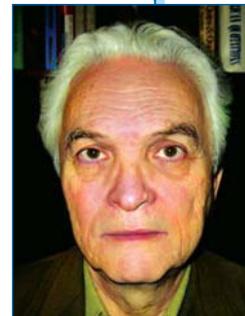


Автокорреляционный метод выделения основного тона речи. Пятьдесят лет спустя

С.П. Баронин,

кандидат технических наук



Проблема выделения основного тона речи рассматривается с позиций теории статистических решений. Основываясь на модели вокализованного речевого сигнала в виде периодического процесса с неизвестными параметрами, искажённого шумом, и принимая в качестве модели сигнала основного тона нестационарный Марковский процесс, предложена разновидность автокорреляционного выделителя основного тона. Моделирование алгоритма и испытания на тестовых сигналах подтвердили высокую точность измерения периодов основного тона. Приводятся также данные о помехоустойчивости описанного алгоритма.

Немного истории

Пятьдесят лет назад учёный с мировым именем Андрей Андреевич Пирогов получил авторское свидетельство на изобретение устройства для определения частоты основного тона (ОТ) по методу автокорреляционного анализа речевого сигнала [1]. В то же время в руководимой А.А. Пироговым лаборатории начаты работы по теоретическому и экспериментальному исследованию автокорреляционного метода выделения ОТ речи [2]. В 1959 году создан первый макет устройства. Испытания макета подтвердили правильность основной идеи метода, хотя аппаратная реализация на элементной базе того времени представляла большие трудности. Интегральных схем ещё не было, транзисторы с весьма посредственными характеристиками только появлялись и были дефицитом, так что в первом макете использовались лампы, диоды и множество катушек. Линия задержки содержала 150 звеньев LC, параметры которых для исключения отражений приходилось подбирать с высокой точностью. Вместо умножителей использовались ключевые схемы на диодах, так что вместо автокорреляционной функции использовалась функция взаимной корреляции речевого сигнала (РС) и клиппированного РС. В 1960–1967 годы были созданы макеты на транзисторах с использованием элементов цифровой техники. Однако широкое использование автокорреляционных выделителей основного тона (ВОТ) в составе вокодеров началось с появлением БИС и микропроцессоров [3–5].



Первые макеты автокорреляционных ВОТ работали удовлетворительно, правда, при подавлении в РС частот ниже 300 Гц (коммутируемый телефонный канал) количество ошибок резко возрастало. Попытки улучшить работу путём различных усовершенствований схем и алгоритмов не давали желаемых результатов. Выявилась необходимость разработки общего подхода к проблеме выделения ОТ с позиций статистической теории оценивания параметров случайных сигналов [6–10].

Статистический подход к проблеме выделения ОТ

Речевой сигнал представляет собой нестационарный случайный процесс с неизвестными параметрами, один из которых — основной тон. Обычно для удобства анализа РС разбивается на квазистационарные сегменты длительностью 10–45 мс и на каждом сегменте параметры сигнала считаются постоянными. РС на вокализованном сегменте можно представить в виде суммы гармоник частоты основного тона с априори неизвестными амплитудами и фазами. Отклонения реального РС от такой математической модели, обусловленные нестационарностью РС в пределах сегмента и внешними помехами, будем считать шумом. Задача анализа РС сводится, таким образом, к статистической задаче оценки неизвестных параметров случайного процесса, искажённого шумом. При определённых допущениях [6, 8] оптимальным алгоритмом оценки периода ОТ на рассматриваемом сегменте является вычисление автокорреляционной функции (АКФ) РС и определение аргумента, при котором АКФ максимальна.

Достоверное определение значения ОТ по одному сегменту иногда оказывается невозможным. Существенное повышение точности может быть достигнуто при учёте зависимостей значений периодов ОТ на соседних сегментах. Обычно периоды ОТ на соседних сегментах близки, хотя бывают и исключения. Оптимальный алгоритм оценки периода ОТ на данном сегменте должен учитывать оценки периодов ОТ на соседних сегментах и вероятности изменений периодов ОТ.

Простейшей моделью движения ОТ является Марковский процесс, для такой модели алгоритм обработки значений АКФ оказывается достаточно простым и сводится к фильтрации значений АКФ линейным фильтром с импульсной реакцией, определяемой функцией вероятностей изменений периодов ОТ на соседних сегментах. Результаты фильтрации на соседних сегментах используются в качестве априорной информации при определении периода ОТ на данном сегменте [6–10].

