

Теория сигналов и систем

УДК 004.934

Программный инструментарий для тестирования систем обработки речевых сигналов. Часть 1. Моделирование сигналов и систем

А.Н. Продеус, д-р техн. наук

Национальный технический университет Украины «Киевский политехнический институт»,
пр. Победы, 37, Киев-56, 03056, Украина.

Предложен состав программного инструментария для исследований и оптимизации алгоритмов систем коррекции и кодирования речевых сигналов, подверженных различного рода искажениям, таким как шумовая и реверберационная помехи, ошибки кодирования. В первой части данной работы рассмотрен инструментарий, позволяющий решать такие задачи как создание корпусов зашумленной речи и коррекция зашумленных речевых сигналов. Произведен анализ достоинств и недостатков инструментариев FaNT и VoiceBox, выработаны рекомендации по компенсации указанных недостатков средствами Matlab. Показана целесообразность создания обобщенного инструментария, путем объединения инструментариев FaNT и VoiceBox. Кроме того, показана целесообразность дополнения такого инструментария программными модулями, обеспечивающими моделирование оценивания качества систем обработки речевых сигналов. Рассмотренные примеры свидетельствуют о работоспособности и правильности функционирования данного инструментария. Библ. 18, рис. 5, табл. 1.

Ключевые слова: программный инструментарий, оптимизация алгоритмов, создание речевых корпусов, коррекция зашумленных сигналов.

Введение

При разработке и испытаниях математического обеспечения систем цифровой обработки речевых сигналов, таких как системы коррекции речевых сигналов, системы кодирования и автоматического распознавания речи (АРР), необходимо располагать программным инструментарием, позволяющим не только эффективно решать поставленные задачи, но и оперативно изменять план экспериментальных исследований при возникновении новых задач.

В соответствии с обобщенной схемой тестирования систем обработки речевых сигналов, представленной на рис. 1, такой программный инструментарий состоит из трех частей:

- программы создания искаженных речевых сигналов;
- программы моделирования тестируемых систем;
- программы оценивания качества тестируемых систем.
- вимірювального пристрою.

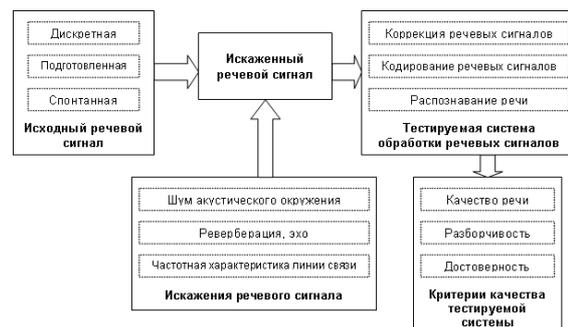


Рис. 1. Схема тестирования систем обработки речевых сигналов

Примеры создания фрагментов подобного «испытательного стенда» известны. В частности, для моделирования искаженных речевых сигналов разработан программный инструментарий FaNT (Filtering and Noise Adding Tool) [7,8,13]. Его недостатком является отсутствие алгоритмов коррекции сигналов и АРР, а также отсутствие алгоритмов оценивания качества тестируемых систем. Указанный недостаток частично компенсирует инструментарий VoiceBox [15], среди наиболее интересных компонентов которого - программы, позволяющие моделировать процедуру коррекции речевых сигналов [14]. Вместе с тем, в библиотеке VoiceBox весьма неполно представлены программные модели систем кодирования и АРР, а также модели систем акустической экспертизы

[17,18], хотя именно такие системы в последнее время вызывают наибольший интерес специалистов сотовых и компьютерных систем связи.

Таким образом, следует констатировать фрагментарный характер существующего программного инструментария, что накладывает существенные ограничения на перечень решаемых с его помощью задач. Цель данной работы состоит в расширении состава и возможностей существующего программного инструментария, позволяющем решать значительно более сложные задачи комплексного характера. Примерами задач такого рода являются сопоставление и оптимизация алгоритмов шумоподавления, входящих в состав математического обеспечения систем цифровой связи и систем автоматического распознавания речи.

1. Создание тестовых сигналов

При разработке и оптимизации систем обработки речевых сигналов, искаженных шумами, реверберацией и иными мешающими факторами, тестовыми сигналами служат искаженные соответствующим образом речевые сигналы. При акустической экспертизе каналов речевой коммуникации зачастую используют синтезированные речевые сигналы либо стационарный (или нестационарный) шум, спектрально-временные свойства которого подобны, хотя и не идентичны таковым для речевых сигналов [17].

