



УДК 621.391.8

ИССЛЕДОВАНИЕ ДВУХКАНАЛЬНОГО АЛГОРИТМА MVDR ДЛЯ ВЫДЕЛЕНИЯ РЕЧИ ИЗ КОГЕРЕНТНОГО ШУМА

М.Б. Столбов, Чонг Тхе Куан

Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация

Адрес для переписки: stolbov@speechpro.com

Информация о статье

Поступила в редакцию 08.12.18, принята к печати 25.12.18

doi: 10.17586/2226-1494-2019-19-1-180-183

Язык статьи – русский

Ссылка для цитирования: Столбов М.Б., Чонг Тхе Куан. Исследование двухканального алгоритма MVDR для выделения речи из когерентного шума // Научно-технический вестник информационных технологий, механики и оптики. 2019. Т. 19. № 1. С. 180–183. doi: 10.17586/2226-1494-2019-19-1-180-183

Аннотация

Предмет исследования. Рассмотрены особенности реализации двухканального алгоритма минимума дисперсии шума (MVDR) для выделения речи из когерентного шума с использованием двухэлементных микрофонных решеток.

Методы. Исследование выполнено с использованием аналитических моделей и записей, сделанных в безэховой камере. **Основные результаты.** Алгоритм MVDR может быть представлен как комбинация дифференциального алгоритма с адаптивной настройкой нуля в направлении источника когерентной помехи и эквалайзера, выравнивающего пространственно-частотный отклик в направлении целевого источника. **Практическая значимость.** Полученные результаты могут быть применены при проектировании систем с большим числом микрофонов.

Ключевые слова

двухэлементные микрофонные решетки, алгоритм MVDR, подавление когерентных шумов

Благодарности

Исследования выполнены за счет стартового финансирования Университета ИТМО в рамках НИР № 618278 «Синтез эмоциональной речи на основе генеративных состязательных сетей».

STUDY OF MVDR DUAL-MICROPHONE ALGORITHM FOR SPEECH ENHANCEMENT IN COHERENT NOISE PRESENCE

M.B. Stolbov, T.The Quan

ITMO University, Saint Petersburg, 197101, Russian Federation

Corresponding author: stolbov@speechpro.com

Article info

Received 08.12.18, accepted 25.12.18

doi: 10.17586/2226-1494-2019-19-1-180-183

Article in Russian

For citation: Stolbov M.B., T.The Quan. Study of MVDR dual-microphone algorithm for speech enhancement in coherent noise presence. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2019, vol. 19, no. 1, pp. 180–183 (in Russian). doi: 10.17586/2226-1494-2019-19-1-180-183

Abstract

Subject of research. This paper considers a dual-channel algorithm implementation, namely, Minimum Variance Distortionless Response (MVDR), for speech enhancement in the presence of coherence noise using dual-microphone arrays. **Methods.** The study was performed using analytical models and recordings in an anechoic chamber. **Main results.** It is shown that the MVDR algorithm can be represented as a combination of a differential algorithm with an adaptive null steering adjustment in the directions of the coherent noise sources and an equalizer that equalizes the spatial-frequency response in the direction of the target source. **Practical relevance.** The advantage of the MVDR algorithm is automatic null steering in the directions of noise sources. The results are applicable in the design of systems with a large number of microphones.

Keywords

dual-microphone arrays, MVDR algorithm, coherent noise reduction

Acknowledgements

This work was partially financially supported by the Government of the Russian Federation (Grant 618278) and by the initial funding from the ITMO University.

Двухэлементные микрофонные решетки (MP2) широко применяются благодаря своей простоте и возможности компактного размещения. Вопросы обработки сигналов MP2 рассмотрены, например, в [1–3]. В настоящей работе рассмотрена задача отделения речевых сигналов (РС) от широкополосного когерентного шума с использованием MP2.

Для подавления когерентного шума в MP2 выполняют, в частности, формирование фиксированного нуля диаграммы направленности MP2 в направлении источника шума [4–7]. В работе [5] показана эквивалентность алгоритма формирования нуля алгоритму минимума дисперсии шума (Minimum Variance Distortionless Response, MVDR). В работе [7] сравниваются алгоритмы формирования нуля и алгоритмы обработки сигналов дифференциальных MP2 (ДМП2).

В условиях изменяющейся шумовой обстановки алгоритмы с фиксированным положением нуля неэффективны, поскольку положение источников шума может меняться. Указанную проблему позволяют преодолеть адаптивные алгоритмы формирования нуля, в частности, MVDR [8]. Работа посвящена исследованию адаптивного алгоритма MVDR для MP2 и его взаимосвязи с алгоритмами обработки сигналов ДМП2.

