

---

УДК 534.4; 534.83; 004.4

**И. В. Косяк**, канд. техн. наук  
Ин-т проблем регистрации информации НАН Украины  
(Украина, 03113, Киев, ул. Шпака, 2,  
тел: (044) 454-21-85, E-mail: kivfox@cki.ipri.kiev.ua)

## **Устранение паразитной частотной модуляции в звуковом сигнале**

*(Статью представил д-р техн. наук В. П. Симоненко)*

Проведен анализ решения проблемы временных искажений сигнала-детонации в фонограммах аналоговых носителей записи. Рассмотрена возможность восстановления звукового сигнала, поврежденного паразитной частотной модуляцией. Предложен алгоритм реконструкции деградированного сигнала.

Проведено аналіз вирішення проблеми часових викривлень сигналу-детонації у фонограмах аналогових носіїв запису. Розглянуто можливість відновлення звукового сигналу, ушкодженого паразитною частотною модуляцією. Запропоновано алгоритм реконструкції деградованого сигналу.

*К л ю ч е в ы е с л о в а: детонация, преобразование частоты дискретизации, частотные треки, спектр сигнала, сонограмма.*

Сохранение и восстановление звуковых записей в настоящее время вызывает значительный интерес во всем мире. В механической звукозаписи, особенно старой, невозможно получить постоянную скорость носителя записи вследствие ограниченной точности механического привода. В результате неравномерного движения носителя записи возникают различные формы искажений, например паразитная частотная модуляция, которую особенно трудно устранить.

Паразитная частотная модуляция в воспроизведенном сигнале, или детонация, это форма деградации сигнала, которая обычно наблюдается в фонограммах аналоговых носителей записи — фоноцилиндрах, грампластинках, магнитных лентах, киноплёнках. Такие искажения сигнала на слух воспринимаются как плавание и дробление звука или как изменения высоты тона.

Есть несколько закономерных причин появления детонации. Одна из них — изменение частоты вращения носителя записи на протяжении процесса регистрации, другая — эксцентриситет деталей механизмов и са-

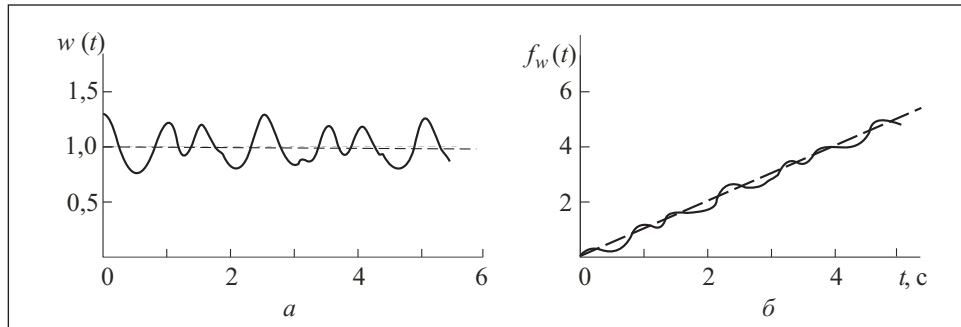


Рис. 1. Кривые колебаний скорости носителя записи

мого носителя записи и, наконец, деформация в течение времени относительно мягкого материала фоноцилиндра.

Следует заметить, что еще не созданы коммерческие программные продукты по восстановлению дефекта детонации, наблюдаемого в звуковом сигнале.

Некоторые результаты исследований по устранению детонации и несколько алгоритмов описаны в работах [1—9]. Однако эта проблема не изучена всесторонне и не найдено единого универсального решения данной задачи. В приведенных примерах по восстановлению сигналов рассмотрены ограниченные характерные проявления детонации, такие как ее величина и продолжительность во времени. Искаженные спектры первоначальных звуков и, как следствие, низкое качество звучания ухудшают восприятие записей, так как человеческое ухо слишком чувствительно к колебаниям высоты тона.

