УДК. 004.934

# НЕЭТАЛОННАЯ ОЦЕНКА РАЗБОРЧИВОСТИ ЗАШУМЛЕННЫХ РЕЧЕВЫХ СИГНАЛОВ

***Топников А.И.**, к.т.н., ассистент кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: topnikov@gmail.com;*
***Нестеров М.С.**, аспирант кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: maximator89@mail.ru;*
***Новоселов С.А.**, к.т.н., научный сотрудник отдела верификации и идентификации диктора, «Центр Речевых Технологий», e-mail: novoselov@speechpro.com;*
***Приоров А.Л.**, д.т.н., доцент кафедры динамики электронных систем Ярославского государственного университета им. П.Г. Демидова, e-mail: andcat@yandex.ru.*

*Предлагается методика неэталонной оценки значений показателя разборчивости зашумленной речи SNR loss, а также ее модификация, отличающаяся меньшими вычислительными затратами. Достоверность полученных неэталонных оценок сравнивается путем сопоставления со значениями, полученными с использованием эталонного метода.*

**Ключевые слова:** речевой сигнал, разборчивость, неэталонная оценка, шумоподавление, методика, парная регрессия.

## Введение

Одной из важнейших составляющих акустической экспертизы каналов речевой коммуникации является измерение разборчивости речи [1]. Под разборчивостью понимается степень, с которой определенные речевые единицы (фонемы, слоги, слова, фразы) могут быть верно восприняты слушателем. Различные искажающие факторы (шумы, реверберация, нелинейные искажения) могут привести к значительному снижению разборчивости и потери части информации. Вместе с разборчивостью снижается и качество речевого сигнала [2], однако эти понятия стоит разграничивать.

Наиболее общим подходом при классификации методов оценки разборчивости является разделение их на субъективные и объективные. Отличительная особенность субъективных методов – непосредственное участие дикторов и аудиторов (слушателей) в процессе оценки. Эти методы обладают высокой достоверностью, но ресурсоемки и непригодны для проведения экспертизы каналов коммуникации в режиме реального времени. Поэтому примерно с середины 50-х годов прошлого века активно развиваются объективные методы измерения разборчивости.

В настоящее время выделяются две наиболее крупные группы этих методов: формантные и модуляционные. К первой группе относятся зарубежные методы: AI (Articulation Index) и SII (Speech Intelligibility Index), их предшественники и модификации [1, 3-8], а также отечественные методы, предложенные Ю.С. Быковым, Н.Б. Покровским, М.А. Сапожковым [1, 9-12]. Модуляционные методы представлены различными модификациями метода STI (Speech Transmission Index), включая методы RASTI, STIPA и STITEL [1, 13-15]. Предпринимаются попытки и по объединению двух подходов [16, 17]. Общей чертой указанных выше мер разборчивости является их эталонность – потребность в «чистом» (незашумленном, неискаженном) сигнале, без которого процедура оценки невозможна. Отсутствие таких сигналов в реальных условиях приводит к необходимости разработки неэталонных методов [18–21]. Это одна из наиболее сложных и актуальных задач в рассматриваемой области. До сих пор ни один неэталонный метод не включен в известные стандарты, регламентирующие оценку разборчивости в помещениях или телекоммуникационных системах.

Целью данной работы является исследование возможности получения неэталонных оценок разборчивости зашумленных речевых сигналов с использованием эталонного метода и подсистемы, восстанавливающей эталонный (квазиэталонный) сигнал из искаженного. Для случая, когда основным искажающим фактором является аддитивный шум, роль такой системы может выполнять шумоподавитель, реализуемый по методу [22]. В данной работе рассматривается техника, основанная на применении меры разборчивости речи SNR loss [7] и метода шумоподавления Скалара (Scalart) [23]. Выбранный подход, заключающийся в неэталонной оценке значений уже существующего эталонного метода, позволяет существенно упростить процесс разработки. Так как исходный эталонный метод SNR loss верифицирован (сравнен по достоверности с субъективными оценками), то достоверность предлагаемого неэталонного подхода можно оценивать по близости полученных оценок к истинным значениям SNR loss.

