

Бояров А.Г.

старший государственный судебный эксперт
лаборатории судебной экспертизы видео- и звукозаписей
ФБУ РФЦСЭ при Минюсте России

ОБНАРУЖЕНИЕ И АНАЛИЗ СЛЕДОВ СЖАТИЯ ЗВУКОВЫХ СИГНАЛОВ КОДЕКАМИ MP3, AAC, WMA И VORBIS

(МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ДЛЯ ЭКСПЕРТОВ)

В работе описывается метод выявления следов кодирования звуковых сигналов кодеками MP3, AAC, WMA и Vorbis. Приводятся примеры использования данного метода для обнаружения признаков различных изменений фонограмм – монтажа, смены частоты дискретизации и повторного кодирования.

Ключевые слова: криминалистическая экспертиза видео- и звукозаписей, поиск следов кодирования, психоакустические кодеки, MP3, AAC, WMA, Vorbis.

Boyarov Alexander, Senior Forensic Examiner of Laboratory of Forensic Video and Audio Examinations of the Russian Federal Center of Forensic Science of the Russian Ministry of Justice

DETECTION AND ANALYSIS OF MP3, WMA, OGG AND VORBIS CODECS TRACES IN AUDIO SIGNAL

The article describes the method of MP3, WMA, OGG and Vorbis codecs traces detection in audio signal. The method reveals digital audio editing, sampling rate changing and traces of multiple coding.

Keywords: forensic analysis of digital audio and video, codecs traces detection, psychoacoustic codecs, MP3, AAC, WMA, Vorbis.

Введение

В настоящее время достаточно часто объектами экспертизы видео- и звукозаписей становятся цифровые фонограммы или видеофонограммы, звуковой сигнал которых подвергался сжатию. Целью сжатия, как правило, является или сокращение голосового трафика в каналах связи, или уменьшение объёма хранимых данных. Устройства и программы, реализующие алгоритмы сжатия аудио- и видеосигналов, называются кодеками.

Большее признание в области цифровой записи и хранения звуковых сигналов получили так называемые психоакустические кодеки, которые обеспечивают сжатие сигнала за счёт удаления из него неслышимых человеком спектральных компонент (частотная и временная маскировка) [2]. Использование подобных кодеков значительно снижает объём памяти, необходимый для представления сигнала, оставляя качество звучания на приемлемом для бытового использования уровне, поэтому психоакустические кодеки получили широкое распространение в медиаиндустрии.

Наиболее известным и распространённым представителем семейства психоакустических кодеков является MPEG 1/2/2.5 Layer 3, чаще называемый MP3-кодеком. Разработанный более 20 лет назад MP3-кодек сегодня реализован практически в любом устройстве с функцией записи и воспроизведения фонограмм или видеофонограмм на программном или аппаратном уровне.

В последнее десятилетие все большее распространение получают психоакустические кодеки, использующие более совершенные психоакустические модели: Advanced Audio Codec (AAC), WMA (Windows Media Audio) и Ogg Vorbis (OGG).

Теоретические основы

Анализируя динамическую спектрограмму сигнала, прошедшего кодирование с использованием психоакустических кодеков, зачастую легко заметить выпадающие области прямоугольной формы (рис. 1), что является одним из признаков применения одного из психоакустических кодеков.

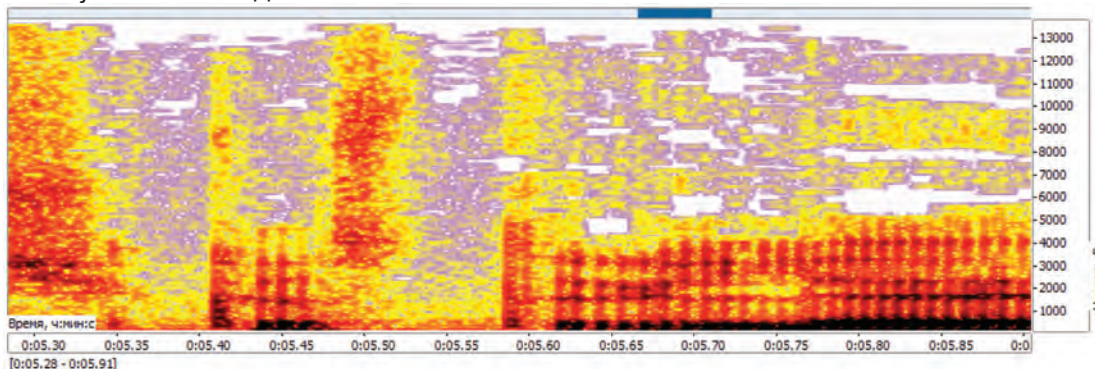


Рис. 1. Динамическая спектрограмма со следами кодирования MP3-кодеком

Данные выпадения являются результатом кодирования сигнала психоакустическим кодеком, описание работы которого приводится ниже – на примере MP3-кодека.

