

УДК 004.934.2

**И.О. Архипов**

канд. техн. наук, доцент,  
заведующий кафедрой «Программное обеспечение»,  
ФГБОУ ВПО «Ижевский государственный  
технический университет им. М.Т. Калашникова»

**М.А. Голубев**

магистрант, кафедра «Программное обеспечение»,  
ФГБОУ ВПО «Ижевский государственный  
технический университет им. М.Т. Калашникова»

## АЛГОРИТМ ОПРЕДЕЛЕНИЯ ИНТОНАЦИОННОГО КОНТУРА ФРАЗЫ

**Аннотация.** Необходимость изучения иностранного языка становится всё более и более актуальной. Интонация, с которой произносится та или иная фраза имеет большое значение и должна учитываться при разработке систем электронного обучения. В работе описан алгоритм, позволяющий осуществить вычисление интонационного контура фразы.

**Ключевые слова:** интонационный контур фразы, автокорреляционная функция, частота основного тона.

**I.O. Arkhipov, Kalashnikov Izhevsk State Technical University**

**M.A. Golubev, Kalashnikov Izhevsk State Technical University**

### ALGORITHM FOR ESTIMATION INTONATION CONTOUR OF PHRASES

**Abstract.** The necessity to study foreign language become more and more ubiquitous. Intonation, which any phrase is pronounced with, does matter quite a lot in case of learning a foreign language, that's why this fact should be taken into account during designing e-learning system. In this paper is described the algorithm, which can provide the intonation tracking.

**Keywords:** intonation tracking, autocorrelation function, fundamental frequency.

Задача определения интонационного контура фразы в цифровой обработке сигналов может быть решена различными способами, одним из которых является метод на основе автокорреляционной функции сигнала [1]. Общая схема работы алгоритма представлена на рисунке 1. Речевой сигнал проходит предварительную обработку фильтром верхних частот с частотой среза 50 Гц для удаления низкочастотных составляющих, обусловленных возможным использованием низкокачественной аппаратуры.

Предлагаемый алгоритм, помимо автокорреляционной функции (АКФ) сигнала, также использует энергию сигнала, нормированный коэффициент корреляции с единичной задержкой (далее  $R(1)$ ) и энергию сигнала в полосе частот. Также в процессе расчёта частоты основного тона используется статистика по двум предыдущим участкам сигнала для снижения вероятности случайного сбоя. Финальным этапом алгоритма является медианное сглаживание полученного интонационного контура.

Для определения признака тональности речи используется метод, в основе которого лежит анализ трёх свойств сигнала: автокорреляционной функции (АКФ), энергии в полосе частот и нормализованного коэффициента корреляции с единичной задержкой.

АКФ рассчитывается по формуле (1) [1].

$$j(k) = \sum_{m=0}^N x(m) \times x(m+k). \quad (1)$$

где  $x$  – элементы массива с исходным сигналом.

На рисунке 2 видно, что АКФ тонального сигнала имеет ярко выраженную периодичность, в то время как АКФ нетонального сигнала имеет произвольную форму. Таким образом, чтобы считать участок сигнала тональным, необходимо, чтобы АКФ была квазипериодической.

Энергию сигнала в полосе частот вычисляли после обработки сигнала фильтром нижних частот с частотой среза 600 Гц. Затем производится расчёт энергии по формуле [1]:

$$E = \sum_{i=0}^N |x[i]| \quad (2)$$

На нетональных участках значение энергии ниже, чем на тональных [1], поэтому для определения тональности достаточно сравнивать вычисленное значение энергии с определённой пороговой величиной.

Нормированный коэффициент корреляции с единичной задержкой рассчитывается по формуле (3).

$$R(1) = \frac{\sum_{i=0}^{N-1} x[i] \cdot x[i+1]}{\sum_{i=0}^{N-1} x[i]^2} \quad (3)$$