При использовании кратковременного преобразования Фурье сигнал $S(f, k)$ акустического целевого источника с направления θ_s (угол от оси MP2) и когерентный широкополосный шум $V(f, k)$ с направлений $\theta_v(f)$ формирует вектор сигналов:

$$\mathbf{X}(f, k) = S(f, k)\mathbf{D}(f, \theta_s) + V(f, k)\mathbf{D}(f, \theta_v(f)) = \mathbf{S}(f, k) + \mathbf{V}(f, k),$$

где f, k – индексы частоты и номера кадра, $\mathbf{D}(f, \theta_x) = \mathbf{D}_x(f) = [e^{+j\Phi_x}, e^{-j\Phi_x}]^T$ – вектор фазовых сдвигов сигналов микрофонов относительно центральной точки между ними, $\Phi_x(f)$ – фазовый сдвиг:

$$\Phi_x(f) = \pi d \cos(\theta_x) / \lambda = \pi f \tau_0 \cos(\theta_x),$$

где d – расстояние между микрофонами; $\tau_0 = d/c$ – время прохождения звука MP2 между микрофонами, c – скорость звука в воздухе; θ_x – направление прихода сигнала.

Сигнал на выходе MP2 формируется взвешиванием входных сигналов комплексными весовыми коэффициентами:

$$Z(f, k) = \mathbf{W}^H(f)\mathbf{X}(f, k),$$

где $\mathbf{W}(f)$ – вектор коэффициентов; $()^H$ – символ сопряжения Эрмита; $Z(f, k)$ – комплексная амплитуда сигнала на k -м кадре.

В основе работы алгоритма MVDR лежит условие неискаженного приема сигнала с целевого направления θ_s и минимизации мощности шума на выходе MP2. Оптимальные веса MP2 можно найти, решив следующую оптимизационную задачу:

$$\min_{\mathbf{W}} E \left\{ |Z(f, k)|^2 \right\} = \min_{\mathbf{W}} \left[\mathbf{W}^H(f) \mathbf{P}_{vv}(f) \mathbf{W}(f) \right] \text{ при } \mathbf{W}^H(f) \mathbf{D}_s(f) = 1,$$

где $\mathbf{P}_{vv}(f)$ – ковариационная матрица спектров шума на микрофонах; $E\{\cdot\}$ – символ математического ожидания.

Ковариационная матрица шума оценивается для каждой частоты:

$$\mathbf{P}_{vv}(f) = E \left\{ \mathbf{V}(f, k) \mathbf{V}^H(f, k) \right\}.$$

Решив оптимизационную задачу, получим следующее соотношение для вектора оптимальных коэффициентов [1]:

$$\mathbf{W}_0(f) = \frac{\mathbf{P}_{vv}^{-1}(f) \mathbf{D}_s(f)}{\mathbf{D}_s^H(f) \mathbf{P}_{vv}^{-1}(f) \mathbf{D}_s(f)}.$$

В случае помехи, поступающей с направления θ_v , MP2 имеет следующий пространственный отклик [5]:

$$H_{\text{MVDR}}(f, \theta_x, \theta_s, \theta_v) = \left| \mathbf{W}_0^H(f, \theta_v, \theta_s) \mathbf{D}(f, \theta_x) \right| = \left| \frac{\sin(\Phi_x - \Phi_v)}{\sin(\Phi_s - \Phi_v)} \right|.$$

Отклик в направлении источника помехи $H_{\text{dif}}(f, \theta_x = \theta_v, \theta_v) = 0$, отклик в направлении источника целевого сигнала $H_{\text{MVDR}}(f, \theta_s, \theta_v) = 1$.

Пространственно-частотный отклик MVDR можно представить в виде произведения передаточной функции эквалайзера и передаточной функции ДМП2:

$$H_{\text{MVDR}}(f, \theta_x, \theta_s, \theta_v) = H_{\text{eq}}(f, \theta_s, \theta_v) H_{\text{dif}}(f, \theta_x, \theta_v).$$

Форма пространственного отклика MVDR совпадает с формой пространственного отклика ДМР2 и достигает максимальной величины в направлении $\theta_s = 0$ (рис. 1).

В адаптивном алгоритме MVDR ковариационная матрица и оптимальные веса оцениваются (в паузах целевого сигнала) по поступающим сигналам микрофонов:

$$\mathbf{P}_{vv}(f, k) = (1 - \beta)\mathbf{P}_{vv}(f, k - 1) + \beta \mathbf{X}(f, k)\mathbf{X}^H(f, k),$$

где β – коэффициент сглаживания.

