

Программные средства шумоочистки записей речи

Столбов М.Б.,

Зубов Г.Н.

В статье приводятся обзор и классификация программных средств обработки зашумленных речевых сигналов применительно к задаче снижения уровня помех и повышения разборчивости речи. Дается прогноз их развития.

• шумоочистка • разборчивость речи • фильтрация сигналов • звукозапись

This paper presents the overview and classification for audio filtering software. The materials presented in this tutorial result from real-world problems in the domain of speech engineering. The forecast of speech filtering growing along with advancement in science and digital processing technologies is given.

• low enforcement software • speech enhancement • speech intelligibility
• digital filters • audio recording

Введение

Средства видео- и звукозаписи активно используются в различных приложениях с целью документирования возникающих ситуаций. Это обусловлено, в числе прочего, высокой информативностью записанного сигнала и особенностями самих средств регистрации, среди которых можно отметить следующие:

- невысокая стоимость;
- простота и удобство использования;
- возможность дистанционной записи и в автоматическом режиме.

Одним из наиболее важных требований, предъявляемых к качеству звукозаписи, является понимание смыслового содержания записанной речи. Однако обычная практика проведения звукозаписи (вне звукозаписывающих студий; воздействие акустических помех различной интенсивности, удаленность от источника речи и др.) далеко не всегда позволяет обеспечить требуемое качество. Поэтому возникает потребность в специальной обработке звукозаписи



с целью повышения разборчивости речи путём снижения присутствующих помех и искажений речи. Эффективность такой обработки во многом зависит от уровня развития программных средств шумоочистки (далее – ПСШ), а также умения их грамотно применять.

Следует отметить, что публикации по данной тематике носят эпизодический характер [1–9], а подавляющая часть литературы, в которой содержится информация о методах обработки речевых сигналов на фоне помех и приёмах их использования, малоприменима для конкретных ситуаций по следующим причинам:

- значительная часть методических материалов, посвящённых ПСШ, относится к тематике мастеринга студийных звукозаписей [10–13];
- большинство публикаций по алгоритмам фильтрации речи посвящены либо самым общим, либо частным задачам (см., например, библиографию [14]) и не дают пользователям систематизированной информации по данной тематике. При этом имеет место значительный терминологический разброс, и одни и те же алгоритмы обработки сигнала в разных программных продуктах имеют разные названия;
- имеющаяся у пользователя информация (публикации, методические материалы) о применении ПСШ во многом устарела и не соответствует настоящему состоянию дел в этой области. Это объясняется, в частности, переходом к цифровым средствам записи, кодирования, передачи и хранения звука [15];
- описание функциональных возможностей ПСШ зачастую носит поверхностный либо рекламный характер, или ограничивается перечислением органов управления и элементов графического интерфейса.

В связи с этим представляется актуальным анализ текущего состояния дел в области программных средств обработки аудиосигналов, а также потенциала их дальнейшего развития.

Задача оценки сравнительной эффективности ПСШ при этом не является главной (обсуждение этих вопросов можно найти, например, в материалах Centre for Law Enforcement Audio Research – CLEAR [16]).

Основные типы помех и искажений

Записи речевых сообщений, передаваемых по каналам телефонной связи общего пользования, имеют, как правило, удовлетворительную разборчивость речи, так как микрофон находится вблизи источника речи. Это позволяет в подавляющем большинстве случаев обойтись без дополнительной обработки речевого сигнала. Снижение качества некоторых записей возможно, в частности, для каналов сотовой связи, что связано, главным образом, с ошибками низкоскоростного кодирования и некорректной работой детектора речи в условиях воздействия интенсивных акустических помех.

Значительно большую проблему представляют так называемые «акустические» фонограммы, записанные в помещениях, в салоне автомобиля, на улице или местах скопления людей. Они характеризуются чрезвычайно широким разбросом качества записанной речи ввиду воздействия множества факторов, начиная с большого удаления микрофона от источника полезного сигнала и заканчивая несоответствием условий звукозаписи техническим характеристикам используемых для этого устройств.