При сопоставлении результатов, полученных различными экспериментаторами, а также при сопоставлении эффективности конкурентных алгоритмов удобно использовать тестовые сигналы в форме речевых корпусов [16], представляющих собой базы данных в виде звуковых и текстовых файлов с соответствующими транскрипциями. Транскрипции звуковых файлов необходимы при испытаниях алгоритмов АРР, тогда как при тестировании алгоритмов шумоподавления и кодирования, можно ограничиться звуковой частью речевых корпусов.

Создавая звуковую часть речевых корпусов, целесообразно учитывать ряд факторов: виды моделируемой речи; разновидности искажений, которым подвергаются речевые сигналы; различие в назначении и технических характеристиках тестируемых систем; разнообразие и особенности критериев и методов оценивания качества тестируемых систем; необходимость сравнения результатов, полученных разными исследователями.

Коротко прокомментируем некоторые из указанных факторов. Учет видов моделируемой речи на практике означает, как минимум, необходимость учета пауз для последующего корректного добавления шума к первичному «чистому» речевому сигналу.

Учет разновидностей искажений сигналов означает, что для отработки алгоритмов подавления шума акустического окружения достаточно использовать модель речевого сигнала $y(t)$ в виде аддитивной смеси чистого речевого сигнала $x(t)$ и шумового процесса $n(t)$:

$$y(t) = x(t) + n(t), \quad (1)$$

тогда как при учете влияния характеристик помещения, характеризуемого импульсной характеристикой $h_r(t)$, модель речевого сигнала имеет вид:

$$y(t) = x(t) \otimes h_r(t) + n(t), \quad (2)$$

где \otimes - символ свертки. При учете влияния фильтрующих свойств канала передачи, характеризуемого импульсной характеристикой $h_c(t)$, модель речевого сигнала описывается более сложным выражением:

$$y(t) = (x(t) \otimes h_r(t) + n(t)) \otimes h_c(t). \quad (3)$$

Следствием разнообразия отлаживаемых или тестируемых систем является разнообразие дополнительных требований к модели искаженного тестового сигнала. К ним относятся:

- полоса частот сигнала (узкая - 30-3400 Гц; широкая - 50-7000 Гц; сверхширокая - 50-14000 Гц) [11,12];
- «априорное» отношение сигнал-шум $\xi = 10 \lg(D_x/D_n)$ (D_x и D_n - дисперсии сигнала $x(t)$ и шума $n(t)$, соответственно);
- характер окрашенности шума (для «синтезированных» шумов);
- вид шума естественного происхождения;
- параметры нестационарной шумовой помехи;
- наличие дополнительных видов помех и искажений (реверберация, эхо, нелинейные искажения, особенности частотной характеристики тракта, способ кодирования сигнала и т.п.);
- наличие международных стандартов и рекомендаций относительно характеристик фильтров и кодеков речи, используемых в цифровых линиях связи.

При разработке корпусов тестовых сигналов также целесообразно учитывать, по каким кри -

териям будет оцениваться тестируемая система, поскольку объективным (инструментальным) методам акустической экспертизы присущ ряд ограничений [17]. Наконец, может оказаться весьма удобным хранить отдельно записи «чистого» речевого сигнала, записи шума и импульсной характеристики канала связи.

Инструментарий FaNT для формирования тестовых сигналов

Примером весьма известного современного программного инструментария, позволяющего быстро формировать речевые корпуса с тестовыми сигналами, является упомянутая выше система FaNT [7,8,13]. К достоинствам данного инструментария относятся следующие возможности: моделирование влияния шумовой и реверберационной помех; моделирование частотных характеристик линий связи; возможность учета особенностей восприятия звука слуховой системой человека (путем частотного А-взвешивания); моделирование искажений, свойственных системам кодирования в цифровых линиях связи; наличие детектора голосовой активности, позволяющего корректно учитывать паузы в речевых сигналах.

Вместе с тем, инструментарий FaNT свойственен ряд недостатков. Например, отсутствует возможность работы с шумами из библиотеки пользователя, а к некоторым из предлагаемых образцов шумов можно предъявить серьезные претензии. Так, шум метро имеет выраженный нестационарный характер, причем параметры этой нестационарности в документации не описаны (рис. 2). Кроме того, в системе FaNT невозможна работа со сверхширокополосными (14 кГц) речевыми сигналами, хотя в современных стандартах ITU-T такой вид сигналов предусмотрен [12].

