

УДК 621.395

СПОСОБЫ ОБНАРУЖЕНИЯ ПАУЗ НА ОСНОВЕ НЕЛИНЕЙНОЙ ОБРАБОТКИ С ИСПОЛЬЗОВАНИЕМ ВСПОМОГАТЕЛЬНЫХ СИГНАЛОВ

В. А. Волченков, к.т.н., доцент кафедры ТОР РГРТУ, Рязань, Россия;
orcid.org/0000-0002-6656-3418, e-mail: volchenkov.rzn@yandex.ru

Рассматривается задача детектирования пауз, так как одной из актуальных проблем радиомониторинга и детектирования активности речи остается повышение вероятности правильного обнаружения относительно коротких пауз длительностью менее 40 мс, что позволит повысить вероятность правильного детектирования пауз и таким образом увеличить эффективность использования канального ресурса. Целью работы является повышение вероятности правильного обнаружения и выделения пауз на основе сегментации с применением вспомогательных сигналов.

Ключевые слова: обнаружение пауз, детектор активности речи.

DOI: 10.21667/1995-4565-2023-86-32-37

Введение

Рассматривается решение задачи повышения вероятности правильного обнаружения и выделения пауз на основе сегментации с применением вспомогательных сигналов. Приводится информация о разработанных способах обнаружения пауз и их модификаций, а также результаты исследования их работы при различных параметрах и сравнение с некоторыми существующими способами обнаружения пауз. На речевые сигналы накладывается транспортный шум с разным уровнем ОСШ, и в них производится обнаружение пауз.

Теоретическая часть

Предлагаемые способы обнаружения пауз основаны на использовании идеального амплитудного ограничителя (ИАО) при обработке сигнала. При этом учитываются следующие его свойства:

1. Для ИАО свойственно приблизительно линейное изменение ОСШ на выходе в зависимости от изменения ОСШ на входе. При идеальном ограничении ОСШ на выходе ограничителя отличается от ОСШ на его входе в $(0,7854 \div 2)$ раза [1].

2. На выходе ИАО суммарная мощность сигнала и шума остается всегда постоянной, несмотря на изменения ОСШ на входе [1].

Данное свойство может быть записано как

$$\sigma^2 + U_{\text{зо}}^2 = C^2 = \text{const}, \quad (1)$$

где σ^2 – мощность сигнала шума на выходе ИАО, $U_{\text{зо}}^2$ – мощность звукового сигнала на выходе ИАО.

3. Из (1) следует, что с увеличением амплитуды звука речевого сигнала $U_{\text{зо}}$ второй сигнал, обозначенный σ , будет уменьшаться. При прохождении через ИАО двух сигналов сильный сигнал подавляет слабый.

При добавлении вспомогательного детерминированного сигнала $\varphi_i(t)$ с амплитудой $U_{\text{вс}}(t)$ и мощностью $U_{\text{вс}}^2(t)$ такого, что

$$U_{\text{вс}}^2(t) \gg \sigma^2,$$

в паузах, в отсутствие звукового сигнала и при наличии только фонового шума $n(t)$ мощностью σ^2 станет возможно выделить вспомогательный сигнал $\varphi_i(t)$ с выхода идеального амплитудного ограничителя.

В случае присутствия активного звукового сигнала, так как

$$U_{30}^2(t) \gg U_{вс}^2(t),$$

в свою очередь звуковой сигнал $s_{30}(t)$ будет выделен с выхода идеального амплитудного ограничителя.

Сигнал с выхода идеального амплитудного ограничителя будет состоять из суммы сигналов звука, имеющих мощность $U_{30}^2(t)$, и вспомогательных сигналов, имеющих мощность $U_{вс}^2(t)$, причем суммарная мощность указанных сигналов равна постоянной величине:

$$U_{30}^2(t) + U_{вс}^2(t) = C^2.$$

Задавая $C = 1$, получаем ортонормированную систему

$$U_{30}^2(t) + U_{вс}^2(t) = 1,$$

в которой неизвестной является длительность сигналов, которую нужно определить.