Таким образом, проблема устранения паразитной частотной модуляции в звуковом сигнале остается открытой и ее решение является важной и актуальной задачей, в частности при восстановлении архивных звукозаписей.

Рассмотрим принципы детектирования и компенсации детонации и исследуем возможности восстановления звукового сигнала, поврежденного паразитной частотной модуляцией.

**Восстановление деградированного сигнала.** Детонация — это частный случай временных искажений сигнала. Физический механизм проявления эффекта детонации эквивалентен неравномерному изменению масштаба оси времени. На рис. 1, *a* представлен типичный пример колебаний скорости носителя записи  $w(t)$ . Обозначив неискаженную форму волны сигнала во времени функцией  $s(t)$ , а деформированную ось времени — монотонно возрастающей функцией  $f_w(t)$ , искаженный сигнал запишем в виде

$x(t) = s(f_w(t))$ . Если функция изменения масштаба времени  $f_w(t)$  известна (рис. 1, б), то возможно восстановление неискаженной формы волны сигнала:  $s(t) = x(f_w^{-1}(t))$ .

Таким образом, устранение детонации соответствует нахождению функции  $f_w(t)$ , что эквивалентно функции изменения высоты тона, т.е. функции детонации

$$p_w(t) = \frac{d(f_w(t))}{dt}.$$

В дискретной области получаем отсчеты  $x[n] = x(nT)$ , где  $T$  — период дискретизации. Если функция изменения высоты тона  $p_w[n]$ , соответствующая каждому моменту выборки, известна, то возможно нахождение неискаженного сигнала с применением операции передискретизации.

Существуют различные подходы к решению проблемы временных искажений сигнала-детонации, в частности нахождение функции изменения масштаба времени  $f_w(t)$ . Известна техника записи с пилотом-тоном (контрольным сигналом) на магнитных лентах [8], на киноплёнках [9]. В этом случае дополнительный пилот-тон записывается одновременно со звуковым сигналом отдельной дорожкой, а затем для восстановления изменения скорости записи используется демодуляция. Однако в подавляющем большинстве случаев архивные записи были выполнены без пилота-тона с отдельно записанной дорожкой.

Предлагается также использовать в качестве пилота-тона сигнал высокочастотного подмагничивания частотой от 40 до 435 кГц, используемый при записи магнитных лент [6, 7]. Недостатком такого подхода является то обстоятельство, что такой высокочастотный сигнал очень трудно записать на аппаратуре магнитной записи звукового сигнала. Поэтому существует ограниченное число записей архивных звуковых фонограмм с таким уровнем записи.

В работе [7] использован тот факт, что в каждой записи в той или иной мере присутствует низкочастотная помеха сети частотой 50 Гц и ее можно рассматривать как пилот-тон. Однако запись при этом должна быть электрической.

Как известно, первые записи звукового сигнала были выполнены акустическим способом на фоноцилиндрах (восковых валиках Эдисона) и грампластинках. Именно они имеют наибольшую степень повреждения паразитной частотной модуляцией. Поэтому к таким носителям записи предложенные алгоритмы невозможно применить.

В работах [1—4] использован более общий подход для выявления и исправления дефекта детонации в записях. В них применен алгоритм, основанный на технике анализа (синтеза) синусоидальной модели, извест-

ной как метод МакАулау и Кватиери [10]. В этом случае проведение оценки изменения высоты тона в деградированных данных сигнала возможно при наличии хороших статистических моделей для неискаженного звукового сигнала и процесса, генерирующего детонацию [3]. При этом предполагается, что музыкальный сигнал представлен как совокупность сочетания основных тонов и их гармоник всех воспроизведенных музыкальных нот, а процесс детонации моделируется как плавно изменяемая форма волны без крутых разрывов. Оценка проводится в частотной области.