## Описание оригинального метода SNR loss

Метод SNR loss является развитием формантного метода AI и в отличие от него учитывает влияние на разборчивость речи работы нелинейных методов, применяемых для подавления шума в речи. Кратко рассмотрим метод SNR loss, подробно описанный в [7].

Вычисления осуществляются для отдельных временных интервалов (кадров). Для каждой $j$-й критической полосы каждого $m$-го кадра вычисляется следующая величина:

$$L(j,m) = SNR_X(j,m) - SNR_{\hat{X}}(j,m) =$$

$$=10\lg\frac{X^2(j,m)}{D^2(j,m)}-10\lg\frac{\hat{X}^2(j,m)}{D^2(j,m)}=10\lg\frac{X^2(j,m)}{\hat{X}^2(j,m)},$$

где $SNR_X(j,m)$ – входное отношение сигнал/шум (ОСШ) в полосе $j$ и интервале $m$, $SNR_{\hat{X}}(j,m)$ – ОСШ после обработки в полосе $j$ и интервале $m$, $X(j,m)$ – $j$-я составляющая спектра незашумленного (эталонного) сигнала, вычисленного с учетом критических полос для $m$-го временного интервала, $\hat{X}(j,m)$ – $j$-я составляющая спектра обработанного сигнала, вычисленного с учетом критических полос для $m$-го временного интервала, $D(j,m)$ – $j$-я составляющая спектра шума, вычисленного с учетом критических полос для $m$-го временного интервала. Очевидно, что при $\hat{X}(j,m)=X(j,m)$ величина $L(j,m)=0$. В целом значение $L(j,m)$ может быть как положительным, так и отрицательным.

Значения $L(j,m)$ ограничиваются в пределах определенного диапазона $\left[-SNR_{\lim}, SNR_{\lim}\right]$:

$$\hat{L}(j,m)=\min\left(\max\left(L(j,m),-SNR_{\lim}\right),SNR_{\lim}\right).$$

Полученные на предыдущем этапе значения $\hat{L}(j,m)$ масштабируются на интервал [0, 1]:

$$SNR_{LOSS}(j,m)=\begin{cases}-\dfrac{C_-}{SNR_{\lim}}\hat{L}(j,m), & если\ \hat{L}(j,m)<0\\ \dfrac{C_+}{SNR_{\lim}}\hat{L}(j,m), & если\ \hat{L}(j,m)\geq 0,\end{cases}$$

где $C_-$ и $C_+$ – параметры масштабирующей функции.

Для каждого временного интервала осуществляется усреднение $SNR_{LOSS}(j,m)$ по всем критическим полосам:

$$fSNR_{LOSS}(m)=\frac{\sum_{j=1}^{K}W(j)SNR_{LOSS}(j,m)}{\sum_{j=1}^{K}W(j)},$$

где $W(j)$ – весовая функция, учитывающая психоакустические закономерности восприятия речевых сигналов.

Вычисляется среднее значение $\overline{SNR}_{LOSS}$ путем усреднения $fSNR_{LOSS}(m)$ по всем окнам:

$$\overline{SNR}_{LOSS}=\frac{1}{M}\sum_{m=0}^{M-1}fSNR_{LOSS}(m),$$

где $M$ – число окон, на которые разделен сигнал. Получившаяся величина, для удобства обозначенная как SNRloss, изменяется в интервале от 0 до 1. Нулевое значение соответствует идеальной разборчивости, единичное – ее полному отсутствию. Графики, характеризующие взаимосвязь разборчивости в процентах со значениями SNRloss, представлены в [7].

Стоит пояснить, что в тексте для удобства используется два варианта написания: SNR loss – для названия метода и SNRloss – для обозначения значений, полученных при помощи этого метода.

**Методика неэталонной оценки разборчивости**

Идея, лежащая в основе предлагаемой неэталонной версии критерия SNR loss, состоит в том [22], чтобы вычислять значение SNRloss для оцениваемого (зашумленного) сигнала и сигнала, полученного шумоподавлением с использованием метода, предложенного Скаларом и Филхо в работе [23]. При этом сигнал на выходе шумоподавителя считается чистым (эталонным). В данном случае это допустимо, так как сигнал на выходе шумоподавителя является оценкой чистого (незашумленного) сигнала.