Для «акустических» фонограмм характерны следующие типы помех и искажений:

- акустические помехи от различных внешних источников (бытовой и офисной техники, компьютеров, систем кондиционирования, радио- и ТВ-приемников и т.п.);
- фоновая речь других людей в помещении или общественных местах;
- реверберация вследствие отражений звука в помещении;
- наводки от источников электромагнитных излучений (ЭМИ), включая от мобильных телефонов;
- искажения, вносимые устройствами звукозаписи (ограничение полосы частот, недостаточный динамический диапазон, нелинейные искажения сигнала из-за ограничения амплитуды, некорректное кодирование речи при воздействии интенсивных акустических помех).

Для аудиозаписей, сделанных на улице, характерны:

- помехи транспортного потока;
- различные звуковые сигналы и сирены;
- шум ветра, листвы и др.

В табл. 1 представлены оценки значимости типовых шумов и искажений, определенные как произведение показателей, отражающих частоту их встречаемости, степень негативного влияния на разборчивость речи и сложность их снижения.

Таблица 1

Типы шумов и искажений	Частотность встречаемости	Степень негативного влияния, сложность снижения	Оценка значимости
Искажения, вносимые устройствами звукозаписи	3.7	4.0	14.8
Речь других людей	2.0	4.7	9.4
Музыкальные помехи (радио, ТВ и т.п.)	1.7	4.7	8,0
Шумы и искажения в каналах передачи	4,5	1.5	6.7
Реверберация	2.3	2.7	6.2
Наводки от мобильных телефонов	2.0	2.7	5.4
Офисные и бытовые шумы	2.0	2.3	4.6
Уличные помехи	2.3	1.7	3.9
Наводки от мощных источников ЭМИ	1.0	1.7	1.7

Полученные оценки позволяют сделать вывод, что при обработке аудиозаписей наиболее актуальной является борьба с шумами и искажениями, приведёнными в первых трёх строках таблицы. При этом следует учитывать, что некоторые искажения сигнала, вносимые устройствами звукозаписи, являются необратимыми и не могут



быть компенсированы последующим применением ПСШ: например, потери относительно слабых составляющих речи вследствие недостаточного динамического диапазона и/или использования психоакустических алгоритмов сжатия.

Типовой порядок повышения фильтрации речи

Выбор конкретных алгоритмов обработки низкокачественных аудиозаписей речи и последовательности их применения обусловлен как типом присутствующих шумов и искажений, так и особенностями слуха человека. Последнему присущ так называемый эффект маскировки. Он связан с процессом взаимодействия сигналов, при котором происходит снижение чувствительности слуха к речи в присутствии маскирующей помехи. При этом относительно громкие участки шума делают практически неслышимыми звуки речи, следующие за ними, а в некоторых случаях – даже предшествующие им.

С учётом изложенного, а также опыта работы с зашумленными речевыми сигналами можно предложить следующий типовой порядок шумочистки одно- и двухканальных аудиозаписей, схемы которых показаны на рисунках 1 и 2.