Сначала находят тональные составляющие, взяв за основу алгоритм дискретного преобразования Фурье, с помощью которого определяют пиковую амплитуду сигнала. Затем извлекают информацию о девиации частоты (изменении тональности) из частотно-временных треков, подобных изображенным на сонограммах. Основная проблема заключается в том, что информация об изменении частоты может содержать значительную ошибку, вследствие предположения об изменении тональности (созданного исполнителем) как вибрато или тремоло. В основе решения этой проблемы лежит априорное знание о том, что типичный музыкальный сигнал не имеет импульсных изменений высоты тона.

**Алгоритм устранения паразитной частотной модуляции.** Задача устранения паразитной частотной модуляции в звуковом сигнале решается в два этапа: нахождение функции изменения высоты тона и компенсация изменений высоты тона с помощью операции передискретизации.

Весь процесс по восстановлению звукового сигнала, поврежденного паразитной частотной модуляцией, можно представить в виде блок-схемы, изображенной на рис. 2, из которой видно, что процедура детектирования детонации, т.е. нахождение функции изменения высоты тона  $\tilde{p}_w(t)$ , состоит из четырех шагов.

На первом шаге применяем разложение сигнала на спектральные составляющие с помощью кратковременного преобразования Фурье [11]:

$$X(n) = \sum_{k=0}^{N-1} x(k) w(k) e^{-j \frac{2\pi nk}{N}},$$

где  $x(k)$  — последовательность входных отсчетов;  $n$  — индекс преобразования в частотной области;  $N$  — длина весового окна;  $w(k)$  — размер весового окна.

Длину весового окна выбираем достаточно малой, чтобы частотные компоненты в отдельном блоке данных были почти одинаковыми. Типичные значения длины окна при звуковом сигнале частотой дискретизации 44,1 кГц составляют 512, 1024 и 2048 отсчетов, при частоте дискретизации

