



Новые возможности анализа сигнала для определения положения формант в АПК «САПФИР»

Лобанова М.А.

В настоящее время АПК «САПФИР» активно используется для проведения фоноскопических экспертиз в лабораториях МВД. В то же время постоянно ведутся работы по его модернизации, включая не только дальнейшую автоматизацию работы эксперта, но и разработку новых методов проведения различных видов анализа сигнала. В данной статье будет дан краткий обзор некоторых новых возможностей программы, предлагаемых эксперту для определения положения формант исследуемого речевого сигнала.

Введение

В настоящее время при тесном сотрудничестве с ЭКЦ МВД России нами продолжается дальнейшее развитие созданного в 2007 году аппаратно-программного комплекса (АПК) для автоматизации проведения фоноскопической экспертизы «САПФИР» ([1]).

Сделаем краткий экскурс в историю создания данного комплекса.

К разрабатываемому АПК «САПФИР» предъявлялись следующие требования:

- возможность выполнения основных этапов типового технологического процесса производства экспертизы;
- возможность использования результатов уже выполненных работ на каждом последующем этапе выполнения экспертизы;
- автоматизация работы эксперта с целью сокращения временных трудозатрат и облегчения выполнения различных видов работ.

При разработке АПК «САПФИР» учитывалась принятая в настоящее время в МВД России методика проведения фоноскопической экспертизы, а именно методика «Диалект». Соответственно данной методике была разработана модульная структура программного обеспечения АПК «САПФИР», позволяющая не только удобным образом следовать методике «Диалект» (в том

числе обеспечить её освоение начинающими экспертами), но и разделить работу по проведению экспертизы между различными экспертами. Также в результате консультаций и плодотворного сотрудничества с экспертами ЭКЦ были определены наиболее важные направления автоматизации процесса производства экспертизы, что и было реализовано в конечном программном продукте.

В настоящее время АПК «САПФИР» активно используется для проведения фоноскопических экспертиз в лабораториях МВД. В то же время нами постоянно ведутся работы по его модернизации, включая не только дальнейшую автоматизацию работы эксперта, но и разработку новых методов проведения различных видов анализа сигнала. В статье дан краткий обзор некоторых новых возможностей программы, предлагаемых эксперту для определения положения формант исследуемого речевого сигнала.

Задача оценивания положения формант

Следует сразу отметить, что описываемые ниже способы определения положения формант не являются автоматическими и требуют активного участия эксперта.

Задача оценивания положения формант встаёт перед экспертом при проведении сравнительного лингвистического анализа образцов речи, а также при проведении сравнительного акустического микроанализа. При проведении микроанализа, заключающегося в поиске в образцах речи и в спорных фонограммах сопоставимых триад звуков (согласный–гласный–согласный) и их последующем параметрическом описании, требуется особенно точная оценка частотного положения формант.

Для оценивания значений формант эксперт, пользуясь возможностями имеющегося в его распоряжении ПО, обычно проводит вычисление спектрограммы с последующим анализом её изображения. Среди основных задач, которые при этом решает эксперт, можно назвать следующие:

- выбор оптимального размера спектрального окна для вычисления спектрограммы для исследуемого речевого фрагмента (с учётом значения основного тона голоса);
- выделение формантных траекторий на изображении вычисленной спектрограммы;
- оценка значений положений формант для выбранного в речевом сигнале момента времени.

Возможность и сложность выделения формантных траекторий по изображению спектрограммы, конечно, во многом зависят от характеристик исследуемого речевого сигнала: от качества сигнала (уровня шума и его амплитудно-частотных характеристик), от речевых навыков и манеры диктора, от экстралингвистических факторов, оказавших влияние на речевой процесс. Вместе с тем задачу выделения формантных траекторий можно облегчить программными средствами, среди которых, конечно, следует назвать возможность выбора для каждого конкретного сигнала оптимального диапазона цветопередачи спектрограммы, а также возможность вводить подъём амплитуды высоко-частотных составляющих сигнала при их отображении.

В АПК «САПФИР», кроме перечисленных выше способов улучшить изображение спектрограммы и проявить на ней формантные полосы, реализованы дополнительные методы, которые условно можно разбить на три группы:

- методы, позволяющие одновременно (синхронно по частоте и по времени) проводить изменения в окнах, представляющих различные данные (результаты разных видов анализа);



- методы, основанные на сравнительном анализе спектрограмм идентичных речевых фрагментов;
- методы, основанные на различных способах отображения спектрограмм: например, дифференциальные спектрограммы или отображение спектрограммы с помощью «спектральных профилей» (см. рис.1).

