



ЦРТ

ГРУППА
КОМПАНИЙ

WWW.SPEECHPRO.RU

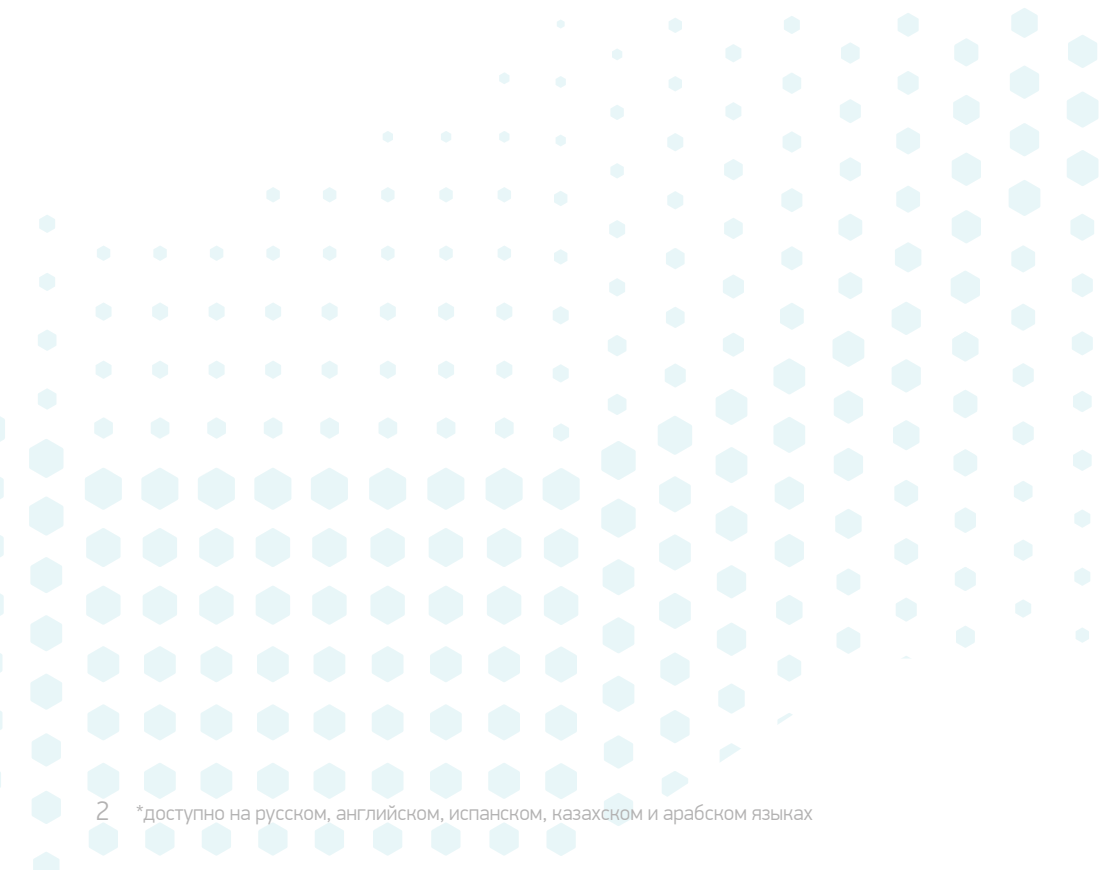


ИКАР Лаб 3

Новое поколение

АПК для криминалистической экспертизы аудиозаписей

Исследование подлинности | Идентификация личности | Распознавание речи* | Шумоочистка



Новое поколение ИКАР Лаб 3

Отвечает мировым требованиям ведущих экспертов-криминалистов: обеспечивает высокую надежность доказательной базы, максимально автоматизирует, ускоряет и упрощает работу.

Все операции выполняются на одном рабочем месте в едином удобном интерфейсе. Это позволяет эксперту сфокусироваться на принятии решения и оперативно предоставить оценку.



Надежные методы оценки аутентичности и применимости фонограммы для экспертизы



Ускоренная и надежная идентификация человека, круга лиц и реплик каждого



Высокая скорость и точность выполнения шумоочистки



Быстрый доступ к фонограммам видеозаписей



Работа с **большими объемами фонограмм**

Основные компоненты

SIS

Специализированный звуковой редактор

5

Sound Cleaner

Программа шумоочистки речевых сигналов

22

Цезарь-Р

Программный модуль преобразования речи в текст

24

STC-H246

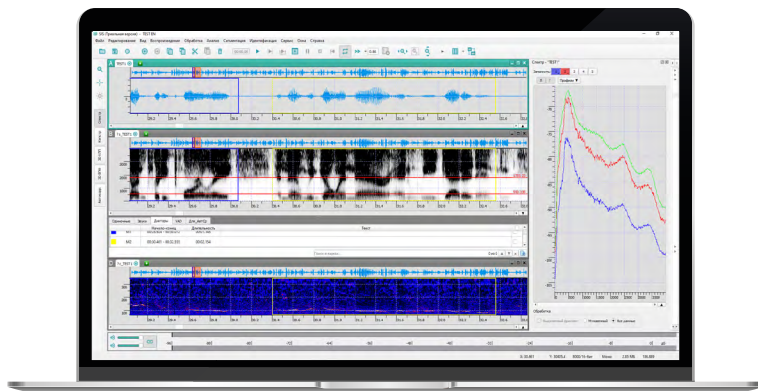
Устройство ввода-вывода звука

25



Специализированный звуковой редактор

Является центральным компонентом и включает в себя мощные инструменты исследования речевых сигналов, превосходные возможности отображения и анализа видимой речи, сегментации и текстовой расшифровки, автоматические и полуавтоматические средства идентификации, а также многие другие функции.



