

ПО Форвард Т

SLAutoDetectTest



Тестирование и настройка параметров
распознавания звуковых отбивок

*Дата выпуска:
04 июля 2007 г.*

Руководство пользователя



Содержание

ПО AutoDetect	3
1. Общие сведения	3
2. Состав	3
3. Схема использования ПО для обнаружения звуковых отбивок	3
Распознавание звуковой отбивки.....	4
1. Эталонный фрагмент	4
2. Процедура распознавания	4
3. Коэффициент сходства	5
4. Временной ряд коэффициентов сходства.....	5
5. Параметры распознавания	5
Программа SLAutoDetectTest. Общие сведения	7
Главное окно программы SLAutoDetectTest	8
1. Назначение элементов главного окна	8
2. Загрузка и сохранение настроек.....	8
3. Протоколирование	9
4. Настройка параметров.....	9
5. Выполнение тестирования	10
6. Визуальная оценка качества распознавания.....	10
Порядок работы с программой SLAutoDetectTest.....	12
1. Общая схема работы	12
2. Общие рекомендации по настройке параметров	13
3. Настройка Sample rate	13
4. Настройка Analysis digit	13
5. Интервал не подходит.....	14
6. Длительное тестирование.....	14



ПО AutoDetect

1. Общие сведения

ПО AutoDetect – дополнительная опция к ПО Форвард Т. ПО AutoDetect предназначено для автоматизации управления вещанием на основе распознавания звуковых сигналов. Например, может использоваться для автоматической врезки рекламных блоков по звуковым отбивкам, а именно: DTMF-меткам, джинглам, музыкальным заставкам.

Программы, входящие в состав ПО AutoDetect, осуществляют мониторинг входящего звукового сигнала с целью выявления звуковых меток. При обнаружении метки формируется условный сигнал, аналогичный сигналу при нажатии кнопки GPI.

2. Состав

ПО AutoDetect включает в себя: программу конфигурирования отбивок, модули распознавания, специальные программы.

Программа конфигуратор – SLAutoDetectConfig.exe, – служит для настройки GPI-команд.

Модуль распознавания ведет наблюдение за входящим сигналом и, обнаружив заданные звуковые метки, сообщает системе.

Специальная программа – SLAutoDetectTest.exe, – предназначена для настройки параметров распознавания.

3. Схема использования ПО для обнаружения звуковых отбивок

1. Записать отбивку, сохранить в файл (программа FDCapture).
2. Вырезать из нее фрагмент – эталон (программа Forward AV Studio).
3. Настроить GPI-команду: установить соответствие между образцом звуковой отбивки и действием, которое следует выполнять при обнаружении сходного звукового сигнала на входе. На один GPI может быть назначена коллекция образцов, т.е. одно и то же действие может быть выполнено при обнаружении нескольких разных отбивок (программа SLAutoDetectConfig).
4. Настроить параметры распознавания (программа SLAutoDetectTest).
5. Вставить в расписание FDonAir GPI-команды (программа FDonAir).

Распознавание звуковой отбивки

Для обнаружения звуковой отбивки модуль распознавания постоянно наблюдает за входящим сигналом, производя его сравнение с эталонным фрагментом.

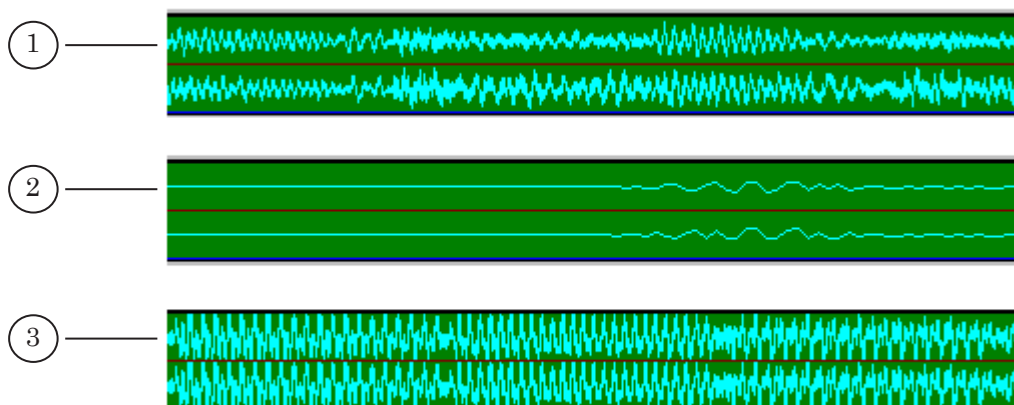
1. Эталонный фрагмент

Эталонный фрагмент (образец) – это фрагмент звуковой отбивки, заранее подготовленный пользователем для распознавания.