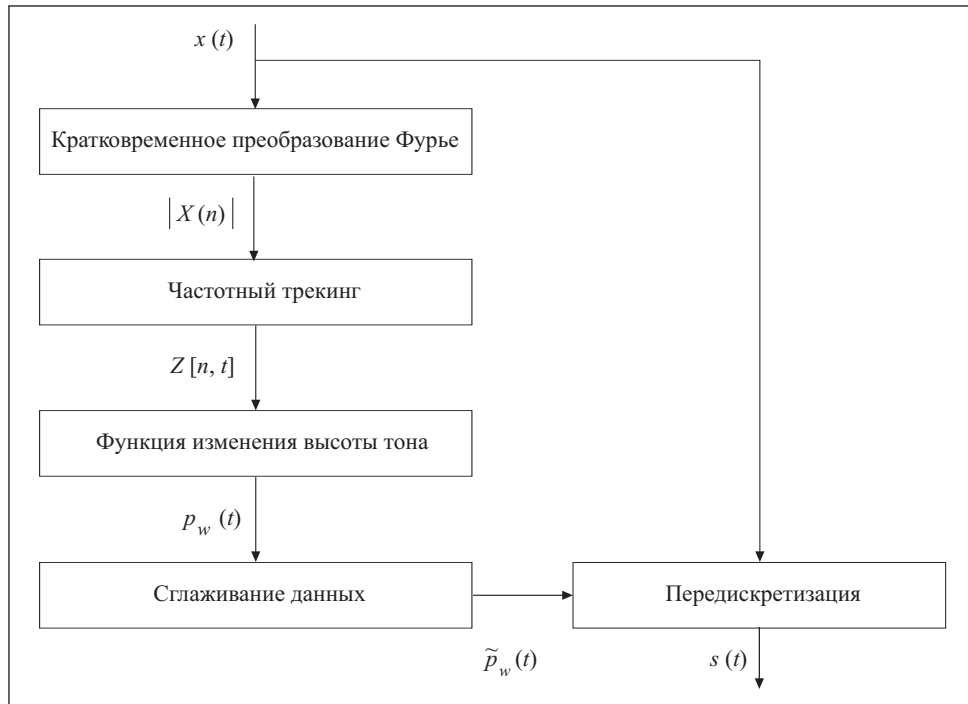


Рис. 2. Блок-схема алгоритма устранения паразитной частотной модуляции

22,05 кГц — 256, 512 и 1024 отсчетов. Затем проводим анализ максимальных значений вычисленного спектра для оценки мгновенной частоты тональных компонентов в каждом блоке данных. Для этого используем синусоидальный анализ (синтез) для обработки голоса [10]. При использовании метода МакАулау и Кватиери звук может быть представлен, как совокупность синусоидальных составляющих (так называемых треков), каждый из которых в течение времени изменяет амплитуду и частоту.

Максимальное значение амплитуды определяется как локальный максимум значений модуля спектра по правилу: наибольшее из двух соседних отсчетов, т.е.  $|X(n-1)| < |X(n)| > |X(n+1)|$ .

В значениях спектра дискретного преобразования Фурье локальный максимум отображается на синусоиде рядом с настоящим максимумом, который расположен в непрерывном спектре. Точность определения частоты должна быть 0,1 % вершины sinc-функции. Поэтому к идентифицированным в частотном спектре амплитудным максимумам применяем параболическую интерполяцию для получения более точного приближе-