Исследование фонограммы для идентификации диктора по образцам речи (анализ формант и основного тона речевого сигнала)

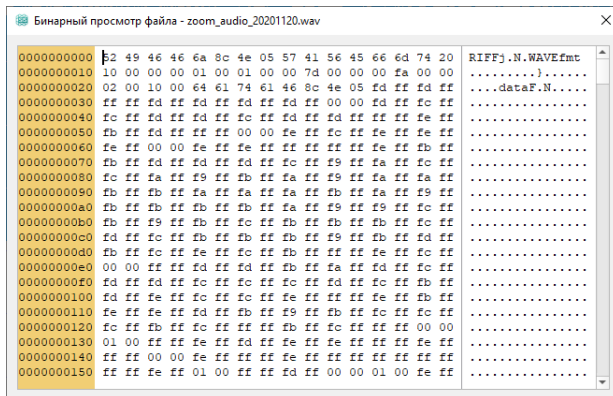
Методы

- Бинарный просмотр файла
- Поиск слов в разметке сигналов
- Модуль «Речь в текст»
- Подключение Sound Cleaner в SIS
- Визуализация
- Редактирование и обработка
- Детектирование речи и шумов
- Текстовая расшифровка и сегментация
- Разделение дикторов в диалоге/полилоге
- Многооконный интерфейс
- Сравнение сигналов
- Расчет характеристик сигнала
- Управление проектами, создание отчетов
- Идентификация
- Автоматическое сравнение
- Сравнение речевых признаков
- Сравнение формант
- Сравнение основного тона
- «Методика»
- Общее решение
- Анализ извлеченной из видеофайла фонограммы
- EdiTracker и модуль диагностики

Бинарный просмотр файла

НОВОЕ

Данная операция позволяет исследовать бинарное содержимое заголовков звуковых файлов, записанных в виде шестнадцатеричных чисел.

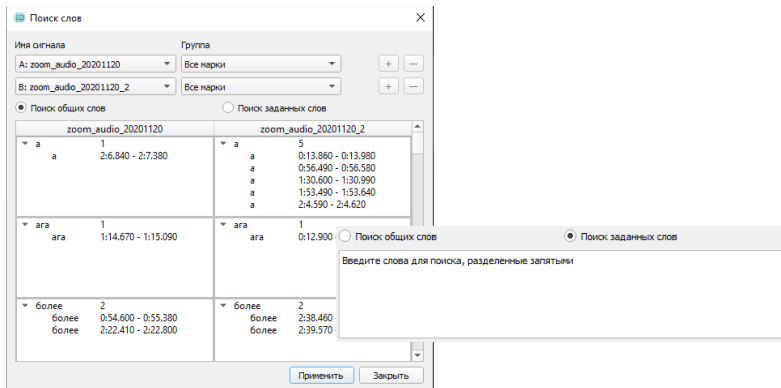


SIS | Методы

Поиск слов в разметке сигналов

НОВОЕ

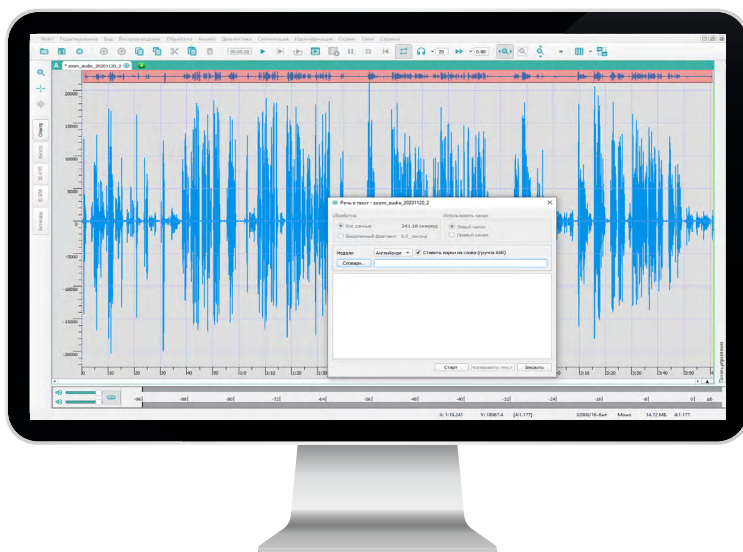
Данная операция позволяет в ходе работы с сигналами произвести поиск либо всех одинаковых слов в двух заданных сигналах, либо найти слова из заданного оператором списка.



Модуль «Речь в текст»

Наличие пользовательских словарей позволяет эксперту, знакомому с фонетикой, подключить к модулю распознавания речи дополнительные словари транскрипций. Это полезно в том случае, когда речь на фонограмме насыщена сленгом, т.е. словами, которые не встречаются в обычной речи. В этом случае автоматическое распознавание «не знает», как произносятся эти слова. В этом может помочь дополнительный словарь.

НОВОЕ



SIS | Методы

Подключение Sound Cleaner в SIS (через VST плагин)

В программе SIS в меню «Обработка» появилась операция «Sound Cleaner». Это позволяет воспользоваться схемами обработки сигналов в Sound Cleaner непосредственно для работы с сигналом в SIS.

НОВОЕ

Визуализация

Используемые алгоритмы спектрального представления сигнала обеспечивают максимально возможное качество и четкость видимой речи. Пользователь «на лету» подбирает оптимальные параметры отображения или использует пресеты для различных видов спектрального анализа.

- Осциллограммы
- БПФ и КЛП спектрограммы
- Средний и мгновенный спектр
- Кепстрограмма
- Автокоррелограмма
- Выделитель основного тона
- Выделитель формант
- Энергия
- Гистограмма и корреляция гистограмм

Редактирование и обработка

SIS дает эксперту широкий выбор инструментов редактирования и обработки сигналов для улучшения разборчивости и подготовки фонограмм к дальнейшему исследованию.

- Нормализация амплитуды
- Линейное преобразование
- Удаление постоянной составляющей
- Микширование
- Модуляция
- Темпокоррекция*
- Изменение частоты дискретизации
- Изменение битности
- Разделение стереосигнала и объединение двух моносигналов в стерео
- Изменение фазы
- Адаптивный инверсный фильтр
- Адаптивный тональный фильтр
- Адаптивный широкополосный фильтр

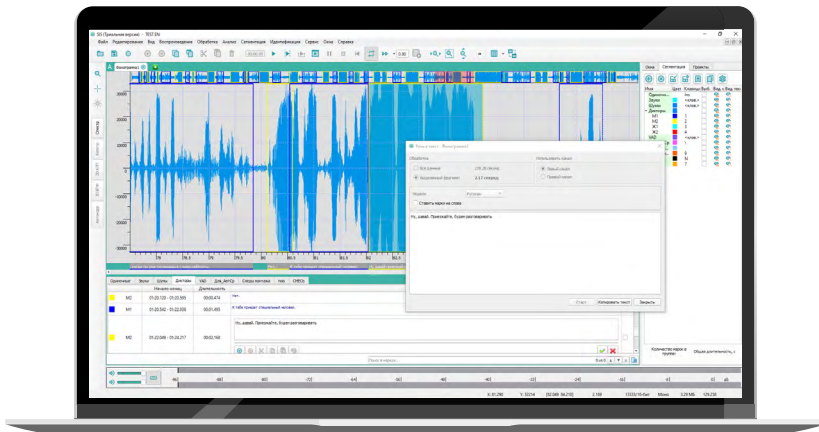
Детектирование речи и шумов

Детектор речи автоматически размечает фонограмму на речевые фрагменты, пригодные для идентификации. Модуль может быть также настроен на детектирование шумовых участков: телефонные гудки, клиппирование, щелчки.