Проанализируем взаимосвязь значений SNRloss, вычисленных таким образом, (обозначим их как SNRloss') с истинными значениями SNRloss, полученными при использовании чистого сигнала в качестве эталона. Для моделирования использовалось шесть незашумленных речевых фрагментов, каждый из которых зашумлялся аддитивным белым гауссовским шумом (АБГШ). Для каждого фрагмента формировались 230 зашумленных версий: 5 реализаций шума для каждого ОСШ из диапазона от –15 до 30 дБ [22]. В ходе моделирования значения SNRloss измерялись для 1380 сигналов (6 фраз по 230 вариантов зашумления каждая). Диаграмма рассеяния между истинными значениями SNRloss и значениями SNRloss' представлена на рис. 1. Видно, что при изменении ОСШ в пределах от –15 до 30 дБ (рис. 1) истинное (т.е. измеренное эталонно – с использованием чистого сигнала) значение критерия SNRloss изменяется в пределах примерно от 0,4 до 1. Множество точек на рис. 1 распределено таким образом, что позволяет сделать предположение о возможности линейной аппроксимации зависимости SNRloss от SNRloss'. На основе имеющихся данных и метода наименьших квадратов взаимосвязь может быть описана уравнением парной линейной регрессии:

$$\mathrm{SNRloss}=b_1*\mathrm{SNRloss'}+b_0. \quad (1)$$

Установлено, что для данного типа шума коэффициенты уравнения (1) принимают следующие значения: $b_1$=0,8909; $b_0$=0,043. Даже при использовании линейной регрессии наблюдается высокое значение коэффициента детерминации (выше 0,99), характеризующего связь между истинными значениями SNRloss и неэталонными оценками SNRloss, полученными путем подстановки значений SNRloss' в (1).

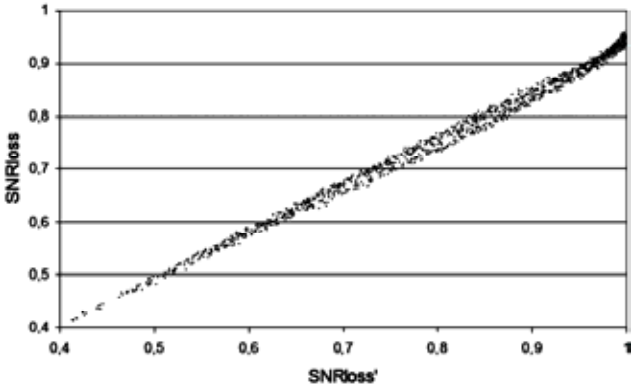


*Рис. 1. Диаграмма рассеяния значений SNRloss' и SNRloss для АБГШ*

Применение степенной или полиномиальной функций для описания исследуемой зависимости приводит к дополнительному незначительному повышению точности аппроксимации. Выбор в пользу линейной функции сделан с учетом возможности достижения высокой точности и сохранения простоты.

Для проверки предложенной неэталонной методики оценки значений SNRloss проведено дополнительное моделирование. Для обеспечения достоверности выбран набор речевых фрагментов, отличный от того, который использовался в первой части исследований. Для 24 речевых фрагментов было измерено в общей сложности 5520 пар значений SNRloss и SNRloss' (для разных реализаций шума и ОСШ). Полученные значения SNRloss' подставлялись в выражение (1). Таким образом было найдено 5520 неэталонных оценок значений SNRloss. Сравнение неэталонных оценок с истинными значениями критерия, полученными эталонным методом, позволяет судить о достаточно высокой точности предложенного метода неэталонной оценки (рис. 2).

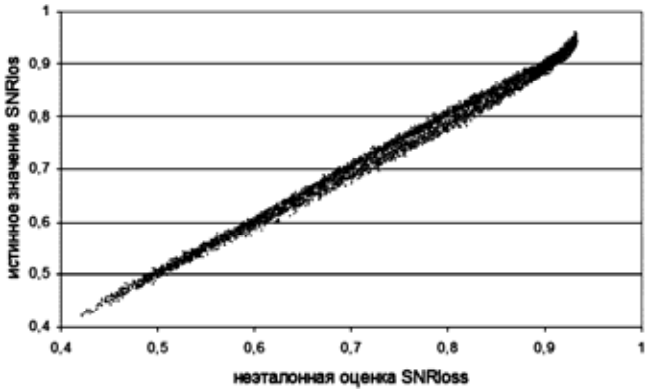


*Рис. 2. Диаграмма рассеяния истинных значений SNRloss и их неэталонных оценок для АБГШ*