Опишем кратко каждую группу методов.

Методы, позволяющие одновременно (синхронно по частоте и по времени) проводить измерения в окнах, представляющих различные данные (результаты разных видов анализа)

Данная группа методов базируется на возможностях пользовательского интерфейса АПК «САПФИР», позволяющего измерять синхронно в разных окнах значения частоты и времени.

На рис. 1 представлены виды окон осциллограммы, спектрограммы и мгновенного спектра сигнала, в которых отображены связанные между собой данные.

Для вычисления широкополосного спектра выбраны следующие параметры: размер спектрального кадра — 128 отсчётов сигнала, смещение между кадрами — 2 отсчёта.

В окне спектра отображены два мгновенных спектра, вычисленные для временных участков сигнала, имеющих общее начало, но разную длину (512 и 128 отсчётов сигнала). Начало этих временных участков соответствует положению поставленного в окне спектрограммы (осциллограммы) маркера.

Совершаемые экспертом, например, по окну спектрограммы перемещения указателя мышки отображаются также в окне осциллограммы и в окне спектра. В окне осциллограммы выделяются границы временного окна, для которого был вычислен спектральный срез, соответствующий указателю мышки в окне спектрограммы. В окне спектра отображается маркер, соответствующий частоте указателя мышки окна спектрограммы.

Возможность синхронного измерения временных и частотных координат данных, представленных в различных формах, позволяет эксперту принять решение о наличии форманты в той или иной частотной области и оценить её значение.

На рисунке показано проведение синхронного по частоте и времени анализа данных, представленных в разных окнах.

Методы, основанные на сравнительном анализе спектрограмм идентичных речевых фрагментов

Принятие решения по спектрограмме сигнала о том, где проходят формантные траектории, осложняется иногда не только плохим качеством фонограммы, но и ненормативными (в плане положения формант) особенностями речи.

