

СИСТЕМИ ТА МЕТОДИ ОБРОБКИ ІНФОРМАЦІЇ

УДК 000.684

В.И. Соловьев,

кандидат технических наук, доцент

ВЫЯВЛЕНИЕ СЛЕДОВ ЦИФРОВОЙ ОБРАБОТКИ СИГНАЛА В ЗАДАЧАХ МОНТАЖА АУДИОЗАПИСИ

Приведены результаты исследований задачи монтажа аудиофайлов. Рассматриваемое направление исследований базируется на комплексном подходе к выявлению и локализации монтажа. Подход использует статистические характеристики различных участков пауз между словами.

Ключевые слова: монтаж аудиозаписи, выявление и локализация монтажа, звуковой файл.

Наведено результати досліджень завдання монтажу аудіофайлів. Розглянутий напрям досліджень базується на комплексному підході до виявлення і локалізації монтажу. Підхід використовує статистичні характеристики різних ділянок пауз між словами.

Ключові слова: монтаж аудіозапису, виявлення та локалізація монтажу, звуковий файл.

The research of a problem of audiophiles' installation is resulted. Considered direction of researches is based on the complex approach to revealing and localisation of an assembly. An approach uses statistical characteristics of various sites of pauses between words.

Keywords: audio record installation, revealing and localization of installation, sound file.

В последнее десятилетие существенное внимание уделяется вопросам разработки методологии и средств автоматизации решения задачи выявления монтажа аудиозаписи [1]. Предложен ряд подходов. Однако пока что, несмотря на большое количество исследований, неясно, в какой степени разрешима эта задача и при каких условиях возможно обнаружение и локализация места возможного монтажа аудиосигнала.

Рассматривается следующая постановка задачи. В звуковом файле с записью речевого сигнала осуществлена вставка фрагментов записи из другого звукового файла или осуществлена перестановка фрагментов записи из различных частей одного и того же файла. Имеется в виду запись речи одного и того же человека. Необходимо цифровыми методами обработки звукового сигнала локализовать вероятное место монтажа.

Ниже проводится анализ эффективности различных подходов и излагаются результаты исследования возможностей нового направления.

Задачи выявления монтажа аудиозаписи, с точки зрения возможности эффективного применения методов спектрального анализа, имеют весьма важную специфику, существенно усложняющую решение. Монтаж, осуществляемый с

использованием современных программных продуктов (например, программная система CoolPro), реализуется на весьма малом временном интервале. Этот интервал зависит от частоты дискретизации файла и, в зависимости от наличия цифровой обработки сигнала, может лежать в диапазоне от 0.0001 до 0.001 секунды для частоты дискретизации 44,1 кГц и 0.05–0.1 секунды для частоты – 8 кГц. Важно, что по порядку величины этот интервал соизмерим с $1/F$, где F – частота дискретизации.

На рис. 1 и рис. 2 приведены характерные примеры практически “необработанного” монтажа двух фрагментов звукового файла.

Если в первом случае (на рис. 1) визуально видно изменение характеристик аудиосигнала, то во втором (рис. 2) – практически никаких визуальных отличий нет. Во втором случае были специально подобраны фрагменты речи с близкими визуальными характеристиками пауз. При прослушивании второго фрагмента наличие монтажа проявляется сразу, ввиду логического несоответствия последовательности слов речи. Однако наша задача состоит в разработке автоматизированных средств локализации мест вероятного монтажа. Поэтому эта иллюстрация чрезвычайно показательна. Простая вставка полностью визуально идентичных фрагментов не имеет никаких явных визуальных признаков монтажа. Применение различных тонких методов спектрального анализа для выявления следов монтажа, которые неразличимы визуально, правомочно, и в ряде конкретных случаев, возможно, дает результат. Однако человеческий глаз, как инструмент визуального восприятия и анализа, обладает весьма высокими характеристиками по выявлению “аномалий”, которые могут быть в месте локализации монтажа.