Текстовая расшифровка и сегментация

Плагин текстовой расшифровки позволяет **автоматически получить текстовое содержание** речевого сигнала, записанного на фонограмме, на русском, английском, испанском, казахском и арабском языках. Дополнительно расшифровка сопровождается **сегментацией**, указывающей расположение произнесенных слов. Данный функционал дает возможность эксперту эффективно работать с большими объемами фонограмм.

В ручном режиме выделенный аудиофрагмент может быть отмечен группой марок (например, диктор, звук, шум) с текстовыми комментариями, а общий текст — экспортирован в MS Word. При наличии текстовой расшифровки для двух файлов программа автоматически ищет все совпадающие слова в сравниваемых фонограммах.



Исследование фонограммы для идентификации диктора по образцам речи (анализ формант и основного тона речевого сигнала)

Разделение дикторов в диалоге/полилоге

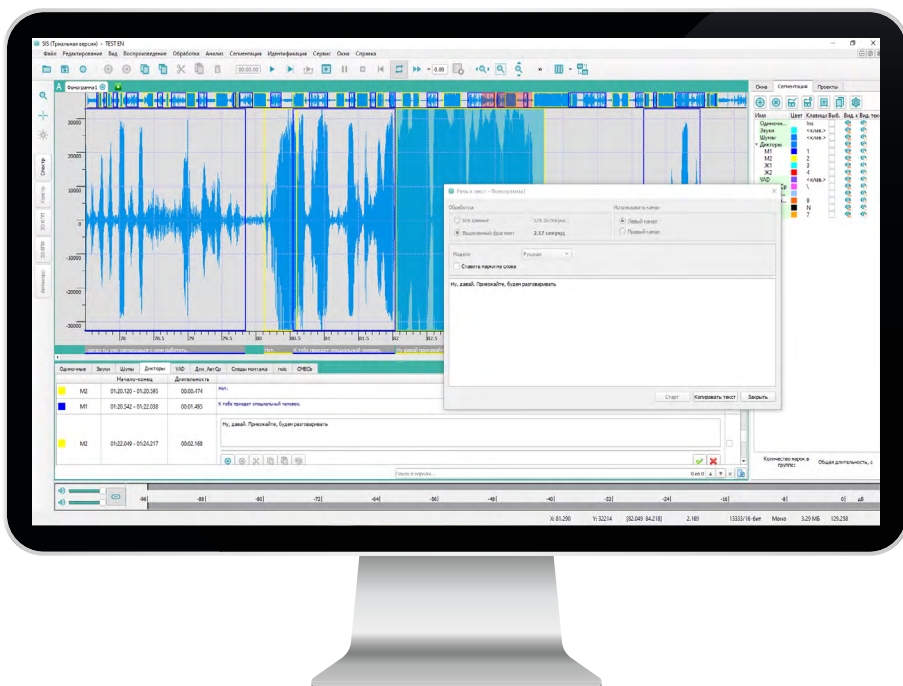
Модуль автоматически маркирует реплики по принадлежности дикторам. Надежность его работы достигает 95% при отношении сигнал-шум не менее 20 дБ и длительности речи каждого диктора на фонограмме не менее 16 секунд.

С помощью встроенных алгоритмов модуль позволяет разделять реплики до 5 дикторов, участвующих в разговоре.

Многооконный интерфейс

SIS позволяет открывать много файлов с различным представлением сигнала в одном или разных окнах. Окна могут быть удобно и быстро расположены под различные задачи: вертикально для идентификации и горизонтально для сравнения копий фонограмм или вариантов шумоочистки.

Сигналы могут открываться несколькими слоями в одном окне, а их цвета и прозрачность изменяться для лучшего восприятия.



Экспертное исследование фонограмм в многооконном интерфейсе

Сравнение сигналов

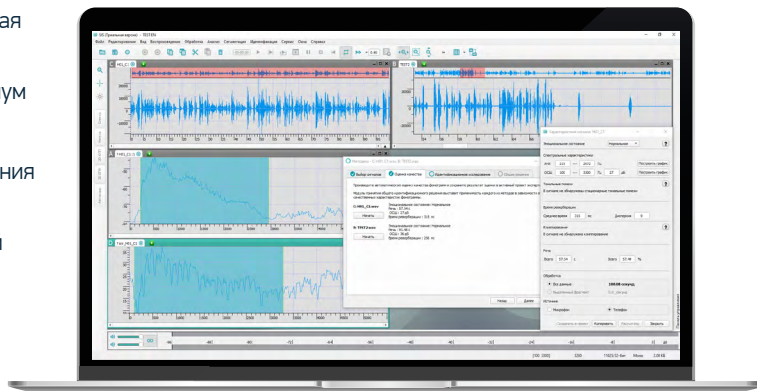
Окна могут связываться во временной и частотной областях для удобного измерения с помощью вертикального и горизонтального курсора. Мгновенные спектры с формантами могут накладываться для визуального сравнения.

Возможно также сопоставление гистограмм основного тона дикторов визуально или численно по значениям минимума, максимума, медианы, асимметрии и общей корреляции.

Расчет характеристик сигнала

SIS автоматически рассчитывает характеристики сигнала, основываясь на которых эксперт делает вывод о степени пригодности аудиозаписей к проведению идентификационной экспертизы.