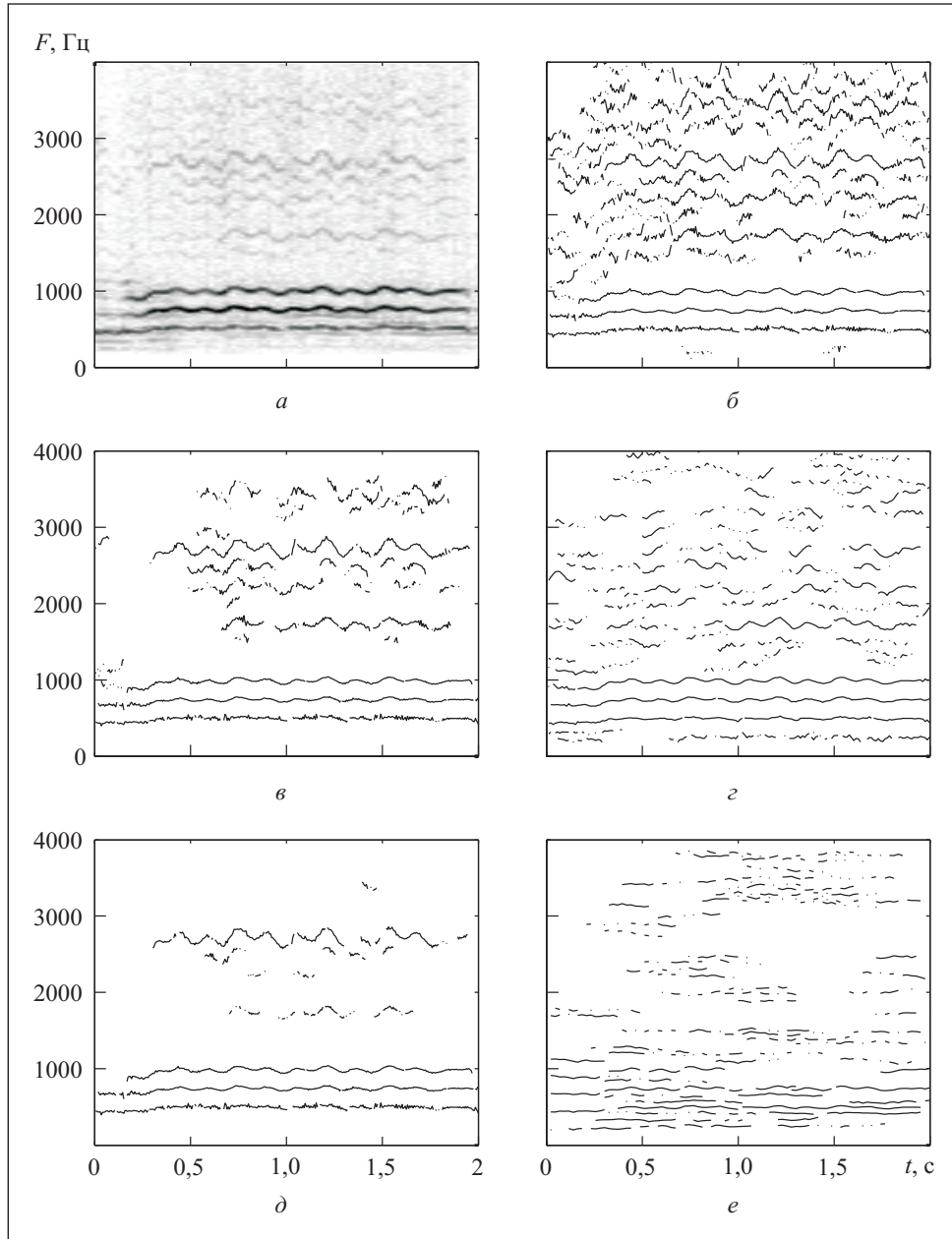


Рис. 3. Изображения частотных треков с различными параметрами: *a* — исходная сонограмма сигнала; *б* —  $w(k) = 256$  отсчетов,  $\Delta = 5$ ,  $\gamma = 0,01$ ; *в* —  $w(k) = 256$  отсчетов,  $\Delta = 0,5$ ,  $\gamma = 0,01$ ; *г* —  $w(k) = 512$  отсчетов,  $\Delta = 5$ ,  $\gamma = 0,01$ ; *д* —  $w(k) = 256$  отсчетов,  $\Delta = 0,5$ ,  $\gamma = 0,02$ ; *е* —  $w(k) = 1024$  отсчетов,  $\Delta = 5$ ,  $\gamma = 0,01$

ния частотных максимумов спектра. Эти максимумы наиболее заметны в звучании. Максимумы в соседних блоках данных сравниваются, и если их частоты подобны, то они соединяются. Приближенность к значениям сходства частоты определяется параметром контроля отклонений от максимума. Бесперывная цепь соединенных максимумов называется частотным треком. Максимум, который не соединяется, представляет собой начало или конец трека.

Процесс создания частотных треков происходит при следующих условиях:

1) если следующий максимум согласуется с интервалом отклонений, то траектория трека продлевается;

2) если следующий максимум не согласуется с интервалом отклонений, то трек заканчивается;

3) если трек находит согласование с максимумом, уже присутствующим в другом треке, то из двух треков будет продлен тот, к которому этот максимум ближе, а другой трек закончится.

Просчитанные таким образом частотные треки сохраняются в матрице данных  $Z$ . Колонки матрицы  $Z$  соотносятся с колонками матрицы развернутого во времени спектра (сонограммы), однако они содержат информацию, описывающую только синусоидальный состав сигнала. Сонограмма представляет собой диаграмму распределения спектральной энергии сигнала в координатах частоты и времени, на которой по вертикали откладывается частота, по горизонтали — время, а цвет (или оттенок серого) указывает на амплитуду спектра.

На рис. 3, *a* представлена сонограмма сигнала продолжительностью 2 с, а на рис. 3, *b—e* — соответствующие ей частотные треки с различными интервалами отклонений  $\Delta$  от максимума, различной длиной весового окна  $w(k)$  и порогом чувствительности  $\gamma$ . Был исследован фрагмент фонограммы звукового сигнала с частотой дискретизации 22,05 кГц, считанный из фонографического цилиндра, поврежденного паразитной частотной модуляцией. Как видим, наиболее соответствуют сонограмме частотные треки, изображенные на рис. 3, *b, в, г*.