Структурная схема, реализующая способ, производящий обнаружение пауз посредством введения дополнительного вспомогательного сигнала, приведена на рисунке 1 [2].

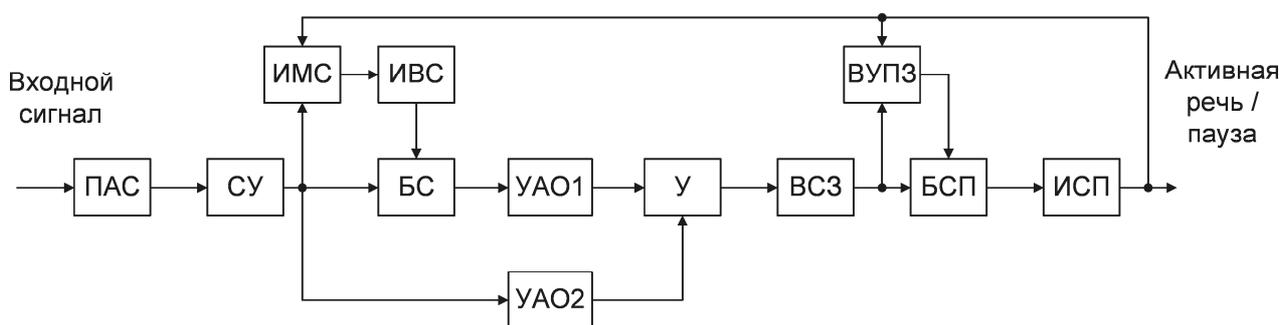


Рисунок 1 – Способ, производящий обнаружение пауз посредством введения дополнительного вспомогательного сигнала
Figure 1 – Block diagram implementing a method that detects pauses by introducing an additional auxiliary signal

В состав устройства входят: приёмник акустического сигнала (ПАС), селективный усилитель (СУ), измеритель мощности сигнала (ИМС), источник вспомогательного сигнала (ИВС), блок сложения (БС), усилители – амплитудные ограничители (УАО 1) и (УАО 2), умножитель (У), вычислитель среднего значения (ВСЗ), вычислительное устройство порогового значения (ВУПЗ), блок сравнения с порогом (БСП), источник сигнала «пауза» (ИСП).

С помощью селективного усилителя происходит усиление речевого сигнала с выхода приёмника акустического сигнала. Сигнал с выхода селективного усилителя подается на измеритель мощности сигнала и вход блока сложения. Вспомогательный сигнал с выхода источника вспомогательного сигнала подаётся на второй вход блока сложения. Так как на интервале первых 100 мс активный сигнал обычно отсутствует, то в течение этого времени в измерителе мощности сигнала происходит вычисление значения необходимой мощности вспомогательного сигнала. Если устройство обнаружения пауз принимает решение о присутствии паузы у подряд идущих фреймов на интервале 500 мс, то значение необходимой мощности вспомогательного сигнала снова пересчитывается. Сигнал с выхода блока сложения поступает на вход усилителя – амплитудного ограничителя 1, где сигнал усиливается и ограничивается по амплитуде. Сигнал с выхода блока селективного усилителя поступает на вход усилителя – амплитудного ограничителя 2, где сигнал усиливается и ограничивается по амплитуде. На первый вход умножителя подается сигнал с выхода идеального амплитудного ограничителя 1, а на второй – с выхода идеального амплитудного ограничителя 2. С выхода умножителя сигнал поступает на вход вычислителя среднего значения сигнала, которое затем сравнивается в блоке сравнения с порогом, предварительно установленным с помощью вычислительного устройства порогового значения путем анализа интервала длительностью 50 мс после изменения мощности вспомогательного сигнала. С выхода блока сравнения с

порогом сигнал поступает на вход источника сигнала «пауза», который в случае принятия решения о присутствии паузы формирует сигнал «пауза» длительностью 10 мс.