- Амплитудно-частотная характеристика
- Отношение сигнал/шум
- Время реверберации
- Наличие клиппирования и тональных помех
- Длительность чистой речи



Оценка технических характеристик сигнала

Управление проектами и создание отчетов

ИКАР Лаб 3 эффективно организует рабочий процесс эксперта. Проект запускает из интерфейса SIS файлы, относящиеся к экспертизе (аудио, текст, фото и видео), структурировано хранит данные файлы, а также результаты идентификации, и создает отчеты в формате MS Word. Отчет может быть дополнен информацией о настройках видимой речи и иллюстрациями: скриншотами рабочего экрана или его области.

Идентификация

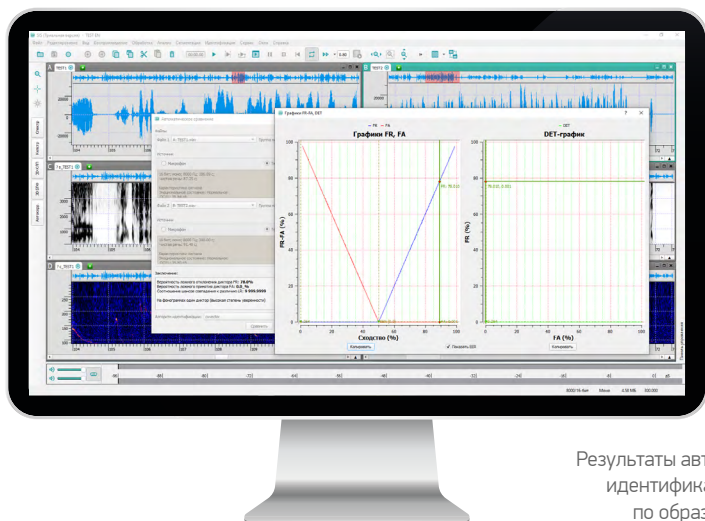
В основе уникального инструмента — биометрические алгоритмы и экспертные модули, которые автоматизируют и формализуют процессы фоноскопического идентификационного исследования: поиск одинаковых слов или звуков, отбор сравниваемых звуковых и мелодических фрагментов, сравнение дикторов по формантам и основному тону, речевой анализ. Результаты представляются в виде численных показателей общего идентификационного решения.

Автоматическое сравнение

Модуль позволяет сравнивать речевые сигналы 1:1, применяя те или иные методы в зависимости от характеристик речевых сигналов исследуемых фонограмм. Результаты сравнения строятся на выделении идентификационно значимых признаков и вычислении меры близости полученных наборов признаков между собой.

Используется больше методов сравнения: `sxvector` (развитие `xvector`) — в качестве основного, дополнительно к нему — `smart-speaker` и `genb-v3` (когда содержание чистой речи в фонограмме от 1,5 до 5 сек.). Новый функционал обеспечивает ускоренную и надежную идентификацию.

Модуль обучался на большой выборке фонограмм: десятки тысяч дикторов, различные каналы и множество сессий звукозаписи, разнообразные типы речевого материала. Сегодня высокая надежность биометрического движка подтверждена испытаниями NIST.



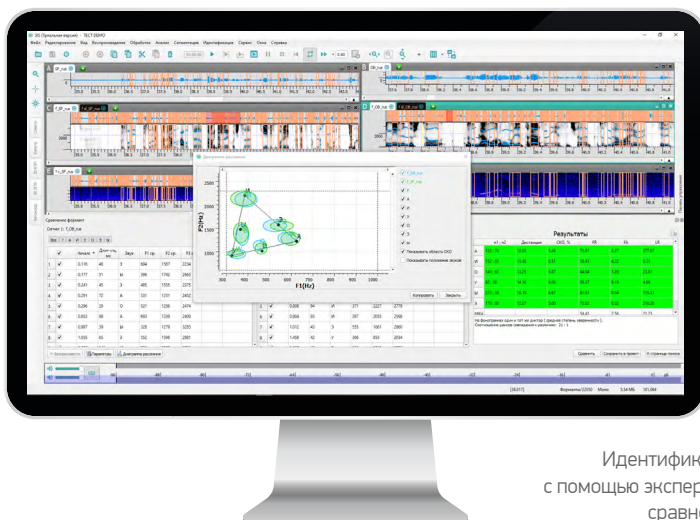
Сравнение речевых признаков

Модуль реализован в виде списка из 45 базовых признаков и расширенного списка признаков с примерами их звучания в современном русском языке. Для каждого признака эксперт может отметить его выраженность и пример реализации на фонограмме. После заполнения таблицы признаков для двух дикторов, модуль покажет общую статистику по несовпавшим, полностью совпавшим и частично совпавшим.

Сравнение формант

Процесс идентификации с использованием модуля сравнения формант осуществляется в два этапа.

1. Поиск и отбор опорных звуковых фрагментов для известного и неизвестного дикторов:
 - по диаграмме рассеяния с треугольником Фанта путем выделения области поиска
 - указанием частотных диапазонов поиска формант
 - по установленным горизонтальным маркам с указанием допуска в герцах и процентах
 - по всем сравниваемым гласным звукам
2. Экспертное сравнение. Модуль автоматически рассчитывает FR, FA и LR для выбранных звуков и принимает решение о положительной/отрицательной идентификации или выдает неопределенный результат



Идентификация диктора с помощью экспертного метода сравнения формант

Дополнительные возможности модуля:

- Визуальное сравнение распределения выбранных звуков на скаттерграмме
- Сравнение усредненных значений формант для выбранных звуков двух дикторов
- Текстовые комментарии для опорных фрагментов – возможность указания слова или триады
- Экспорт таблиц опорных фрагментов и результатов в MS Word

Сравнение основного тона

Модуль автоматизирует процесс идентификации дикторов с помощью метода анализа мелодического контура. Предусмотрено 18 типов фрагментов контура и 15 параметров их описания (значения минимума и максимума, среднее значение, значение скорости изменения тона, эксцесса, скаса и др.).

Модуль возвращает результаты сравнения в виде процентного совпадения для каждого из параметров и принимает решение о положительной/отрицательной идентификации или выдает неопределенный результат. Все данные могут экспортироваться в текстовый отчет.



Идентификация диктора с помощью экспертного метода сравнения параметров основного тона

«Методика»

Модуль систематизирует порядок работы эксперта при проведении идентификационной экспертизы: показывает этапы исследования и результаты по каждому.