Инструментарий Matlab для формирования тестовых сигналов

Наличие указанных особенностей инструментария FaNT означает, что в ряде случаев пользователи должны создавать собственный инструментарий для формирования тестовых сигналов. Весьма удобно использовать при этом возможности среды Matlab [4]. Запись «чистых» речевых сигналов можно осуществить с помощью стандартной функции `wavrecord()`, используя качественные микрофон и звуковую карту, в заглушенном помещении со временем реверберации не более 0,3 с.

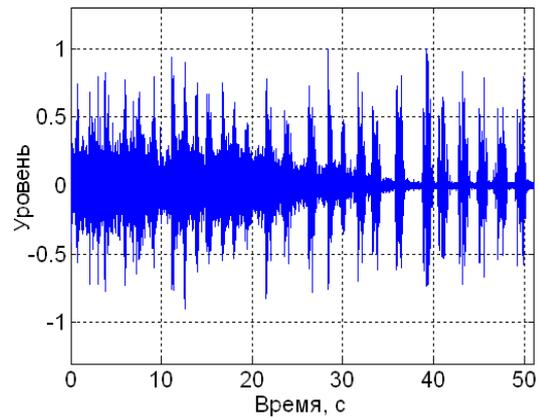


Рис. 2. Сформированная системой FaNT смесь речевого сигнала и шума метро

При записи дискретной речи целесообразно создавать вспомогательные звуковые файлы с удаленными паузами, что позволит корректно определить уровни речевых сигналов (файлы с удаленными паузами могут понадобиться и при оценивании разборчивости речи формантным методом). Удаление пауз можно осуществить с помощью детектора голосовой активности `vadsohn()` из библиотеки VoiceBox. С этой же целью можно использовать соответствующую функцию программных звуковых редакторов, имеющихся в распоряжении пользователя. В конце концов, удаление пауз можно выполнить «вручную» с помощью тех же звуковых редакторов. Поскольку «ручное» удаление пауз является более трудоемким, его следует использовать при создании небольших по объему, но очень качественных, корпусов. Примеры удаления пауз различными способами приведены на рис. 3 (слово «листочек»). Как показано на рис. 3,б и 3,в, при неудачной регулировке параметров системы автоматического удаления пауз относительная ошибка в определении длины сигнала может достигать 20%, что приводит к значительной, до 20%, ошибке оценки эффективного уровня сигнала.

Генерирование искусственного белого либо окрашенного шума $n_1(t)$ выполняют с помощью стандартной функции `gandn()`, с последующим применением соответствующих фильтров [17]. После вычисления дисперсии D_{n1} искусственного или естественного шума $n_1(t)$ выполняют вычисление дисперсии D_x сигнала без пауз, после чего оценивают «имеющееся» отношение сигнал-шум $\xi_1 = 10 \lg(D_x/D_{n1})$. Поскольку зна -

чение ξ_1 , как правило, не соответствует требуемому отношению ξ_0 , далее следует вычислить коэффициент k коррекции мгновенных значений шума: $k = 10^{0,05(\xi_1 - \xi_0)}$.