Марковский процесс первого порядка (случайные блуждания типа броуновского движения) представляет весьма грубую модель изменений периодов ОТ реального РС. Тем не менее даже такая простая модель позволяет существенно повысить точность измерений. Дальнейшее повышение точности может быть достигнуто увеличением порядка модели. Модель второго порядка, учитывающая не только предыдущее значение ОТ, но и тенденцию изменений ОТ, больше соответствует характеристикам РС. Изменения периодов ОТ на соседних обычно невелики, хотя встречаются и значительные изменения, так что вероятности переходов плохо описываются гауссовским законом, более точное описание даёт сумма двух гауссовских распределений с разными дисперсиями [7, 8]. Вероятности больших изменений увеличиваются на границах вокализованных участков речи

и при изменениях характера звука, так что более совершенная модель движения ОТ представляется нестационарным процессом с изменяющейся дисперсией функции вероятностей переходов. Учёт этих особенностей позволяет повысить точность работы ВОТ.

Критерии периодичности

Общим определением периодичности процесса является повторяемость формы сигнала. Минимальный временной сдвиг, при котором достигается повторяемость формы, называется периодом. Посторонние шумы и нестационарность РС исключают возможность точного повторения формы через период ОТ. В связи с этим встаёт вопрос о критерии сравнения исходного и сдвинутого во времени фрагментов РС при определении периодичности. Для шумов с гауссовским распределением следует использовать среднеквадратический критерий, для этого критерия оптимальна автокорреляционная обработка. Однако внешние шумы (комнатные, шумы телефонной линии) обычно отличаются от гауссовских. Для выяснения свойств внутренних шумов, учитывающих отклонение реального РС от модели периодического процесса, проведён следующий эксперимент.

На вокализованных участках вычислялись абсолютные значения разностей между отсчётами РС, разнесёнными на период ОТ. Частота дискретизации 16 кГц получена интерполированием отсчётов РС [11] с частотой дискретизации 8 кГц. Для исключения влияния вариаций уровня на результаты измерений сигналы нормировались таким образом, чтобы средние уровни сигналов на сравниваемых фрагментах были одинаковы. На рис. 1 чёрным цветом представлена функция распределения PR абсолютных значений разностей отсчётов, отстоящих на период ОТ. По оси абсцисс отложены величины DS разностей отсчётов, нормированные относительно среднего абсолютного значения сигнала на периоде ОТ.

Полученное распределение хорошо аппроксимируется функцией вида $pr(ds) = \alpha \times \exp(-\beta \times |ds|)$, где α, β — постоянные коэффициенты (красная кривая). Функция гауссовского распределения $pr(ds) = \lambda \times \exp(-\mu \times ds^2)$ (зелёная кривая) хуже представляет это распределение (ошибка аппроксимации при оптимальных значениях параметров λ, μ больше в 5,8 раза). Из этого можно заключить, что при оценке периода ОТ лучше пользоваться не критерием минимизации среднего квадрата отклонений (что эквивалентно максимизации АКФ), а критерием минимизации средней величины абсолютных значений отклонений. Соответствующий вариант разностной (сдвиговой) функции $S(t, \tau)$ [12, 13], по максимуму которой определяется период ОТ, может быть записан в виде

$$S(t, \tau) = am / \left(\sum_{i=1}^{\tau} |x(t+i) - gain \times x(t+\tau+i)| + \delta \times am \right), \quad (1) \text{ где}$$
$$am = \sum_{i=1}^{\tau} |x(t+i)|, \quad gain = am / \sum_{i=1}^{\tau} |x(t+\tau+i)|.$$

Параметр δ (в наших экспериментах $\delta=0,0625$) позволяет регулировать соотношение между максимумами функции $S(t, \tau)$.

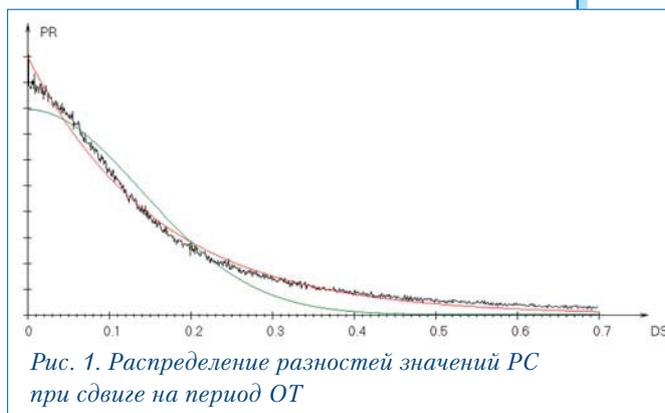


Рис. 1. Распределение разностей значений РС при сдвиге на период ОТ

На рис. 2, 5 показан вид разностных (сдвиговых) функций (РСФ) $S(t, \tau)$ и АКФ $R(t, \tau)$. Функции вычислены для одного и того же сегмента сигнала $x(t)$, отсчёты по t и τ берутся через $1/16$ кГц = 0,0625 мс. Пики у РСФ более острые, чем у АКФ, поэтому увеличивать интервал между отсчётами до часто используемого значения $1/8$ кГц = 0,125 мс нежелательно. Положение пиков на оси τ у этих функций отличается обычно незначительно, но соотношение амплитуд пиков разное. Пик, соответствующий периоду ОТ, у РСФ чаще всего более выражен. К достоинствам РСФ можно отнести также меньшие вычислительные затраты — многочисленные умножения, производимые при вычислении АКФ, заменяются более простыми операциями сложения и отбрасывания знака у результата сложения.