Ниже приведены примеры сигналов с выходов некоторых блоков в случае присутствия активного звукового сигнала (а) и паузы (б) (рисунок 2).

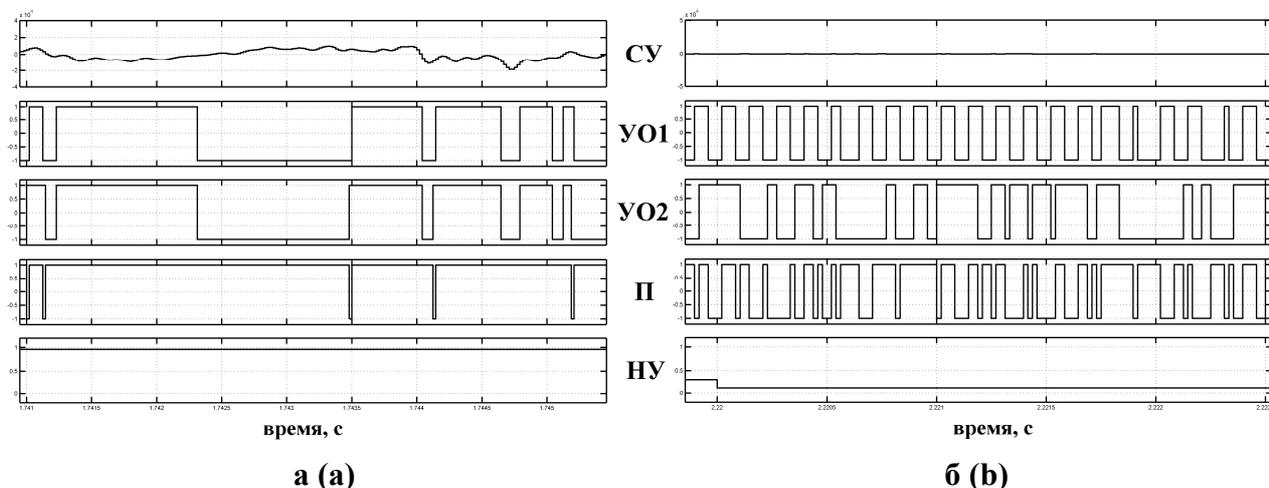


Рисунок 2 – Примеры сигналов с выходов некоторых блоков в случае присутствия активного звукового сигнала (а) и паузы (б)
Figure 2 – Examples of signals from the outputs of some blocks in the presence of active audio signal (a) and pause (b)

На основе вышеприведённого способа был разработан способ обнаружения пауз, базирующийся на применении двух вспомогательных сигналов. Схема реализации данного способа приведена на рисунке 3.