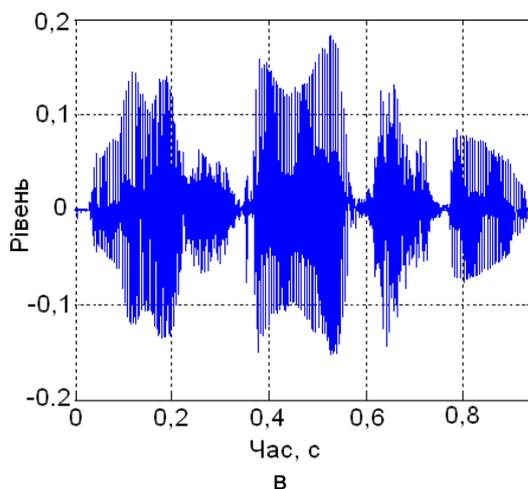
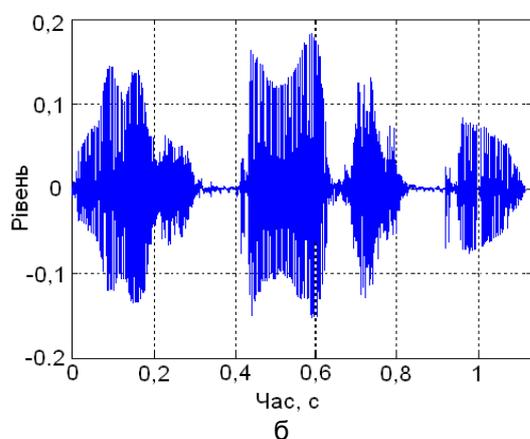
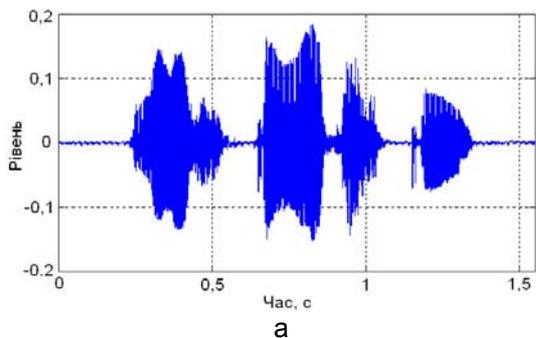


Рис. 3. Речевой сигнал до (а) и после удаления пауз вручную (б) и автоматически (в)

Формирование смеси (1) с требуемым отношением сигнал-шум ξ_0 выполняются по формуле

$$y(t) = x(t) + kn_1(t).$$

Формирование смесей по формулам (2)-(3) требует выполнения процедуры свертки, для реализации которой в библиотеке Matlab имеется ряд стандартных функций (`conv()`, `filter()`, `xcorr()`, `fftfilt()` и др.).

2. Моделирование систем обработки речевых сигналов

Моделирование кодеков мобильных линий связи

Влияние на речевой сигнал такой разновидности систем обработки речи как кодеки мобильных линий связи GSM и UMTS, удобно моделировать с помощью системы FaNT. Высокое качество моделирования обеспечено использованием библиотек программных модулей, рекомендованных Международным Союзом Телекоммуникаций (International Telecommunication Union - ITU), а также использованием программных модулей речевого кодека AMR, стандартизированных Европейским Институтом Телекоммуникационных Стандартов (European Telecommunications Standards Institute - ETSI) и консорциумом «3rd Generation Partnership Project» (3GPP) [11].

Моделирование систем коррекции речевых сигналов

При моделировании алгоритмов коррекции речевых сигналов, в качестве «образцовых» целесообразно использовать набор функций инструментария VoiceBox (табл. 1). Базовые функции этого набора, `submmse()` и `specsub()`, соответствуют двум направлениям в теории коррекции речевых сигналов – методу оптимальной фильтрации и методу спектрального вычитания [4,5,14].

Несмотря на различие указанных методов, структурная схема процедуры коррекции речевого сигнала, реализуемой с помощью функций табл. 1, может быть унифицирована и представлена рис. 4 [2-6,9,10].

Обработка сигнала начинается с разбиения его (с помощью функции `enframe()`) на перекрывающиеся и взвешенные окном Хэмминга сегменты (по умолчанию степень перекрытия 50%). Каждый из сегментов подвергается спектральной обработке (с помощью функции `rfft()`), результатом которой является набор спектров сегментов.

Табл. 1. Функции коррекции речевых сигналов

Функция	Действие функции
<i>ssubmmse()</i>	коррекция речи методами MMSE или logMMSE
<i>specsub()</i>	коррекция речи методом спектрального вычитания
<i>estnoiseg()</i>	оценка спектра помехи по зашумленному сигналу методом MMSE
<i>estnoisem()</i>	оценка спектра помехи по зашумленному сигналу методом минимальной статистики
<i>enframe()</i>	разбиение сигнала на сегменты
<i>rfft()</i>	преобразование Фурье сегмента сигнала
<i>irfft()</i>	обратное преобразование Фурье спектра сегмента сигнала

Процедура оценивания спектра помехи является весьма ответственной и сложной, поскольку, по существу, при ее реализации выполняется автоматическое детектирование голосовой активности. При оценивании (по умолчанию) спектра помехи методом минимальной статистики [9] используется функция *estnoisem()*, однако при желании пользователь может использовать альтернативную функцию *estnoiseg()*, реализующую вычисления в соответствии с методом MMSE [3, 6].