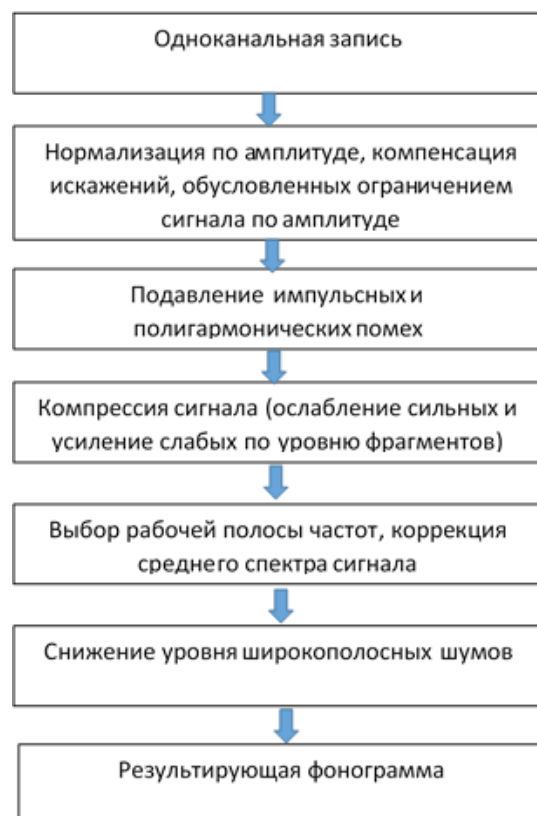


Рис. 1. Типовой порядок обработки одноканальной записи

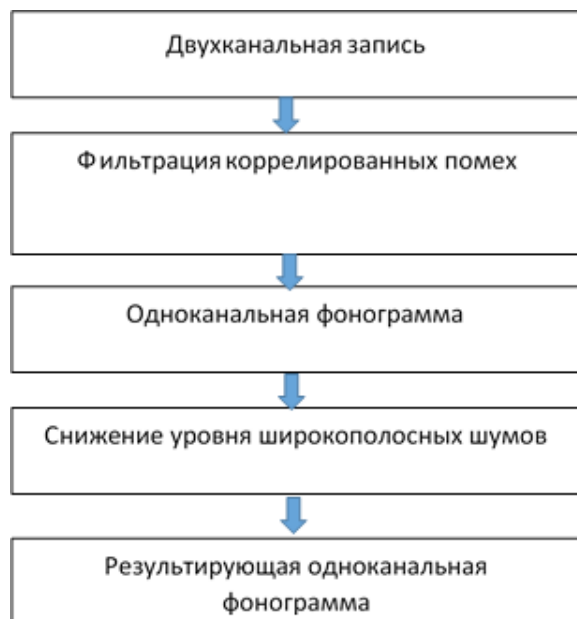


Рис. 2. Типовой порядок обработки двухканальной фонограммы

Два различных варианта обработки одноканальной записи можно объединить в двухканальную фонограмму, а небольшая задержка между сигналами в каналах позволит создать эффект псевдостерео.

Обработка двухканальных записей обычно проводится с использованием взаимных корреляционных связей между сигналами и помехами в каждом канале. Намного большие возможности для шумоочистки предоставляют n-канальные записи (см., например, [17]).

Критериями эффективности обработки аудиозаписей могут являться:

- повышение качества слухового восприятия и, прежде всего разборчивости речи, в типовых помеховых ситуациях;
- невысокая трудоёмкость (временные затраты на обработку должны быть соизмеримы с длительностью аудиозаписи).

Классификация программных средств шумоочистки

Хотя и не существует общепринятой классификации ПСШ, условно их можно разделить на ПСШ общего назначения и специализированные.

ПСШ общего назначения получили широкое распространение, поскольку являются достаточно универсальными, имеют, как правило, удобный интерфейс, широкий набор средств анализа, представления и обработки сигналов. Недостатком, являющимся обратной стороной их универсальности, является сложность обработки из-за множества доступных «эффектов» и фильтров. Поэтому для эффективного использования ПСШ общего назначения требуется более высокая квалификация пользователей, что в обычной практике сложновыполнимо.

В табл. 2 приведён перечень ПСШ общего назначения, разработанных рядом известных зарубежных компаний.



Таблица 2

Компания	Наименования продуктов
Acon Digital Media http://www.acondigital.com	Studio Clean, AcousticA
Adobe www.adobe.com	Adobe's Audition
Andrea Electronics www.andreaelectronics.com	SoundMAX Superbeam Array Microphone, Pure Audio
Audacity www.AudacityStore.com	AUDACITY Звуковой редактор с открытым кодом
Clarity www.clarityco.com	Clear Voice Capture
Cortologic www.cortologic.de	Echo Suppressor (CES), Noise Suppressor (CNS)
DART www.dartpro.com	Digital Audio Restoration Technology - Dart Pro
DSP Factory Ltd. www.dspfactory.com	Voice Focus, Signal Klara, Signal Enhancer Algorithms Series: ClearMessenger, Antishocker, VoiceGuard™
Extell Technology Corp. www.extell.com	Noise Elimination Speech Stream (NESS)
Kay Elemetrics Corp. www.kayelemetrics.com	Analysis-Synthesis Lab (ASL), Computerized Speech Lab (CSL), Signal Enhancement Program, Model 5142
Microsoft http://research.microsoft.com	Speech enhancement, Array processing, Dereverberation
NCT www.nctclearspeech.com	Clear Speech, Adaptive Speech Filter
Phillips Signal Processing www.speech.Phillips.com/vc	VoCon Product Family Modules
Sonic Solutions www.sonifoundry.com	Sound Forge, NoNoise, Sonic Studio - набор DirectX-плагинов
Steinberg Media Technologies GmbH www.steinberg.de	WaveLab sound editor, Clean, Cubase. Virtual Studio Technology (VST) – программный пакет
Wavemakers Inc. www.wavemakers.com	Clear Stream Technology, Smart AGC
Waves Audio Ltd www.waves.com	Native Power Pack, производитель плагинов для обработки аудио