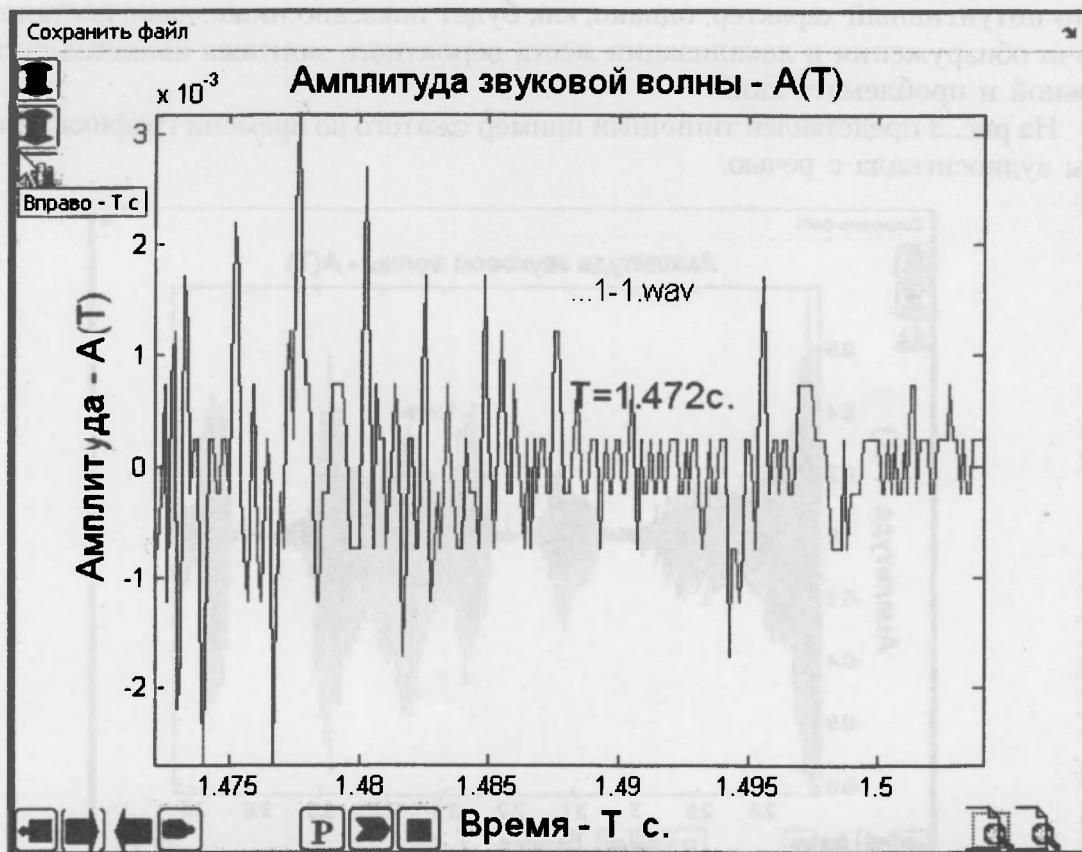


Рис. 1. Иллюстрация монтажа в паузе (место монтажа – $T=1.487$ с)