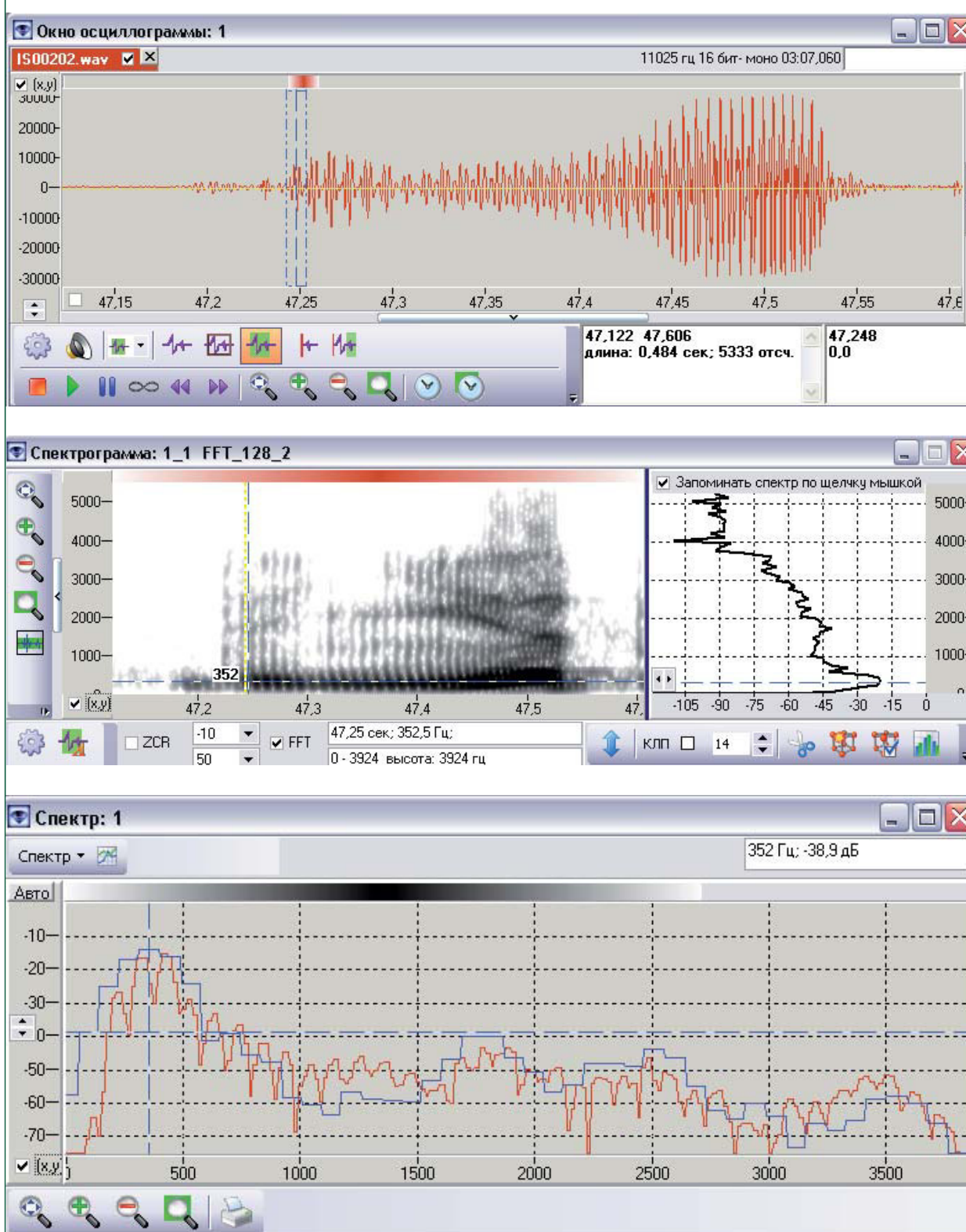


Рис.1. Окно осциллограммы с участком звукового сигнала; окно спектрограммы для данного сигнала; окно спектра с изображениями двух мгновенных спектров (широкополосного и узкополосного), вычисленных для временных кадров, имеющих общее начало

Мы имеем в виду наличие на спектрограмме сигнала выраженных затемнённых полос в тех частотных диапазонах, в которых их быть не должно согласно известным акустическим характеристикам (для присутствующих в речи фонетических единиц).

В этом случае провести формантный анализ может помочь сравнительный анализ спектрограмм идентичных речевых фрагментов (триад звуков) известного и спорного дикторов.

Одним из направлений развития АПК «САПФИР» явилась разработка возможностей проведения такого рода сравнений. На рис. 2 представлены наложенные друг на друга фрагменты спектрограмм идентичных речевых фрагментов, произнесённых разными дикторами. Пользовательский интерфейс позволяет эксперту сдвигать накладываемые фрагменты относительно друг друга по времени и по частоте, а также настраивать изображения, изменяя параметры цветопередачи для каждого из фрагментов отдельно.

На рис. 3 представлены наложенные друг на друга фрагменты кепстрограмм. Наложение кепстрограмм может помочь эксперту принять решение о возможности сравнения речевых сигналов (например, для определения идентичности эмоционального состояния сравниваемых дикторов, речь которых представлена на фонограммах).