В качестве образца рекомендуется выбирать наиболее характерный участок отбивки. Длительность такого фрагмента должна быть не менее 1 с. Громкость звука не должна быть слишком слабой или наоборот сильной.

На рисунке ниже представлены три фрагмента звуковых отбивок:

1. Оптимальный;
2. Не подходит для распознавания, т.к. звук очень слабый;
3. Не подходит для распознавания, т.к. звук очень сильный.



2. Процедура распознавания

Программа выполняет сравнение следующим образом:

1. Выделяет фрагмент входящего звукового сигнала;
2. Подсчитывает коэффициент сходства пары фрагментов – поступившего для распознавания и эталонного;
3. Принимает решение о сходстве входящего сигнала и образца: если коэффициент больше заданного порогового значения, то выбранный фрагмент считается сходным с образцом, если меньше – отличным от него.
4. Фрагмент, похожий на образец, идентифицируется как звуковая отбивка.

3. Коэффициент сходства

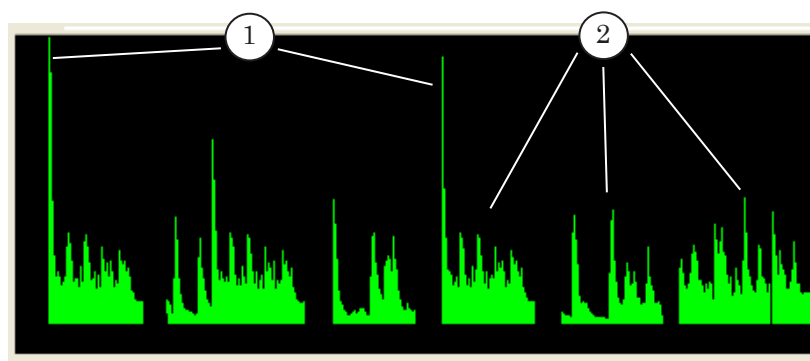
Коэффициент сходства является мерой сходства пары звуковых фрагментов. Программа распознавания использует этот коэффициент для принятия решения о принадлежности входящего сигнала к звуковым отбивкам.

Коэффициент сходства может принимать значения от 0 до 100%. Чем больше сходство фрагментов, тем выше должно быть значение коэффициента.

4. Временной ряд коэффициентов сходства

Так как входной сигнал изменяется с течением времени, то программа постоянно выбирает новые фрагменты для сравнения с эталоном. Таким образом, попарное сравнение выполняется многократно. В результате для каждого момента времени вычисляется свой коэффициент сходства входящего сигнала с образцом.

Результирующий набор коэффициентов сходства, упорядоченных по времени, можно представить в виде графика.



1. В эти моменты времени коэффициент сходства высокий — близко к 100%;
2. В эти моменты времени значение коэффициента сходства низкое — примерно 30% и ниже.

5. Параметры распознавания

Распознавание звуковых отбивок регулируется следующими параметрами:

1. Частота дискретизации, используемая при обработке сигнала;
2. Количество сэмплов эталонного фрагмента, выбираемых для распознавания;
3. Пороговый уровень сходства, необходимый для принятия решения о сходстве сигнала с образцом;
4. Параметры низкочастотного фильтра. НЧ фильтр используется для подавления шумов от «промышленной» частоты 50 Гц.



Настройки этих параметров влияют на качество и скорость распознавания.



Программа SLAutoDetectTest. Общие сведения

Программа SLAutoDetectTest входит в состав программного обеспечения AutoDetect. Утилита предназначена для оценки качества звукового образца (пригодности для распознавания) и подбора оптимальных параметров распознавания звуковых отбивок.