Следующим шагом процесса устранения детонации является определение функции изменения высоты тона  $p_w(t)$  после соответствующих вычислений матрицы данных  $Z$ , в которой содержится информация о частотных треках. Вначале проводится нормирование  $Z$  для того, чтобы значения величин матрицы были соизмеримы со значениями величин функции изменения высоты тона. Процедуру нормирования  $Z$  записываем в виде

$$R(n, t) = \frac{Z(n, t)}{\max \{Z(n)\}}, \quad n=1 \dots N,$$

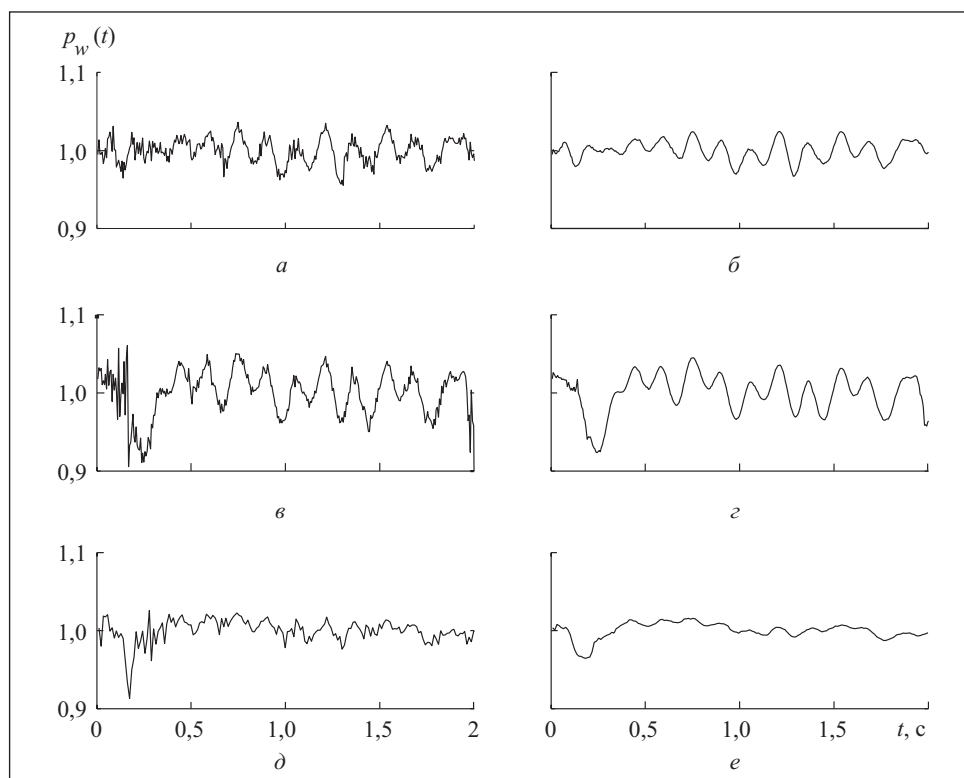


Рис. 4. Графическое изображение функции  $P_w(t)$ , соответствующее частотным трекам, представленным на рис. 3, б (а), рис. 3, в (в), 3, г (г); б, г, е — сглаженные функции  $P_w(t)$ , соответствующие частотным трекам на рис. 3, а, в, д

где  $n$  — номер частотного трека (номер строки матрицы  $Z$ );  $t$  — временной индекс (номер колонки матрицы  $Z$ );  $N$  — число строк матрицы  $Z$ ;  $\max \{Z(n)\}$  — максимальное значение в строке  $n$  матрицы  $Z$ .