Более эффективную обработку и для сложных акустических условий позволяют осуществлять специализированные ПСШ (iZotope, CEDAR, Sound Cleaner и др.) или средства анализа речи со встроенными функциями шумочистки (CSL, SIS, «Сапфир»). Во многом это связано с тем, что разработчики и поставщики этих ПСШ проводят обучение пользователей своей продукции.

В табл. 3 приведён перечень наиболее известных специализированных ПСШ, разработанных как зарубежными, так и отечественными компаниями.

Таблица 3

Компания	Наименования продуктов
Acoustic Magic Inc. www.acousticmagic.com	Voice Tracker I™ since 2002 (линейные микрофонные решётки) Voice Tracker II™ (плоские микрофонные решётки)
Algoritmix GmbH www.algoritmix.com	Renovator, CreamWare, Sound Laundry, OSIRIS XP
Andrea Electronics www.andraelectronics.com	SoundMAX Superbeam Array Microphone, Pure Audio
Audio Intelligence Inc www.intelligentdevicesinc.com http://intelligentdevices.com	SES (Speech Extracting System), Law Enforcement Software), EARS (Expedient Audio Recovery System)
Cedar Audio Ltd. www.cedar-audio.com	CEDAR (Computer Enhanced Digital Audio Restoration), SADIe, Pyramix, Soundscape, CSR (Cambridge Sound Restoration Studio)
Clarifying Technologies Inc http://www.clarifyingtech.com	Audio and video FORENSIC filtering
CLEAR labs http://www.clear-labs.com	CLEAR (Centre for Law Enforcement Audio Research),
CubeTech International GmbH www.cube-tec.com	Voice Restoration Software for Forensics Experts, Forensic Audio and Speech Analysis Package
Digital Audio Corp. www.dacaudio.com	QuickEnhance, CARDINAL MiniLab Suite
Diamond Cut Productions http://diamondcut.com	DIAMONDCUT (DC): DC-32, DC-50, DC-Millennium, DC-Live, DC-Five, DC-Six, Restoration Combo, Restoration Doo.
Enhance Audio Corp. www.enhancedaudio.com	DC Audio FORENSIC software
iZotope Inc www.izotope.com	iZotope RX1, RX2, RX3, iZotope music&speech cleaner, ANR-B
GritTec Laboratory www.grittec.ru	Технологии 1-, 2-канального шумоподавления
Иста Системс www.istasystems.ru	«Сапфир»
Центр речевых технологий www.speechpro.com	Sound Cleaner II, IKAR Lb, ANF II



Алгоритмы обработки, используемые в ПСШ

Алгоритмы обработки аудиозаписей, используемые в ПСШ, можно разделить на следующие три группы:

1. Алгоритмы реставрации аудиозаписей

- полосовые фильтры (ФНЧ, ФВЧ и др.);
- фильтры широкополосного шума;
- фильтры тресков (последовательности коротких импульсов);
- эквалайзеры или корректоры спектра сигнала.

2. Алгоритмы фильтрации речи на фоне слабых помех

- фильтры широкополосного шума;
- фильтры коррелированных помех;
- динамические фильтры;
- компенсаторы амплитудного ограничения сигнала;
- фильтры спектрального выравнивания.

3. Алгоритмы фильтрации речи на фоне помех повышенной интенсивности

- фильтры для снижения реверберационных искажений;
- фильтры для подавления шумов высокого уровня;
- фильтры для подавления сложных наводок ЭМИ;
- фильтры для выделения речи целевого диктора в речевом хоре;
- двухканальные фильтры подавления нестационарных помех.

Отдельно следует выделить алгоритмы «ручного» или «визуального» удаления/ослабления различного типа помех (например, в Adobe's Audition) посредством следующих преобразований: исходный сигнал → изображение динамической спектрограммы → редактирование (усиление/ослабление выделенных областей сигнала средствами графического редактора) → обработанный сигнал. Такие алгоритмы, с одной стороны, отличаются высокой универсальностью, а с другой – низкой производительностью.