Оптимизировать процедуру означает:

1. Уменьшить затраты ресурсов на узнавание отбивок:
 - а) времени (чем быстрее определяется, тем лучше);
 - б) ресурсов системы.
2. Не потерять при этом точность узнавания, т.е. свести к минимуму количество неузнаваний и ложных узнаваний.

Запуск приложения выполняется путем вызова на исполнение файла: SLAutoDetectTest.exe

С помощью программы можно оценить качество распознавания двумя способами: визуально и на основе статистики, накапливаемой в файле протокола.

Главное окно программы SLAutoDetectTest

1. Назначение элементов главного окна

Главное окно программы содержит управляющие элементы, с помощью которых можно выполнить следующие операции:

- провести экспериментальную регулировку параметров;
- визуально оценить результат изменения настроек;
- выполнить длительное тестирование с протоколированием действий по распознаванию отбивок;
- сохранить оптимальные настройки для дальнейшей работы.

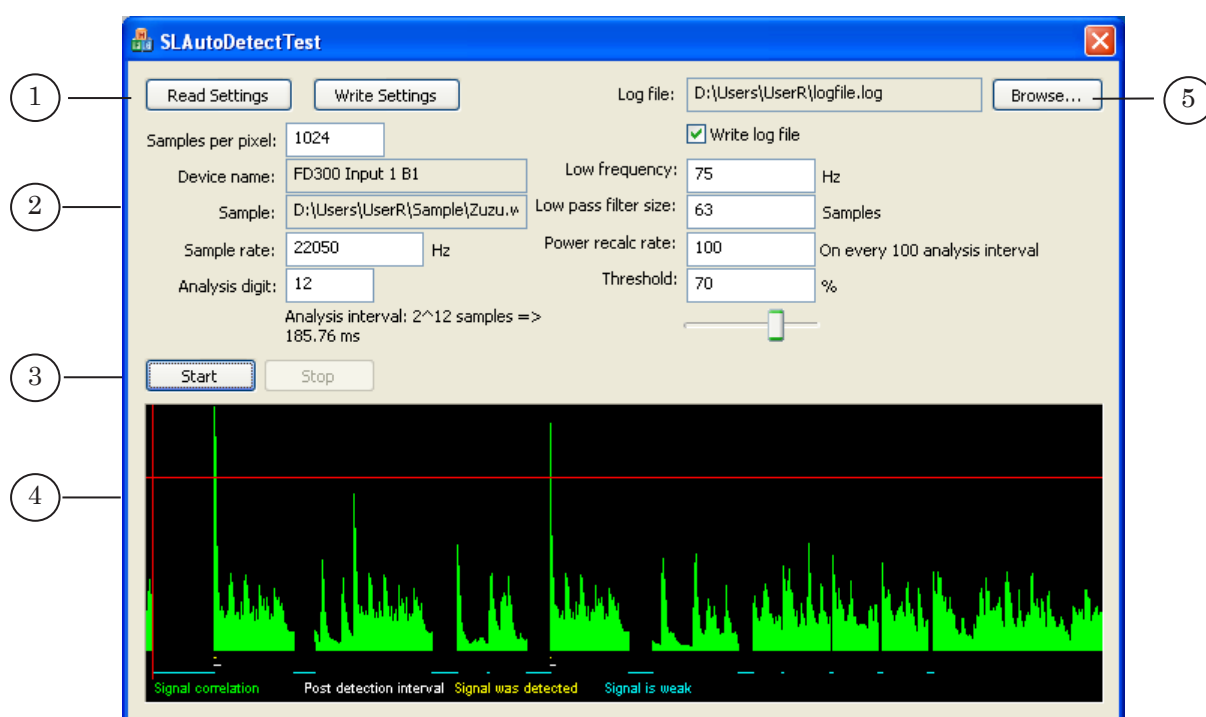


Рисунок 1. Основное окно приложения

1 – кнопки для загрузки и сохранения настроек распознавания; 2 – текстовые поля для отображения и настройки параметров; 3 – кнопки запуска/остановки тестирования; 4 – график коэффициентов сходства сигналов ; 5 – элементы для настройки протоколирования.