При использовании линейной функции средняя величина абсолютной ошибки составляет примерно 0,008, а ее максимальная величина равна 0,036. Средняя величина относительной ошибки составляет примерно 1,05 %, а ее максимальное значение равно 4,72 %. Стоит отметить, что величины ошибок могут быть дополнительно снижены благодаря применению полиномиальной или кусочно-линейной аппроксимаций. Полученные результаты свидетельствуют о перспективности применения предложенного подхода к неэталонной оценке значений показателя SNR loss.

Следующим шагом исследования является проверка применимости данного подхода для других видов шумов, наиболее часто встречающихся при решении практических задач. С этой целью использовалась специализированная база шумов Noisex-92, состоящая из записей реальных шумов различных типов. Так как частота дискретизации тестовых речевых сигналов составляла 8 кГц, то частота дискретизации записей шумов также была понижена до 8 кГц. В работе приводятся оригинальные названия записей шумов на английском языке.

Для осуществления неэталонной оценки значений SNRloss в случае воздействия на сигнал шума произвольного типа возможно три подхода: использование регрессионного выражения (1) с параметрами, полученными для АБГШ (подход 1); использование этого регрессионного выражения с параметрами, полученными для смешанной выборки, состоящей из речевых сигналов, зашумленных шумами разных типов (подход 2); нахождение новых параметров зависимости (1), индивидуальных для каждого конкретного типа шума (подход 3). Первые два подхода более универсальны и удобны на практике, однако третий подход потенциально способен обеспечить большую точность.

Для реализации второго и третьего подхода необходима оценка новых коэффициентов уравнения (1). При нахождении коэффициентов для реализации второго подхода использовалась смешанная выборка, состоящая из речевых сигналов, искаженных шумами различных типов. Установлено, что в этом случае коэффициенты уравнения (1) принимают следующие значения: $b_1$=0,8879; $b_0$=0,0479. При этом наблюдается высокое значение коэффициента детерминации $R^2$=0,9927.

Для реализации третьего подхода необходимо найти коэффициенты уравнения регрессии, индивидуальные для каждого типа шума. Как и в случае с АБГШ, линейная регрессия хорошо подходит для описания взаимосвязи между значениями SNRloss и SNRloss'. Для нахождения коэффициентов проведено моделирование. Для каждого типа шума формировалось 1380 сигналов (6 фраз по 230 вариантов зашумления каждой фразы). На основе этих выборок для каждого типа шума найдены коэффициенты линейной регрессии, связывающей значения SNRloss и SNRloss', а также оценено среднеквадратическое отклонение (СКО) и коэффициент детерминации $R^2$ (табл. 1). Значения коэффициента детерминации, близкие к единице, свидетельствуют о применимости линейной модели и тесной взаимосвязи между значениями SNRloss и SNRloss'. Установлено, что наибольшее СКО и наиболее низкое значение коэффициента детерминации наблюдается в случае нахождения параметров линейной регрессии для шума «Speech babble».

Далее проводилось сравнение указанных подходов по точности. С этой целью для каждого типа шума, как и в случае с АБГШ, для 24 фраз измерялось по 5520 пар значений SNRloss и SNRloss' (для разных реализаций шума и ОСШ). Полученные значения SNRloss' подставлялись в выражение (1) с соответствующими коэффициентами, зависящими от применяемого подхода. После этого полученные оценки значений SNRloss сравнивались с истинными значениями, полученными при помощи эталонного метода. Вычислялись значения средней абсолютной ошибки, корня из среднеквадратической ошибки, относительной ошибки, выраженной в процентах:

$$\mathrm{MAE}=\frac{1}{\mathrm{N_{зн}}}\sum_{l=1}^{\mathrm{N_{зн}}}\left|\mathrm{SNRloss}_l^{оцен}-\mathrm{SNRloss}_l^{ист}\right|,$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{\mathrm{N_{зн}}}\sum_{l=1}^{\mathrm{N_{зн}}}\left(\mathrm{SNRloss}_l^{оцен}-\mathrm{SNRloss}_l^{ист}\right)^2},$$

$$\mathrm{MAPE}=\frac{100\,\%}{\mathrm{N_{зн}}}\sum_{l=1}^{\mathrm{N_{зн}}}\frac{\left|\mathrm{SNRloss}_l^{оцен.}-\mathrm{SNRloss}_l^{ист.}\right|}{\mathrm{SNRloss}_l^{ист.}},$$