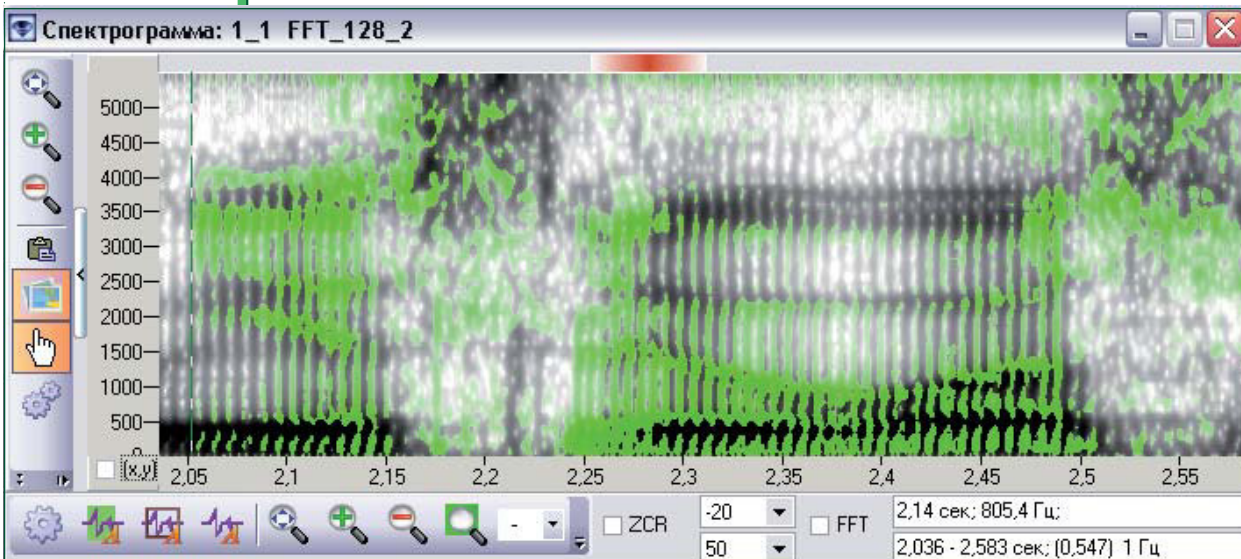


Рис. 2. Наложение друг на друга спектрограмм идентичных речевых фрагментов в АПК «САПФИР»

Методы, основанные на различных способах отображения спектрограмм (например, дифференциальные спектрограммы или отображение спектрограммы с помощью «спектральных профилей»)

Большую помощь эксперту при анализе сигнала, как показала практика, могут оказать дифференциальные спектрограммы, включённые в АПК «САПФИР» как один из возможных видов представления спектрограммы сигнала (см. [2]). Дифференциальные спектрограммы оказываются очень

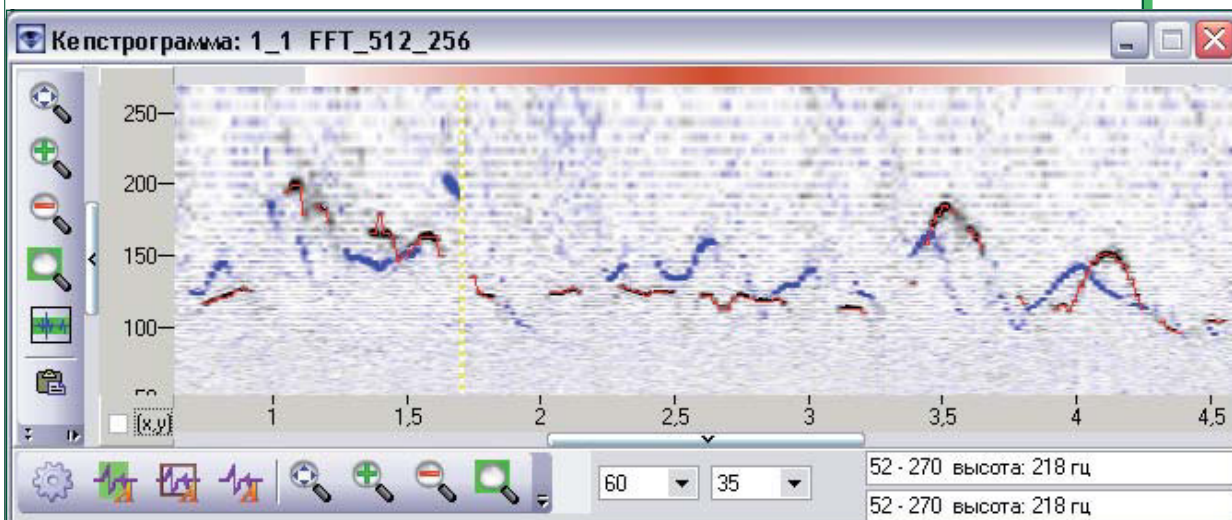


Рис. 3. Наложение друг на друга кепстрограмм одинаковых фраз, произнесённых разными дикторами

полезными при определении границ между звуковыми единицами потока речи, а также для выделения на спектрограмме формантных траекторий.

Одним из направлений развития АПК «САПФИР» является реализация возможности отображения данных, представляемых на спектрограмме (вклада частотных составляющих в общий спектр сигнала, а также изменения этого вклада во времени), в виде «спектральных профилей».

Идея состоит в том, чтобы представлять изменение значения каждой частотной составляющей в виде отдельной линии на общем для всех частотных составляющих графике. На рис. 4 представлен такой график для нескольких соседних частотных составляющих.

