

## НОВЫЙ АЛГОРИТМ ДЛЯ ДЕТЕКТИРОВАНИЯ И ОБРАБОТКИ СИТУАЦИИ ДВОЙНОГО РАЗГОВОРА

Назаровский А.Е., Тараканов А.Н.

Ярославский Государственный Университет им. П.Г.Демидова  
150000, Россия, Ярославль, ул. Советская, 14, Тел. (0852) 79-77-75. [dcslab@uniyar.ac.ru](mailto:dcslab@uniyar.ac.ru)

Явление эхо в телефонных сетях возникает из-за отражения электрических сигналов. Часть составляющих принимаемого сигнала проникает в линии передачи и возвращается обратно к говорящему абоненту в виде эхосигналов, которые при больших задержках, что имеет место в междугородней, международной и спутниковой связи, являются серьезной помехой для разговора [1,2].

Для борьбы с эхосигналами в настоящее время используются эхокомпенсаторы. Они строятся на основе адаптивных фильтров (АФ) с конечной импульсной характеристикой (КИХ), а для подстройки коэффициентов применяются алгоритмы на основе метода наименьших квадратов (МНК) [3]. Преимущество данного подхода заключается в его простоте и низкой сложности в вычислительном плане. Требования к характеристикам эхокомпенсаторов определены в рекомендациях Международного Союза Электросвязи ITU-T G.165 и G.168 [4,5].

При ситуации, когда говорят оба абонента одновременно (далее – ситуация двойного разговора), работа адаптивного алгоритма нарушается. Для обнаружения ситуации двойного разговора и формирования управляющего сигнала для алгоритма адаптации в состав эхокомпенсатора вводится детектор двойного разговора (ДДР). Однако, работе ДДР присущи недостатки, во-первых, задержки срабатывания, приводящие к ухудшению подавления эхосигнала, а во-вторых, ложные срабатывания [6]. Одной из причин ложного срабатывания ДДР может быть резкое изменение импульсной характеристики эхотракта. ДДР. Например, известный алгоритм Гейгеля [7] не различает резкого изменения импульсной характеристики эхотракта от ситуации двойного разговора.

Исходя из этого, представляется актуальной разработка новых алгоритмов детектирования и обработки ситуации двойного разговора учитывающих вышеприведенные недостатки.

Основой разработки нового алгоритма является использование следующих соображений:

- для реальных эхотрактов длина импульсной характеристики обычно существенно превосходит длину так называемой активной части импульсной характеристики, т.е. части, которая вносит наибольший вклад в энергию эхо сигнала. Поэтому, АФ подстроившийся под эхотракт, на каждом шаге будет иметь большое количество коэффициентов почти равных нулю, а суммарная энергия коэффициентов АФ будет примерно постоянной.

- кроме того, поскольку сигнал от ближнего абонента является некоррелированной помехой для алгоритма адаптации, коэффициенты АФ при наличии сигнала от ближнего абонента уже за несколько шагов расходятся от оптимальных значений [8]. При этом резко возрастает суммарная энергия коэффициентов АФ.

Для перестройки весовых коэффициентов АФ используется нормализованный метод наименьших квадратов (НМНК). Здесь и далее введены следующие обозначения:  $x(n)$  - сигнал дальнего абонента,  $v(n)$  - сигнал ближнего абонента,  $d(n)$  - эхо-сигнал,  $y(n)$  - сумма эхо-сигнала и сигнала от ближнего абонента,  $\hat{y}(n)$  - оценка эхо-сигнала, получаемая с выхода адаптивного фильтра,  $\hat{e}(n)$  - ошибка оценки эхо-сигнала,  $e(n)$  - остаточное эхо.

Подстройка коэффициентов АФ производится согласно НМНК

$$\hat{h}_k(n) = \hat{h}_k(n-1) + \mu(n) \cdot x(n-k) \cdot \hat{e}(n), \quad k = 0, 1 \dots L-1, \quad (1)$$

величина шага адаптации, соответственно, задается выражением

$$\mu(n) = \frac{\alpha}{\beta + \mathbf{x}^T(n) \cdot \mathbf{x}(n)}, \quad \alpha \in (0,2), \beta \geq 0, \quad (2)$$

где,  $\alpha, \beta$  - константы,  $\mathbf{x}$  - вектор отсчетов сигнала дальнего абонента длины  $L$ . В дальнейшем моделировании используется  $L = 128$ .

Эхотракт моделируется КИХ-фильтром, длиной 128 коэффициентов, и соответствует импульсной характеристике дифференциальной системы, задаваемой рекомендацией G.168. Используемые в моделировании сигналы от ближнего и дальнего абонентов являются дискретизированными с частотой 8 КГц.

