

## К вопросу о векторизации графических спектрограмм

© А.М. Бонч-Бруевич, С.Б. Козлачков

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

*Для повышения точности оценки параметров математических моделей речевого сигнала предлагается использовать методы контурного анализа изображений применительно к графическим спектрограммам. Использование векторизации обеспечивает возможность выделения индивидуальных признаков речи на основе статистического анализа характеристик контуров на спектрограмме. Каждый из выделенных контуров является незамкнутой линией, содержащей информацию об изменении мгновенной частоты и амплитуды компонента речевого сигнала (речевого вокализма). При этом учитываются следующие особенности речевого сигнала: средняя скорость изменения частоты основного тона, минимальное расстояние между отдельными гармониками, взаимная корреляция уровня сигнала для соседних следов фонообъектов. В настоящей работе рассмотрен подход к формированию массива данных, описывающего графическую спектрограмму речевого сигнала в векторном формате. Показано, что представление графической спектрограммы как массива отдельных контуров дает новые возможности в области анализа, синтеза, обработки и классификации речевых сигналов.*

**Ключевые слова:** спектрограмма, векторизация, контурный анализ, шумоочистка, обработка сигналов.

При оценке защищенности речевой информации, обработке акустических сигналов или идентификации диктора по голосу необходимо выделить информационные признаки в речевом сигнале. Как правило, для выделения используются следующие признаки: частотные (спектр, частота основного тона, спектральная плотность мощности); временные (огibaющая, частота переходов через ноль); частотно-временные (графическая спектрограмма в виде полутонного изображения). Также применяются различного рода математические модели сигналов и помех. Использование математических моделей в ряде случаев оказывается эффективнее, чем непосредственный анализ отсчетов сигнала.

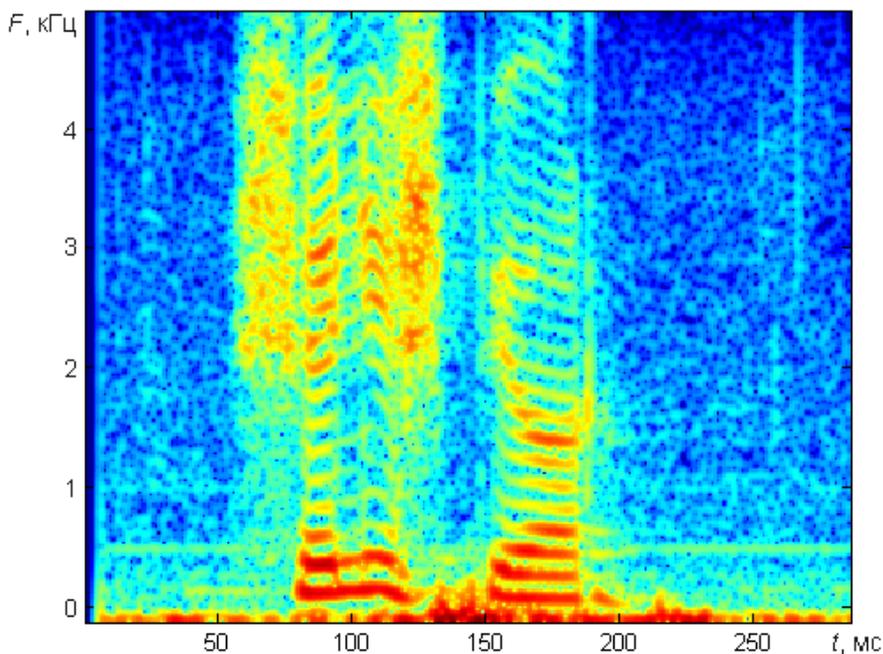
Как показано в [1], речевой сигнал обладает сложной полимодуляционной структурой, что затрудняет создание адекватной модели. Существующие математические модели речевого сигнала основаны на описании речевого тракта человека дифференциальными уравнениями, на совокупности мгновенных, экстремальных или усредненных характеристик речи, полученных различными методами анализа, на линейном или нелинейном предсказании, а также на полигармонической ап-

проксимации сигнала. Известна модель [2], описывающая речевой сигнал в виде совокупности гармонических составляющих:

$$y(t) = \sum_{i=1}^N A_i(t) \cos(2\pi F_i(t - \tau_i)),$$

где  $A_i(t)$  — изменение амплитуды гармоники во времени;  $F_i(t - \tau_i)$  — изменение частоты гармоники с учетом фазового сдвига гармоники во времени.

Известные математические модели описывают речевой сигнал с высокой точностью, однако на практике применение большинства таких моделей ограничивается сложностью определения их параметров. Определение параметров осложняется из-за действия помех, искажающих структуру сигнала. Чтобы повысить точность оценки параметров математических моделей речевого сигнала, можно использовать методы обработки изображений применительно к графической спектрограмме речевого сигнала.