В табл. 4 перечислены наиболее распространённые алгоритмы шумоочистки, реализованные в ПСШ разных компаний.

В целом существующие ПСШ при умелом использовании способны заметно повышать качество речи во многих практических случаях. Однако следует отметить, что они всё-таки недостаточно эффективны для подавления музыкальных и речеподобных помех, снижения реверберационных искажений и амплитудного ограничения сигнала. Кроме того, они требуют от пользователя значительных временных затрат на подбор оптимальных параметров для получения приемлемого результата обработки.

Наблюдается также явное отставание методического обеспечения ПСШ от уровня развития технологий шумоочистки. Обработка аудиосигналов остаётся во многом «искусством», опирающимся на личный опыт конкретного пользователя, хотя некоторые компании (Adobe, CLEAR, DAC, ЦРТ) проводят обучение пользователей применению ПСШ в форматах обучающих курсов, Educational Webinar series и др. (см., например, [18–21]).

Таблица 4

Компания	АРУ	Эквалайзеры	Подавление тональных помех	Фильтр широко-полосного шума	Фильтр импульсных помех	Спектр.выравнивание	Ослабление ампл. ограничения	Ослабление реверберации	2-канальная фильтрация	Многоканальная фильтрация
Acon										
Digital Media		+		+	+		+			
Acoustic Magic Inc.										+
Aculab									+	
Adobe	+	+	+	+	+	+	+		+	
Algoritmix				+	+		+			
Andrea Electronics										+
Cedar Audio Ltd.	+	+	+	+	+		+		+	
Clarity	+	+	+		+				+	
Cortologic				+					+	
CubeTec		+	+	+			+			
DAC	+	+	+	+		+			+	
DART		+	+	+	+					
DGI Corp.	+	+	+							
DSP Factory Ltd.					+	+				+
Enhance Audio Corp.	+	+	+	+	+		+	+		
Extell Technology Corp.	+	+	+							+
iZotop, Inc	+	+	+	+				+		
RWTH Institute, Aachen	+	+	+						+	
Kay Elemetrics Corp.		+	+	+	+				+	
NCT	+	+	+	+					+	
Sonic Solutions	+	+	+	+	+					
SONY	+	+		+	+					
Steinberg		+		+	+					
Voice Make Technologies Inc.										
Wavemakers Inc.									+	
GritTec Laboratory				+					+	+
Иста Системс		+	+	+	+		+			
Центр речевых технологий	+	+	+	+	+	+	+	+	+	+



Заключение

Далеко не всегда благоприятные условия проведения звукозаписи не оставляют сомнений в том, что ПСШ будут востребованы и в будущем.

Исходя из вышеизложенных результатов, можно сформулировать требования к перспективным направлениям развития ПСШ:

1. Создание специализированных ПСШ с максимальным уровнем автоматизации обработки, в том числе с использованием субоптимальных вариантов, понятных неквалифицированному пользователю (например: «шумы офиса», «на улице с интенсивным движением», «в гулком помещении», «разговор нескольких человек» и т.п.).
2. Создание средств совместной обработки аудио- и видеосигналов в реальном масштабе времени.
3. Создание облачных высокопроизводительных сервисов с возможностью удалённого доступа через WEB-интерфейс.
4. Разработка технологий многоканальной пространственной обработки зашумлённой речи.

При этом ПСШ общего назначения будут применяться в очень ограниченных масштабах по причине их избыточной функциональности, сравнительно высоких требований к квалификации пользователей и отсутствия методической поддержки со стороны разработчиков.

Рост вычислительных мощностей компьютеров, успехи в разработке новых алгоритмов цифровой обработки сигналов и прогресс в понимании процессов восприятия речи создают предпосылки для более широкого применения технологий шумоочистки в различных приложениях.

