

Денисов Максим Владимирович

**ВЫДЕЛЕНИЕ КЛЮЧЕВЫХ СЛОВ ИЗ ПОТОКА РЕЧИ**

Адрес статьи: [www.gramota.net/materials/1/2011/11/17.html](http://www.gramota.net/materials/1/2011/11/17.html)

Статья опубликована в авторской редакции и отражает точку зрения автора(ов) по рассматриваемому вопросу.

Источник

**Альманах современной науки и образования**

Тамбов: Грамота, 2011. № 11 (54). С. 56-58. ISSN 1993-5552.

Адрес журнала: [www.gramota.net/editions/1.html](http://www.gramota.net/editions/1.html)

Содержание данного номера журнала: [www.gramota.net/materials/1/2011/11/](http://www.gramota.net/materials/1/2011/11/)

**© Издательство "Грамота"**

Информация о возможности публикации статей в журнале размещена на Интернет сайте издательства: [www.gramota.net](http://www.gramota.net)

Вопросы, связанные с публикациями научных материалов, редакция просит направлять на адрес: [almanac@gramota.net](mailto:almanac@gramota.net)

архитектурно-градостроительные приемы береговых укреплений заслуживают использования в современной проектной практике. Весь комплекс фортификационных сооружений может быть включен в туристско-рекреационное использование.

#### Список литературы

1. **Амирханов Л. И.** Морская крепость Императора Петра Великого. СПб.: Иванов и Лещинский, 1995. 100 с.
2. **Госс А.** Береговая артиллерия [Электронный ресурс]. URL: <http://www.nortfort.ru/coastal/> (дата обращения: 15.06.2010).
3. **Кириллов В. В.** Архитектура «северного модерна». М.: Эдиториал УРСС, 2001. 110 с.
4. **Сильваст П.** Сооружения крепости императора Петра Великого на полуострове Ханко // Цитадель. 1996. № 2. С. 26-41.

УДК 004.522

*Максим Владимирович Денисов*  
*Волгоградский государственный технический университет*

### ВЫДЕЛЕНИЕ КЛЮЧЕВЫХ СЛОВ ИЗ ПОТОКА РЕЧИ<sup>©</sup>

#### Введение

В настоящее время все чаще используется распознавание речи. Данная технология практически незаменима для создания дополнительного канала управления в различных автоматизированных системах, таких, как человек-машинный интерфейс взаимодействия, средства для людей с ограниченными возможностями, в сфере информационной безопасности, в сфере образования для проведения тестирования и многих других.

В настоящее время в большинстве случаев используются скрытые марковские модели. Скрытые марковские модели основаны на вероятностном соотношении порядка фонем в слове. Они хорошо работают при небольшом размере словаря, но при размере словаря больше определенного предела данный метод практически перестает работать. Таким образом, разработка метода, способного распознавать ключевые слова в потоке речи с большим размером словаря является актуальной задачей. Одним из важнейших этапов распознавания речи является предварительный анализ звукового сигнала с целью выделения признаков, которые впоследствии будут поданы на вход модели распознавания ключевых слов. Далее к данной задаче возможно применить рекуррентные нейронные сети, обучаемые генетическим алгоритмом. Применение генетического алгоритма при обучении рекуррентной нейронной сети позволит повысить размер применяемого словаря.

#### Механизм распознавания ключевых слов

Для распознавания ключевых слов из потока речи в большинстве случаев используется метод скрытых марковских моделей, но при таком подходе при большом размере словаря данный механизм практически перестает работать. Поэтому была предпринята попытка применить коннективистскую модель, основанную на нейронных сетях.

В данной работе используется рекуррентная нейронная сеть с генетическим алгоритмом обучения для построения модели распознавания речи. Рекуррентные нейронные сети - это наиболее сложный вид нейронных сетей, в которых имеются обратные связи. Наличие обратных связей позволяет запоминать и воспроизводить целые последовательности реакций на один стимул. Использование механизмов генетической эволюции для обучения нейронных сетей кажется естественным, поскольку модели нейронных сетей разрабатываются по аналогии с мозгом и реализуют некоторые его особенности, появившиеся в результате биологической эволюции. Такие особенности потенциально предоставляют множество возможностей для моделирования биологических нейронных сетей [1].