Весовые коэффициенты фильтра MVDR вычисляются по текущим значениям ковариационной матрицы:

$$\mathbf{W}_0(f, k) = \frac{\mathbf{P}_{vv}^{-1}(f, k)\mathbf{D}_s(f)}{\mathbf{D}_s^H(f)\mathbf{P}_{vv}^{-1}(f, k)\mathbf{D}_s(f) + \sigma},$$

где σ – положительная константа, предохраняющая от деления на нуль.

Сигнал на выходе вычисляется с использованием текущих значений коэффициентов:

$$\mathbf{Z}(f, k) = \mathbf{W}_0^H(f, k)\mathbf{X}(f, k).$$

Вектор отсчетов сигнала $\mathbf{Z}(k)$ на k -м кадре вычисляется с помощью дискретного обратного преобразования Фурье (Inverse Discrete Fourier Transform, IDFT):

$$\mathbf{Z}(k) = \text{IDFT}\{\mathbf{Z}(f, k)\}.$$

Последовательность векторов отсчетов $\{\mathbf{Z}(k)\}$ преобразуется в выходную последовательность отсчетов сигнала на основе процедуры пересечения и суммирования (OverLap-and-Add, OLA): $z(t) = \text{OLA}\{\mathbf{Z}(k)\}$.

Адаптивный алгоритм MVDR исследовался на сигналах, записанных в безэховой камере с помощью микрофонов (мик₁, мик₂), разнесенных на расстояние $d = 5$ см. Широкополосная помеха с акустической колонки поступала с направлений φ_v ($0 - 90^\circ$) относительно нормали МР2 с шагом 5° . Каждое положение акустической колонки озвучивалось диктором, находившимся приблизительно в направлении $\varphi_s = -30^\circ$ (рис. 2).

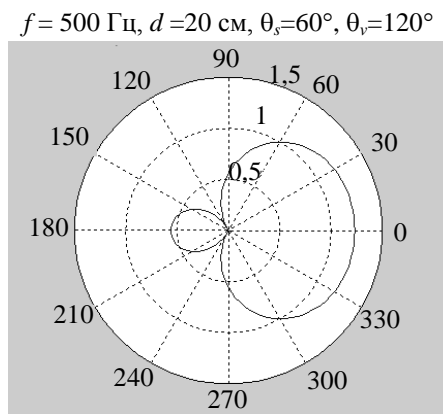


Рис. 1. Пространственный отклик MVDR МР2: $\theta_s = 60^\circ$, $\theta_v = 120^\circ$

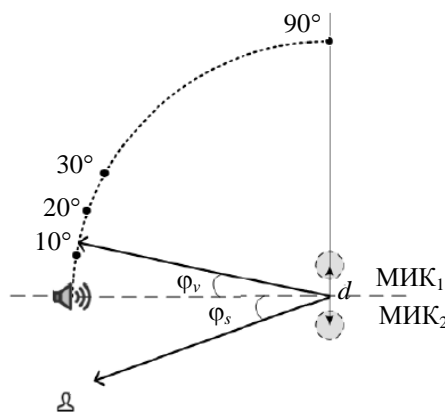


Рис. 2. Схема эксперимента в безэховой камере

Обработка проводилась со следующими параметрами: частота дискретизации $F_s = 16 \text{ кГц}$, размер кадра $N = 512$, $\beta = 0,5$. Результаты эксперимента представлены на рис. 3.

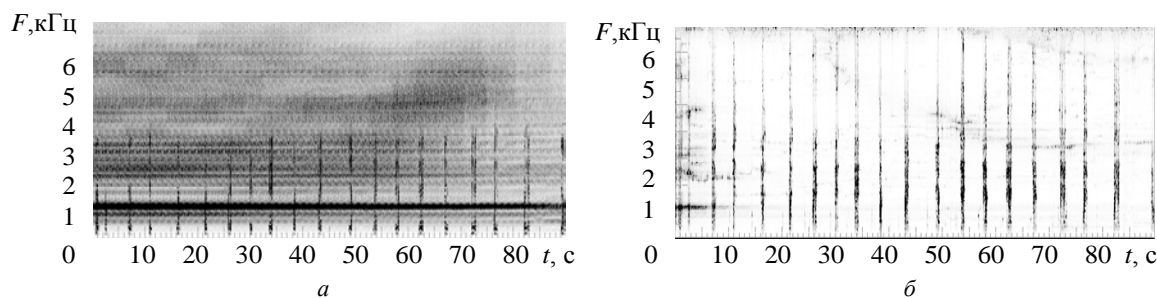


Рис. 3. Спектрограммы сигнала микрофона (а), сигнала на выходе MVDR (б)

Из эксперимента в безэховой камере можно сделать следующие выводы.