В качестве пороговой статистики для обнаружения ситуации двойного разговора предлагается использовать функцию от суммы квадратов коэффициентов АФ (3), в предположении, что импульсная характеристика эхотракта в течение разговора изменяется незначительно.

$$\xi(n) = \ln \left( \sum_{i=0}^{L-1} \hat{h}_i^2(n) \right), \quad (3)$$

Здесь  $L$  - длина АФ,  $\hat{h}_i(n)$  - значение  $i$ -го коэффициента АФ,  $n$  - номер обрабатываемого отсчета.

Обычно при обнаружении детектором ситуации двойного разговора адаптивный алгоритм должен приостановить подстройку коэффициентов, чтобы оценка эхотракта далее не ухудшалась. Но так как данный алгоритм основан на вычислении суммы квадратов коэффициентов АФ на каждом отсчете, очевидно, что приостановка процесса адаптации недопустима. Кроме того, решающая статистика  $\xi$  после окончания двойного разговора, сходится к обычному значению довольно медленно, поскольку коэффициенты АФ за время двойного разговора сильно разошлись от оптимальных, и это ухудшает обнаружение завершения ситуации двойного разговора.

Решением данной проблемы может служить модифицированная структура с двумя фильтрами (рис. 1). Оригинальная структура описана ранее в [9]. Фильтр  $A$  – используется для формирования оценки эхотракта и для обнаружения ситуации двойного разговора. Фильтр  $B$  – служит для собственно подавления эхо сигнала. Модификацией оригинальной структуры является связь фильтра  $B$  с фильтром  $A$ . Применение модифицированной схемы с двумя фильтрами позволяет оперативно обнаруживать начало и конец ситуации двойного разговора, и сохранять уровень подавления эхосигнала во время двойного разговора.

Рассмотрим два режима работы системы. Далее для простоты принимаем, что длины фильтров  $A$  и  $B$  одинаковы. В обычном режиме (нет двойного разговора) адаптивный фильтр  $A$  на основании вычисленной ошибки  $\hat{e}(n)$  производит подстройку своих коэффициентов. Через заданное число отсчетов  $\tau$  вычисляется статистика  $\xi$  на основе суммарной энергии коэффициентов фильтра  $A$ , и сравнивается с заданным пороговым значением  $Q$ . Если  $\xi$  не превысила порогового значения ( $\xi < Q$ ), принимаем, что в течение заданного промежутка  $\tau$  двойного разговора не было обнаружено, и коэффициенты из фильтра  $A$  копируются в фильтр  $B$ . В противном случае, декларируется режим двойного разговора.

В режиме двойного разговора, напротив, в начале каждого интервала  $\tau$ , коэффициенты копируются из фильтра  $B$  в фильтр  $A$ . Через заданное число отсчетов  $\tau$  опять вычисляется статистика  $\xi$  на основе суммарной энергии коэффициентов фильтра  $A$ . Если она превышает порог ( $\xi > Q$ ), то ситуация двойного разговора считается продолженной, на следующий интервал времени, коэффициенты снова копируются из  $B$  в  $A$  и данный шаг повторяется. В противном случае, режим двойного разговора отменяется, и система начинает работать в обычном режиме. Вследствие нестационарности речевого сигнала целесообразно перед переходом в обычный режим сделать задержку еще на некоторый промежуток времени  $T_{hold}$ . Согласно [10] адекватным значением для речевых сигналов является  $T_{hold} \approx 75мс$ , что соответствует задержке в 600 отсчетов.

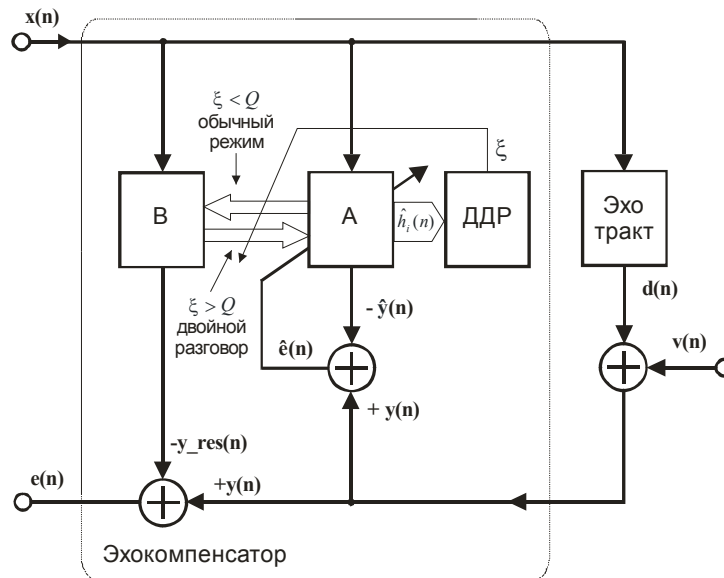


Рис.1 Модель эхокомпенсатора с модифицированной структурой с двумя фильтрами и детектором на основе суммарной энергии коэффициентов АФ