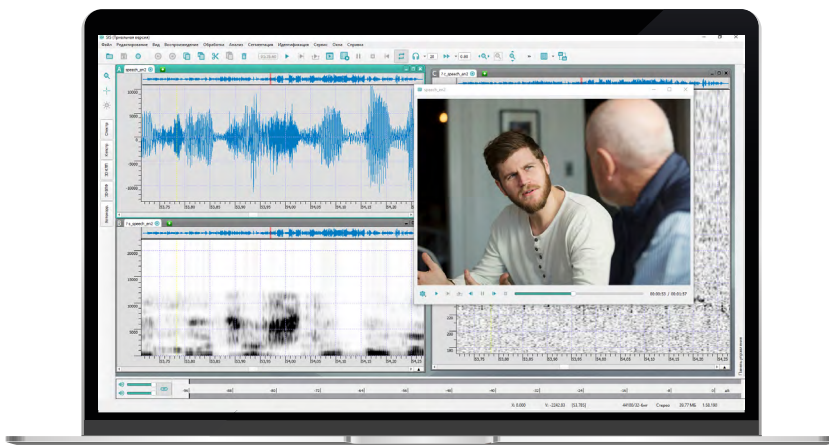
Общее решение

Результаты работы каждого метода исследования могут быть сохранены экспертом в проект. Основываясь на результатах работы модулей, программа позволяет принять общее идентификационное решение. При этом либо пользователь настраивает применимость каждого из методов, изменяя его вес в общем решении, либо она выставляется автоматически по результатам расчета качественных и количественных характеристик сравниваемых фонограмм. По результатам проведенного исследования эксперт может автоматически сгенерировать подробный отчет.

Анализ извлеченной из видеофайла фонограммы

Используя новый метод SIS, эксперт без каких-либо дополнительных редакторов получает быстрый доступ к звуковой дорожке, сопровождающей видеофайл. Для этого необходимо загрузить видеофайл как любую другую фонограмму, после чего SIS автоматически выделит из видео звуковую дорожку и загрузит ее в отдельное окно.

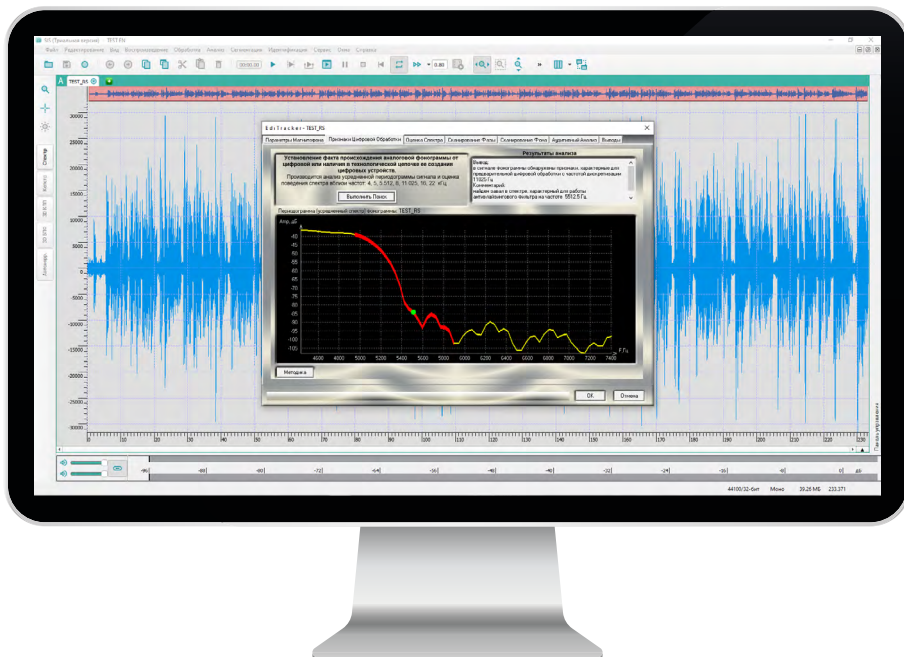
Модуль обеспечивает параллельную работу с видео в видеоплеере и с фонограммой в редакторе: видео и звук синхронизируются, а при редактировании фонограммы видео автоматически корректируется.



Анализ извлеченной из видеофайла фонограммы

EdiTracker

Модуль проводит оперативную диагностику достоверности аналоговых и цифровых фонограмм, вместе с SIS значительно упрощает экспертный анализ, предоставляя пользователю ручные и автоматические методы анализа.



Проверка аутентичности фонограммы на предмет применения предварительной цифровой обработки

Методы анализа EdiTracker

- Исследование технических характеристик звукозаписывающего устройства
- Обнаружение следов цифровой предобработки сигнала
- Обнаружение стационарных гармоник и сканирование их фазы
- Сканирование фонового шума
- Аудитивный анализ

Исследование технических характеристик звукозаписывающего устройства

У каждого аналогового устройства аудиозаписи есть свои особенные технические характеристики, такие как амплитудно-частотная характеристика, коэффициент нелинейных искажений, детонация, эффективный частотный диапазон, отклонение скорости и др.

Используя тестовый сигнал, EdiTracker автоматически оценивает характеристики записывающего устройства. Несоответствие параметров звукозаписывающего устройства характеристикам аудиосигнала, который был предположительно на него записан, может являться признаком монтажа аудиозаписи.

Обнаружение следов цифровой предобработки сигнала

Цифровая обработка аналоговых сигналов всегда подразумевает определенную частоту дискретизации. Во время оцифровки аналогового сигнала может наблюдаться элайзинг-эффект: высокочастотные составляющие накладываются на низкочастотные — аудиосигнал искажается. Чтобы это

избежать, большинство аналого-цифровых и цифро-аналоговых преобразователей используют антиэлайзинговые фильтры. EdiTracker автоматически анализирует спектр сигнала и ищет следы применения антиэлайзинговых фильтров, что может говорить о предшествующей цифровой обработке сигнала.