Функцию  $p_w(t)$  находим как вектор-строку, которая содержит значения срединных элементов каждой колонки матрицы  $R(n, t)$ :

$$p_w(t) = \text{median}(R(1, t) R(2, t) \dots R(N, t)).$$

При  $p_w(t)=1$  изменения высоты тона не происходит. Поэтому для правильной оценки функции  $p_w(t)$  необходимо провести коррекцию уровней данных вокруг единицы. Вычисленные таким способом функции  $p_w(t)$ , соответствующие частотным трекам, изображенным на рис. 3, графически представлены на рис. 4. Для удаления возможных выбросов и резких изме-



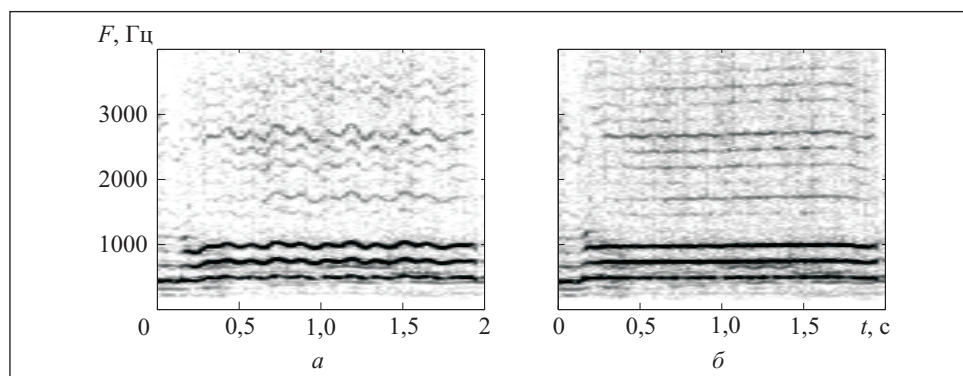


Рис. 5. Сонограммы звукового сигнала, поврежденного детонацией (а) и восстановленного сигнала (б)

нений высоты тона проводим сглаживание данных методом скользящего среднего, что способствует устойчивой оценке детонации (см. рис. 4, б, з, е).

Полученная таким образом функция  $\tilde{p}_w(t)$  позволяет реконструировать деградированный сигнал, в котором присутствует детонация. Поскольку деградированный сигнал можно рассматривать как неискаженный звуковой сигнал с неравномерной дискретизацией, его можно восстанавливать, используя особую технику передискретизации [12], т.е. реконструировать равномерные отсчеты оцифрованного сигнала.

Следует заметить, что изменения высоты тона происходят очень медленно по сравнению с частотой дискретизации. Поэтому в каждый момент времени можно аппроксимировать неравномерные отсчеты входного сигнала как равномерные с частотой дискретизации  $F' = F \tilde{p}_w(t)$ .

После нахождения функции  $\tilde{p}_w(t)$  переходим ко второму этапу решения задачи устранения паразитной частотной модуляции в звуковом сигнале — операции передискретизации. Для выполнения этой операции использован способ динамического преобразования частоты дискретизации на основе полифазного фильтра, подробно описанный в работе [13].

Для проведения экспериментальных исследований была разработана программа в интегрированной среде Matlab. Результаты исследований представлены в виде сонограммы сигнала с имеющейся паразитной частотной модуляцией (рис. 5, а) и сонограммы восстановленного сигнала, в котором почти отсутствует паразитная частотная модуляция (рис. 5, б). В качестве сигнала управления изменением частоты использована функция  $\tilde{p}_w(t)$ , изображенная на рис. 4, з, а также фрагмент фонограммы звукового сигнала, считанный из фонографического цилиндра.

Для оценки степени подавления паразитной частотной модуляции воспользуемся косвенным методом. Измеряя на сонограмме девиацию частоты какого-либо частотного трека, можно получить коэффициент паразитной частотной модуляции. На рис. 5, *а* отклонения частотного трека находятся в пределах 953—1031 Гц. Следовательно, полное изменение частоты равно 78 Гц, или 7,56 %. После восстановления деградированного сигнала отклонения того же частотного трека будут находиться в пределах 985—992 Гц (рис. 5, *б*). При этом полное изменение частоты равно 7 Гц, или 0,704 %. Полученные результаты свидетельствуют о подавлении паразитной частотной модуляции более чем в 10 раз.