Пороговое значение  $Q$  должно задаваться из соображений, что система не должна переходить в режим двойного разговора при резких изменениях импульсной характеристики эхотракта. Промежуток  $\tau$  не должен быть слишком малым, в противном случае коэффициенты фильтра  $A$  не успеют разойтись на достаточные для обнаружения двойного разговора значения. Однако задание больших интервалов  $\tau$  тоже не рекомендуется вследствие ухудшения характеристик алгоритма, т.к. во-первых, аттенюация эхо сигнала становится скачкообразной, что заметно на слух, а во-вторых, ухудшается скорость отслеживания изменений эхотракта.

В фильтре  $A$  применяется адаптивный алгоритм НМНК. Остальные параметры алгоритма следующие:  $\alpha = 1$ ,  $\beta = 0.01$ ,  $Q = -3$ ,  $\tau = 100$  отсчетов,  $T_{hold} = 600$  отсчетов. В качестве характеристики свойств

нового алгоритма используется величина подавления эхосигнала  $ERLE$ , задаваемая следующим выражением:

$$ERLE(n) = 10 * \log_{10} \left( \frac{E[y(n)^2]}{E[e(n)^2]} \right), \quad (4)$$

а также нормированная расстройка весовых коэффициентов АФ относительно импульсной характеристики эхотракта

$$\varepsilon(n) = 20 * \log_{10} \left( \frac{\|h - \hat{h}(n)\|}{\|h\|} \right), \quad (5)$$

На рис. 2(а) и рис. 2(б) приводится сравнение  $ERLE$  и расстройки для трех случаев: нового алгоритма, при использовании ДДР Гейгеля[7] и в отсутствии ДДР.

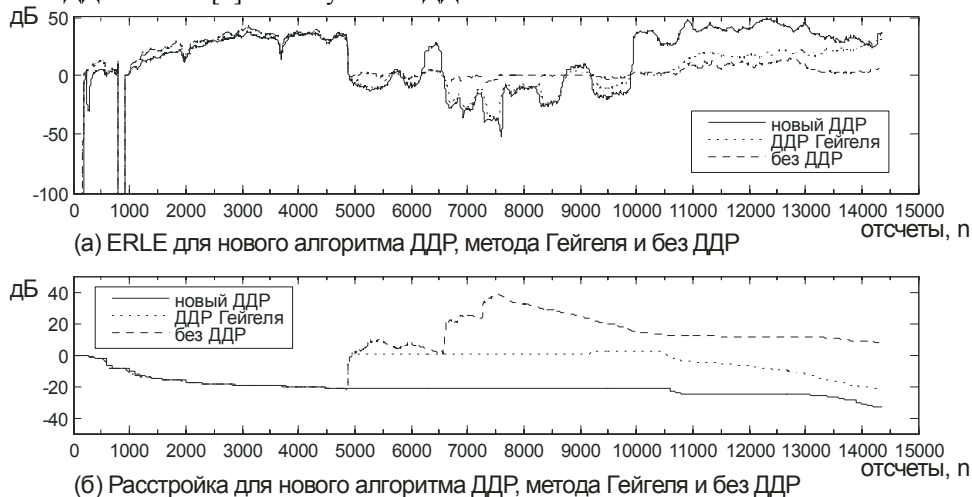


Рис.2. Сравнение  $ERLE$  и расстройки при использовании нового алгоритма, ДДР Гейгеля и при отсутствии ДДР

Уменьшение расстройки весовых коэффициентов адаптивного фильтра при использовании нового алгоритма по сравнению с алгоритмом Гейгеля и без ДДР составляет 20 дБ и 60 дБ соответственно, что говорит о значительном улучшении характеристик эхокомпенсатора в ситуации двойного разговора.

Предложен новый способ обнаружения и обработки ситуации двойного разговора на основе вычисления суммарной энергии коэффициентов адаптивного фильтра и модификации структуры с двумя фильтрами. Моделирование на речевых сигналах показало успешность данного подхода по сравнению с использованием ДДР на основе алгоритма Гейгеля.