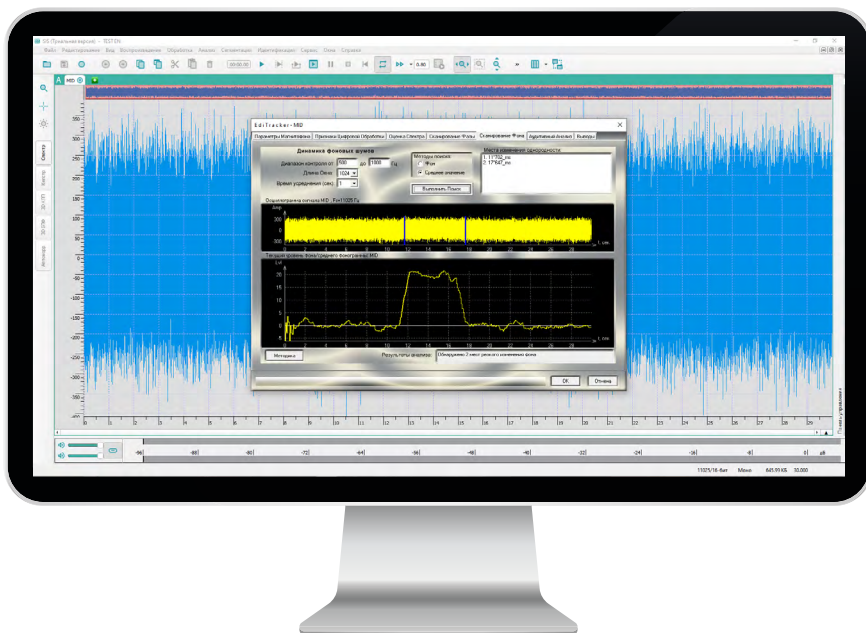
Обнаружение стационарных гармоник и сканирование их фазы

EdiTracker автоматически сканирует аудиосигнал на наличие стационарных узкополосных сигналов — гармоник, которые обычно производятся электрической сетью, аккумуляторными батареями, электрическими приборами.

После обнаружения подобного рода сигналов модуль сканирует их фазу на предмет ее непрерывности. Неоправданный разрыв фазы гармоник может быть интерпретирован как возможный след редактирования аудиозаписи.

Сканирование фонового шума

Сканирование позволяет обнаружить места резкого изменения спектра фоновых шумов. Эти изменения могут быть ситуационно-оправданными или являться признаками монтажа. Также данный метод анализирует постоянную составляющую сигнала и отмечает места резкого изменения ее значения.



Проверка аутентичности фонограммы на предмет скрытого редактирования по признаку равномерности фоновых шумов

Аудитивный анализ

Во время звучания оригинального аудиосигнала совокупность коммуникативных обстоятельств образуют целостную картину взаимосогласованных звуковых и речевых событий (речевое и неречевое поведение дикторов, характеристики звучания дополнительных источников звука). Слуховой анализ данных событий с учётом известных экспертам характеристик

использованной аппаратуры и способов звукозаписи позволяет выявить нарушения данной целостности и установить места, факты и способы таких нарушений. EdiTracker включает в себя расширенный список аудитивно-лингвистических признаков нарушения подлинности фонограмм, который в дальнейшем можно использовать в текстовом отчете эксперта.

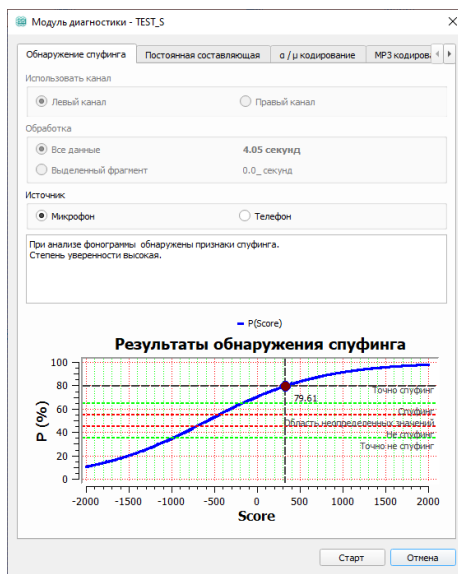
Модуль диагностики

Это новый модуль SIS для более надежной оценки аутентичности и применимости фонограммы для экспертизы. Модуль выявляет в сигнале разнообразные признаки, объясняющие природу его происхождения или возможные способы обработки, которые в свою очередь могут быть либо неизвестны, либо умышленно скрыты. Дополняя EdiTracker, он выявляет применение тех или иных операций с сигналом с помощью следующих методов:

- детектирование спуфинга (Spoofing Detector)
- анализ постоянной составляющей (Constant Component)
- анализ следов А/μ-кодирования (Codecs)
- анализ следов MP3-кодирования (MP3 Coding)

Детектирование спуфинга (Spoofing Detector)

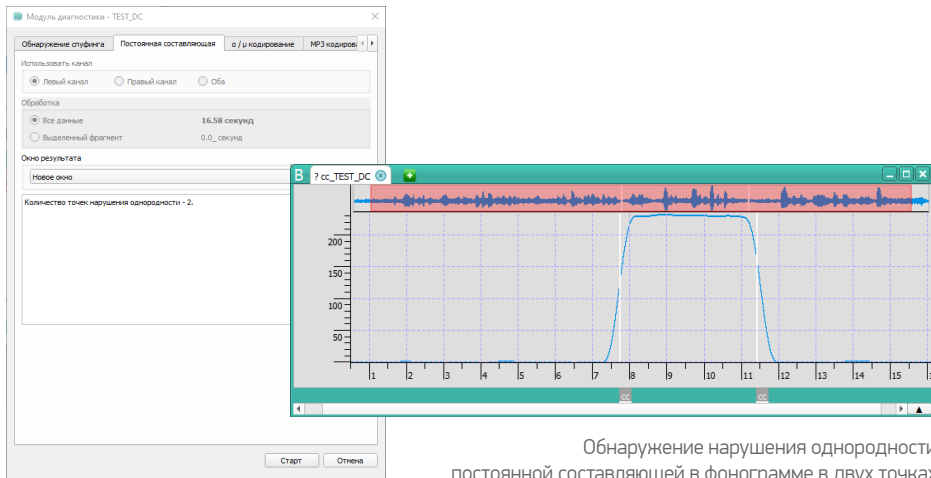
Модуль анализирует фонограмму на предмет следов применения спуфинг-атак: повторное воспроизведение, синтез речи, преобразование голоса. Алгоритм работы основан на использовании нейронной сети, обученной на базе спуфинг-атак различной природы. В результате работы алгоритма делается вывод о том, является ли фонограмма попыткой подменить аутентичную запись диктора или нет.