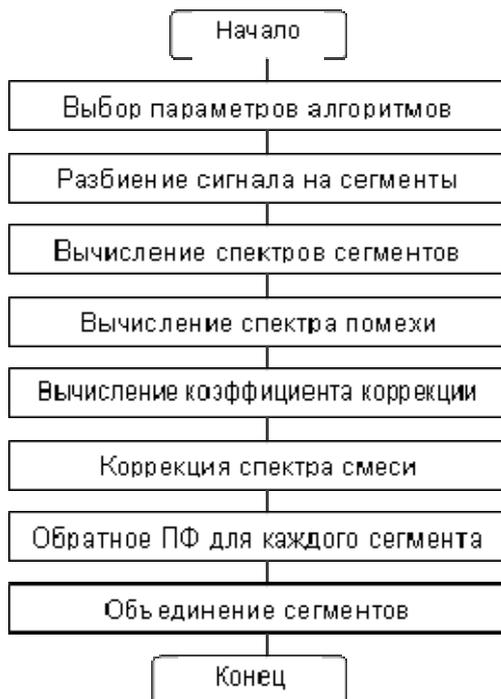


Рис. 4. Процедура коррекции речевого сигнала

Ядром алгоритмов шумоподавления является расчет корректирующего коэффициента $G(\xi_k, \gamma_k)$, при этом откорректированная амплитуда k -й составляющей спектра \hat{A}_k рассчитывается по формуле

$$\hat{A}_k = G(\xi_k, \gamma_k) R_k, \quad (4)$$

где $R_y(k)$ - оценка k -й составляющей спектра зашумленного сигнала (1).

При использовании метода logMMSE корректирующий коэффициент $G(\xi_k, \gamma_k)$ рассчитывают в соответствии с выражением [5]:

$$G(\xi_k, \gamma_k) = \frac{\xi_k}{1 + \xi_k} \exp \left\{ \frac{1}{2} \int_{\nu_k}^{\infty} \frac{e^{-t}}{t} dt \right\}, \quad (5)$$

где $\xi_k = \lambda_x(k)/\lambda_n(k)$ означает *априорное* отношение сигнал-шум для k -го спектрального компонента ($\lambda_x(k)$ - спектр мощности чистого сигнала, а $\lambda_n(k)$ - спектр мощности шума);

$\nu_k = \frac{\xi_k}{1 + \xi_k} \gamma_k$, где $\gamma_k = R_y(k)/\lambda_n(k)$ - *апостериорное* отношение сигнал-шум k -го спектрального компонента.

При расчете корректирующего коэффициента методом MMSE [4] используется соотношение:

$$G(\xi_k, \gamma_k) = \frac{\sqrt{\pi \nu_k}}{2 \gamma_k} \exp \left(-\frac{\nu_k}{2} \right) \times \left[(1 + \nu_k) I_0 \left(\frac{\nu_k}{2} \right) + \nu_k I_1 \left(\frac{\nu_k}{2} \right) \right], \quad (6)$$

где $I_0(\cdot)$ и $I_1(\cdot)$ - модифицированные функции Бесселя нулевого и первого порядка, соответственно.

Откорректированные, в соответствии с (4), спектры сегментов далее подвергаются обратному преобразованию Фурье (с помощью функции *irfft()*) и объединяются, образуя конечный продукт - откорректированный сигнал.

На рис. 5,а показан речевой сигнал (слово «листочек»), замаскированный аддитивным белым шумом (отношение сигнал-шум 5 дБ), а на рис. 5,б показан результат его восстановления с помощью метода logMMSE, реализованного с помощью описанной выше процедуры.