#### Литература

1. Беллами Дж. Цифровая телефония: Пер. с англ. Э.Б. Ершовой, Э.В. Кордонского – М.: Радио и связь, 1986. 544 с.
2. Уидроу Б., Стирнз С. Адаптивная обработка сигналов: Пер. с англ. под ред. В.В. Шахгильдяна – М.: Радио и связь, 1989. 440 с.
3. М. М. Sondhi. An adaptive echo canceler // The Bell Systems Technical Journal. 1967. vol. XLVI, no. 3, pp. 497–510.
4. ITU-T Recommendation G.165. Echo cancellers. 1993.
5. ITU-T Recommendation G.168. Digital network echo cancellers. 1997.
6. Cho J.H., Morgan D.R., Benesty J. An Objective Technique for Evaluating Doubletalk Detectors in Acoustic Echo Cancellers // IEEE Transactions on Speech and Audio Processing. 1999. V. 7. №. 6. pp. 718-724.
7. Duttweiler D.L. A twelve-channel digital echo canceller // IEEE Transactions on Speech and Audio Processing. 1978. vol. 26, №5, pp.647-653
8. Брюханов Ю.А., Тараканов А.Н. Усовершенствование адаптивного алгоритма эхокомпенсации // Электросвязь. 2003. №9. С. 38-39.
9. Ochiai K., Araseki T., Ogihara T. Echo Canceller with Two Echo Path Models // IEEE Transactions on Communications. 1977. vol. 25, №6, pp. 589 - 595
10. D. Messerschmitt, D. Hedberg, C. Cole, A. Houi, P. Winship. Digital Voice Echo Canceller with a TMS32020 // Digital Signal Processing Applications. Texas Instruments. 1986. Vol.1 , pp.415-454.



## A NEW METHOD FOR DETECTION AND HANDLING OF DOUBLE TALK SITUATION IN LINE ECHO CANCELLATION

Nazarovsky A., Tarakanov A.

Yaroslavl State University  
150000, Yaroslavl, Sovetskaya 14, Russia.  
Phone: (0852)79-77-75, e-mail: [dcslab@uniyar.ac.ru](mailto:dcslab@uniyar.ac.ru)

The line echo (also known as hybrid or electric echo) in international and satellite telephony is caused by the interrelations of receiving and transmitting wire lines. As a result, part of the signal in the send line is reflected back to the receive line. The general approach to line echo cancellation is using adaptive filtering methods [1]. The Finite Impulse Response (FIR) adaptive filters implementing Least Mean Squares (LMS) methods are commonly in this task. The advantage of this approach is its low computational cost and quite good performance. Tests and demands on line echo cancellers are defined in ITU-T recommendations G.165 and G.168.

In double talk situation (i.e. when both subscribers talk simultaneously) the work of adaptive echo canceller is corrupted. Common way to handle this problem is to use the Double Talk Detector (DTD). DTD should detect such situation and generate control signal for the adaptive algorithm.

A new method of double talk detection and handling the double talk situation is proposed. Double talk detection is carried out using summary energy of adaptive filter coefficients and double talk situation is handled by a modified two path structure.

Since near-end signal acts as a high level uncorrelated noise to the adaptive algorithm, the adaptive filter coefficients in double talk case very quickly drift from the estimated echo path impulse response. Close to zero coefficients are also diverging, so in double talk case the total energy of adaptive filter coefficients is quickly increasing. This fact is proposed to use for double talk detection.

We use the Normalized Least Mean Squares (NLMS) algorithm and the Finite Impulse Response (FIR) filter. Echo path is modeled by a FIR-filter and is defined according G.168 recommendation. Far-end and near-end signals used in the modeling are discretized at 8 KHz frequency.

As a decision function for the double talk detection we use the function from the total sum of squares of adaptive filter coefficients. However common method of double talk handling i.e. suspension of adaptive process during double talk is not applicable here, because decision function after detection would be “frozen” forever. Moreover detection of double talk ending is difficult, because adaptive filter coefficients have strongly diverged.

To solve these problems we suggest modification of two filter structure. Original structure was proposed in [2]. Filter *A* is used for echo path estimation and double talk detection, and filter *B* – for echo cancellation. New in this approach is link between filter *B* and filter *A*. System has two modes: normal mode and double talk mode.

As a result the new method of double talk detection and double talk handling was proposed, and modeling on speech signals has shown the advantages of the new method over Geigel DTD. New method is quite simple and has comparatively low computational complexity. However, the proper choice of new method’s parameters requires further research.

### References

1. B. Widrow and S. D. Stearns. Adaptive Signal Processing. New Jersey: Prentice-Hall, Inc., 1985.
2. Ochiai K., Araseki T., Ogihara T. Echo Canceller with Two Echo Path Models // IEEE Transactions on Communications. 1977. vol. 25, №6, pp. 589 – 595.