*Таблица 1. Параметры линейной регрессии для взаимосвязи SNRloss и SNRloss'*

| Тип шума | $b_1$ | $b_0$ | *СКО* | $R^2$ |
|---|---|---|---|---|
| АБГШ (AWGN) | 0,891 | 0,043 | 0,010 | 0,995 |
| Car interior noise | 0,894 | 0,039 | 0,010 | 0,997 |
| Destroyer engine room noise | 0,892 | 0,042 | 0,015 | 0,991 |
| Destroyer operations room noise | 0,879 | 0,056 | 0,011 | 0,995 |
| F-16 cockpit noise | 0,891 | 0,040 | 0,011 | 0,995 |
| Factory floor noise 1 | 0,877 | 0,068 | 0,015 | 0,990 |
| Factory floor noise 2 | 0,872 | 0,057 | 0,012 | 0,995 |
| Jet cockpit noise 1 | 0,891 | 0,042 | 0,010 | 0,995 |
| Jet cockpit noise 2 | 0,889 | 0,046 | 0,010 | 0,995 |
| Military vehicle noise | 0,870 | 0,057 | 0,011 | 0,996 |
| Speech babble | 0,889 | 0,070 | 0,022 | 0,982 |
| Tank noise | 0,869 | 0,055 | 0,010 | 0,997 |

*Таблица 2. Точность неэталонной оценки значений SNRloss при использовании разных подходов*

| Тип шума | *Средняя относительная ошибка, %* | | |
|---|---|---|---|
| | *Подход 1* | *Подход 2* | *Подход 3* |
| АБГШ (AWGN) | 1,05 | 1,07 | 1,05 |
| Car interior noise | 1,87 | 2,43 | 1,60 |
| Destroyer engine room noise | 1,58 | 1,65 | 1,57 |
| Destroyer operations room noise | 1,27 | 1,17 | 1,14 |
| F-16 cockpit noise | 1,18 | 1,34 | 1,13 |
| Factory floor noise 1 | 2,11 | 1,87 | 1,56 |
| Factory floor noise 2 | 1,42 | 1,40 | 1,28 |
| Jet cockpit noise 1 | 1,06 | 1,12 | 1,06 |
| Jet cockpit noise 2 | 1,09 | 1,09 | 1,08 |
| Military vehicle noise | 1,35 | 1,36 | 1,17 |
| Speech babble | 3,55 | 3,22 | 2,20 |
| Tank noise | 1,43 | 1,48 | 1,26 |
| В среднем | 1,58 | 1,60 | 1,34 |

где $\mathrm{SNRloss}_l^{оцен}$ – неэталонная оценка значения SNRloss для $l$-го сигнала, $\mathrm{SNRloss}_l^{ист}$ – истинное значение SNRloss для $l$-го сигнала, $N_{зн.}$ – общее число сигналов, для которых измеряются значения SNRloss.

Результаты, полученные с использованием моделирования, позволяют сравнить точность предложенных подходов (табл. 2). Как и ожидалось, третий подход обеспечивает наилучшую точность. Первый и второй подходы обладают сопоставимой точностью. Версию метода, использующую индивидуальные коэффициенты уравнения регрессии, стоит использовать в случаях, когда есть возможность оценить коэффициенты уравнения (1), соответствующие конкретным шумовым условиям. В случаях, когда шумовые условия заранее не известны или тип шума может существенно меняться с течением времени, применение третьего подхода невозможно и рекомендуется применять первый или второй подход.