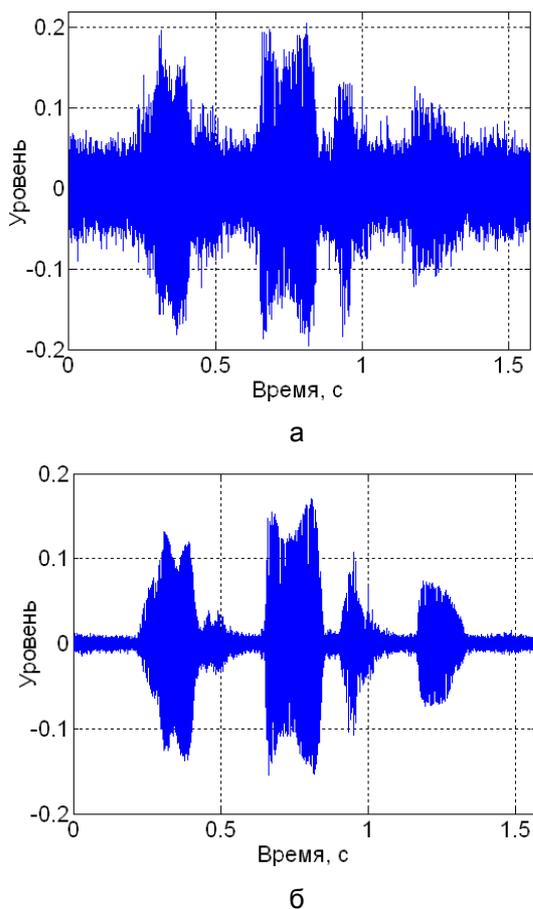


Рис. 5. Смесь сигнала и белого шума с отношением сигнал-шум 5 дБ (а) и результат восстановления сигнала (б)

В данном примере оценка выигрыша в отношении сигнал-шум оказалась близкой 6 дБ, что соответствует литературным данным о потенциальных возможностях метода logMMSE [14]. Приведенный пример свидетельствует о работоспособности и правильности функционирования алгоритмов коррекции речевых сигналов данного программного инструментария, что позволяет использовать их в качестве своеобразных «эталонов» при испытаниях модернизированных или новых алгоритмов обработки речевых сигналов.

Выводы

Программный инструментарий, позволяющий осуществлять тестирование и оптимизацию алгоритмов систем обработки речевых сигналов, подверженных различного рода искажениям, весьма востребован специалистами в области цифровой обработки сигналов.

Первая часть такого программного инструментария, представленная в данной работе, позволяет решать такие задачи как оперативное создание корпусов зашумленной речи и опти-

мизацию параметров алгоритмов коррекции зашумленных речевых сигналов. Рассмотренные примеры свидетельствуют о работоспособности и правильности функционирования данного программного инструментария.

Во второй части работы предполагается рассмотреть набор компонентов программного инструментария, обеспечивающий построение моделей систем акустической экспертизы.

Литература

1. *Beerends J.G.* Extension of ITU-T Recommendation P.862 PESQ towards Measuring Speech Intelligibility with Vocoders / Beerends J.G., van Wijngaarden S., van Buuren R. // [Электронный ресурс]. – Режим доступа к ресурсу: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA454414>
2. *Cappe O.* Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor / Cappe O. // IEEE Trans Speech Audio Processing. – 1994. – Vol. 2. – No 2. P. 345–349.
3. *Erkelens J.* A Data-Driven Approach to Optimizing Spectral Speech Enhancement Methods for Various Error Criteria / Erkelens J., Jensen J., Heusdens R. // Speech Communication. – 2007. – No 49. – P. 530–541.
4. *Ephraim Y.* Speech Enhancement Using a Minimum-Mean Square Error Short-Time Spectral Amplitude Estimator / Ephraim Y., Malah D. // IEEE Trans Acoustics Speech and Signal Proc. – 1984. – Vol. 32. – No. 6. – P. 1109–1121.
5. *Ephraim Y.* Speech Enhancement Using a Minimum Mean-Square Error log-Spectral Amplitude Estimator / Ephraim Y., Malah D. // IEEE Trans Acoustics Speech and Signal Proc. – 1985. – Vol. 33. – No 2. – P. 443–445.
6. *Gerkmann T.* Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay / Gerkmann T., Hendriks R. // IEEE Trans Audio, Speech, Language Proc. – 2012. – No 20. – P. 1383–1393.
7. *Hirsch H.-G.* FaNT - Filtering and Noise Adding Tool. [Электронный ресурс]. – Режим доступа к ресурсу: <http://dnt.kr.hsnr.de/>
8. *Hirsch H.-G.* The Simulation of Realistic Acoustic Input Scenarios for Speech Recognition Systems / Hirsch H.G., Finster H. // 9th European Conf on Speech Communication and Technology. – Lisboa, September 2005. – P. 1–4.
9. *Martin R.* Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics / Martin R. // IEEE Trans.