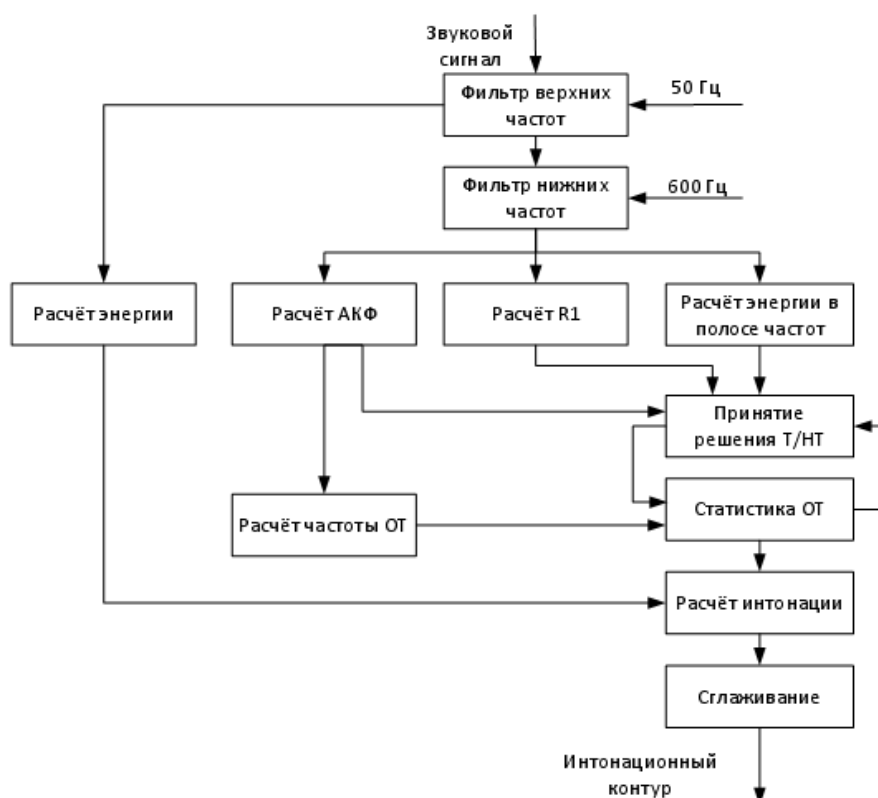


Рисунок 1 – Общая схема работы алгоритма: АКФ – автокорреляционная функция; ОТ – основной тон; Т/НТ – признак тональности сигнала; R1 – нормированный коэффициент корреляции с единичной задержкой

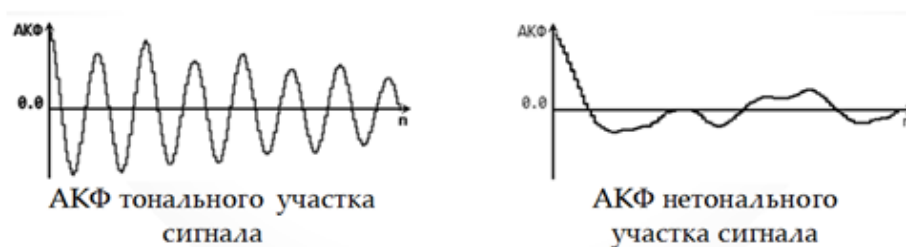


Рисунок 2 – АКФ тонального и нетонального участков сигнала

Как видно из рисунка 3, значение коэффициента  $R(1)$  на тональных участках сигнала стремится к единице, в то время как на нетональных оно значительно ниже.

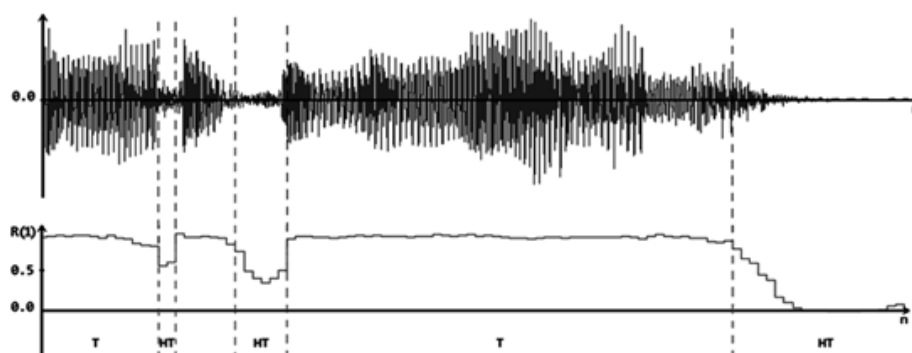


Рисунок 3 – Нормированный коэффициент корреляции с единичной задержкой

Для определения тональности участка сигнала достаточно сравнить значение коэффициента  $R(1)$  с определённой пороговой величиной. Алгоритм определения тональности участка сигнала использует все три описанные выше характеристики. На первом шаге определяется периодичность АКФ. Затем, если АКФ периодична, подсчитываются коэффициент  $R(1)$  и энергия сигнала в полосе частот.