**Дальнейшее совершенствование предложенной методики**

После того как моделирование с различными типами шумов подтвердило жизнеспособность предлагаемой неэталонной методики, встал вопрос ее дальнейшего совершенствования. Анализ количества операций сложения и умножения, необходимых для неэталонной оценки разборчивости речевого сигнала, показал, что значительная их часть приходится на быстрое преобразование Фурье (БПФ) и обратное БПФ (ОБПФ) [24]. Однако используемая схема вычислений (рис. 3) содержит операции, которые можно исключить. При этом крайне важным является приведение параметров спектрального преобразования в блоке шумоподавления к тем, которые используются в методе SNR loss, так как они влияют на достоверность оценки разборчивости. Схема предлагаемой модификации, основанной на исключении повторных и взаимоисключающих операций, имеет более простую структуру и содержит один блок вычисления БПФ (рис. 4).

Предложенная модификация проанализирована и сравнена с исходной методикой по числу требуемых вычислительных операций. Выявлено, что предложенная модификация позволяет уменьшить количество операций примерно на 30,5 % по сравнению с показателями исходной методики [24].

Также проводилось сравнение скорости работы двух версий методик на основе моделирования в среде Matlab. Для этого использовался персональный компьютер следующей конфигурации: Intel(R) Pentium(R) D 930 CPU 3.00GHz, RAM 3,00GB 400MHz, 64-bit operating system Windows 8. В результате выигрыш предлагаемой модификации по времени составил 29,9 % [24].

Как показало моделирование, предложенная модификация обладает несколько меньшей точностью. Результаты сравнения по показателям MAE, RMSE и MAPE представлены на рис. 5. На графиках приводятся усредненные по всем типам шумов характеристики.