Приведённые на рис. 4–6 изображения показывают возможности использования «спектральных профилей»:

- для определения положения формант (путём изучения графика трёх-четырёх соседних частотных компонент, соответствующих максимуму в спектре выбранного момента времени);
- для определения среднего и мгновенного значения основного тона голоса (по присутствующей в спектральных профилях амплитудной модуляции с частотой основного тона);
- для определения границ между звуковыми единицами речевого потока.

На графике «спектральных профилей» хорошо видно изменение энергии частотных характеристик, соответствующее происходящим в голосовом тракте говорящего перестройкам. Также на «спектральных профилях» хорошо видна периодическая структура, соответствующая работе голосовых связок.

На приведённых «спектральных профилях» в середине речи виден всплеск энергии частотных характеристик, соответствующий произнесению звука «р». На «спектральных профилях» хорошо видна разница в скорости нарастания (более пологая) и скорости уменьшения (более крутая) спектральной энергии частотных характеристик звука «р».

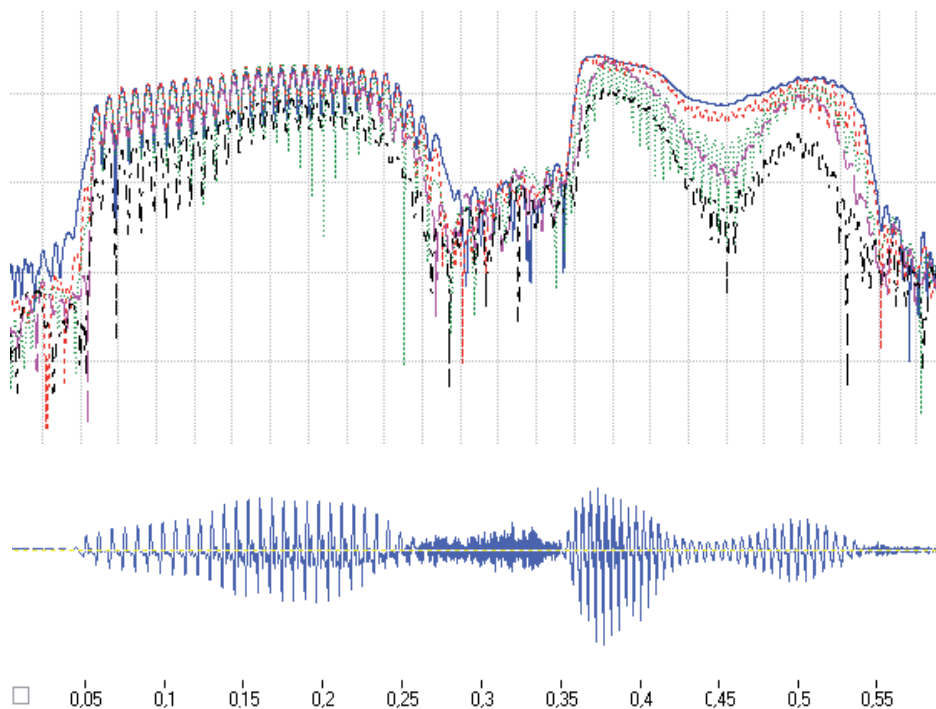


Рис. 4. График «спектральных профилей» соседних частотных компонент (верхнее окно): по горизонтальной оси отложено время, по вертикальной — значение спектральной энергии; в нижнем окне — осциллограмма того же сигнала