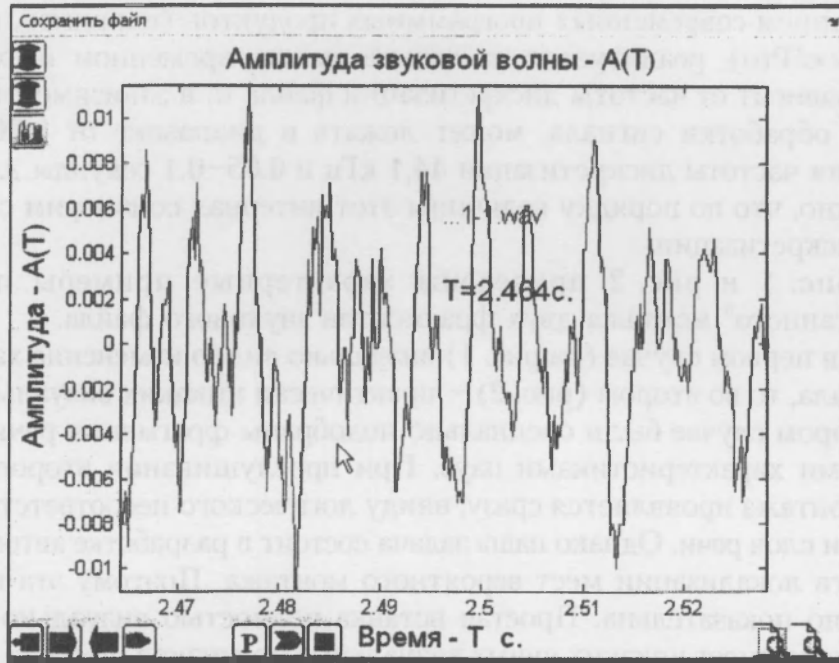


Рис. 2. Иллюстрация монтажа в паузе (Место монтажа – $T=2.482$ с)

Маловероятно, что при отсутствии явно выраженных визуальных “аномалий” в месте монтажа тонкие спектральные (или иные) методы статистического анализа выявят вероятное наличие монтажа аудиозаписи. Эти замечания носят чисто интуитивный характер, однако, как будет показано ниже, даже постановка задачи обнаружения и локализации места вероятного монтажа является весьма сложной и проблематичной.

На рис. 3 представлен типичный пример сжатого во времени графика амплитуды аудиосигнала с речью.

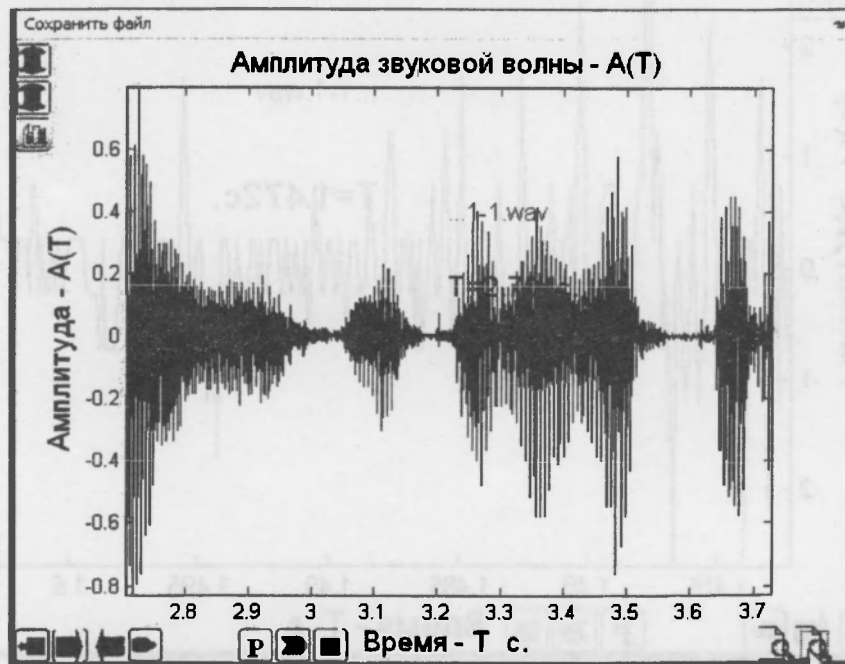


Рис. 3. График амплитуды речевого сигнала

На рис. 3 представлений типичний характер різних пауз. Аналіз подібних сигналів показує, що в більшості випадків характеристики сигналу в паузах візуально суттєво відрізняються, принаймні по одному параметру – середній амплітуді сигналу в паузах. Фізично це обумовлено достатньою варіабельністю шумів фону і в ряді випадків нестабільністю характеристик апаратури запису аудіосигналу на інтервалах часу декількох речевих фрагментів. На інтервалах часу однієї паузи ці характеристики в цілому достатньо стабільні в межах однієї паузи. Дані постулати несуть чисто евристичний характер на постановочній частині задачі дослідження і є базовими гіпотезами, які далі підтверджуються в процесі дослідження.