2. Загрузка и сохранение настроек

При нажатии кнопки Read Settings текущие настройки параметров распознавания считываются из реестра и отображаются в соответствующих текстовых полях главного окна.

В том числе, в следующих полях выводятся справочные сведения:

- Device name – название входного звукового устройства;
- Sample – полный путь к эталонному фрагменту.

Информацию, размещенную в этих полях, невозможно изме-



нить с помощью приложения SLAutoDetectTest.

✓ **Важно:** Программа работает только с GPI-событием, идущим первым в списке программы SLAutoDetectConfig (т.е. из реестра считывается информация о настройках для первого файла-образца)!

Значения других параметров, отображаемых в главном окне, могут быть отредактированы. Изменение настроек влияет только на экспериментальное распознавание и никак не отражается на действующих значениях параметров.

Для того чтобы изменить текущие настройки модуля распознавания необходимо сохранить новые значения параметров в реестре. Для этого следует нажать кнопку Write Settings.

3. Протоколирование

Протокол – это файл, в котором автоматически фиксируются все моменты срабатывания детектора звуковых отбивок, с указанием времени и значения коэффициента сходства. Протокол можно использовать при длительном тестировании для принятия решения о качестве распознавания.

Текстовое поле Log file служит для указания полного пути к файлу протокола.

Для ведения протокола тестирования необходимо установить флажок Write log file.

4. Настройка параметров

Для подбора значений параметров распознавания служат следующие элементы главного окна:

1. Sample rate – частота дискретизации, используемая при распознавании.

✓ **Важно:** Частота дискретизации звукового устройства должна быть кратна частоте, указанной в поле Sample rate! Например, для частоты звука 48000 Гц можно задать значения 8000, 16000 или 24000 Гц.

2. Analysis digit – значение, которое определяет количество сэмплов эталонного фрагмента, выбираемых для распознавания.

$2^{**} \text{Analysis digit}$ (2 в степени Analysis digit) первых сэмплов каждого звукового образца используются для распознавания.

Зная значения Sample rate и Analysis digit, можно рассчитать длительность фрагмента, используемого для распознавания. Значение отображается в строке Analysis interval.



➡ **Пример:** Например, если Analysis digit=12 и Sample rate=8000 Hz, то для распознавания будет использоваться 4096 сэмплов (2×12), а длительность интервала (Analysis interval) составит 0,512 с ($4096/8000 = 0,512$ с).

✓ **Важно:** Значения параметра Analysis digit следует подбирать так, чтобы длительность анализируемого интервала была не меньше 0.3 с. Кроме того следует учитывать, что длительность не может быть больше 1 с.

3. Threshold – пороговый уровень сходства звуковых фрагментов.

Если значение коэффициента сходства больше заданного порогового значения, то выбранный фрагмент считается похожим на образец, если меньше – отличным от него.

Чем выше пороговый уровень, тем меньше ложных срабатываний при распознавании. Однако, при слишком высоком пороговом уровне часть звуковых отбивок может не распознаться.

Снижение порога означает снижение требования к сходству фрагментов, т.е. метка будет улавливаться даже при наличии сильных шумов, но в этом случае увеличивается вероятность принять за отбивку посторонний сигнал.

Не рекомендуется задавать пороговое значение менее 40-50%.

4. Параметры низкочастотного фильтра (используется для подавления шумов от «промышленной» частоты 50 Гц):
- Low frequency – частота среза;
 - Low pass filter size – размер фильтра.
5. Power recalc rate – регулярность пересчета нормирующих коэффициентов.

