

Чуркин Вадим Сергеевич

## **ИСПОЛЬЗОВАНИЕ ЭФФЕКТА МАСКИРОВКИ ДЛЯ РАСПОЗНАВАНИЯ РЕЧЕВЫХ СИГНАЛОВ**

В текущей статье исследуются такая особенность слухового восприятия человека как эффект маскировки и влияние этого эффекта на проблему распознавания речевых сигналов. Для проведения данного исследования автором статьи разработана аппаратно-программная система, которая имитирует эффект маскировки, производит работу со словарем и осуществляет распознавание слов. В заключение представлены результаты, которые удалось получить на текущий момент.

Адрес статьи: [www.gramota.net/materials/1/2013/7/48.html](http://www.gramota.net/materials/1/2013/7/48.html)

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

Источник

### **Альманах современной науки и образования**

Тамбов: Грамота, 2013. № 7 (74). С. 152-154. 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/1/2013/7/](http://www.gramota.net/materials/1/2013/7/)

### **© 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)

состояние художника. Основное требование при составлении глоссария: для раскрытия понятия должны использоваться символы, которые отражают его наиболее характерные черты с точки зрения учащегося. Так, составляя словарь-глоссарий по теме «Разделенные колючей проволокой», учащиеся должны использовать лексику, характеризующую категорию «мир – опасность».

Таким образом, формирование образной картины мира в старших классах целесообразно осуществлять с использованием актуальной для этого возраста категории мировосприятия «мир – опасность», развивая навыки критического мышления, осмысления и выработки собственного отношения к миру, основанного на моральных ценностях гражданского общества.

#### Список литературы

1. Асмус В. Чтение как труд и творчество // Вопросы литературы. 1961. № 2. С. 36-46.
2. Бахтин М. М. Эстетика словесного творчества. М.: Искусство, 1979. 424 с.
3. Борев Ю. Б. Эстетика. Теория литературы: энциклопедический словарь терминов. М.: АСТ; Астрель, 2003. 575 с.
4. Глушенков О. В. Формирование у школьников экологической картины мира в процессе проведения полевого практикума по биологии: автореф. дисс. ... к. пед. н. М., 2001. 18 с.
5. Левидов А. М. Автор – образ – читатель. Л.: Изд-во Ленинградского университета, 1977. 358 с.
6. Никанова Н. В. Изучение художественного мира Ф. М. Достоевского в 10-х классах гуманитарного профиля: автореф. ... к. пед. н. М., 2008. 22 с.
7. Программа по литературе для 5-9-х классов средней школы / под ред. чл.-корр. РАО, проф. В. Г. Маранцмана. СПб.: СпецЛит, 2000. 301 с.
8. Тарасов С. В. Образ мира: опыт изучения категориальных структур мировосприятия школьников: научное издание. СПб.: РАО, Ин-т образования взрослых, 1996. 75 с.
9. Тарасов С. В. Теоретико-методологические основы становления мировосприятия школьников в условиях современной социокультурной среды: автореф. дисс. ... д. пед. н. СПб., 2001. 47 с.
10. Художественное восприятие. Основные термины и понятия: словарь-справочник. Тверь, 1991. 90 с.

УДК 004.934.1'1

#### Технические науки

*В текущей статье исследуются такая особенность слухового восприятия человека как эффект маскировки и влияние этого эффекта на проблему распознавания речевых сигналов. Для проведения данного исследования автором статьи разработана аппаратно-программная система, которая имитирует эффект маскировки, производит работу со словарем и осуществляет распознавание слов. В заключение представлены результаты, которые удалось получить на текущий момент.*

*Ключевые слова и фразы:* эффект маскировки; спектрограмма; словарь; динамический алгоритм; распознавание; быстрое преобразование Фурье.

**Чуркин Вадим Сергеевич**

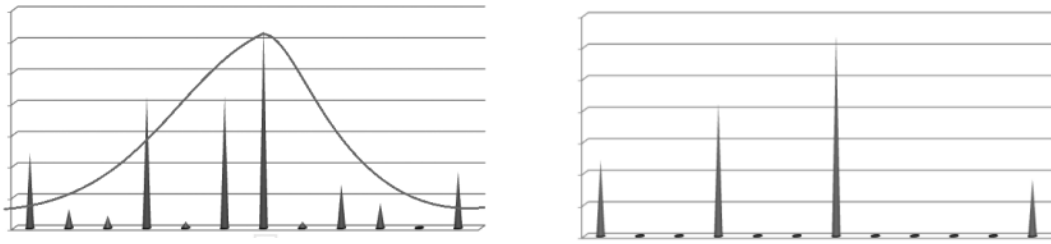
Новосибирский государственный университет

ChurkinVS@mail.ru

#### ИСПОЛЬЗОВАНИЕ ЭФФЕКТА МАСКИРОВКИ ДЛЯ РАСПОЗНАВАНИЯ РЕЧЕВЫХ СИГНАЛОВ<sup>©</sup>

Эффект маскировки – это характерная черта человеческого восприятия звуков, согласно которой для распознавания речевых сигналов используется не полное спектральное описание сигнала, а значения частот формантных максимумов. Данные, полученные в целом ряде экспериментов, проведенных различными исследователями [4; 7], подтверждают тот факт, что для принятия фонемного решения человек использует не весь спектр, а лишь его локальные признаки (особенности). В наши дни этот эффект используется в стандарте MPEG для сжатия информации: при неощутимом снижении качества размер звукового файла снижается приблизительно в десять раз [6].