На первом этапе MP3-кодирования производится вычисление спектра сигнала при помощи Модифицированного дискретного косинусного преобразования (МДКП). Далее на основе психоакустической модели частотной и временно й маскировки неслышимые компоненты МДКП-спектра обнуляются. Затем спектр сигнала квантуется и кодируется при помощи метода Хаффмана.

Для упрощения дальнейшего изложения описание стадии кодирования, связанной с полосовой фильтрацией и понижением частоты дискретизации «полос» сигнала перед вычислением МДКП-спектров, не приводится, так как это не имеет значения в рассматриваемом контексте. В связи с этим упрощением в работе размеры окон анализа будут указываться для исходного сигнала, а не для «полос» сигналов, как указывается в спецификациях [3,4].

Для удобства MP3-спектрами названы МДКП-спектры, которые вычисляются тем же образом, что и при MP3-кодировании. Вычисление MP3-спектров может производиться с использованием четырёх видов окон анализа (см. рис. 2): со стандартным окном размером 1152 отсчёта (обозначено синим цветом), малым окном размером 384 отсчёта (обозначено красным цветом) и двумя видами переходных окон (обозначено зелёным цветом). При этом размеры окон не зависят от частоты дискретизации исходного сигнала. В процессе кодирования исходный сигнал разбивается на пересекающиеся фрагменты с шагом 576 отсчётов (шаг окна MP3-кодирования). Размер фрагмента в зависимости от типа окна может варьироваться от 1152 отсчётов для стандартного окна, 960 для переходного окна и 768 для малого окна (три малых окна с пересечением 50%), однако шаг между «центрами» фрагментов во всех случаях равен 576 отсчётам (см. рис 2).

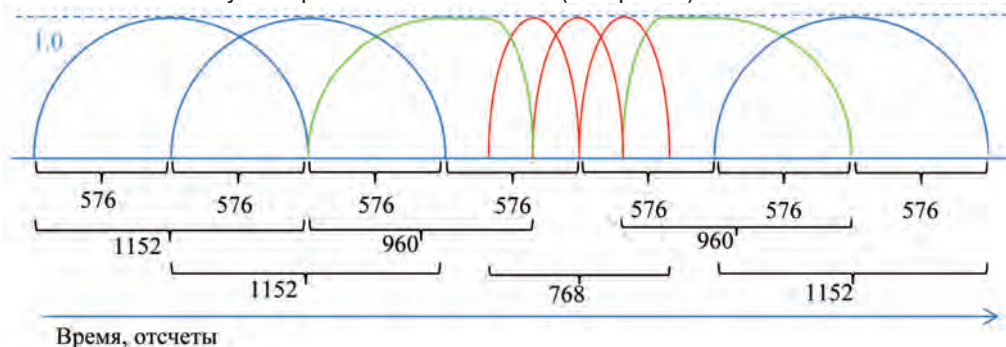


Рис. 2. Схематичное изображение оконных функций при MP3-кодировании

Если для фрагмента сигнала, прошедшего MP3-кодирование, вычислить MP3-спектр в тех же границах окна, как и при MP3-кодировании, то на полученном спектре будут отчетливо видны следы квантования [5,6] (см. рис. 3, изображение «б»).

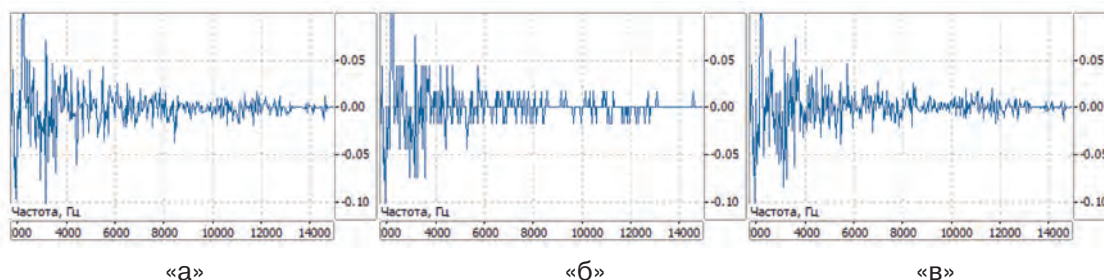


Рис. 3. MP3-спектры трёх фрагментов сигнала

Спектр на Рис. 3 «б» получен для фрагмента, границы которого совпадают с границами окна MP3-кодирования. На изображениях «а» и «в» представлены MP3-спектры фрагментов, сдвинутых на один отсчёт «вправо» и «влево» относительно фрагмента на изображении «б». Данные рисунки демонстрируют, что следы квантования в спектрах обнаруживаются лишь для фрагментов сигнала, границы которых в точности соответствуют окнам MP3-кодирования.