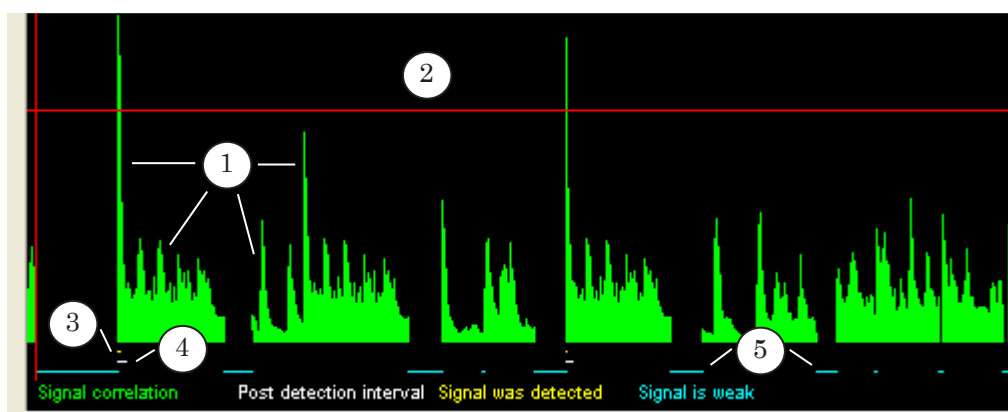
5. Выполнение тестирования

Кнопка Start предназначена для запуска тестового распознавания. Кнопка Stop – для остановки тестирования.

6. Визуальная оценка качества распознавания

Ход процедуры распознавания отображается в главном окне программы в виде графика коэффициентов сходства, упорядоченных по времени. График снабжен дополнительными маркерами.

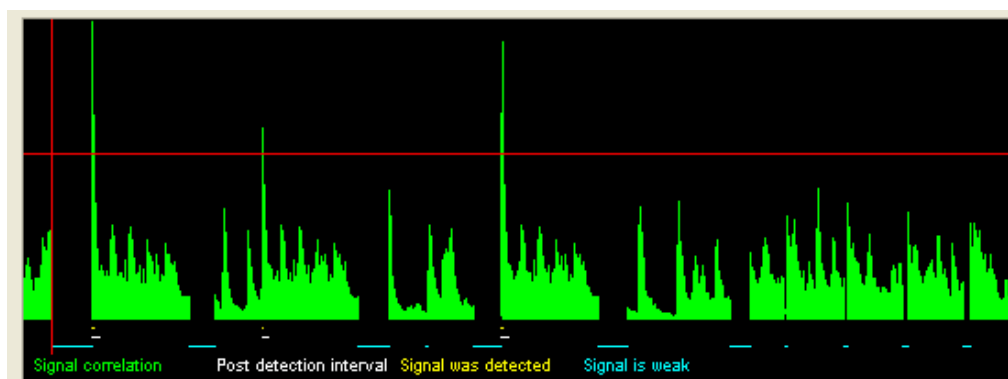
Масштаб оси времени для графика коэффициентов сходства можно установить в текстовом поле Samples per pixel.



1. Signal correlation – значения коэффициента сходства.
2. Красная горизонтальная линия – пороговый уровень узнаваемости (все фрагменты, для которых сходство больше – звуковые отбивки).
3. Метка Signal was detected – сигнал определен как отбивка.
4. Метка Post detection interval – GPI сигнал активен (замкнут).
5. Метка Signal is weak – сигнал слишком слабый, невозможно использовать для определения.

График позволяет оценить качество выявления звуковых отбивок и влияние регулировки параметров на распознавание.

На следующем рисунке видно, что при снижении порогового значения количество фрагментов, опознанных как звуковая отбивка, увеличилось.





Порядок работы с программой SLAutoDetectTest

1. Общая схема работы

1. Запустите утилиту SLAutoDetectTest.exe. Откроется главное окно программы.
2. Загрузите текущие настройки из реестра. Для этого нажмите кнопку Read Settings.
3. Убедитесь, что выбран нужный эталонный фрагмент: проверьте имя файла в поле Sample.
4. Подайте на вход звуковой сигнал.
5. Запустите тестирование – нажмите кнопку Start.
6. На графике будут отображаться результаты тестового распознавания звуковых отбивок.
7. Остановите тестирование – нажмите кнопку Stop.
8. Изучив график, отрегулируйте параметры распознавания и опять запустите процедуру тестирования.

Повторите шаги 5 – 8 столько раз, сколько нужно.

9. Если считаете, что подобраны оптимальные настройки параметров распознавания, то запишите их в реестр. Для этого нажмите кнопку Write Settings.



Важно: Новые настройки распознавания станут действующими только после того, как будут записаны в реестр.

При необходимости записывать результаты тестирования в



файл протокола:

1. Настройте путь к файлу протокола в поле Log file. Для этого используйте кнопку Browse....
2. Установите флажок Write log file.