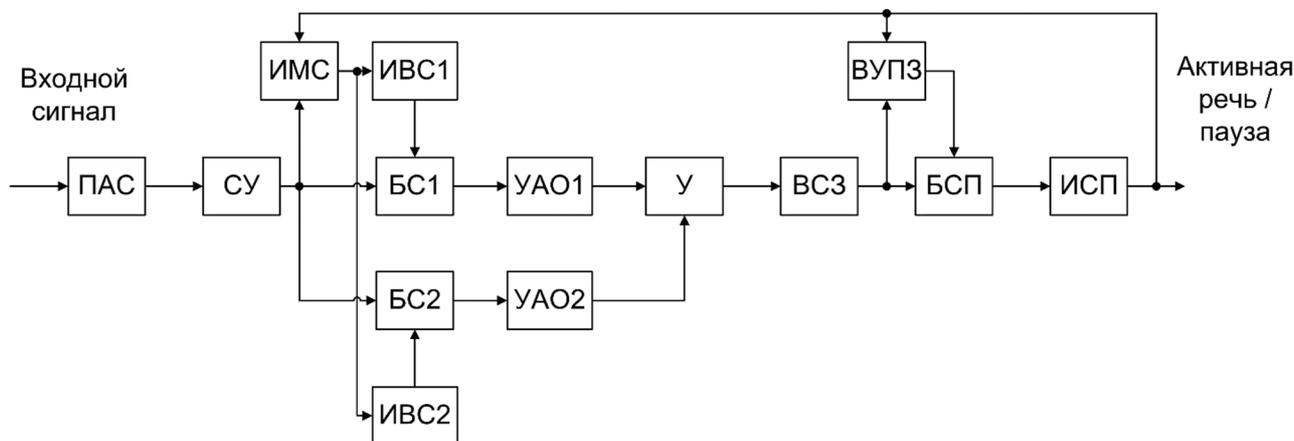


Рисунок 3 – Структурная схема, реализующая способ, производящий обнаружение пауз посредством введения двух дополнительных вспомогательных сигналов
Figure 3 – Block diagram implementing a method that detects pauses by introducing two additional auxiliary signals

В состав устройства помимо элементов, приведённых в первом способе, включены источник вспомогательного сигнала (ИВС 2) и блок сложения (БС 2).

Принцип работы предлагаемого способа схож с приведенным в начале способом с тем лишь отличием, что вместо одного измерительного сигнала используются два; частота измерительного сигнала с ГИС 2 в 2 раза выше частоты измерительного сигнала, поступающего с ГИС 1.

Для увеличения эффективности второго способа при заданном критерии обнаружения были разработаны 2 модификации:

- с добавлением 10 мс активности в конце активного участка речи;
- с добавлением 10 мс активности в конце активного участка речи, а также в начале (с введением 10 мс задержки).

Экспериментальные исследования

Результаты сравнения эффективности приведённых выше способов с алгоритмом VAD на основе отношения правдоподобия (VAD LR – Likelihood-Ratio-Based VAD) [3,4], а также со способом обнаружения пауз в кодере G.729B [4-6] приведены на рисунках 4 и 5. В качестве тестового сигнала разными дикторами был записан речевой сигнал длительностью 108 секунд. Затем на запись было произведено воздействие транспортным шумом с ОСШ от 5 до 25 дБ с шагом 5 дБ [3].

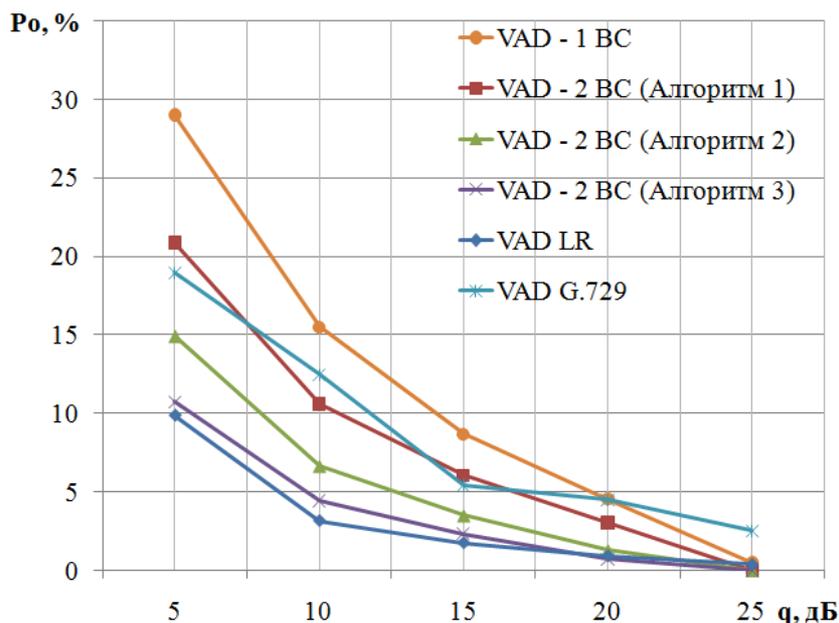


Рисунок 4 – Сравнение ошибки определения речи в условиях воздействия транспортного шума с разным уровнем ОСШ на речевой сигнал в результате обработки рассматриваемыми алгоритмами VAD
Figure 4 – Comparison of speech detection error under the conditions of exposure to vehicle noise with different SNR levels on speech signal as a result of processing by considered VAD algorithms