Для автоматизации процесса обнаружения следов квантования можно использовать функцию, позволяющую вычислить количество нулевых и близких к нулю значений в MP3-спектре. Если вычислять данную функцию для сигнала, сдвигая начало окна анализа на каждом шаге на один отсчёт, то получится график, представленный на рис. 4.

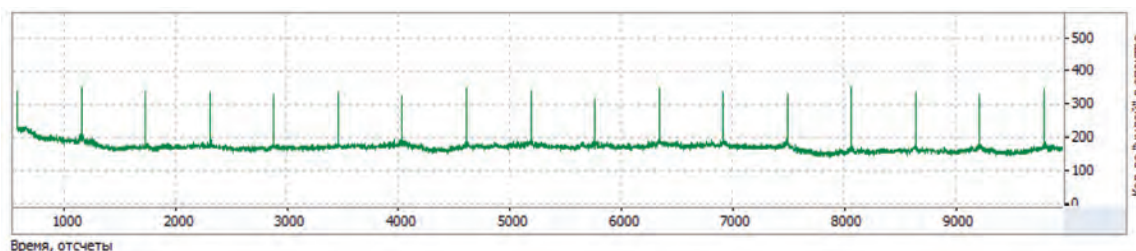


Рис.4. График количества нулей в MP3-спектре

На графике, изображенном на рис. 4, видны периодически повторяющиеся локальные максимумы, соответствующие местам совпадения положения окон анализа и кодирования. Период появления данных максимумов составляет 576 отсчётов, что соответствует шагу окна MP3-кодирования. Таким образом, MP3-кодек в процессе кодирования/декодирования оставляет в сигнале след, который можно использовать при оценке аутентичности фонограмм. Например, на рис. 5 приводится график количества нулей в MP3-спектре участка исследуемой фонограммы, на котором было произведено удаление фрагмента сигнала. Расположение точки нарушения непрерывности (около 7000 отсчетов) соответствует моменту нарушения периода следования максимумов.

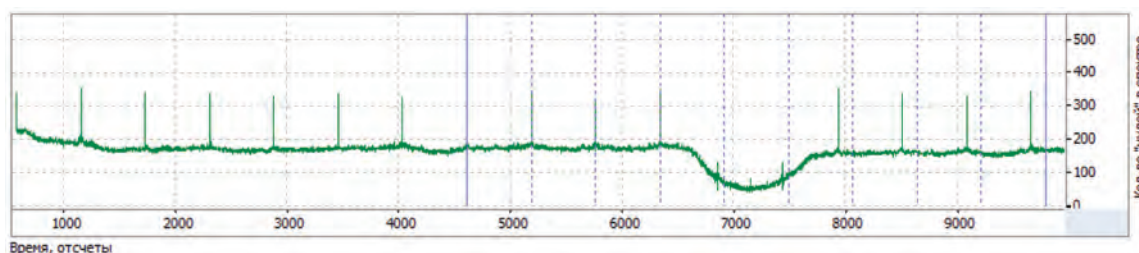


Рис. 5 График количества нулей в MP3-спектре участка фонограммы с нарушением непрерывности

Поиск признаков монтажа

Визуальный анализ графика количества нулей с целью обнаружения нарушения периодичности появления максимумов для фонограмм большой продолжительности является достаточно трудоёмкой задачей. В связи с этим для повышения эффективности работы целесообразно использовать график смещения границ окон кодирования. Для его построения график нулей МРЗ-спектра разбивается на участки длиной 576 отсчётов. Далее на каждом таком участке устанавливается место расположения максимума (от начала участка). В результате получается график, построенный с шагом 576 отсчётов. Каждый отсчёт данного графика принимает значение от 0 до 575 в соответствии с позицией максимума. В случае отсутствия нарушений периодичности на графике количества нулей в МРЗ-спектре график смещения будет иметь вид прямой линии (рис. 6).

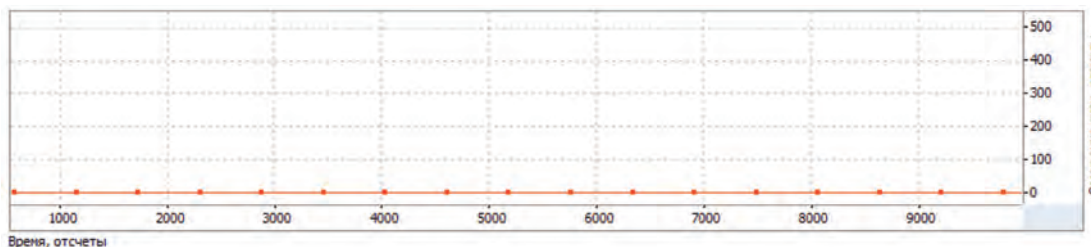


Рис.6. График смещения окон для сигнала без нарушения непрерывности

Если на графике количества нулей в МРЗ-спектре будет однократное нарушение периодичности, то на графике смещения появится «ступенька» (рис. 7).