2. Общие рекомендации по настройке параметров

Оптимизация процедуры распознавания заключается в увеличении точности узнавания звуковых отбивок.

Для этого необходимо решить две задачи:

1. Добиться, чтобы значения коэффициента сходства «хорошо» отражали реальную ситуацию: в момент звучания отбивки значение коэффициента сходства должно быть как можно больше, а при прохождении других звуков – как можно ниже.

Для этого рекомендуется регулировать параметры:

- Sample rate;
 - Analysis digit.
2. Подобрать оптимальный порог срабатывания детектора отбивок (параметр Threshold): при котором будет сведено к минимуму количество неузнаваний и ложных срабатываний.

Оценить точность распознавания можно с помощью графика коэффициентов сходства в главном окне программы.

3. Настройка Sample rate

Sample rate – частота дискретизации, используемая при обработке сигнала.

Чем выше значение этого параметра, тем больше амплитуда колебаний значений коэффициента сходства: более высокие при прохождении звуковых отбивок, и более низкие для других звуковых фрагментов.

Однако, чем меньше значение частоты дискретизации, тем быстрее выполняются вычисления.



Важно: Частота дискретизации звукового устройства должна быть кратна частоте, указанной в поле Sample rate! Например, для частоты звука 48000 Гц можно задать значения 8000, 16000 или 24000 Гц.

4. Настройка Analysis digit

Analysis digit – значение, которое определяет количество сэмплов эталонного фрагмента, выбираемых для распознавания, т.е. длину анализируемого интервала.

Чем больше значение этого параметра, тем точнее распознавание, но и больше времени проходит от начала отбивки до



момента срабатывания GPI.

Чем меньше значение этого параметра, тем быстрее выполняются вычисления, быстрее происходит распознавание, но точность узнавания снижается.



Важно: Значения параметра следует подбирать так, чтобы длительность анализируемого интервала была не меньше 0.3 с. Кроме того следует учитывать, что длительность не может быть больше 1 с.

Длительность анализируемого интервала отображается в строке Analysis interval.

5. Подбор длительности анализируемого интервала

Может возникнуть ситуация, когда интервал, выбранный для распознавания, не подходит для этого – сигнал на данном участке слишком слабый. В этом случае программа выдаст предупреждение, и пользователю следует:

- выбрать другой эталонный фрагмент;
- или изменить длительность анализируемого интервала, регулируя настройки параметров:
 - Sample rate;
 - Analysis digit.

6. Длительное тестирование

1. Рекомендуется установить порог сходства ниже оптимального на 10-20%.
2. Настройте ведение протокола: установите флажок Write log file.
3. Укажите путь к файлу протокола в поле Log file. Для этого используйте кнопку Browse....
4. Сбор статистики.
Запустите тестовое распознавание на сутки.
В течение всего времени тестирования фиксируйте появление в эфире анализируемой звуковой отбивки: записывайте время поступления отбивки на вход.
5. Анализ протокола.
Сравните протокол программы с собственными записями и выясните:
 1. Все ли звуковые отбивки распознались.
Если хотя бы один раз детектор не сработал, то значит реальный уровень сходства меньше порога срабатывания, и, следовательно параметры распознавания подобраны неправильно.
 2. Есть ли ложные срабатывания.
Если есть, то нужно изменить пороговое значение уровня



сходства. Для этого следует выполнить следующие действия:

- а) найти максимальное значение из всех коэффициентов сходства, при которых произошли ложные срабатывания;
- б) найти минимальное значение из всех коэффициентов сходства, при которых происходили истинные срабатывания;
- в) в качестве порогового уровня сходства задать среднее значение между найденными минимальным (а) и максимальным (б) значениями.

Чем больше разница между значениями (а) и (б), тем выше надежность автоматического определения отбивки.



Полезные ссылки

Линейка продуктов ФорвардТ: описание, загрузка ПО, документация, готовые решения

<http://www.softlab-nsk.com/rus/forward>

Техподдержка

e-mail: forward@sl.iae.nsk.su

forward@softlab-nsk.com

Форумы

<http://www.softlab-nsk.com/forum>