Экспертное исследование фонограммы на предмет обнаружения спуфинга

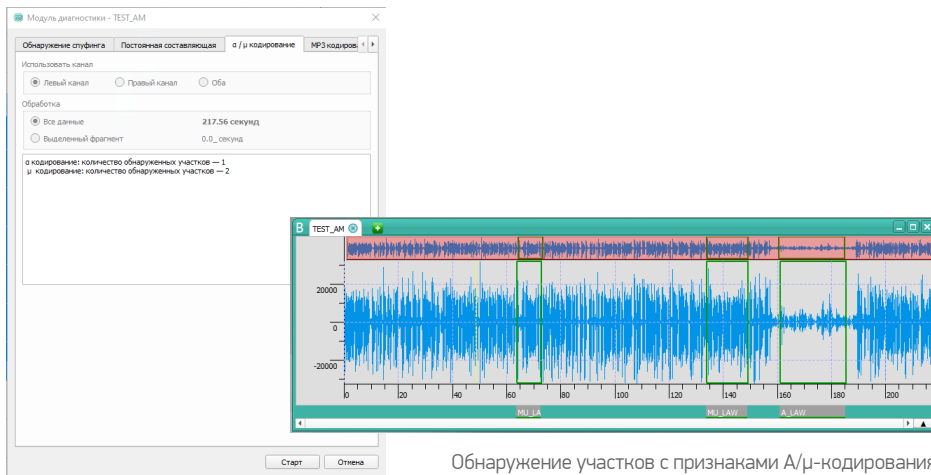
Анализ постоянной составляющей (Constant Component)

Модуль анализирует фонограммы на предмет выявления участков с резким и значительным изменением постоянной составляющей, что может являться одним из признаков нарушения аутентичности. В случае обнаружения таковых модуль выдает отмеченные марками участки.



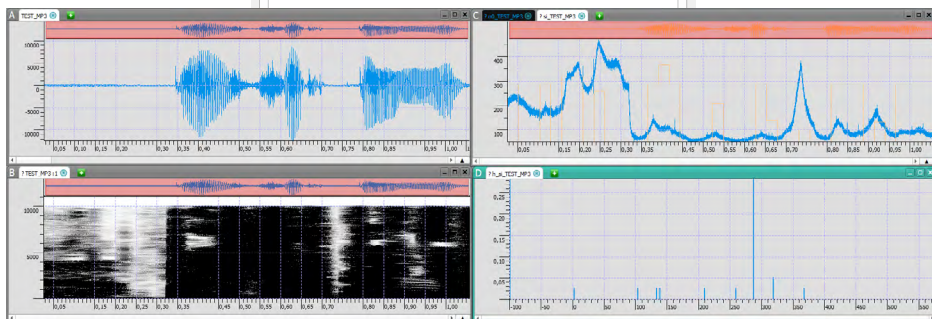
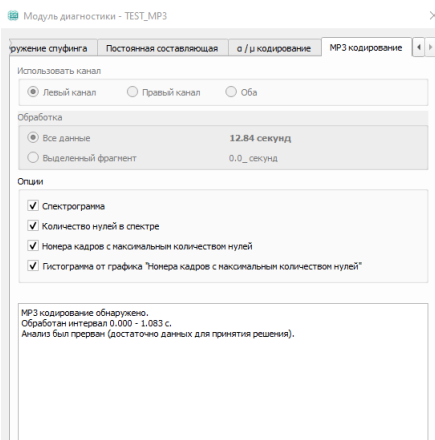
Анализ следов А/μ-кодирования (Codecs)

Модуль исследует фонограммы на предмет выявления участков с признаками кодирования А/μ-кодеками. В формате изучаемых фонограмм не указана возможность обработки данными кодеками. В случае обнаружения модуль выдает отмеченные марками участки или всю фонограмму.



Анализ следов MP3-кодирования (MP3 Coding)

Модуль анализирует фонограммы на предмет обнаружения признаков кодирования MP3-кодеком. В формате исследуемых фонограмм не указана возможность обработки данным кодеком. В случае обнаружения модуль выдает информационное сообщение о наличии или отсутствии признаков MP3-кодирования. Дополнительно отображаются спектрограммы, графики и гистограммы, иллюстрирующие принятое алгоритмом решение.



Sound Cleaner

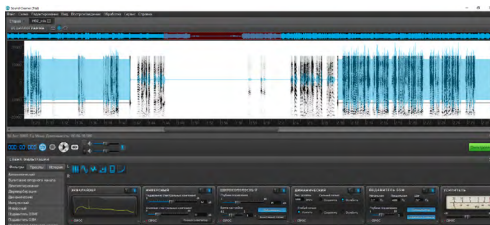
Программа шумоочистки речевых сигналов

Большая часть аудиоматериалов на экспертизе требует установления дословного содержания — стенограммы записи речи. Поскольку полученные оперативным путем фонограммы могут быть записаны в сложных условиях и зачастую не очень разборчивы, подготовительным этапом проводится шумоочистка. Для этого комплекс ИКАР Лаб 3 опционально комплектуется Sound Cleaner. Он включает современные алгоритмы обработки сигнала и успешно подавляет широкополосные шумы, тональные помехи, импульсы, проводит коррекцию АЧХ, выравнивает амплитуду сигнала и др.

Sound Cleaner



Осциллограмма и схема обработки стереосигнала с заданными экспертом последовательностью фильтров и параметрами



Одновременное отображение спектрограммы и осциллограммы для удобной, быстрой оценки участков сигнала и частот, где требуется обработка

Для определения характеристик шумов и помех предусмотрена возможность построения спектрограмм, в том числе в формате 3D БПФ. Это повышает скорость и точность выполнения шумоочистки.

DNN-шумоочистка

НОВОЕ

В программе Sound Cleaner появился новый фильтр, использующий технологию глубоких нейронных сетей для подавления различных видов шумов, включая реверберацию. Фильтр имеет всего один элемент управления - степень подавления, с его помощью эксперт сможет контролировать степень подавления шумов.

Все фильтры работают в режиме реального времени — результат слышен сразу после включения фильтра в цепочку обработки и пользователь может осуществлять подбор оптимальных параметров на слух.