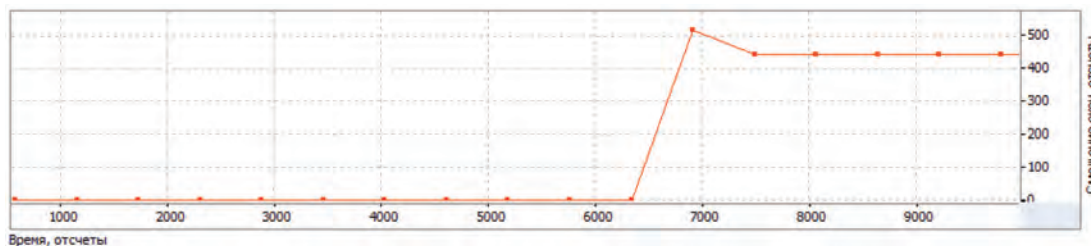


Рис.7. График смещения окон МРЗ-кодирования для сигнала с однократным нарушением непрерывности

Используя график смещения окон, можно построить гистограмму (рис. 8), по вертикальной оси которой откладывается величина смещения окна, а по горизонтали – процент встречаемости значения от общего количества. Гистограмма смещения окон кодирования для сигнала, не имеющего нарушений периодичности границ окон кодирования, будет содержать один ярко выраженный локальный максимум (рис. 8, изображение «а»). При этом относительная частота встречаемости данного значения близка к 100%. Для сигнала с однократным нарушением периодичности границ окон кодирования гистограмма будет выглядеть как на изображении «б» на рис. 8, т.е. будет содержать два ярко выраженных локальных максимума. Один максимум будет соответствовать значению смещений окон кодирования до точки нарушения непрерывности, а другой – значению смещения после точки нарушения непрерывности.

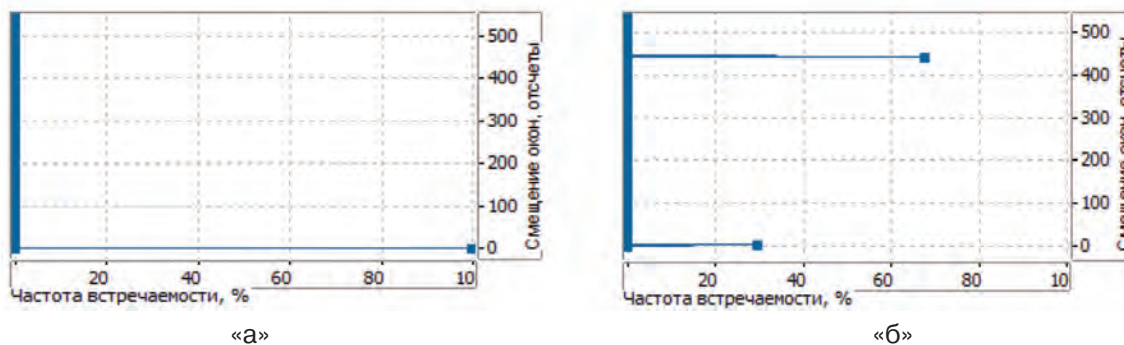


Рис. 8. Гистограммы смещения окон кодирования

Присутствие нескольких ярко выраженных локальных максимумов в гистограмме означает наличие на фонограмме следов кодирования с различным смещением окна. Если участки с различным смещением окна следуют друг за другом, как на Рис. 7, то это может свидетельствовать о наличии признаков монтажа фонограммы. Если участки с различным смещением перемешаны по времени (см. рис. 12), то это может свидетельствовать о перекодировании записи или микшировании сигналов, каждый из которых подвергнулся MP3-кодированию и имеет разное смещение MP3-окон от начала исследуемого сигнала.