Далі викладаються результати досліджень по локалізації передбачуваного монтажу аудіозапису, засновані на порівнянні статистичних характеристик двох рівних частин паузи. Досліджуваний аудіофайл розбивається на сегменти, що складаються з пауз мови. Алгоритм виділення пауз дуже суттєвий при подальших статистичних оцінках. Цей алгоритм базується на підході, заснованому на виділенні речевих фрагментів за проявами основного тону голосу людини. Метод базується на спеціальних дослідженнях, викладених в результаті яких є предметом окремої статті. Необхідно зазначити, що точність розділення меж мови і пауз в аудіофайлах має сильне вплив на статистичні характеристики пауз при будь-якому методі дослідження.

Після розбиття аудіофайлу на фрагменти, що складаються з пауз, здійснюється розбиття кожної паузи на дві рівні за часом частини. На кожній з частин розраховуються дві статистичні характеристики.

Перша – середнє значення абсолютної величини амплітуди звукової хвилі (і дисперсія цієї величини).

Вторая – середнє значення характеристики деформації скейлограми на основі вейвлет-преобразовання з комплексним базисом Морле і дисперсія цієї величини [2, 3, 4].

Після розрахунку вказаних характеристик проводиться порівняння середніх значень відповідних характеристик для двох половин кожної паузи з метою перевірки статистичної гіпотези про рівність середніх значень за критерієм Стюдента [2].

Відповідно, автоматично приймається рішення про наявність можливого монтажу аудіозапису, при нерівності середніх, хоча б по одній з характеристик.

Така загальна концепція підходу, викладеного далі.

Нижче розглядається послідовність оцінок по кожній з характеристик. Перша характеристика, як показують дослідження, чутлива до змін амплітудних характеристик фонових сигналів при переході з паузи в паузу, відповідно, при монтажі аудіосигналу з різних частин пауз, можна виявити ці відмінності.

Вторая характеристика, яку буде розглянуто детально, виявляє можливі спектральні зміни при монтажі в паузах.

Ці дві характеристики охоплюють більшість варіантів зміни характеристик сигналу при монтажі. Винятком є випадки, коли внаслідок монтажу характеристики є практично незмінними, але можлива

какая-либо локальная аномалия в точке монтажа. Однако, на наш взгляд, подобная аномалия, при условии ее выявления, вряд ли может свидетельствовать о наличии монтажа. Среди этих исключений есть один практически важный вариант, который также не может быть учтен развиваемым направлением. В частности, часто при записи аудиосигнала на некачественной аппаратуре аудиозаписи, например, на ноутбуке, в шумах записи четко проявляются наводки от сети на частоте 50 Гц (и частотах, кратных этой частоте). При реализации монтажа и состыковке различных фрагментов из разных пауз визуально может просматриваться фазовая (точечная) аномалия в месте монтажа. Однако, несмотря на достаточно простое визуальное выявление этой аномалии, ее локализация регулярными математическими методами весьма проблематична.

Здесь и далее рассматривается направление выявления предполагаемого монтажа аудиозаписи, которое базируется на статистически достоверных оценках различных параметров аудиосигнала. Точечные аномалии, в принципе, также выявляемы, например, методами вейвлет-анализа, однако их использование при решении статистических по своей сути задач весьма проблематично.

Необходимо также отметить, что проведение рассматриваемых исследований за приемлемое время в принципе невозможно без использования специализированных программных средств автоматизации фоновскопических исследований. Излагаемые результаты, получены с применением системы автоматической идентификации аппаратуры аудиозаписи – “Фрактал” [5].

Рассмотрим детально вначале вторую характеристику – среднее значение характеристики деформации скейлограммы на основе вейвлет-преобразования с комплексным базисом Морле .