Стоит отметить, что так как разработанные способы производят обнаружение пауз посредством введения дополнительных вспомогательных сигналов, то с ростом мощности шума растёт и мощность самих дополнительных вспомогательных сигналов, которые тем самым подавляют шум. Вследствие этого с ростом мощности шума ошибка определения пауз становится ниже.

Исследовалось восприятие речевого сигнала на слух после обработки разработанными способами обнаружения пауз. Для этого после обработки был записан аудиосигнал блоком аудиозаписи, управление которым осуществляли предлагаемые обнаружители пауз. Временные окна, которые способы обнаружения пауз определяли как речь, записывались без изменений. Во фреймы, которые способы обнаружения пауз определяли как паузы, записывался ноль. Качество записанного речевого сигнала ухудшилось. Слова были различимы, но воспринимались на слух значительно тяжелее. Во втором случае в интервалы, которые способы обнаружения пауз определяли как паузы, был записан комфортный шум с параметрами шума из исходной записи. Полученный аудиосигнал по качеству соответствовал исходному. Слова были хорошо различимы. Речь в записанном аудиосигнале на слух можно было воспринять достаточно легко.

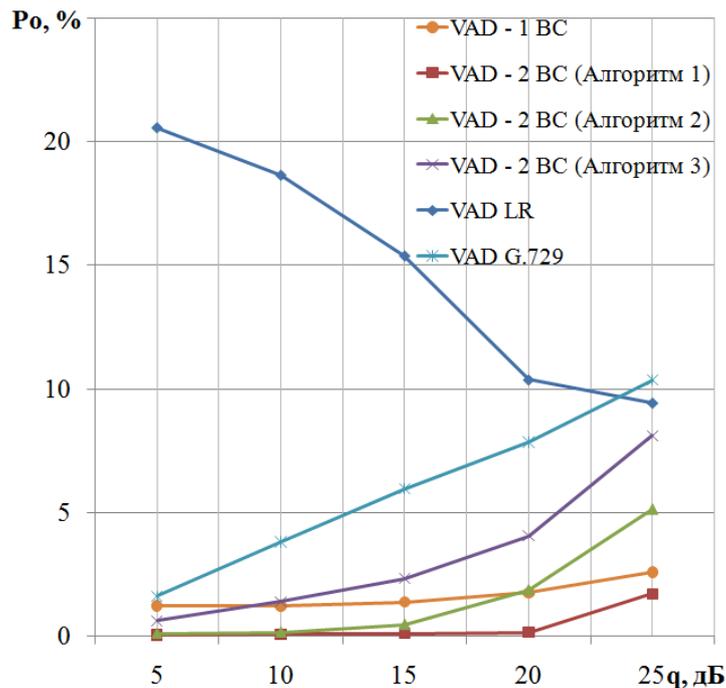


Рисунок 5 – Сравнение ошибки определения пауз в условиях воздействия транспортного шума с разным уровнем ОСШ на речевой сигнал в результате обработки рассматриваемыми алгоритмами VAD
Figure 5 – Comparison of silence detection error under the conditions of exposure to vehicle noise with different SNR levels on speech signal as a result of processing by the considered VAD algorithms

Заключение

В статье предложены способы обнаружения пауз и их модификации, а также результаты исследования их работы при различных параметрах в сравнении с некоторыми существующими способами обнаружения пауз. Представлены способы и их модификации, позволяющие повысить вероятность правильного детектирования пауз и таким образом увеличить эффективность использования канального ресурса.