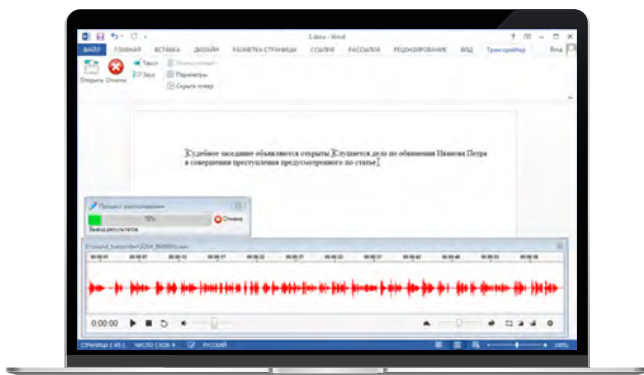
DNN - фильтр Существенное понижение уровня наиболее распространенных типов шумов с помощью одного регулятора.	Подавление GSM-наводки Ослабление помех от входящих звонков мобильных телефонов с характерными прерывистыми звуками.
Широкополосный фильтр Снижение уровня шума помещений и улиц, помех каналов связи или записывающей аппаратуры. По звучанию они соответствуют шипению и не устраняются другими методами, так как спектр помехи пересекается/совпадает со спектром полезного сигнала.	Инверсный фильтр Выравнивание амплитудно-частотной характеристики канала связи, в котором сделана аудиозапись. Две настройки: усиление слабых спектральных компонент сигнала и ослабление мощных (приближение усредненного спектра к плоскому).
Тональный фильтр Подавление узкополосных стационарных и регулярных помех (вибрации, сетевые наводки, шумы бытовых приборов, медленная музыка, звук проезжающего автомобиля, шум воды или зала, реверберация и т.д.).	Дереверберация Увеличение разборчивости речи, снижение уровня реверберации в аудиозаписях и утомляемости пользователя за счет комфортного восприятия полезного ревербированного речевого сигнала при наличии аддитивного шума.
Импульсный фильтр Восстановление в автоматическом режиме речевых или музыкальных сигналов, искаженных импульсными помехами (щелчками, помехами радиозэфира, стуками, треском и т.д.).	Динамический фильтр Повышение разборчивости при больших перепадах уровня сигнала. Например, усиление тихого и ослабление громкого сигнала для выравнивания амплитуды выходного сигнала.
Эквалайзер 4096-полосный графический эквалайзер со встроенным спектрографом для детальной коррекции спектра в искаженных аудиозаписях.	Фильтр деклиппирования Восстановление перегруженных фрагментов аудиозаписи за счет реконструирования волновой формы данных участков.
Стереофильтрация по опорному каналу Вычитание из основного канала шумов, которые присутствуют в опорном (например, телевизионная или радиотрансляция, музыка).	Подавитель DTMF-сигналов Обработка сигналов набора телефонного номера, то есть последовательности коротких прямоугольных импульсов с двухчастотным заполнением.

Sound Cleaner сохраняет результаты обработки в формате WAV и автоматически создает текстовый протокол с фиксацией всей информации о произведенной обработке сигнала. Также программа совместима с любыми звуковыми редакторами по формату VST 3.

Цезарь-Р

Программный модуль преобразования речи в текст

Модуль предназначен для голосового ввода текстовой информации в ПК и установления дословного содержания фонограмм речи. В первом режиме текст набирается в любом активном окне ПО Windows с функцией ввода текста. Во втором режиме Цезарь-Р подключается как плагин к текстовому редактору MS Word и автоматически конвертирует звук в текст. Распознанный или набранный с клавиатуры текст фонограммы автоматически связывается со звуком. Это упрощает последующий поиск соответствующего тексту аудиофрагмента и корректировку текста. Таким образом, автономный режим воспроизведения и стенографирования в едином интерфейсе упрощает работу эксперта.



Автоматическая текстовая расшифровка фонограммы

Параметр	Значение
Язык распознавания	Русский
Объем базового словаря	353 364 слов и словоформ
Словари	Базовый: общая тематика и разговорный Тематические: юридический, финансовый
Точность распознавания	Диктовка: 86%* Аудиозаписи с медийного канала: до 77%**
Скорость распознавания	Диктовка: 1 RT*** Аудиозаписи: до 1 RT
Форматы аудиофайлов	WAV, MP3, AIFF, WMA, FLAC, OGG и др.

STC-H246

Устройство ввода-вывода звука

Для получения гарантированно высокого качества входного и выходного сигналов комплексы ИКАР Лаб 3 комплектуются профессиональным устройством ввода-вывода сигнала STC-H246.

STC-H246 идеально подходит для организации рабочей станции по оцифровке аналоговых фонограмм, так как является надежным устройством для измерения характеристик и формирования электрических сигналов в звуковом диапазоне частот.



STC-H246

Параметр	Значение
Поддерживаемые частоты дискретизации, кГц	8-200
Формат преобразования данных, бит на отсчёт	16, 24
Соотношение сигнал/шум в сквозном канале, в полосе частот от 20 до 20 000 Гц, дБ	105
Типы разъемов входных/выходных каналов	XLR, RCA, S/PDIF, TRS 6,3
Количество каналов	2
Питание	110/220В 60/50Гц
Корпус	Металл
Габариты, мм	111x166x190
Операционные системы	Windows 7, 8, 10

О компании

Группа компаний ЦРТ — глобальный разработчик продуктов и решений на основе интеллектуальных речевых технологий, голосовой и лицевой биометрии. Технологический эксперт в области искусственного интеллекта и машинного обучения с 30-летним опытом.

Мы занимаем лидирующие позиции в мировых рейтингах NIST, ASVspoof Challenge, VOICES, SHIME, и на данный момент реализовали проекты более чем для 5000 клиентов в 75 странах.

Многие решения компании находят успешное применение в государственном и коммерческом секторе, от небольших экспертных лабораторий до сложных систем безопасности национального масштаба по всему миру, включая США, Латинскую Америку, Ближний Восток, Европу.



Санкт-Петербург

194044, Выборгская наб., 45, лит. Е
Тел.: +7 812 325 8848
Факс: +7 812 327 9297
stc-spb@speechpro.com

Москва

109004, Земляной Вал, 59, стр. 2
Тел.: +7 495 669 7440
Факс: +7 495 669 7444
stc-msk@speechpro.com