Преобразование Морле осуществлялось на основе комплексного вейвлета Морле [3]:

$$C_{mor}(t) = (\pi * F_b)^{-0.5} * \exp(2 * j * \pi * F_c * t) * \exp(-(t^2) / F_b), \quad (1)$$

где

$\pi = 3.14\dots$,

j – комплексная единица,

F_b – параметр ширины вейвлета,

F_c – центральная частота вейвлета.

Выбранная половина паузы делилась на несколько равных по времени отрезков. Для каждого из отрезков – временных интервалов выбранной половины паузы, вычислялись дискретные свертки вейвлета Морле с аудиосигналом и в каждой временной точке рассчитывался модуль амплитуды коэффициентов преобразования при фиксированном параметре F_b , с дискретным варьированием центральной частоты в диапазоне от 200 Гц до 4000 Гц для файлов с частотой дискретизации – 8 кГц, и от 100 Гц до 20000 Гц для файлов с частотой дискретизации – 40 кГц (и выше), с различным шагом дискретизации по частоте. Число разных частот, для которых выполнялись преобразования, равнялось 50. Выбор этой величины обусловлен рядом причин. С одной стороны – высокой вычислительной сложностью, с другой – достаточным для практических целей приближением при построении скейлограммы.

После вычисления модуля амплитуды вейвлет-преобразования с базисом Морле вычислялась “характеристика деформации” скейлограммы:

$$D_m = \sum_{F_c} (A(F_c) * F_c) / (\sum_{F_c} (A(F_c))), \quad (2)$$

где D_m – количественная характеристика деформации скейлограммы;

$A(F_c)$ – модуль амплитуды вейвлет-преобразования для частоты F_c ;

F_c – частота.

Суммирование осуществляется для всего набора частот. Предлагаемая характеристика фактически определяет “геометрический центр тяжести скейлограммы” в определенной временной точке по частоте. При смещении частотных характеристик в ту или иную сторону, соответственно изменяется эта характеристика.

После указанных расчетов в каждой временной точке подинтервалов паузы вычислялось среднее значение этой характеристики и ее дисперсия для каждой половины паузы – D_m , G . Эти величины для каждой половины паузы позволяют на основе, например, статистического критерия Стьюдента осуществить проверку гипотезы об идентичности спектральных характеристик сигнала в двух частях паузы.

На рис. 4 и рис. 5 представлены в качестве иллюстрации пространственные скейлограммы (вейвлет-спектрограммы) для двух участков паузы с наличием монтажа и отличными в двух половинах паузы деформациями спектра (рис. 4) и с отсутствием монтажа (рис. 5).

Данные скейлограммы построены для 8 масштабных факторов, которые на графике пронумерованы в порядке уменьшения центральной частоты вейвлета Морле. На иллюстрациях отчетливо видна разница в двух половинах паузы характеристик для варианта с монтажом (рис. 4). Необходимо еще раз подчеркнуть важный момент для рассматриваемого критерия – деформация скейлограммы. Этот критерий чувствителен по отношению к перераспределению энергии между частотами. Определение величин критерия деформации и его дисперсии для первого и второго варианта иллюстраций и проверка гипотезы равенства этих величин для двух половин каждой из пауз по критерию Стьюдента показывает, соответственно, отличие этих величин и их равенство для первого (рис. 4) и второго (рис. 5) графиков, с вероятностью ошибки второго и первого рода не более 0.025.