Адаптивный алгоритм MVDR позволяет отслеживать направление источника когерентного шума и подавлять этот шум; степень подавления 20–25 дБ ограничивалась, по всей вероятности, различием характеристик микрофонов. Уровень подавленного когерентного шума с помощью алгоритма MVDR приблизительно равен уровню подавленного шума в ДМР2, в направлении нуля пространственного отклика. Время реакции алгоритма MVDR на изменение шумовой обстановки (в нашем варианте приблизительно 30 мс) регулируется параметром сглаживания.

Основным результатом настоящей работы является представление алгоритма MVDR как алгоритма дифференциальной МР2 с адаптивной настройкой нулей в направлениях источников когерентного шума.

Литература

1. Microphone Arrays. Signal Processing Techniques and Applications / Eds. M. Brandstein, D. Ward. Springer-Verlag, 2001. 402 p.
2. Benesty J., Chen J. Study and Design of Differential Microphone Arrays. Springer, 2013. 184 p.
3. Benesty J., Chen J., Pan C. Fundamentals of Differential Beamforming. Springer, 2016. 129 p.
4. Монзинго Р.А., Миллер Т.У. Адаптивные антенные решетки. М: Радио и связь, 1986. 448 с.
5. Столбов М.Б., Перельгин С.В. Алгоритмы двухэлементной микрофонной решетки для выделения речевых сигналов в присутствии когерентных помех // Цифровая обработка сигналов. 2017. № 4. С. 34–39.
6. Buck M., Rößler M. First order differential microphone arrays for automotive applications // Proc. 7th International Workshop on Acoustic Echo and Noise Control, IWAENC. Darmstadt, Germany, 2001. P. 19–22.
7. Столбов М.Б., Тхе К.Ч. Прием речевых сигналов в шумовой обстановке с использованием двухэлементных микрофонных решеток // Научно-технический вестник информационных технологий, механики и оптики. 2018. Т. 18. № 5. С. 850–857. doi: 10.17586/2226-1494-2018-18-5-850-857
8. Lockwood M. et al. Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms // The Journal of the Acoustical Society of America. 2004. V. 115. N 1. P. 379–391. doi: 10.1121/1.1624064

Авторы

Столбов Михаил Борисович – кандидат технических наук, доцент, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, Scopus ID: 56566801600, ORCID ID: 0000-0002-3848-4763, stolbov@speechpro.com

Куан Тхе Чонг – аспирант, Университет ИТМО, Санкт-Петербург, 197101, Российская Федерация, Scopus ID: 57203964397, ORCID ID: 0000-0002-2456-9598, quantrongthe@corp.ifmo.ru

References

1. *Microphone Arrays. Signal Processing Techniques and Applications* / Eds. M. Brandstein, D. Ward. Springer-Verlag, 2001, 402 p.
2. Benesty J., Chen J. *Study and Design of Differential Microphone Arrays*. Springer, 2013, 184 p.
3. Benesty J., Chen J., Pan C. *Fundamentals of Differential Beamforming*. Springer, 2016, 129 p.
4. Monzingo R.A., Miller T.W. *Introduction to Adaptive Arrays*. NY, John Wiley and Sons, 1980.
5. Stolbov M.B., Perelygin S.V. Algorithms of two-element microphone array for speech enhancement in presence of coherent interference. *Digital Signal Processing*, 2017, no. 4, pp. 34–39. (in Russian)
6. Buck M., Rößler M. First order differential microphone arrays for automotive applications. *Proc. 7th International Workshop on Acoustic Echo and Noise Control, IWAENC*. Darmstadt, Germany, 2001, pp. 19–22.
7. Stolbov M.B., The Quan Trong. Speech acquisition in noisy environments using dual microphone arrays. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2018, vol. 18, no. 5, pp. 850–857 (in Russian). doi: 10.17586/2226-1494-2018-18-5-850-857
8. Lockwood M. et al. Performance of time- and frequency-domain binaural beamformers based on recorded signals from real rooms. *The Journal of the Acoustical Society of America*, 2004, vol. 115, no. 1, pp. 379–391. doi: 10.1121/1.1624064

Authors

Mikhail B. Stolbov – PhD, Associate Professor, ITMO University, Saint Petersburg, 197101, Russian Federation, Scopus ID: 56566801600, ORCID ID: 0000-0002-3848-4763, stolbov@speechpro.com

Tchong The Quan – postgraduate, ITMO University, Saint Petersburg, 197101, Russian Federation, Scopus ID: 57203964397, ORCID ID: 0000-0002-2456-9598, quantrongthe@corp.ifmo.ru