразование представляют:

$$W(a,b)=\frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} s(t)\,\psi^{*}\left(\frac{t-b}{a}\right)dt, \qquad (1)$$

где $s(t)\in L^{2}(R)$;

$s(t)$ – исходный сигнал;

ψ – базисная функция;

t – время;

a – масштаб;

b – сдвиг;

$a,b\in R,\ a>0$.

Важной особенностью вейвлет-анализа является то, что в нем можно использовать большое число основных вейвлет-функций, реализующих различные варианты соотношения в сигнале между частотой и его локализацией по времени как показано на рисунке 1.

Имеется возможность выбора между семействами вейвлетных функций и гибкого применения тех, которые наиболее эффективно решают конкретную задачу.

Использовав вейвлет-преобразования в речевом анализе, важно уделить внимание выбору и настройке вейвлетов. Это может существенно влияет на результаты преобразования и возможность дальнейшей интерпретации получаемых результатов, так как при изменении характеристик исследуемого сигнала применив одни и те же настройки для различных вейвлетов можно получить совершенно не предсказуемые результаты, например, искаженную информацию. Таким образом, для успешного анализа тонкой структуры и особенностей анализируемого сигнала, возникает необходимость анализа параметров материнского вейвлета. Для этого на основании анализа гласных

звуков русского языка необходимо обосновать выбор материнского вейвлета с учетом его частотных характеристик.

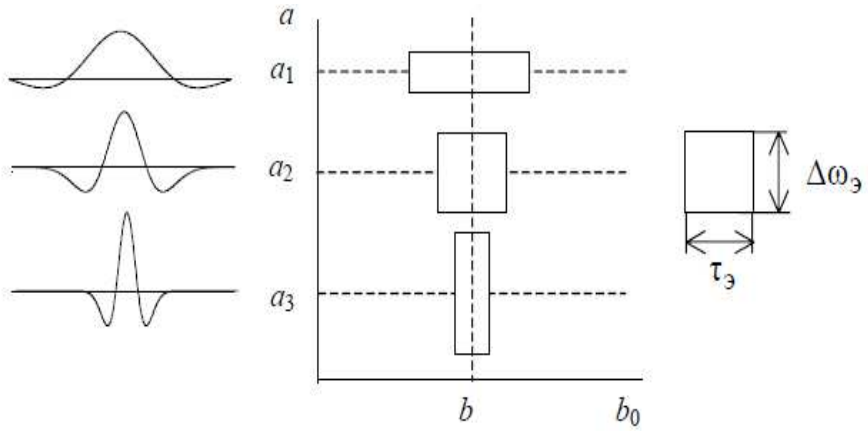

Рис. 1 – Варианты соотношения в сигнале между частотой и его локализацией по времени

В работе [3] были исследованы: вейвлеты Гауссова типа, «Мексиканской шляпы», Морле и Меера. Перечисленные вейвлеты, обладают минимумом свойств, обеспечивающих полноценные возможности в технике преобразования и обладают свойствами частотно-временной локализации.

Из работы [3] можно отметить, что из исследованных вейвлет-функций, комплексный вейвлет Морле имеет более узкий Фурье образ и продолжителен во временной области.

Комплексный вейвлет Морле описывается функцией:

$$\psi(t) = \frac{1}{\sqrt{\pi B}} e^{\frac{-t^2}{B}} e^{i2\pi f_0 t}. \qquad (2)$$

Во временной области комплексный вейвлет Морле представляет собой комплексную экспоненту, модулируемую функцией Гаусса. В частотной же области он имеет форму Гауссова окна с центральной частотой f_0 и шириной B. Таким образом, частотный диапазон, покрываемый окном комплексного вейвлета Морле, ограничен интервалом $[f_0 - B/2, f_0 + B/2]$ (полосой его пропускания), где сосредоточена наибольшая часть его энергии.