#### Подготовка входных данных для включения в модель

Речевой сигнал описывается последовательностью отчетов. После обработки последовательности отчетов сигнала получена матрица характерных признаков, которая поступает в модель. Матрица характерных признаков имеет вид

$$\begin{matrix} x_{11} & x_{12} & \dots & x_{1l} \\ x_{21} & x_{22} & \dots & x_{2l} \\ x_{31} & x_{32} & \dots & x_{3l} \\ \dots & \dots & \dots & \dots \\ x_{Q1} & x_{Q2} & \dots & x_{Ql} \end{matrix},$$

где  $x_{q1} \ x_{q2} \ \dots \ x_{qI}$  - вектор характерных признаков q-го кадра сигнала. Рекуррентная нейронная сеть распознавания речи представляет собой трехслойную сеть с L обратными связями. Трехслойная сеть состоит из I входных, J скрытых и K (K=1) выходных нейронов [2].

#### **Описание рекуррентной нейронной сети распознавания речи**

Входными сигналами является матрица характерных признаков

$$\begin{matrix} x_{11} & x_{12} & \dots & x_{1I} \\ x_{21} & x_{22} & \dots & x_{2I} \\ x_{31} & x_{32} & \dots & x_{3I} \\ \dots & \dots & \dots & \dots \\ x_{Q1} & x_{Q2} & \dots & x_{QI} \end{matrix},$$

где Q - количество векторов характерных признаков, I - размер вектора характерных признаков, и количество нейронов во входном слое.

Количество нейронов в скрытом слое не должно быть более  $(2I+1)$ , где I - количество нейронов во входном слое (по теореме Колмогорова).

Единственный нейрон в выходном слое имеет выходное значение, лежащее в диапазоне  $[0, 1]$ .

Количество обратных связей 3-5.

Начальные весовые коэффициенты нейронов производятся в соответствии с методом инициализации весов Нгуена-Видроу.

Весовые коэффициенты нейронов лежат в диапазоне  $[-50, 50]$ .

В качестве активационной функции используется сигмовидная функция.

При обучении желаемое значение нейронной сети соответствует единице, если выборка соответствующего распознаваемого слова на входе, остальные - нулю [4].

#### **Обучение рекуррентных нейронных сетей**

Основная идея обучения нейронных сетей состоит в подстройке весов для фиксированного набора связей, при которых достигается минимальная ошибка. Наиболее распространенная парадигма обучения нейронных сетей - обратное распространение ошибки. Но его недостаток - возможность преждевременной остановки из-за попадания в область локального минимума. Эту проблему решает использование генетического алгоритма, который благодаря оператору мутации менее склонен к попаданию в локальные минимумы. Генетический алгоритм позволяет сократить время и улучшить качество обучения нейронных сетей. Использование механизмов генетической эволюции для обучения нейронных сетей кажется естественным, поскольку модели нейронных сетей разрабатываются по аналогии с мозгом и реализуют некоторые его особенности, появившиеся в результате биологической эволюции [3].

#### **Кодирование признаков хромосомы**

Для кодирования признаков можно использовать самый простой вариант - битовое значение этого признака. Однако основной недостаток заключается в том, что соседние числа отличаются в значениях нескольких битов, так, например, числа 7 и 8 в битовом представлении различаются в 4-х позициях, что затрудняет функционирование генетического алгоритма и увеличивает время, необходимое для его сходимости. Для того чтобы избежать этой проблемы, лучше использовать кодирование, при котором соседние числа отличаются меньшим количеством позиций, в идеале - значением одного бита. Таким кодом является код Грея.

#### **Оценка качества модели**

База звуковых выборок:

- 5 отдельных слов;
- 11 выборок с различными вариациями произношения одного слов: 6 обучающих и 5 тестовых.

Точность обучения (обучающие выборки):

$$O = \frac{N}{Q} 100(\%),$$

где N - количество распознанных слов из обучающей выборки;

Q - количество слов в обучающей выборке.

Точность обобщения (тестовые выборки):

$$T = \frac{M}{D} 100(\%),$$

где M - количество распознанных слов из тестовой выборки;

D - количество слов в тестовой выборке [Там же].