Библиографический список

1. **Davenport W.B., Jr.** Signal-to-Noise Ratios in Bandpass Limiters // Journal of Applied Physics, Volume 24, Issue 6, p.720-727.
2. **Витязев В.В., Розов В.И., Волченков В.А.** Пат. 2436173 Российская Федерация, МПК G10L 15/00, G10L 11/02, Способ обнаружения пауз в речевых сигналах и устройство, его реализующее; заявитель и патентообладатель Рязанский государственный радиотехнический университет. – № 2010124342/08, заяв. 15.06.10; опубл. 10.12.11, Бюл. 34.
3. **Sohn J., Kim N. S. and Sung W.** A statistical model-based voice activity detection // IEEE Signal Processing Letters. 1999, 6(1):1–3.
4. **Kondoz A.M.** Digital Speech. Coding for Low Bit Rate Communication Systems. John Wiley & Sons, Ltd. 2004. 442 p.
5. ITU-T (1996) A silence compression scheme for G.729 optimised for terminals conforming to ITU-T V.70, ITU-T Rec. G.729 Annex B.
6. ITU-T (1996) Coding of speech at 8 kbit/s using conjugate-structure algebraiccode excited linear prediction (CS-ACELP), ITU-T Rec. G.729.

UDC 621.395

METHODS FOR DETECTING PAUSES BASED ON NONLINEAR PROCESSING USING AUXILIARY SIGNALS

V. A. Volchenkov, Ph.D. (in technical sciences), associate professor, Department of Telecommunications and Radio Engineering Fundamentals, RSREU, Ryazan, Russia;
orcid.org/0000-0002-6656-3418, e-mail: volchenkov.rzn@yandex.ru

*The problem of detecting pauses is considered, as one of the urgent problems of radio monitoring and detecting speech activity is still to increase the probability for correct detection of relatively short pauses lasting less than 40 ms, which will increase the probability of correct detection of pauses thus increasing the efficiency of using channel resource. **The aim of the work is to increase the probability of correct detection and allocation of pauses based on segmentation using auxiliary signals.** The information is provided on the developed methods for detecting pauses and their modifications, as well as the results of studying their operation under various parameters, as well as the comparison of their operation with some existing methods for detecting pauses. Pauses are detected in speech signals when vehicle noise with different levels of SNR is superimposed on them.*

Keywords: pause detection, voice activity detector.

DOI: 10.21667/1995-4565-2023-86-32-37

1. **Davenport W.B., Jr.** *Signal-to-Noise Ratios in Bandpass Limiters*. Journal of Applied Physics, Volume 24, Issue 6, p.720-727.

2. **Vitjazev V.V., Rozov V.I., Volchenkov V.A.** Pat. № RU 2436173 C1, Int. Cl. G10L 15/00, G10L 11/02. *Sposob obnaruzhenija pazv v rechevyh signalah i ustrojstvo ego realizujushhee* (Method of Detecting Pauses in Speech Signals and Device for Realising Said Method), заявитель i patentoobladatel' - Rjazanskij gosudarstvennyj radiotekhnicheskij universitet. Application: 2010124342/08, Date of filing: 15.06.2010, Date of publication: 10.12.2011 Bull. 34. (in Russian).

3. **Sohn J., Kim N. S. and Sung W.** A statistical model-based voice activity detection. *IEEE Signal Processing Letters*. 1999, 6(1):1-3.

4. **Kondo A.M.** *Digital Speech*. Coding for Low Bit Rate Communication Systems. John Wiley & Sons, Ltd. 2004. – 442 p.

5. ITU-T (1996) *A silence compression scheme for G.729 optimised for terminals conforming to ITU-T V.70*, ITU-T Rec. G.729 Annex B.

6. ITU-T (1996) *Coding of speech at 8 kbit/s using conjugate-structure algebraiccode excited linear prediction (CS-ACELP)*, ITU-T Rec. G.729.