Основными преимуществами комплексного вейвлета Морле являются: присутствие доминирующей частоты позволяет варьировать избирательностью вейвлета в частотной области; он обладает частотной локализацией лучшей среди других базисов; на практике при $f_0 \gg 0$ вейвлет Морле может быть использован с минимальной погрешностью; позволяет разделить фазовые и амплитудные компоненты сигнала при выполнении вейвлет-преобразования; имеет близкое сходство с речевыми фрагментами и в отличие от преобразования Фурье и применения других вейвлетов, дает более высокую степень «гладкости» скейлограммы.

Исследования звуков речевого сигнала при использовании комплексного вейвлета Морле [3, 4] показало, что форма нормированных пиков исследованных участков сигнала индивидуальна, что позволяет обнаружить отличительные особенности голоса

диктора. Двумерное представление исследуемого сигнала в частотной области в плоскости частота-положение позволяет разделять крупные и мелкие структурные особенности сигналов, одновременно локализуя их на временной шкале как показано на рисунке 2.

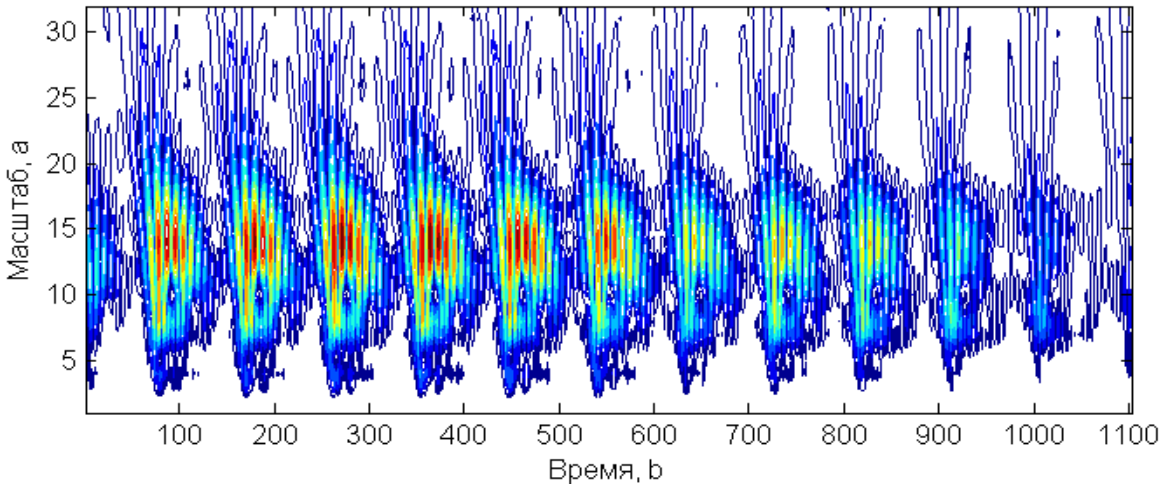

Рис. 2 – Топологическая карта звука речи

Вывод. Применение вейвлет-преобразований позволяет дополнить характеристики сигналов, получаемые обычными методами (в частности спектральными) и расширяет имеющиеся подходы к оценке их скейлинговых параметров, что значительно повышает точность при определении значения частоты основного тона исследуемых речевых сигналов. Практика использования рядов Фурье показала, что базисные функции – синус и косинус – явно неудачны для представлений функций и сигналов с локальными особенностями, например, разрывами, скачками, резкими перепадами и значительно уступает вейвлет-преобразованиям.

Бесконечное число членов в ряде Фурье обусловливает большие погрешности, что исключает вейвлет-переобразование. В отличие от преобразований Фурье, вейвлет-преобразование одномерных сигналов обеспечивает двумерную развёртку, при этом частота и координата рассматриваются как независимые переменные, что даёт возможность анализа сигналов сразу в двух пространствах и с гораздо более высокой точностью представляет локальные особенности речевых сигналов.

Выбор в качестве материнского комплексного вейвлета Морле и результаты анализа экспериментальных исследований параметров гласных звуков, полученных с помощью данного вейвлет-преобразования, позволяют при настройке параметров вейвлета получать тонкую структуру анализируемого речевого сигнала конкретного диктора в реальном масштабе времени.