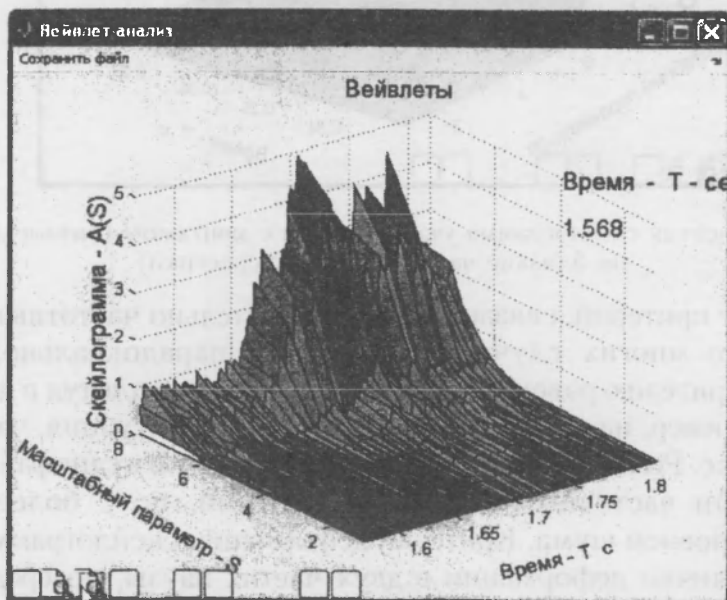


Рис. 4. Пространственная скейлограмма участка паузы с монтажом аудиозаписи

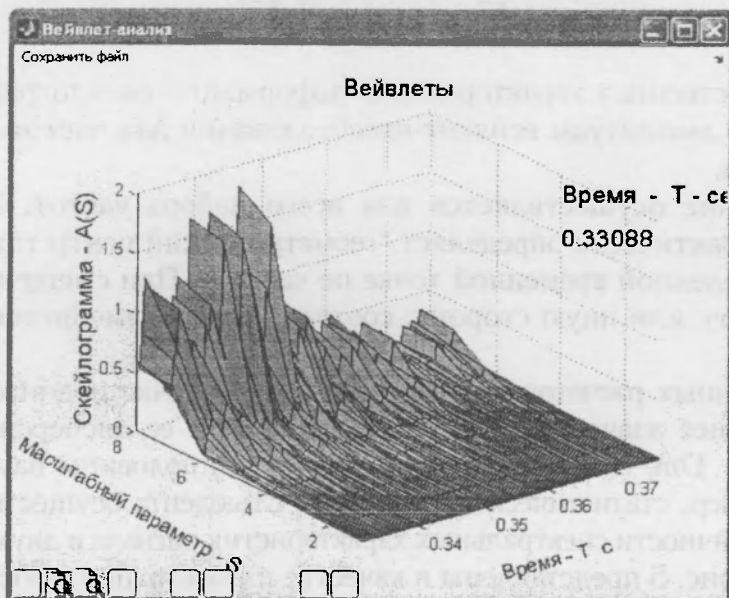


Рис. 5. Пространственная скейлограмма участка паузы без монтажа аудиозаписи

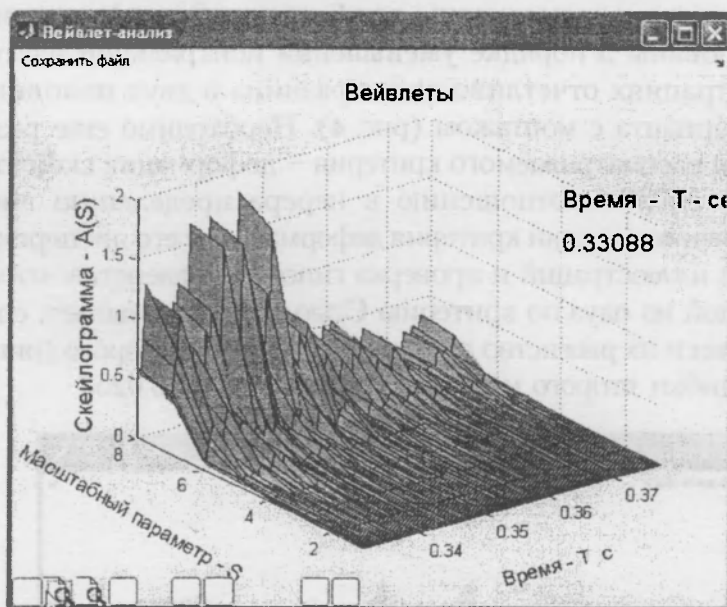


Рис. 6. Пространственная скейлограмма участка паузы с монтажом (разные дисперсии шума, но близкие частотные характеристики)