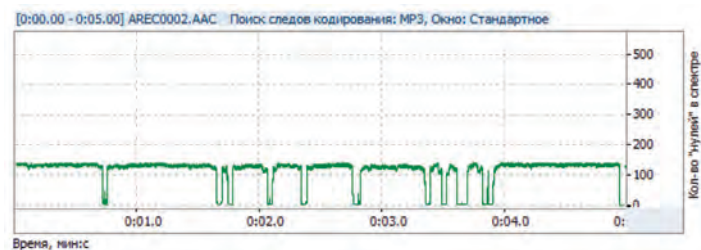
Поиск признаков кодирования

Следы кодирования сигнала могут быть обнаружены как на графиках количества нулей в MP3-спектре (рис. 9 «а»), так и на графиках смещения окон MP3-кодирования (рис. 10 «а»), или гистограммах смещения окон MP3-кодирования (рис. 11 «а»). Однако в случаях, когда следы кодирования проявляются недостаточно чётко и/или замаскированы вследствие модификации сигнала, проведенной после декодирования, наиболее надёжным способом обнаружения следов кодирования является анализ гистограмм смещения окон кодирования. В таком случае, если следы кодирования видны на гистограмме, то далее по графику смещения окон MP3-кодирования можно выделить участки фонограммы с наиболее ярко выраженными проявлениями следов кодирования (см. рис. 15).

Если на гистограмме смещения окон MP3-кодирования отображаются ярко выраженные максимумы, то это однозначно свидетельствует о наличии следов MP3-кодирования. Если сигнал не подвергался MP3-кодированию, то для него гистограмма смещения окон MP3-кодирования не будет иметь ярко выраженных локальных максимумов (рис. 11 «б»). На рис. 9-11 представлены графики, полученные для сигналов, подвергавшихся и не подвергавшихся MP3-кодированию.

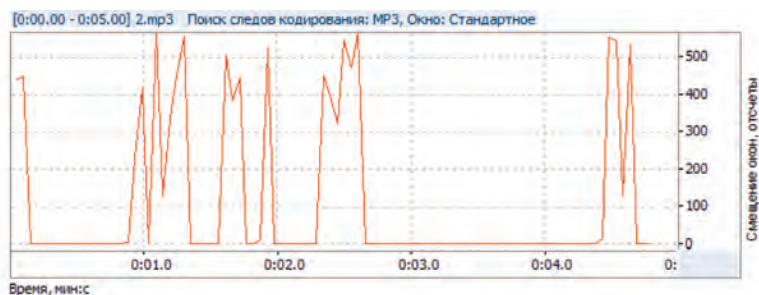


«а»



«б»

Рис. 9. Графики количества нулей в MP3-спектре для сигнала со следами MP3 кодирования («а») и сигнала без следов MP3-кодирования («б»)

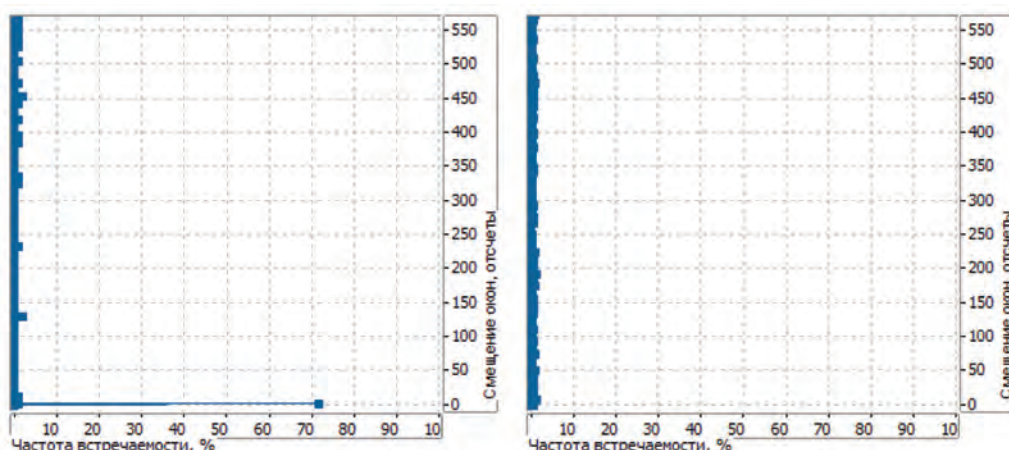


«а»



«б»

Рис. 10. Графики смещения окон MP3-кодирования для сигнала со следами MP3 кодирования («а») и сигнала без следов MP3-кодирования («б»)



«а»

«б»

Рис. 11. Гистограммы смещения окон MP3-кодирования для сигнала со следами MP3 кодирования («а») и сигнала без следов MP3-кодирования («б»)