**Выводы.** Предложенный алгоритм восстановления и реконструкции звукового сигнала в фонограммах аналоговых носителей записи, поврежденных паразитной частотной модуляцией, позволяет восстановить деградированные фонограммы записей. Слуховые тесты показали, что в случаях, когда в звуковом сигнале присутствуют естественные групповые импульсные изменения высоты тона, вызванные вибрато, тремоло или созданные исполнителем, восстановление может быть неоднозначным, с наличием в отдельных фрагментах фонограммы не устраненной детонации. Полученные результаты являются основой дальнейших исследований.

Solution of the problem of time distortions of signal-detonation in phonograms of analogue medium has been analyzed. A possibility of restoration of a sound signal damaged by the parasite frequency modulation has been considered. An algorithm has been offered for the reconstruction of the degraded signal.

1. *Godsill S. J., Rayner P. J. W.* The Restoration of Pitch Variation Defects in Gramophone Recordings // In Proc. IEEE Workshop on Audio and Acoustics. — Mohonk, NY State. — 1993, October. — P. 148—151.
2. *Godsill S. J.* Recursive Restoration of Pitch Variation Defects in Musical Recordings // IEEE Int. Conf. on Acoustics, Speech and Signal Processing 2. — Adelaide, Australia. — 1994, April. — P. 233—236.
3. *Godsill S. J., Rayner P. J. W.* Digital Audio Restoration a Statistical Model Based Approach. — NY : Springer-Verlag, 1998. — 328 p.
4. *Nichols J.* An Interactive Pitch Defect Correction System for Archival Audio // AES 20th Int. Conf. Budapest, Hungary. — 2001, October. Paper № 1941.
5. *Czyzewski A., Maziewski P., Dziubinski M., Kaczmarek A.* Wow Detection and Compensation Employing Spectral Processing of Audio // Proc. 117 AES Convention, San Francisco, USA. — 2005, January. Paper №. 6212.
6. *Howarth J., Wolfe P.* Correction of Wow and Flutter Effects in Analog Tape Transfers // Proc. 117 AES Convention. — San Francisco, USA, 2005, January. Paper № 6213.
7. *Czyzewski A., Ciarkowski A., Kaczmarek A., Kotus J.* DSP Techniques for Determining «Wow» Distortion // Journal of the Audio Engineering Society. — 2007. — Vol. 55, № 4. — P. 266—284.

8. Daniel E. D., Mee C. D., Clark M. H. Magnetic Recording: The First 100 Years. — IEEE Press, NY, 1999. — 341 p.
9. Read P., Meyer M. P. Restoration of Motion Picture Film.— Butterworth Heinemann, Oxford, UK, 2000. — 388 p.
10. McAulay R. J., Quatieri T. F. Speech Analysis/Synthesis Based on a Sinusoidal Representation // IEEE Transactions on Acoustics, Speech, Signal Processing. — 1986. — Vol. 34. — P. 744—754.
11. Рабинер Л. П., Шафер Р. В. Цифровая обработка речевых сигналов. Пер. с англ. М. В. Назарова, Ю. Н. Прохорова. — М. : Радио и связь, 1981. — 312 с.
12. Crochiere R., Rabiner L.R. Multirate Digital Signal Processing. — Englewood Cliffs, NJ: Prentice-Hall, 1983.
13. Косяк І. В. Динамічне перетворення частоти дискретизації на основі поліфазного фільтра // Реєстрація, зберігання і оброб. даних. — 2009. — **11**, № 4. — С. 68—77.

Поступила 04.03.10

*КОСЯК Игорь Васильевич, канд. техн. наук, ст. науч. сотр. Ин-та проблем регистрации информации НАН Украины. В 1990 г. окончил Киевский политехнический ин-т. Область научных исследований — регистрация и обработка информации.*