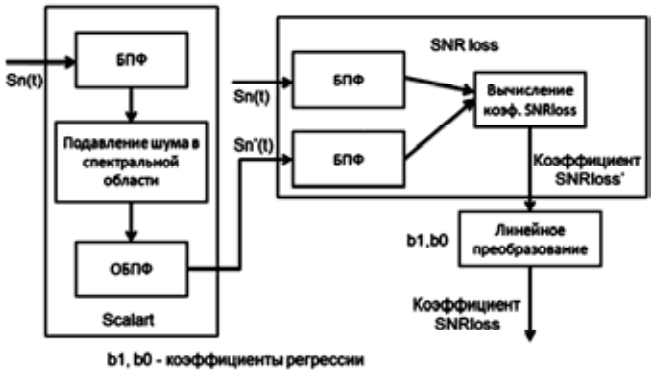


*Рис. 3. Блок-схема исходной неэталонной методики оценки разборчивости*

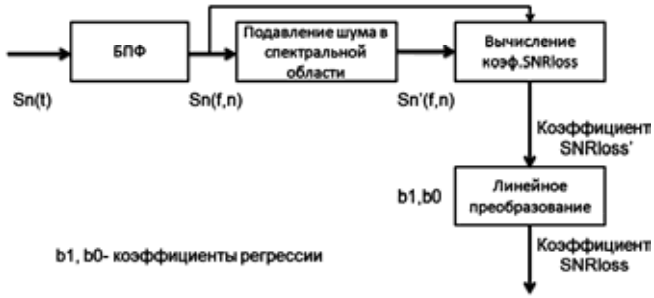


*Рис. 4. Блок-схема модифицированной методики*

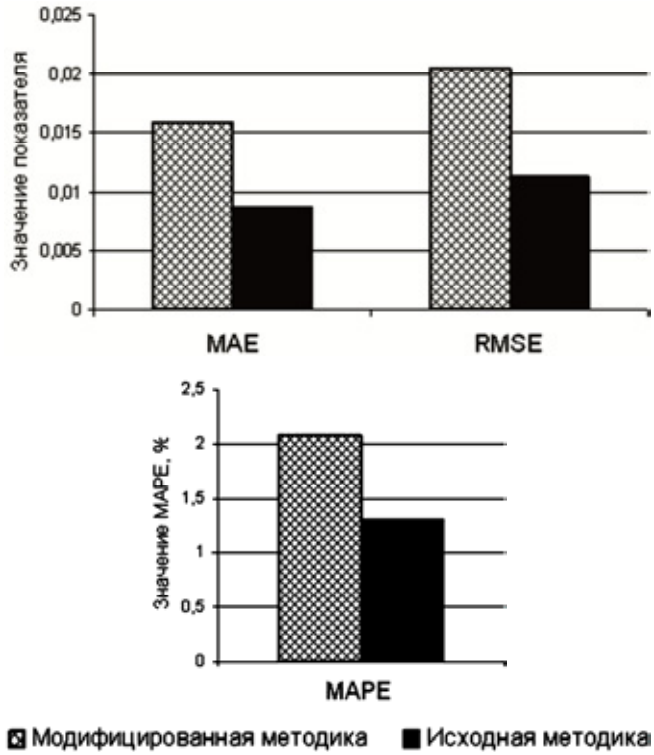


*Рис. 5. Сравнение точности методик на основе показателей MAE, RMSE и MAPE*

Представленные диаграммы демонстрируют, что точность оценок предложенной модификации несколько уступает точности исходной методики. Отношение показателей точности для исходной методики и ее модификации не столь информативны, как абсолютное изменение этих показателей. Например, средняя ошибка, выраженная в процентах (MAPE), выросла примерно с 1,3 % до 2,1 %. Относительное изменение в 1,6 раза кажется значительным, однако на практике указанное ухудшение характеристик вполне приемлемо. Кроме того, не исключено, что дальнейшая доработка предлагаемой модификации позволит ещё сильнее приблизиться по точности к исходной методике.

## Заключение

Таким образом, предложена неэталонная версия меры разборчивости зашумленных речевых сигналов SNR loss, базирующаяся на применении исходной (эталонной) версии метода SNR loss, метода шумоподавления и парной регрессии. Исследована точность предложенной неэталонной меры разборчивости в случае воздействия на речевые сигналы шумов различных типов. Полученные результаты свидетельствуют об относительно высокой точности предложенной методики неэталонной оценки (средняя относительная ошибка составляет 1,05–3,55%).

Также представлена более быстрая версия неэталонной методики. Из результатов исследований следует, что модификация имеет достаточно высокую точность, которая незначительно уступает оригиналу, при этом скорость работы возросла за счет значительного сокращения числа осуществляемых вычислительных операций. Предложенная модифицированная методика может применяться для автоматического управления системами шумоподавления, а так же для выбора режима передачи, обеспечивающего допустимый уровень разборчивости, в системах связи.

Часть результатов, представленных в работе, получена при поддержке гранта РФФИ № 14-07-31335.

## Литература

1. Дидковский В.С., Дидковская М.В., Продеус А.Н. Акустическая экспертиза каналов речевой коммуникации. Монография. – Киев: Имэкс-ЛТД, 2008. 420 с.

2. Новоселов С.А., Топников А.И., Савватин А.И., Приоров А.Л. Подавление шума в речевых сигналах на основе метода нелокального усреднения // Цифровая обработка сигналов. 2011. №4. – С. 23-28.

3. Collard J. A theoretical study of the articulation and intelligibility of a telephone circuit // Electrical Communication. 1929. V. 7. – P. 168.

4. French N.R., Steinberg J.C. Factors governing the intelligibility of speech sounds // The journal of the Acoustical Society of America. 1947. V. 19. Is. 1. – P. 90-119.

5. Kryter K.D. Methods for the calculation and use of the articulation index // The Journal of the Acoustical Society of America. 1962. V. 34. Is. 11. – P. 1689-1697.

6. Kryter K.D. Validation of the articulation index // The Journal of the Acoustical Society of America. 1962. V. 34. Is. 11. – P. 1698-1702.

7. Ma J., Loizou P. SNR loss: a new objective measure for predicting the intelligibility of noise-suppressed speech // Speech Communication. 2011. V. 53. Is. 3. – P. 340-354.

8. Loizou P., Ma J. Extending the articulation index to account for non-linear distortions introduced by noise-suppression algorithms // Journal of the Acoustical Society of America. 2011. V. 130. Is. 2. P. 986–995.

9. Быков Ю.С. Теория разборчивости речи и повышение эффективности радиотелефонной связи. – М.-Л.: Госэнергоиздат, 1959. 350 с.

10. Покровский Н.Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. 390 с.

11. Сапожков М.А. Речевой сигнал в кибернетике и связи. – М.: Связьиздат, 1963. 452 с.