Следы кодирования кодеками AAC, WMA и OGG

Кодеки AAC (Advance Audio Codec), WMA (Windows Media Audio) и OGG (Vorbis Audio) используют аналогичные MP3-кодеку принципы кодирования. Перечисленные кодеки объединяет то, что все они используют психоакустическую модель маскировки для выявления неслышимых компонентов сигнала и схожий подход спектрального анализа кодируемого сигнала. Сходство подхода на спектральном уровне состоит в том, что все эти кодеки используют МДКП для спектрального анализа, а шаг окна кодирования составляет 50% от его размера (для случая одинакового размера соседних окон). Различия методов спектрального анализа, используемого данными кодеками, состоит в том, что применяются окна кодирования различного размера, а также разные типы окон взвешивания сигнала. Так, например, MP3-кодек использует два стандартных размера окна анализа: стандартный размер – 1152 отсчета – и уменьшенный размер – 384 отсчета. AAC-кодек вне зависимости от частоты дискретизации сигнала использует окна размером 2048 и 256 отсчетов. А вот у кодеков WMA и OGG размер окна анализа зависит от частоты дискретизации и может принимать следующие значения: 64; 128; 256; 512; 1024; 2048; 4096; 8192; 16384.

Для поиска следов кодирования WMA-, AAC- и OGG-кодеками используется метод, сходный с методом поиска следов MP3-кодирования, – так же строится график количества нулей в спектре. Отличие состоит в том, что для каждого кодека применяются соответствующие размеры и типы окон.

Из вышесказанного следует, что следы MP3-кодирования являются уникальными, т.к. не существует других широко применяемых алгоритмов кодирования сигнала, использующих размер окна кодирования 1152 отсчета. Таким образом, если на графике нулей в MP3-спектре фрагмента сигнала обнаруживаются скачки с периодичностью 576

отсчётов, то можно с полной уверенностью говорить, что исследуемый фрагмент сигнала подвергался кодированию при помощи MP3-кодека. Не так однозначно обстоят дела с классификацией следов кодеков WMA, AAC и OGG. Размеры окон кодирования этих кодеков могут совпадать, в связи с чем на графиках нулей спектров, полученных для всех трёх алгоритмов, могут обнаруживаться периодические скачки.

В качестве примера рассмотрим следующий случай: в графике нулей МДКП-спектра, вычисленного с окном анализа 2048 отсчётов для сигнала с частотой дискретизации 22050 Гц, обнаруживаются периодические максимумы с шагом 1024 отсчёта. В подобной ситуации следы могут быть отнесены сразу к трём кодекам – WMA, AAC и OGG, т.к. все они могут кодировать сигнал с указанной частотой дискретизации с одним и тем же размером окна. В данном случае для установления конкретного типа кодека необходимо произвести описанную процедуру поиска следов кодирования с окнами других размеров. Если обнаружатся следы кодирования с размером окна, отличным от 256 и 2048, то, соответственно, можно исключить из рассмотрения кодек AAC. Далее можно провести визуальный анализ получаемых при анализе спектров для всех типов кодеков и сравнить значения максимумов для графиков количества нулей в спектрах.

Таким образом, при декодировании кодеки MP3, AAC, WMA и OGG оставляют в сигнале следы, отличающиеся периодичностью повторения, которые могут быть выявлены при исследовании сигнала. Наличие таких следов в сигнале однозначно говорит о том, что при формировании сигнала применялся один из вышеперечисленных кодеков, а после декодирования сигнал не подвергался изменению частоты дискретизации. Анализ данных следов позволяет обнаружить точки монтажа, произведённого после декодирования, а в ряде случаев и нарушения непрерывности записи сигнала.

Поиск следов повторного кодирования

На практике встречаются случаи, когда в гистограмме смещения окон кодирования наблюдается несколько ярко выраженных локальных максимумов, однако фрагменты графика смещения окон с различными значениями чередуются во времени (не обособлены во времени).

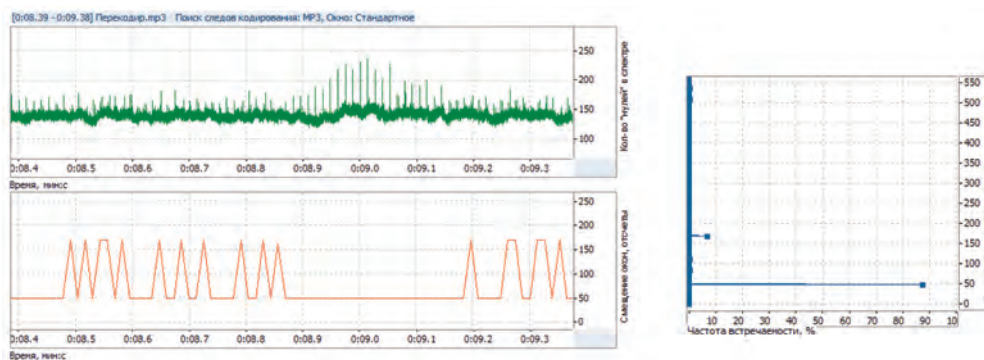


Рис. 12. График количества нулей в MP3-спектре, график смещения окон MP3 кодирования и гистограмма смещения окон для сигнала, дважды прошедшего MP3-кодирование

Результаты анализа подобного сигнала представлены на рис. 12. На графике смещения окон видны чередующиеся фрагменты сигнала со следами MP3-кодирования с различным смещением. Данное явление обусловлено тем, что сигнал дважды подвергался MP3-кодированию.