В психоакустике известны два типа маскировок: в частотной и временной области. При маскировке в частотной области – от каждой гармоники в спектре в сторону высоких частот (прямая маскировка) и низких (обратная маскировка) – строятся кривые маскировки – затухающие и пропорциональные амплитуде маскирующей гармоники. Если амплитуда какой-либо гармоники в спектре оказывается ниже кривой маскировки, то тестируемая гармоника маскируется, т.е. полностью удаляется из спектрального описания. Аналогичны рассуждения и во временной области: маскерами становятся снова все гармоники в спектре, только кривые прямой и обратной маскировки строятся уже во временной области на постоянной частоте. Таким образом, мы получаем форматные максимумы спектрального описания сигнала. Наглядно влияние маскировок (в частотной области) можно оценить на Рисунке 1.



**Рис. 1.** Исходный спектр с кривой маскировки (слева), спектр после маскировки (справа)

Кривые маскировок имеют сложную форму, которая нелинейно изменяется от амплитуды маскера [5], это означает, что математически их можно описать только с определенной долей неоднозначности.

На данный момент точные параметры и форма кривых маскировок касательно проблемы распознавания речи не известны, поэтому нахождение их представляет определенный интерес.

В рассматриваемой работе в качестве аппроксимирующего выражения кривой одновременной маскировки (частотная область) была использована следующая формула [1]:

$$J_j = J_m \cdot \exp\left(-1,44 * Q^2 \left(\frac{\omega_m}{\omega_j} - 1\right)^2\right), \text{ где } J_m - \text{интенсивность маскера, } \omega_m - \text{частота маскера, } J_j - \text{порого-}$$

вая интенсивность тона,  $\omega_j$  – частота тона,  $Q = 1 \div 3$ .

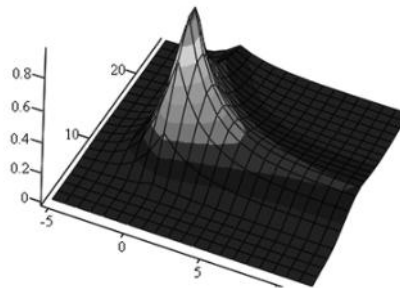
Кривая прямой маскировки аппроксимируется [3] выражением вида:

$$J_j = \beta \cdot J_m \left(1 - \alpha \cdot \left(1 - \exp\left(-\frac{k \cdot \Delta t}{\tau}\right)\right)\right), \text{ где } k = j - m, \alpha = 0,7 \div 0,8, \tau = 50 \text{ мсек}, \beta = 1.$$

А для кривой обратной маскировки было принято следующее выражение:

$$J_j = \beta \cdot J_m \cdot \exp\left(-\frac{k \cdot \Delta t}{\tau}\right), \text{ где } k = j - m, \tau = 25 \text{ мсек}, \beta = 1.$$

Совместное влияние этих кривых представлено на Рисунке 2.



**Рис. 2.** Совместное влияние маскировок

Как уже отмечалась ранее, для исследования была разработана экспериментальная система. Изначально звуковой сигнал после микрофона усиливается и поступает на микросхему – кодек. Кодек осуществляет преобразование аналогового сигнала в цифровой посредством внутреннего аналого-цифрового преобразователя, производит фильтрацию, усиление и отправляет оцифрованный сигнал на микроконтроллер цифровой обработки сигналов. В микроконтроллере каждые 16 мсек мы получаем дискретизированный сигнал из 256 отчетов, к нему применяется окно Хемминга и производится быстрое преобразование Фурье. В итоге получаем спектр из 128 элементов с граничной частотой 8 кГц, такой частотный диапазон достаточен для распознавания речевых сигналов. Полученный спектр поступает на персональный компьютер в разработанную программу, где и производится маскирование.

Для зрительной оценки влияния маскировок в программе предусмотрено окно визуализации, в котором отображается спектрограмма в первоначальном состоянии и после применения к ней маскировок.

Чтобы осуществлять распознавание, необходимо иметь словарь для хранения данных об эталонных словах. Для этой цели применяется база данных Access. Реализация словаря в таком виде выгодна по ряду причин: удобство хранения и транспортировки словаря как файла, удобство модификации и просмотра записей посредством MS Access, высокая скорость работы со словарем, возможность сохранять в таблицу поля с дополнительной информацией, которая поможет в будущем сократить запросы поиска слова в очень большом словаре.

На данный момент таблица словаря состоит из следующих граф: индекс, слово (в текстовом виде), двумерный сериализованный массив чисел с плавающей запятой двойной точности – непосредственно спектрограмма слова.