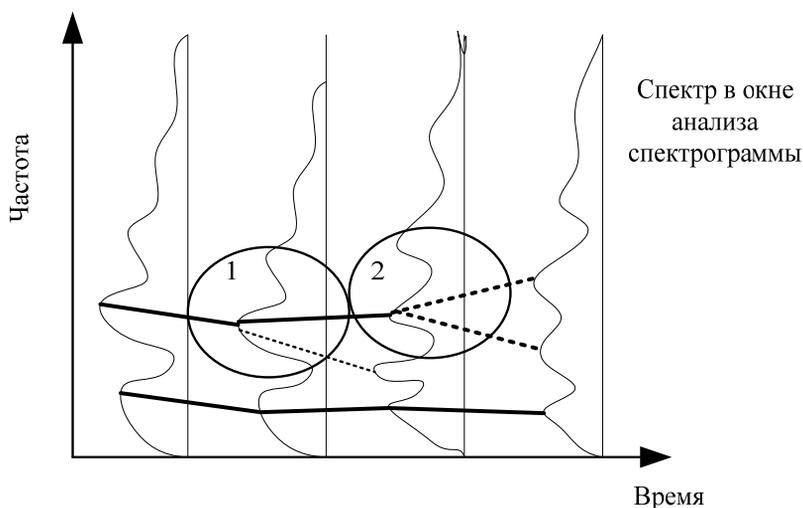
**Рис. 1.** Графическая спектрограмма речевого сигнала

Достаточно эффективным инструментом выделения формантной структуры речевого сигнала по его графической спектрограмме является контурный анализ [3]. Из графической спектрограммы речевого сигнала, показанной на рис. 1, выделяются контуры, соответствующие

щие трекам локальных максимумов изменения спектра в частотно-временной области. Каждый из выделенных контуров является незамкнутой линией, содержащей информацию об изменении мгновенной частоты и амплитуды компонента речевого сигнала (речевого вокализма). При этом учитываются следующие особенности речевого сигнала: средняя скорость изменения частоты основного тона, минимальное расстояние между отдельными гармониками, взаимная корреляция уровня сигнала для соседних следов фонообъектов. Это позволяет повысить точность выделения контуров по спектрограмме даже при высоком уровне шумов (отношение сигнал/шум  $-5$  дБ в полосе частот  $0,3 \dots 3,4$  кГц).

Информативными параметрами, описывающими речевой сигнал, являются величина изменения амплитуды и частоты в пределах одного контура, средняя длительность контура, максимальная скорость изменения амплитуды и частоты, коэффициент корреляции между соседними контурами и др.

Выделение контуров представляет собой задачу векторизации изображения. Некоторые особенности решения этой задачи при обработке спектрограмм показаны на рис. 2.



**Рис. 2.** Особенности формирования контуров по графической спектрограмме

Жирными линиями на рис. 2 отмечены выделенные контуры, описывающие изменение частоты форманты по времени и амплитуде. Кружками обведены характерные ситуации, влияющие на формирование контурного массива. В ситуации 1 на следующем спектральном срезе присутствуют два локальных максимума. В этом случае

для продолжения контура выбирается тот, частота которого меньше отличается от частоты локального максимума спектра в текущем спектральном срезе. В ситуации 2 контур завершается, так как характерного максимума не удастся выделить. Кроме того, при выделении контуров на спектрограмме учитывается, что длительность вокализованных участков речевого сигнала, как правило, не превышает 200 мс. Чем длиннее контур, тем лучше должны быть условия для его продолжения.

На рис. 3 показаны выделенные цветом и интенсивностью контуры, относящиеся к речевому сигналу по характерным признакам — длительности, интенсивности, девиации частоты и пик-фактору [4].