#### **Выводы**

Была построена модель распознавания ключевых слов на основе рекуррентных нейронных сетей. Характерные признаки входного сигнала поступают на вход нейронной сети. Ключевое слово считается распознанным, если на выход нейронной сети попадает в отрезок  $[0,9; 1]$ . Также к данной нейронной сети был

применен генетический алгоритм обучения. В качестве операторов генетического алгоритма используются операторы мутации, скрещивания, селекции, инверсии.

Таким образом, построенный прототип системы распознает ключевые слова с точностью 67%.

#### Список литературы

1. Камаев В. А., Щербаков М. В., Панченко Д. П., Щербакова Н. Л., Бребельс А. Применение коннективистских систем для прогнозирования потребления электроэнергии в торговых центрах // Управление большими системами. М.: ИПУ РАН, 2010. Вып. 31. С. 92-109.
2. Специфика применения интеллектуальных моделей анализа данных для повышения энергетической эффективности / М. В. Щербаков, Н. Л. Щербакова, Д. П. Панченко, А. Бребельс, А. П. Тюков, М. А. Аль-Гунаид // Известия ВолгГТУ. Серия «Актуальные проблемы управления, вычислительной техники и информатики в технических системах». Волгоград, 2010. Вып. 9. № 11. С. 72-76.
3. Щербаков М. В., Камаев В. А. Использование нейросетевых технологий для анализа сложных иерархических систем / М. В. Щербаков, В. А. Камаев, Д. П. Панченко // Вестник Брянского государственного технического ун-та. 2004. № 1. С. 202-207.
4. Shcherbakov M. V., Panchenko D. P. Implementation of Genetic Algorithms for Transit Points Arrangement // Information Technologies & Knowledge: suppl. of "Information Science and Computing". 2009. Vol. 3. № 9. P. 129-131.

УДК 004.522

Максим Владимирович Денисов  
Волгоградский государственный технический университет

### ПРЕДВАРИТЕЛЬНЫЙ АНАЛИЗ ЗВУКОВОГО СИГНАЛА<sup>©</sup>

#### Введение

В настоящее время все чаще используется распознавание речи. Данная технология практически незаменима для создания дополнительного канала управления в различных автоматизированных системах, таких, как человеко-машинный интерфейс взаимодействия, средства для людей с ограниченными возможностями, в сфере информационной безопасности, в сфере образования для проведения тестирования и многих других.

В настоящее время в большинстве случаев используются скрытые марковские модели. Скрытые марковские модели основаны на вероятностном соотношении порядка фонем в слове. Они хорошо работают при небольшом размере словаря, но при размере словаря больше определенного предела данный метод практически перестает работать. Таким образом, разработка метода, способного распознавать ключевые слова в потоке речи с большим размером словаря является актуальной задачей. Одним из важнейших этапов распознавания речи является предварительный анализ звукового сигнала с целью выделения признаков, которые впоследствии будут поданы на вход модели распознавания ключевых слов.

#### Методика выделения характерных признаков речевого сигнала

Для однозначной идентификации ключевого слова необходимо выделить характерные признаки данного ключевого слова. Для решения этой задачи был разработан следующий алгоритм, который разделен на 7 этапов:

1. Выделение амплитуд из исходного сигнала

На вход системы поступает звуковой файл формата "wav". Звуковые файлы данного типа - контейнерные файлы, которые содержат файл формата "pcm". В файле "pcm" хранится последовательность амплитуд.

2. Выделение потока речи из последовательности амплитуд

Весь сигнал разбивается на последовательность кадров длиной в 256 отсчетов.

В любом сигнале первые 150-200 мс являются тишиной. Вследствие этого значения отсчетов в этом временном диапазоне можно считать случайной величиной.

Случайная величина  $x$  имеет нормальное распределение, если её плотность распределения определяется зависимостью:

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},$$

где  $\mu$  - среднее значение случайной величины;

$\sigma$  - нормальное распределение случайной величины.

Среднее значение случайной величины определяется следующей формулой:

$$\mu = \frac{1}{N} \sum_{i=1}^N X_i,$$

где  $N$  - количество случайной величины;