Сам процесс распознавания состоит из следующих этапов:

1. Текущая спектрограмма разбивается по частотным полосам, рекомендованным специалистами по вокодерной технике, они имеют следующие значения (в герцах): 150-400, 400-640, 640-1040, 1040-1520, 1520-2200, 2200-3040, 3040-4200, 4200-7000, 7000-10000.

2. Затем находится максимальный интеграл под кривой маскера для каждой частотной полосы, значение которого приписывают текущей частотной полосе.

3. Из словаря берем спектрограмму слова, которую аналогично разбиваем на частотные полосы. Если задан режим с маскировками, то к спектрограмме слова применяем маскировку.

4. Опционально проводим дополнительные механизмы фильтрации по длине слов. Не прошедшие эти механизмы слова в общем анализе не участвуют.

5. Находим меру сходства между сегментами двух спектрограмм. При расчёте меры сходства каждый сегмент одного слова сравнивается с каждым сегментом другого слова. После этого этапа получаем массив размером  $m$  на  $n$ , где  $m$  – количество сегментов в первом слове,  $n$  – количество сегментов во втором слове. Мера сходства рассчитывается следующим образом:

$$d_{mn} = \frac{\alpha^2}{\alpha^2 + \rho_{mn}^2}$$
 где  $d_{mn}$  – мера близости между двумя сегментами,  $\alpha$  – экспериментально выбираемая константа,  $\rho_{mn}$  – декартово расстояние между сегментами  $m$  и  $n$ .

6. Затем применяем динамический алгоритм, который позволяет сравнивать спектрограммы слов с неоднородной структурой [2].

7. Проводя эти операции с каждым словом в словаре, находим такое, для которого мера близости окажется максимальной.

8. После того как найдется слово с максимальной мерой близости, отображается диалоговое окно, в котором выводятся величина меры близости и выбранное слово.

В текущей реализации система может эффективно распознавать словарь, состоящий из порядка десяти непохожих друг на друга слов с зависимостью от диктора. Максимальные возможности алгоритмов, к сожалению, оценить не удалось, так как они ограничены самым слабым звеном в системе, которым является интерфейс «микроконтроллер – персональный компьютер». Из-за невозможности передать большие объемы данных спектрограмма получилась низкого разрешения, что не позволяет полноценно проанализировать некоторые особенности слов.

Хотя система не является на данный момент идеальной, точно можно заключить то, что использование эффекта маскировки не понижает качества анализа со значительным выигрышем по времени на большом словаре. Стоит отметить, что теоретически возможно даже повышение качества анализа за счёт удаления паразитных составляющих сигнала.

#### Список литературы

1. **Бондаренко В. П.** Разработка и исследование электронной модели периферии слуховой системы человека: автореф. дисс. ... к.т.н. Томск, 1972. 20 с.
2. **Загоруйко Н. Г.** Когнитивный анализ данных. Новосибирск: Гео, 2013. 186 с.
3. **Мушников В. Н., Чистович Л. А.** Проявление адаптации при восприятии речевого сигнала // Вопросы теории и методов исследования восприятия речевых сигналов. 1971. № 2. С. 28-35.
4. **Щупляков В. С.** О тональной высоте звуков (S) и (S') // Механизмы речеобразования и восприятия сложных звуков. Л., 1966. С. 87-95.
5. <http://rus.625-net.ru/archive/z0200/3.htm>
6. <http://rus.625-net.ru/audioproducer/2007/03/techno4.htm>
7. **Lindqvist-Gauffin J., Pauli S.** The Role of Relative Spectrum Levels in Vowel Perception // Speech Transmission Laboratory Quarterly Progress and Status Reports. 1968. Vol. 9. № 4. P. 12-15.

УДК 811.161.1'36

#### Филологические науки

*В статье рассматриваются структурно-семантические отношения дериватов первой ступени словообразовательного гнезда с вершиной «великий», указываются способы их словообразования. Отмечаются морфонологические явления, сопровождающие образование данных дериватов. Описываются моносемантические и полисемантические дериваты первой ступени словообразовательного гнезда.*

*Ключевые слова и фразы:* словообразовательное гнездо; вершина словообразовательного гнезда; дериват; способы словообразования; морфонологические явления.

**Шепырева Ольга Ивановна**

Орловский государственный университет  
vmusatov@inbox.ru

#### СТРУКТУРНО-СЕМАНТИЧЕСКИЕ ОТНОШЕНИЯ В СЛОВООБРАЗОВАТЕЛЬНОМ ГНЕЗДЕ С ВЕРШИНОЙ «ВЕЛИКИЙ»<sup>©</sup>

Словообразовательное гнездо – это упорядоченная отношениями производности совокупность всех однокоренных слов, самая сложная комплексная единица словообразовательной системы, в которую входят все другие комплексные единицы [1, с. 228], поэтому изучение структуры и семантики словообразовательных гнезд является актуальной задачей современного словообразования.