Следы первого и второго кодирования имеют одинаковый период – 576 отсчётов, но разное смещение от начала фонограммы. На разных фрагментах сигнала сильнее проявляются следы, оставленные кодеком как при первом кодировании (смещение 169 отсчётов), так и при втором (смещение 48 отсчётов). При этом степень проявления следов кодирования не связана с порядком кодирования.

На основе экспериментальных данных установлено, что признаки повторного MP3-кодирования можно обнаружить в тех случаях, когда повторное кодирование произ-

водилось с большей информационной скоростью (битрейтом), чем предыдущее. И чем выше значение битрейта повторного MP3-кодирования относительно значений первого, тем отчётливее видны следы первого.

Следует отметить, что в экспертной практике встречаются фонограммы, на которых обнаруживаются следы различных кодеков. Например, динамическая спектрограмма сигнала такой фонограммы представлена на рис. 13.

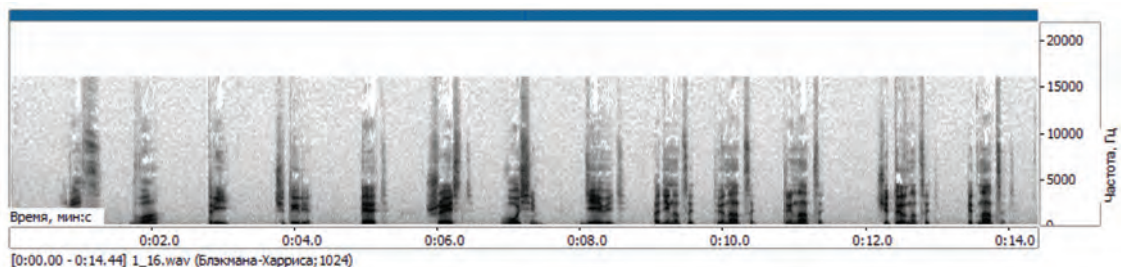


Рис. 13. Динамическая спектрограмма сигнала

На графике смещения окон MP3-кодирования сигнала данной фонограммы (рис. 14 «а») видны следы MP3-кодирования и ступенчатое изменение графика в момент времени 8.8 с.

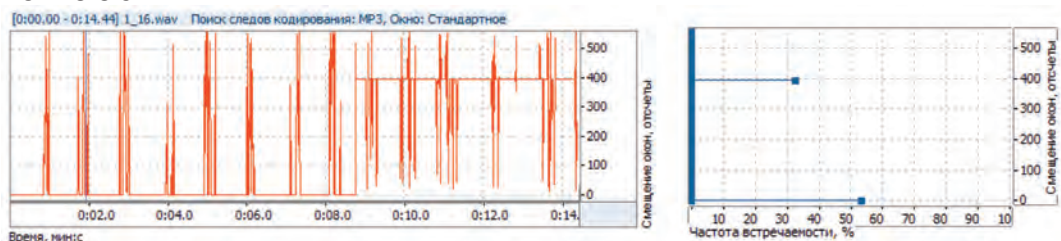


Рис. 14. График смещения окон MP3-кодирования («а») и гистограмма смещения окон MP3-кодирования («б»)

На графике смещения окон AAC-кодирования исследуемого сигнала (рис. 15 «а») видны следы AAC-кодирования и два скачкообразных изменения графика в моменты времени 6.7 с и 8.8 с.

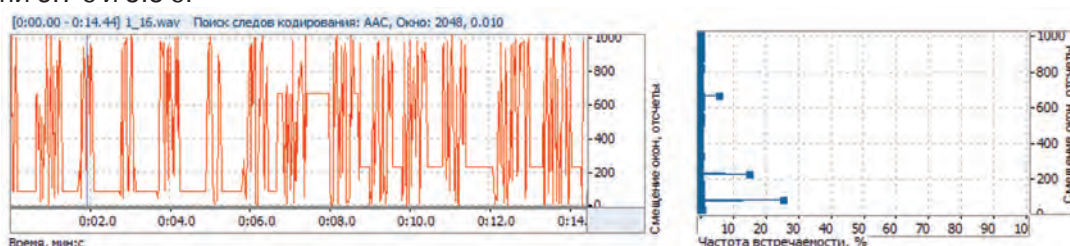


Рис. 15. График смещения окон AAC-кодирования («а») и гистограмм смещения окон AAC-кодирования («б»)