Однако этот критерий, связанный с исключительно частотными свойствами аудиосигнала, во многих случаях, как это ни парадоксально, уступает по эффективности критерию равенства средних значений амплитуд в двух половинах паузы. Так, например, на рис. 6 проиллюстрирована ситуация, часто встречающаяся на практике. Реализован монтаж из двух участков аудиофайла с практически идентичными частотными характеристиками, но с более высоким по интенсивности уровнем шума. Критерий деформации скейлограммы показывает идентичность величин деформации в двух частях паузы, но средние значения звуковой амплитуды в двух половинах паузы существенно отличаются. В этом случае, достаточно часто встречающемся на практике, необходимо применять

первый вариант критерия проверки наличия монтажа – среднее значение абсолютной величины амплитуды звуковой волны (и дисперсию этой величины).

На рис. 7 и рис. 8 в качестве иллюстраций этого подхода продемонстрирована работа “идентификатора монтажа” на участке паузы с монтажом аудиозаписи в рамках системы идентификации аппаратуры аудиозаписи – “Фрактал” [5].

На рис. 7 размещен график изменения средних значений абсолютного среднего значения амплитуд звуковой волны на участках паузы без монтажа. На рис. 8, соответственно, с монтажом.

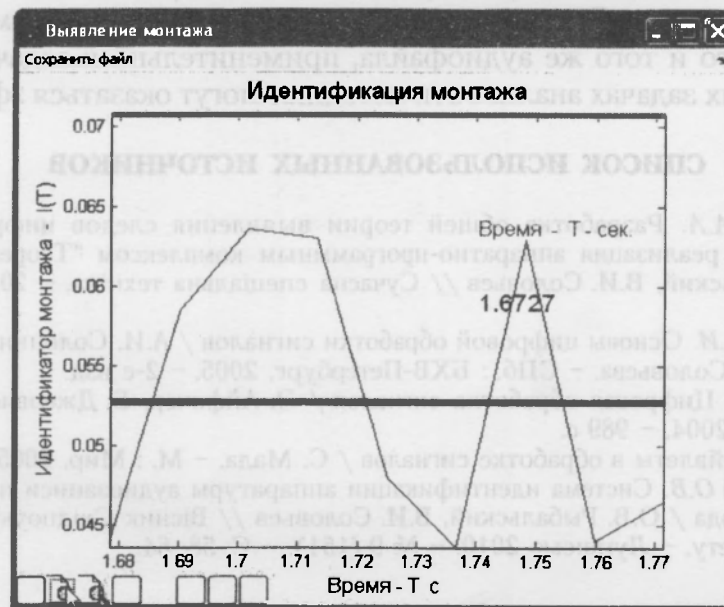


Рис. 7. Изменение среднего абсолютного значения амплитуды звуковой волны (пауза без монтажа)

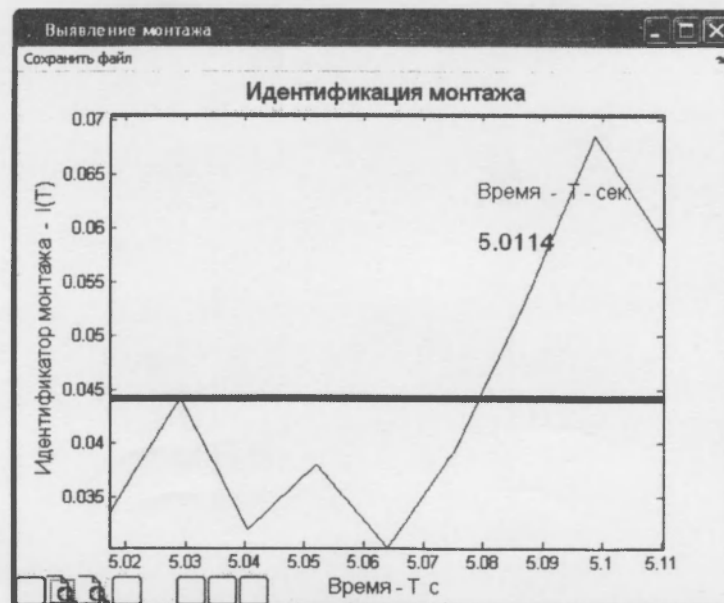


Рис. 8. Изменение среднего абсолютного значения амплитуды звуковой волны (пауза с монтажом)