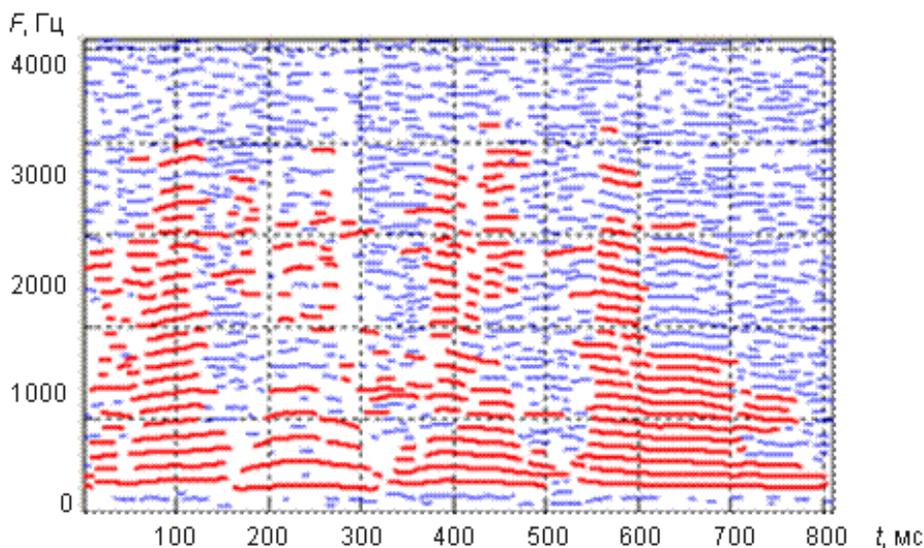


Рис. 3. Контурный массив, описывающий спектрограмму

Поскольку каждый контур обладает уникальными характеристиками, его можно представить в виде независимого сигнала. Для каждого такого сигнала могут быть определены правила оценки принадлежности к речевому сигналу или к помехе, причем подобным образом можно анализировать не только помеху типа «белый шум», но и речеподобные помехи или помехи с характерной структурой (шум транспорта, толпы и т. п.).

При обработке в программном обеспечении информация о совокупности контуров хранится в массиве записей M1, в котором для каждого контура определены векторы изменения амплитуды и частоты, время начала и длительность (последняя может быть получена и из размерности массива амплитуд и частоты дискретизации). Этот массив соотносится с другим массивом, M2, описывающим, какие

контуры из массива M1 существуют в каждый момент времени анализа. Структуры массивов, использованных для хранения контуров, описывающих спектрограмму, приведены в табл. 1 и 2.

Таблица 1

**Структура массива записей M1**

№ п/п	Массив амплитуд	Массив частот	Начало контура
1	$A_1$	$F_1$	$T_1$
2	$A_2$	$F_2$	$T_2$
...	...	...	...
$N$	$A_N$	$F_N$	$T_N$

Таблица 2

**Структура массива записей M2**

Отсчеты времени	Индексы массива M1, соответствующие контурам, определенным в заданный момент
$t_1$	$L_1$
$t_2$	$L_2$
...	...
$t_N$	$L_N$

Благодаря представлению графической спектрограммы в виде массива отдельных контуров появляются новые возможности в области анализа, синтеза, обработки и классификации речевых сигналов, такие как:

- классификация сигналов в графической спектрограмме по длительности, амплитуде, частоте девиации;
- определение сигнала по заданным критериям;
- работа с отдельными контурами (например, воспроизведение только таких контуров, среднее значение которых выше заданного порога).

Кроме того, обработка каждого контура по отдельности повышает качество шумоочистки за счет аппроксимации контура в пространстве частот и амплитуд.

Таким образом, применяя методы контурного анализа изображений графических спектрограмм, можно не только получить формализованное и структурированное математическое описание речевого сигнала, но и реализовать алгоритмы реконструкции структуры сигнала после выполнения шумоочистки.

## ЛИТЕРАТУРА

- [1] Голубинский А.Н., Булгаков О.М. Система аутентификации личности по голосу на основе модуляционных математических моделей речевого сигнала. *Информация и безопасность*, 2011, вып. 1, с. 141–144.
- [2] Дворянкин С.В. *Цифровая шумоочистка аудиоинформации*. Москва, ИП РадиоСофт, 2011, с. 102.
- [3] Фурман Я.А., ред. *Введение в контурный анализ; приложения к обработке изображений и сигналов*. Москва, ФИЗМАТЛИТ, 2003, 592 с.
- [4] Горелик А.Л. Скрипкин В.А. *Методы распознавания*. Москва, Высшая школа, 1977, 222 с.

Статья поступила в редакцию 28.06.2013

Ссылку на эту статью просим оформлять следующим образом:

Бонч-Бруевич А.М., Козлачков С.Б. К вопросу о векторизации графических спектрограмм. *Инженерный журнал: наука и инновации*, 2013, вып. 11. URL: <http://engjournal.ru/catalog/it/hidden/1015.html>

**Бонч-Бруевич Андрей Михайлович** родился в 1982 г., окончил Московский технический университет связи и информатики в 2004 г. Канд. техн. наук, доцент кафедры «Защита информации» МГТУ им. Н.Э. Баумана. Автор 17 научных работ. Области научных интересов: цифровая обработка сигналов, компьютерное моделирование. e-mail: 123andryb@mail.ru

**Козлачков Сергей Борисович** родился в 1955 г., окончил Московский энергетический институт в 1977 г. Канд. техн. наук, старший преподаватель кафедры «Защита информации» МГТУ им. Н.Э. Баумана. Автор 16 научных работ. Области научных интересов: речевые технологии, акустика, защита информации. e-mail: ksb.perovo@mail.ru