Литература

1. Коваль С.Л., Зубов Г.Н. Влияние шумоочистки на результаты последующего идентификационного исследования // «Системы безопасности»: межотраслевое специализированное издание. — 2004. — № 6.
2. Столбов М.Б., Зубов Г.Н. Шумоочистка низкокачественных фонограмм речи с помощью современных средств. Проблемы и решения: Современные методы, технические и программные средства, используемые в криминалистической экспертизе звукозаписей: Методическое пособие. — М.: МЮ РФ, ГУ РФЦСЭ, 2003. — Глава 3, § 2. — С. 155–177.
3. Хитров М.В., Коваль С.Л., Игнатов П.В. и др. Фоноскопическое исследование фонограмм речи: Исследование достоверности фонограмм. — Практическое пособие в 3-х книгах. Книга 1-я. — СПб.: Издательство «Юридический центр-Пресс», 2011.
4. Hilkhuisen G. & Huckvale M. Adjusting a commercial speech enhancement system to optimize intelligibility // AES Conference on Audio Forensics, Copenhagen Denmark, June 2010.
5. Jijul P.V., Singh C.P. and Sharma R.M. Study on the Selection of Specific Filters for Enhancement of Recorded Speech for Speaker Identification // The Open Forensic Science Journal. — 2009. — № 2, P. 29-33.
6. Maher R.C. Audio forensic examination // IEEE Signal Processing Magazine. — 2009. — Vol. 26. — № 2. — P. 84–94. 2009.
7. Koenig B. E. Enhancement of forensic audio recordings // J. Audio Eng. Soc. — 1988. — Vol. 36. — № 11.
8. Tibbits J. and Lu YouBin. Forensic applications of signal processing // Signal

- Processing Magazine, IEEE. — 2009. — Vol. 26. — P. 104–111.
9. *Reid G.* The Selection and Use of Adaptive Filters and Associated Signal Processing // Proc of AES-26, 2005.
 10. *Загуменнов А.П.* Компьютерная обработка звука. — М.: ДМК, 1999. — 384 с.
 11. *Петелин Р.Ю., Петелин Ю.В.* Adobe Audition. Обработка звука для цифрового видео. — СПб.: БХВ-Петербург, 2005.
 12. *Степаненко О.С.* Adobe Audition 3.0. Создание фонограмм и обработка звука. — М.: Диалектика, 2010.
 13. *Кинтцель Т.* Руководство программиста по работе со звуком. Пер. с англ. — М.: ДМК Пресс, 2000. — 432 с.
 14. *Петровский А.А., Азаров И.С., Лихачёв Д.С., Ромашкин Ю.Н. и др.* Шумоподавление на основе перцептуальных алгоритмов спектрального вычитания и обработки сигналов в подпространствах // Речевые технологии. — 2012. — №4. — С. 4–15.
 15. *Галяшин В.Н., Галяшина Е.И.* Цифровые фонограммы как судебное доказательство // Воронежские криминалистические чтения. Сборник научных трудов. — Воронеж: Изд-во Воронеж. ун-та. — 2007. — Вып. 8. — С. 71–99.
 16. CLEAR, Библиография по Law Enforcement Audio Research. Электронный ресурс. Режим доступа: <http://www.clear-labs.com/Tutorial-LitReview/Enhancement>.
 17. *Bertrand A.* Applications and trends in wireless acoustic sensor networks: a signal processing perspective // Proc. IEEE Symposium on Communications and Vehicular Technology (SCVT), Ghent, Belgium, November 2011.
 18. Clarifying Technologies Inc (обучающие курсы). Электронный ресурс. Режим доступа: <http://www.clarifyingtech.com>.
 19. Training courses www.enhancedaudio.com.
 20. Forensic Science Services Inc: материалы по forensic-audio-examinations. Электронный ресурс. Режим доступа: www.forensicservices.ca.
 21. Сайт компании «Центр речевых технологий». Электронный ресурс. Режим доступа: www.speechpro.com.

Сведения об авторах:

Столбов Михаил Борисович,

кандидат технических наук, старший научный сотрудник ООО «ЦРТ-инновации», доцент кафедры речевых информационных систем Национального исследовательского университета Информационных технологий, механики и оптики (НИУ ИТМО). Область научных интересов: шумоочистка, повышение разборчивости речи на фоне помех. E-mail: stolbov@speechpro.com

Зубов Герман Николаевич,

генеральный директор негосударственной экспертной организации «Форенэкс». Область научных интересов: судебная фоноскопия. E-mail: zubov@forenex.biz