Выводы:

1. Автоматическое выявление и локализация монтажа аудиофайла требуют комплексного подхода с анализом различного набора статистических характеристик пауз.
2. Необходим анализ изменения как частотных характеристик аудиосигнала в пределах паузы, так и чисто амплитудных изменений, которые могут не учитываться частотными характеристиками.
3. Применение методик выявления монтажа на основе сравнения статистических характеристик различных пауз, как косвенно указывает данное исследование, малоэффективно. Статистические характеристики пауз весьма вариабельны в пределах одного и того же аудиофайла, применительно к задачам выявления монтажа (в других задачах анализа эти методики могут оказаться эффективными).

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Кобозева А.А. Разработка общей теории выявления следов цифровой обработки сигналов и ее реализация аппаратно-программным комплексом "Теорема-М" // А.А. Кобозева, О.В. Рыбальский, В.И. Соловьев // Сучасна спеціальна техніка. – 2010. – № 1 (20). – С. 5–14.
2. Солонина А.И. Основы цифровой обработки сигналов / А.И. Солонина, Д.А. Улахович, С.М. Арбузов, Е.Б. Соловьева. – СПб. : БХВ-Петербург, 2005. – 2-е изд.
3. Айфичер Э. Цифровая обработка сигналов / Э. Айфичер, Б. Джервис. – М. – СПб. – К. : Радио и связь, 2004. – 989 с.
4. Мала С. Вейвлеты в обработке сигналов / С. Мала. – М. : Мир, 2005. – 671 с.
5. Рыбальский О.В. Система идентификации аппаратуры аудиозаписи на основе мультифрактального подхода / О.В. Рыбальский, В.И. Соловьев // Вісник Східноукраїнського національного університету. – Луганськ, 2010. – № 9 (151). – С. 58–64.

Отримано 20.09.2011

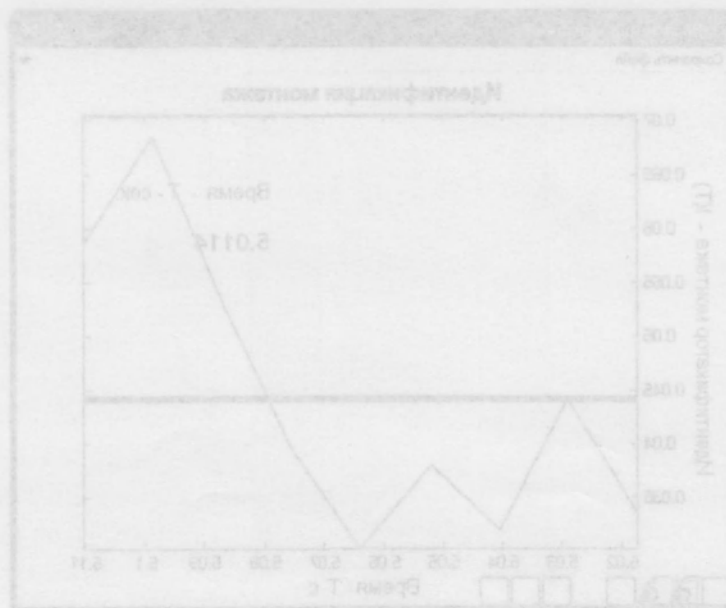


Рис. 7. Намірне зрештлення середнього значення амплітуди звукової форми (пауза без монтажу)