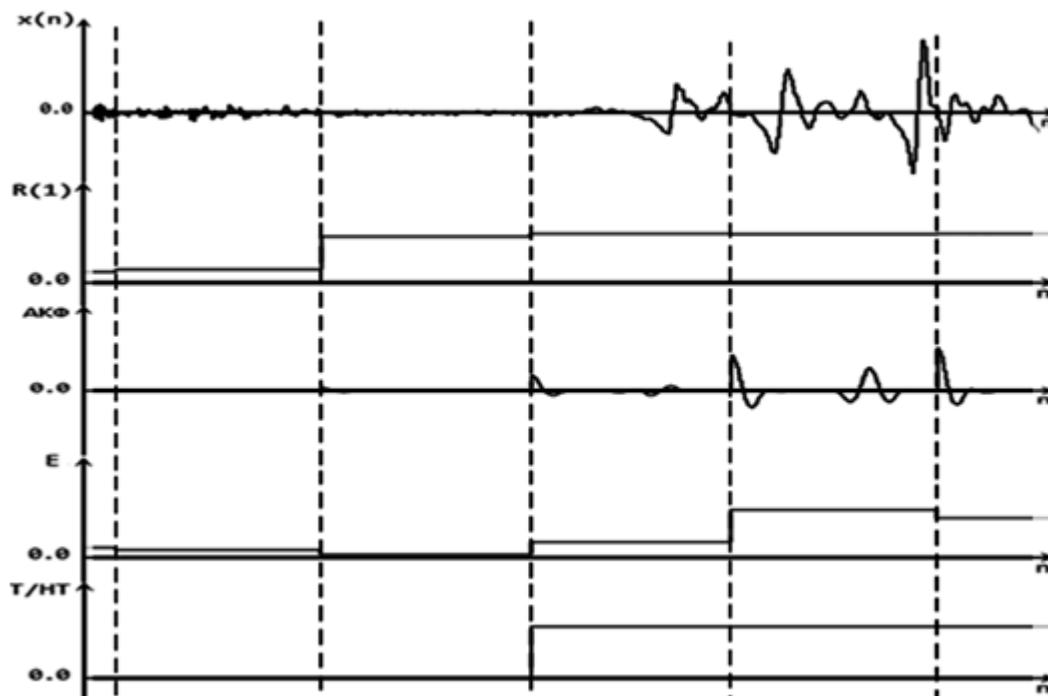


Рисунок 4 – Результат работы алгоритма определения тональности сигнала

Финальным шагом определения тональности является сравнение полученных результатов  $R(1)$  и энергии с пороговыми величинами и принятие решения о тональности текущего участка сигнала. Результат работы алгоритма показан на рисунке 4.

После того, как обрабатываемый участок сигнала определён как тональный, необходимо найти частоту основного тона на данном участке. Найдем индекс главного максимума графика АКФ, исключив при этом побочный максимум, который можно увидеть на рисунке 5:

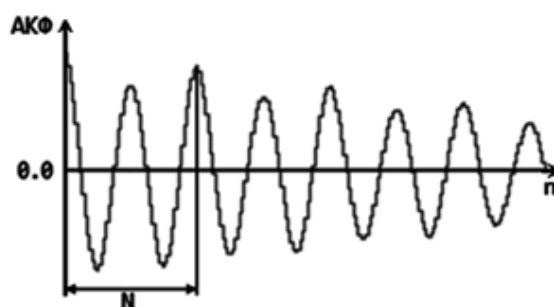


Рисунок 5– Период основного тона на графике АКФ

Частоту основного тона можно рассчитать по формуле (4):

$$F_{от} = \frac{F_{дискр}}{N}, \quad (4)$$

где  $F_{дискр}$  – частота дискретизации сигнала,  $N$  – индекс главного максимума АКФ.

На рисунке 6 представлен результат расчёта периода основного тона для нескольких сегментов сигнала с частотой дискретизации 8 кГц.

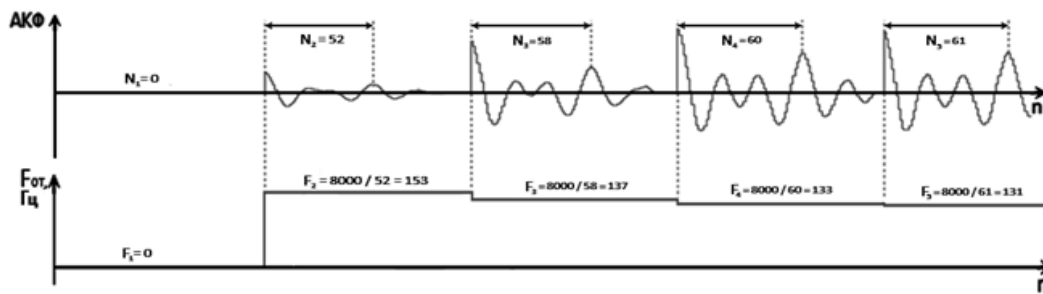


Рисунок 6 – Период основного тона для нескольких сегментов сигнала

После нахождения значения частоты основного тона производится корректировка найденного значения в соответствии со значениями на двух предыдущих участках сигнала – для того, чтобы исключить случайные ошибочные результаты.