Обобщая результаты поиска следов кодирования кодеками MP3 и AAC, можно сделать вывод, что фонограмма подвергалась кодированию и тем и другим кодеком. Причём первоначально фонограмма была закодирована кодеком AAC. После её декодирования был произведён монтаж, следы которого видны на временной отметке 6.7 с. После этого фонограмма была закодирована MP3-кодеком. Далее после декодирования фонограмма была ещё раз смонтирована, о чём свидетельствуют признаки нарушения непрерывности на временной отметке 8.8. В завершении фонограмма была сохранена в формат PCM и представлена на исследование.

Таким образом, используя описанный в статье метод обнаружения и анализа следов кодирования фонограмм, удалось установить процедуры, которым подвергались фонограммы, и последовательность их применения.

Определение исходной частоты дискретизации сигнала

На исследование зачастую поступают фонограммы, частота дискретизации которых была изменена после завершения процесса записи. Подобное изменение может быть следствием как умышленного вмешательства, например, с целью маскировки следов модификации сигнала, так и стандартных процедур экспорта фонограмм из памяти устройства звукозаписи на внешний носитель. Например, в результате сохранения MP3-фонограммы, имеющей частоту дискретизации 48 кГц, на оптическом диске формата Audio-CD сигнал будет преобразован в формат Microsoft PCM 16 бит, 44.1 кГц. И на такой фонограмме, извлеченной из Audio-CD, уже не будут обнаруживаться следы MP3-кодирования. Однако в результате проведенных экспериментов оказалось, что если данную фонограмму извлечь из Audio-CD и пересохранить с исходной частотой дискретизации 48 кГц, то на полученной фонограмме будут обнаруживаться следы MP3-кодирования.

Согласно спецификациям формата MP3 [2], частота дискретизации MP3-фонограммы задается индексом в заголовке MP3-фреймов и может принимать одно из следующих значений (в кГц): 8, 11.025, 12, 16, 22.050, 24, 32, 44.1, 48. С учётом этого фонограмму с любой частотой дискретизации можно проверить на предмет использования в её отношении MP3-кодирования с последующим изменением частоты дискретизации. Для проведения подобной проверки исследуемая фонограмма сохраняется с каждой из указанных 9 частот дискретизации, и полученные таким образом 9 фонограмм проверяются на наличие следов MP3-кодирования. Результатом такой проверки является не только установление факта кодирования фонограммы, но и установление частоты дискретизации сигнала, на которой проводилось MP3-кодирование.

При обнаружении следов кодирования кодеками WMA, AAC и OGG все действия будут аналогичны. Единственным отличием будет лишь более широкий перечень возможных частот дискретизации.

Рекомендуемое программное обеспечение

В настоящее время описанные в данной работе методы исследования в полном объёме реализованы в программном обеспечении OTExpert (начиная с версии 6.0; часть функционала была реализована в версии 5.1). Все представленные в работе результаты и иллюстрации получены с помощью указанного ПО.

Заключение

Описанный метод обнаружения и анализа следов кодирования сигнала позволяет эксперту получить важную для установления аутентичности фонограмм информацию, которая зачастую не может быть получена с использованием других методов анализа. В связи с этим описанный метод является очень важным функциональным дополнением к существующему инструментарию экспертов, позволяющим обеспечить полноту и всесторонность исследования, а также соответствие применяемых методов современному уровню развития средств записи и редактирования звуковых сигналов.

Литература

1. К.Гласман. Формат MP3. Журнал «Звукорежиссер», №№ 2, 3, 4, 5, 7, 8 2005.
2. Особенности технического исследования фонограмм формата MP3 (методические рекомендации для экспертов) / Бояров А.Г. - Теория и практика судебной экспертизы. - М.: ФБУ РФЦСЭ при Минюсте России. № 3 (31) 2013. - С. 70-83.
3. ISO/IEC 11172-3:1993 Information technology -- Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s -- Part 3: Audio.
4. ISO/IEC 13818-3:1998 Information technology -- Generic coding of moving pictures and associated audio information -- Part 3: Audio.
5. J. Herre, S. Moehrs, and R. Geiger: Analysing decompressed audio with the "Inverse Decoder" - towards an operative algorithm. 112th AES Convention, Munich, May 10-13, (2002).
6. Rui Yang, Zhenhua Qu, Jiwu Huang. Detecting Digital Audio Forgeries by Checking Frame Offsets. in Proceedings of the 10th ACM Workshop on Multimedia and Security, ed. by. MM&Sec'08 (ACM, New York, 2008), pp. 21-26.