- Speech and Audio Processing. – 2001. – Vol. 9. – No 5. – P. 504–512.
10. *Martin R.* Statistical Methods for the Enhancement of Noisy Speech. In J. Benesty, S. Makino, and J. Chen, editors, *Speech Enhancement*, chapter 3. - New-York: Springer-Verlag, 2005. – P. 43–64.
 11. *Recommendation ITU-T P.56.* Series P: Terminals and Subjective and Objective Assessment Methods. Objective Measuring Apparatus. Objective Measurement of Active Speech Level. – Geneva: Telecommunication Standardisation Sector of ITU, 2008. – 12. – 24 p.
 12. *Recommendation ITU-T P.863.* Series P: Terminals and Subjective and Objective Assessment Methods. Methods for Objective and Subjective Assessment of Speech Quality. Perceptual Objective Listening Quality Assessment. – Geneva: Telecommunication Standardisation Sector of ITU, 2011. – 1. – 76 p.
 13. *Simulation of Acoustic Environments Including the Transmission over Telephone Channels.* [Електронний ресурс]. – Режим доступу к ресурсу: <http://dnt.kr.hsnr.de/sireac.html>
 14. *Springer Handbook of Speech Processing /* Под редакцией Jacob Benesty, M. Mohan Sondhi, Yiteng Huang. – Berlin: Springer-Verlag, 2008. – P. 1176.
 15. *VoiceBox: Speech Processing Toolbox for MATLAB.* [Електронний ресурс]. – Режим доступу к ресурсу: <http://www.ee.ic.ac.uk/hp/staff/dmb/>
 16. *Кривнова О.Ф.* Речевые корпуса на новом технологическом витке / Кривнова О.Ф. // *Речевые технологии.* – 2008. – № 2. – С. 13–23.
 17. *Продеус А.Н.* Акустическая экспертиза каналов речевой коммуникации. Монография / Продеус А.Н., Дидковский В.С., Дидковская М.В. – К: Имэкс-ЛТД, 2008. – 420 с.
 18. *Продеус А.Н.* О некоторых особенностях развития объективных методов измерений разборчивости речи / Продеус А.Н. // *Электроника и связь, тематический выпуск "Электроника и нанотехнологии".* – 2010. – № 2. – С. 217–223.

УДК 004.934

Програмний інструментарій для тестування систем обробки мовленнєвих сигналів. Частина 1. Моделювання сигналів і систем

А.М. Продеус, д-р техн. наук

Національний технічний університет України «Київський політехнічний інститут»,
пр. Перемоги, 37, Київ-56, 03056, Україна.

Запропоновано склад програмного інструментарію для досліджень і оптимізації алгоритмів систем корекції й кодування мовленнєвих сигналів, що піддані різного роду спотворенням, таким як шумова й ревербераційна перешкоди, помилки кодування. У першій частині даної роботи розглянуто інструментарій, що дозволяє розв'язувати такі завдання як створення корпусів зашумленої мови й корекція зашумлених мовленнєвих сигналів. Виконано аналіз достоїнств і недоліків інструментаріїв FaNT і VoiceBox, вироблені рекомендації з компенсації зазначених недоліків засобами Matlab. Показано доцільність створення узагальненого інструментарію, шляхом об'єднання інструментаріїв FaNT і VoiceBox. Крім того, показано доцільність доповнення такого інструментарію програмними модулями, що забезпечують моделювання оцінювання якості систем обробки мовних сигналів. Розглянуті приклади свідчать про працездатність й правильність функціонування даного інструментарію. Библ. 18, мал. 5, табл. 1.

Ключові слова: програмний інструментарій, оптимізація алгоритмів, створення мовленнєвих корпусів, корекція зашумлених сигналів.

UDC 004.934

Software toolkit for testing of speech signals processing systems. Part 1. Simulation of signals and systems

A.N. Prodeus, Dr.Sc.

National Technical University of Ukraine «Kyiv Polytechnic Institute»,
37 Prospect Peremogy, Kiev 03056, Ukraine.

Structure of the software toolkit for the study and optimization of algorithms of correction and coding systems, which are subject of various distortions, such as noise and reverberation disturbance, coding errors is proposed. The first part of this paper is devoted to consideration of toolkit which permit solve such problems as creating of noisy speech corpora and correction of noisy speech signals. The analysis of advantages and lacks of toolkits FaNT and VoiceBox is made, the guidelines on compensation of the lacks by Matlab resources are developed. The expediency of creation of generalized toolkit, by association of toolkits FaNT and VoiceBox, is shown. Besides the expediency of addition of such toolkit by program modules ensuring simulation of quality estimation of speech processing systems is shown. The examples show efficiency and proper functioning of the proposed toolkit. Reference 18, figures 5, tables 1.