Для того, чтобы более точно реагировать на качество произношения фразы, предлагается при вычислении интонационного контура фразы помимо частоты основного тона также учитывать энергию сигнала, как показано в формуле (5):

$$I[i] = F_{om}[i] \cdot (\lg(E[i] + \varepsilon))^2, \quad (5)$$

где  $I$  – интонационный контур фразы;  $E$  – энергия сигнала;  $\varepsilon$  – положительное малое вещественное число.

На рисунке 7 показан эффект, который оказывает энергия на процесс расчёта интонационного контура фразы.

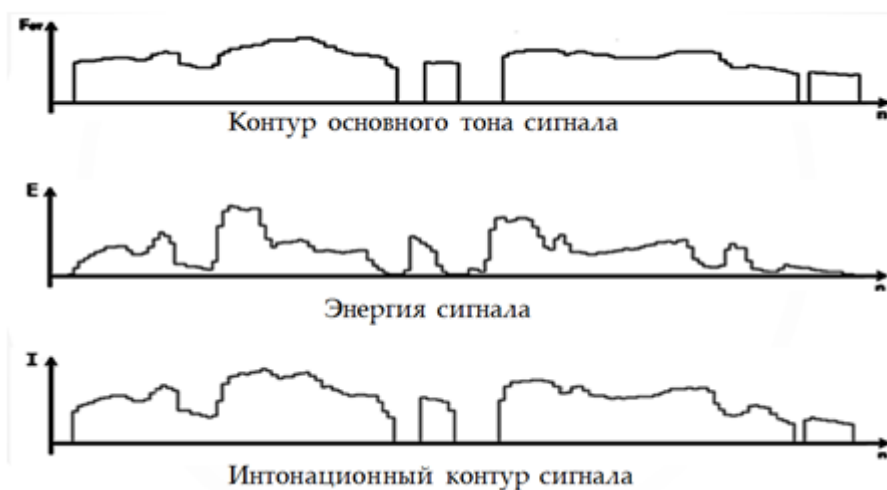


Рисунок 7 – Влияние энергии на интонационный контур фразы

После вычисления интонационного контура, проводится его пост-обработка медианным сглаживанием, которое убирает ошибочные скачки интонации малой длительности. Результат пост-обработки контура представлен на рисунке 8.



Рисунок 8 – Обработка контура медианным сглаживанием

На рисунке 9 представлены интонационные контуры для фразы «Мама мыла Маню» с ударением на первое или на второе слово.

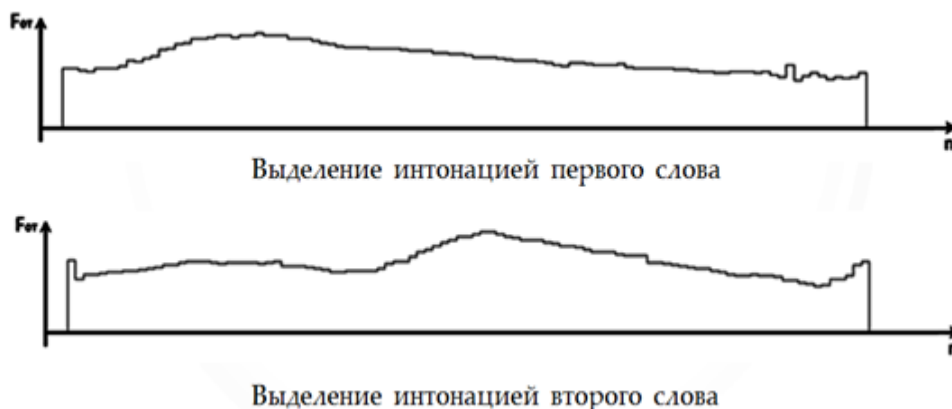


Рисунок 9 – Интонационные контуры фразы «Мама мыла Маню»

Из рисунка 9 видно, что контуры фраз заметно отличаются и имеют пик интонационного контура, который приходится на ударное слово.

По результатам исследований предложен алгоритм формирования интонационного контура речевого сигнала на основе автокорреляционной функции речи. Также показаны промежуточные результаты на всех этапах работы алгоритма. И предложен способ формирования интонационного контура речи как произведение контура основного тона на логарифм энергии сигнала.

**Список литературы:**

1. Рабинер, Л. Цифровая обработка речевых сигналов / Л. Рабинер, Р. Шафер; пер. с англ. под ред. М.В. Назарова и Ю.Н. Прохорова. – М.: Радио и связь, 1981. – 496 с.
2. Голд, Б. Цифровая обработка сигналов / Б. Голд, Ч. Рэйдер; под ред. А.М. Трахтмана. – М.: Совет. радио, 1973. – 368 с.