12. Сапожков М.А., Михайлов В.Г. Вокодерная связь. – М.: Радио связь, 1983. – 248 с.

13. Houtgast T., Steeneken H.J.M. Evaluation of speech transmission channels by using artificial signals // Acta Acustica united with Acustica. 1971. V. 25. Is. 6. P. 355-367.

14. Steeneken H.J.M., Houtgast T. A physical method for measuring speech transmission quality // The Journal of the Acoustical Society of America. 1980. V. 67. Is. 1. P. 318-326.

15. Steeneken H.J.M., Houtgast T. Validation of the revised STIr method // Speech Communication. 2002. V. 38. Is. 3. – P. 413-425.

16. Prodeus A. On possibility of advantages join of formant and modulation methods of speech intelligibility evaluation // Proceedings of VII International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH). 2010. P. 254-259.

17. Prodeus A. Assessment of Speech Intelligibility by Formant-Modulation Method // Journal of Basic and Applied Physics. 2013. V. 2 Is. 5. P. 10-18.

18. Falk T.H., Zheng C., Chan W.Y. A non-intrusive quality and intelligibility measure of reverberant and dereverberated speech // IEEE Transactions on Audio, Speech, and Language Processing. 2010. V. 18. Is. 7. – P. 1766-1774.

19. Santos J.F., Senoussaoui M., Falk T.H. An improved non-intrusive intelligibility metric for noisy and reverberant speech // 14th International Workshop on Acoustic Signal Enhancement (IWAENC). 2014. – P. 55-59.

20. Li F.F. Speech intelligibility of VoIP to PSTN interworking – a key index for the QoS. // Proceedings of International Conference «Telecommunications Quality of Service: The Business of Success». 2004. – P. 104-108.

21. Chen F., Hazrati O., Loizou P. C. Predicting the intelligibility of reverberant speech for cochlear implant listeners with a non-intrusive intelligibility measure // Biomedical signal processing and control. 2013. V. 8. Is. 3. – P. 311-314.

22. Веселов И.А., Новоселов С.А., Топников А.И. Методика неэталонной оценки разборчивости зашумленных речевых сигналов // 15-я международная конференция «Цифровая обработка сигналов и её применение – DSPA-2013». М., 2013. Т. 1. – С. 256-259.

23. Scalart P., Filho J. Speech enhancement based on a priori signal to noise estimation // IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96). 1996. V. 2. – P. 629-632.

24. Нестеров М.С., Топников А.И. Модификация методики неэталонной оценки разборчивости речевых сигналов // 16-я международная конференция «Цифровая обработка сигналов и её применение – DSPA-2014». М., 2014. Т. 1. – С. 208-212.

## NON-INTRUSIVE ESTIMATION OF NOISY SPEECH SIGNAL INTELLIBILITY

***Topnikov A., Nesterov M., Novoselov S., Priorov A.***

It is proposed non-intrusive technique for estimation of noisy speech signals intelligibility. Technique is based on the modern intelligibility measure SNR loss and Scalart speech denoising method. It is provided a comparison of the accuracy of the proposed technique with accuracy of the original reference SNR loss method for speech signals which are noised by different types of noise. The modification of proposed technique is described also that spent allows to reduce the number of computational operations.

## НОВЫЕ КНИГИ

**Вашкевич М.И., Азаров И.С., Петровский А.А.**

**Косинусно-модулированные банки фильтров с фазовым преобразованием: реализация и применение в слуховых аппаратах: научн. изд. – М.: Горячая линия – Телеком, 2014. – 212 с.**

Рассмотрены вопросы построения цифровых слуховых аппаратов. В общем виде описаны технические аспекты коррекции нарушений. Большая часть материала посвящена разработке и эффективной реализации неравнополосных банков фильтров, играющих ключевую роль при проектировании цифровых слуховых аппаратов. Рассмотрены вопросы построения алгоритмов адаптивной компенсации акустической обратной связи, компрессии и шумоподавления, необходимых для реализации основных функций слухового аппарата.

Для научных и инженерно-технических работников, специализирующихся в области проектирования устройств цифровой обработки сигналов, биомедицинской техники, а также студентов и аспирантов соответствующих специальностей.