Keywords: *software toolkit, optimization algorithm, corpora creation, correction of noisy signals.*

References

1. *Beerends J.G.* Extension of ITU-T Recommendation P.862 PESQ towards Measuring Speech Intelligibility with Vocoders / Beerends J.G., van Wijngaarden S., van Buuren R. // [On-line]. – Available: <http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA454414> (Eng) (1.03.2013).
2. *Cappe O.* (2007), [Elimination of the Musical Noise Phenomenon with the Ephraim and Malah Noise Suppressor]. *EEE Trans Speech Audio Processing*. Vol. 2. No 2. Pp. 345–349.
3. *Erkelens J., Jensen J., Heusdens R.* (2007), [A Data-Driven Approach to Optimizing Spectral Speech Enhancement Methods for Various Error Criteria]. *Speech Communication*. No 49. Pp. 530–541.
4. *Ephraim Y., Malah D.* (1984), [Speech Enhancement Using a Minimum-Mean Square Error Short-Time Spectral Amplitude Estimator]. *IEEE Trans Acoustics Speech and Signal Proc.* Vol. 32. No. 6. Pp. 1109–1121.
5. *Ephraim Y., Malah D.* (1985), [Speech Enhancement Using a Minimum Mean-Square Error log-Spectral Amplitude Estimator]. *IEEE Trans Acoustics Speech and Signal Proc.* Vol. 33. No 2. Pp. 443–445.
6. *Gerkmann T., Hendriks R.* (2012), [Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay]. *IEEE Trans Audio, Speech, Language Proc.* No 20. Pp. 1383–1393.
7. *Hirsch H.-G.* FaNT - Filtering and Noise Adding Tool. [On-line]. – Available: <http://dnt.kr.hsnr.de/> (1.03.2013).
8. *Hirsch H.-G., Finster H.* (2005), [The Simulation of Realistic Acoustic Input Scenarios for Speech Recognition Systems]. 9th European Conf on Speech Communication and Technology. Lisboa, September. Pp. 1–4.
9. *Martin R.* (2001), [Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics]. *IEEE Trans. Speech and Audio Processing*. Vol. 9. No 5. Pp. 504–512.
10. *Martin R.* (2005), [Statistical Methods for the Enhancement of Noisy Speech. In J. Benesty, S. Makino, and J. Chen, editors, *Speech Enhancement*, chapter 3]. New-York: Springer-Verlag, Pp. 43–64.
11. Recommendation ITU-T P.56. Series P: Terminals and Subjective and Objective Assessment Methods. Objective Measuring Apparatus. Objective Measurement of Active Speech Level. – Geneva: Telecommunication Standardisation Sector of ITU, 2008. Vol. 12. P. 24.
12. Recommendation ITU-T P.863. Series P: Terminals and Subjective and Objective Assessment Methods. Methods for Objective and Subjective Assessment of Speech Quality. Perceptual Objective Listening Quality Assessment. – Geneva: Telecommunication Standardisation Sector of ITU, 2011. Vol. 1. P. 76.
13. Simulation of Acoustic Environments Including the Transmission over Telephone Channels. [On-line]. – Available: <http://dnt.kr.hsnr.de/sireac.html> (1.03.2013).

14. *Jacob Benesty, M. Mohan Sondhi, Yiteng Huang.* (2008), [Springer Handbook of Speech Processing]. Berlin: Springer-Verlag, P. 1176.
15. *VoiceBox: Speech Processing Toolbox for MATLAB.* [On-line]. – Available: <http://www.ee.ic.ac.uk/hp/staff/dmb/> (1.03.2013).
16. *Krivnova O.F.* (2008), [Speech Corpora on New Technological Way]. *Rechevye tehnologii.* no2. Pp. 13–23. (Rus)
17. *Prodeus A.N., Didkovskiy V.S., Didkovskaya M.V.* (2008), [Acoustic Examination of Speech Communication Channels. Monograph]. Kyiv: Imex-Ltd, P. 420. (Rus)
18. *Prodeus A.N.* (2010), [Some Features of the Development of Objective Methods for Speech Intelligibility Measuring]. *Electronics and Communications, tematicheskij vypusk "Elektronika i nanotehnologii".* Vol. 2. Pp. 217–223. (Rus)

Поступила в редакцию 18 февраля 2013 г.