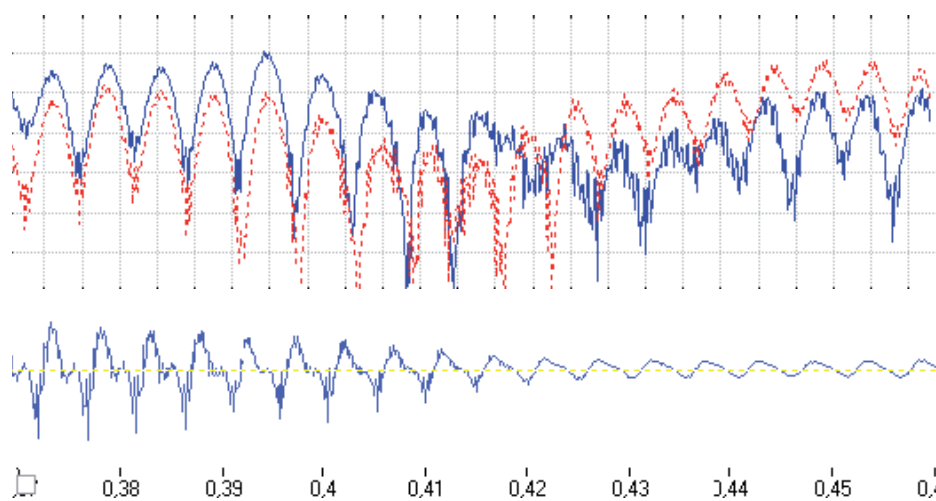


Рис. 5. График «спектральных профилей» двух частотных компонент для небольшого временного окна (верхнее окно); в нижнем окне — осциллограмма того же сигнала

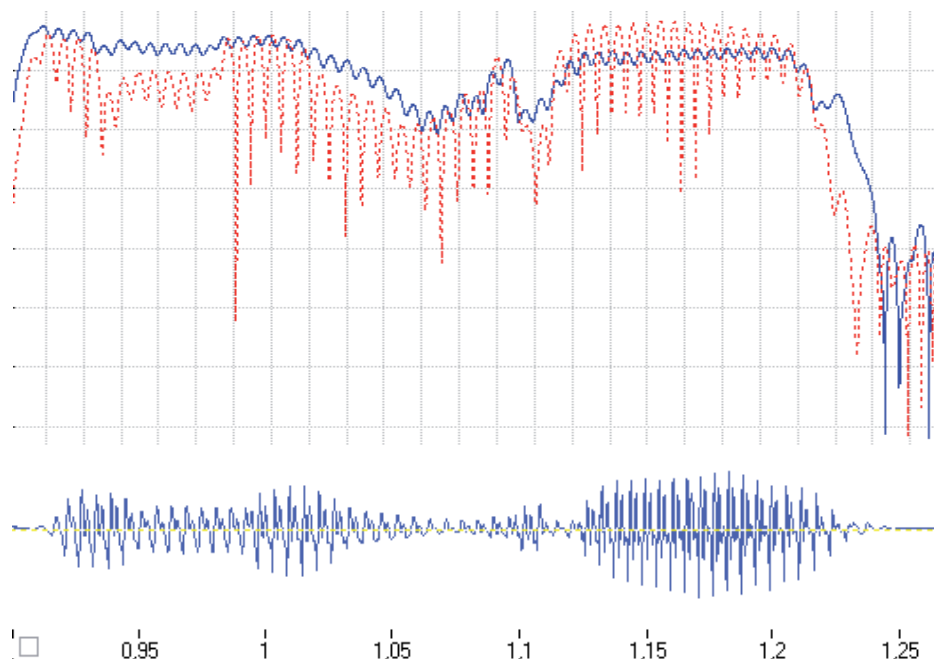


Рис.6. Осциллограмма сигнала (внизу) и «спектральные профили» двух частотных характеристик спектрограммы (наверху)

Заключение

Применение «спектральных профилей» для исследования отклика спектрального анализатора на речевой сигнал можно сравнить с применением микроскопа. На «спектральных профилях» хорошо видны изменения в речевом сигнале не только в энергетически выраженных частотных областях, но и во всех остальных. Как нам кажется, на «спектральных профилях» могут быть хорошо видны артикуляционные жесты говорящего, т.е. его стремление к перестройке артикуляционного тракта согласно речевой программе.

Возможно, изображение спектрограммы в виде «спектральных профилей» окажется интересным и полезным для проведения различных исследований речевого сигнала.

Литература

1. Лобанова М.А., Назарова Т.В. Универсальный комплекс для автоматизации проведения фоноскопической экспертизы (комплекс «САПФИР»). Материалы 15-й Международной научной конференции «Информатизация и информационная безопасность правоохранительных органов», 23–24 мая 2006 г. 313 с.
2. Лобанова М.А. Решение проблемы поиска идентичных речевых фрагментов в универсальном комплексе для автоматизации проведения фоноскопической экспертизы «САПФИР». Построение дифференциальных спектрограмм. Материалы 17-й Международной научной конференции «Информатизация и информационная безопасность правоохранительных органов», 20–21 мая 2008 г. 412 с.

Лобанова М.А.

ЗАО НПП «ИСТА-СИСТЕМС», Санкт-Петербург,
mal@ista.ru.