Список литературы

1. Добеши И. Десять лекций по вейвлетам (пер. с англ.). – Ижевск : НИЦ Регулярная и хаотическая динамика, 2001. – 464 с.

2. Мала С. Вейвлеты в обработке сигналов. –М. : Мир. 2005. – 671 с.

3. Бурачёнок И. Б., Железняк В. К. Анализ вейвлет-преобразованием тонкой структуры гласных звуков речевого сигнала. // Теоретические и прикладные аспекты информационной безопасности : материалы междунар. науч.-практ. конф., Минск, 19 июня 2014 г. / УО «Акад.

М-ва внутр. дел Респ. Беларусь»; редкол.: В. Б. Шабанов (отв. ред.) [и др.]. – Минск, 2015. С. 124–128.

4. Рыбальский О.В., Соловьев В.И., Железняк В.К. Спектральный анализ и современные речевые технологии. // Вестник Полоцкого государственного университета. Серия С. – 2014. – №4. – С. 2-6.

ЭНТРОПИЙНЫЙ АНАЛИЗ КРИПТОГРАФИЧЕСКИХ ГЕНЕРАТОРОВ СЛУЧАЙНЫХ И ПСЕВДОСЛУЧАЙНЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ

В.Ю. ПАЛУХА, Ю.С. ХАРИН

НИИ прикладных проблем математики и информатики БГУ

Введение. Важным структурным элементом средств криптографической защиты информации (криптосистем) являются генераторы случайных и псевдослучайных последовательностей [1]. Стойкость криптосистем зависит от того, насколько близка генерируемая последовательность по своим свойствам к равномерно распределённой случайной последовательности (РРСП).

Для проверки качества криптографических генераторов используются статистические тесты, суть которых заключается в следующем. Наблюдается выходная последовательность криптографического генератора и вводится гипотеза H_* о том, что последовательность является РРСП. Вычисляется некоторая статистика, распределение вероятностей которой при истинной гипотезе H_* известно. На основании значения статистики гипотеза H_* принимается либо отклоняется. В данном докладе рассматривается применение статистической оценки энтропии Шеннона в качестве тестовой статистики для анализа выходных последовательностей криптографических генераторов.

Математическая модель. Пусть на вероятностном пространстве (Ω, F, P) с множеством состояний $\Omega = \{\omega_1, \ldots, \omega_N\}$ определена случайная величина $x = x(\omega) = \omega$ с дискретным распределением вероятностей $P = \{p_k\}$, $p_k = P\{x = \omega_k\}$, $p_k \geq 0$, $\sum_{k=1}^{N} p_k = 1$, $k = 1, \ldots, N$. Энтропия Шеннона определяется формулой [1]

$$H(P) = -\sum_{k=1}^{N} p_k \ln p_k. \tag{1}$$

Пусть имеется случайная последовательность $\{x_t : t = 1, \ldots, n\}$ объёма n из распределения вероятностей $\{p_k\}$. Построим частотные оценки распределения вероятностей $\{p_k: k = 1, \ldots, N\}$:

$$\hat{p}_k = \frac{\nu_k}{n}, \quad \nu_k = \sum_{t=1}^{n} I\{x_t = \omega_k\} \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}, \quad I\{x_t = \omega_k\} = \begin{cases} 1, & x_t = \omega_k; \\ 0, & x_t \neq \omega_k. \end{cases} \tag{2}$$

Как уже было сказано во введении, введём в рассмотрение гипотезу H_* = {{x_t} является РРСП} = {{x_t} – н.о.р.с.в., $p_k = 1/N$, $k = 1, \ldots, N$} и альтернативу $\overline{H_*}$.

Следуя [2], будем полагать, что имеет место схема серий. В таком случае вектор $(\nu_1, \ldots, \nu_N)^T$, составленный из частот ν_k из (2), имеет полиномиальное распределение вероятностей Pol(n, N, p_1, …, p_N), а каждая из компонент распределена по биномиальному