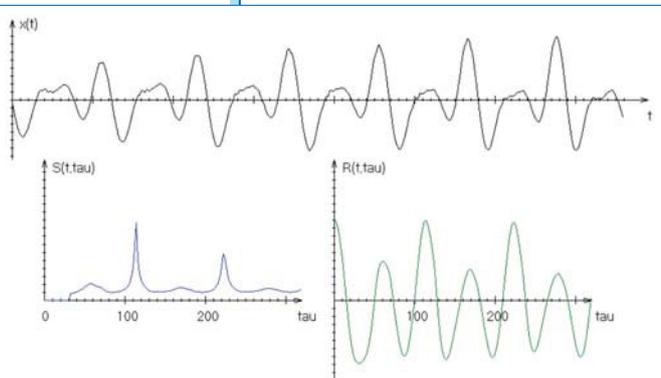


Рис. 2. Пример функции $S(t, \tau)$ и $R(t, \tau)$ сигналах $x(t)$

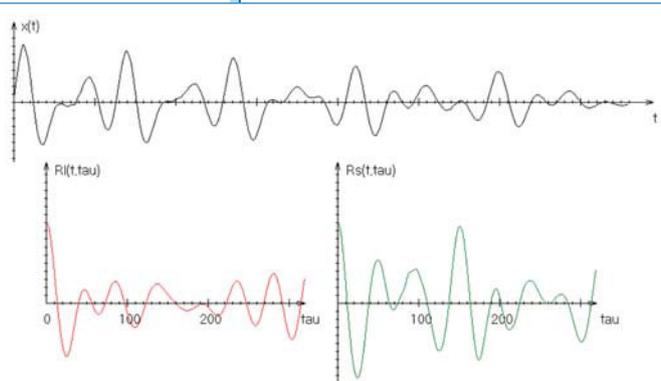


Рис. 3. Пример функции $RI(t, \tau)$ и $Rs(t, \tau)$ при быстром увеличении периодов ОТ сигнала $x(t)$

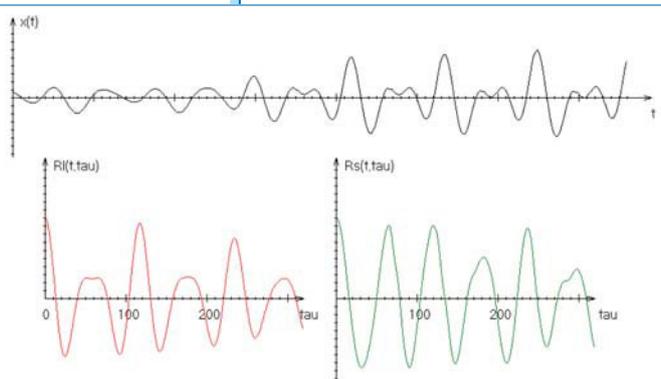


Рис. 4. Пример функции $RI(t, \tau)$ и $Rs(t, \tau)$ при резком изменении формы сигнала $x(t)$

При вычислении АКФ время усреднения произведений $x(t) \times x(t+\tau)$ обычно постоянно и составляет 10–45 мс. В данной работе при вычислении АКФ и РСФ время усреднения переменное и равно предполагаемому периоду ОТ, для которого вычисляется значение функции. Благодаря этому появляется возможность измерять не среднее значение периода на сегменте, а значения каждого периода ОТ. На нестационарных участках при быстрых изменениях периода ОТ разница этих двух методов существенна. На рис. 3 приведён пример АКФ $RI(t, \tau)$ (время усреднения 45 мс) и АКФ $Rs(t, \tau)$ (время усреднения для каждого отсчёта функции равно предполагаемому значению периода ОТ) для РС $x(t)$ с быстрым изменением ОТ. У функции $Rs(t, \tau)$ пик, соответствующий периоду ОТ, явно выражен.

Время усреднения

При большом времени усреднения значения АКФ более устойчивы. Если период ОТ за время усреднения изменяется незначительно, а форма речевой волны резко изменяется (рис. 4), то пик функции $RI(t, \tau)$ при τ равном периоду ОТ более выражен, чем пик функции $Rs(t, \tau)$. Тем не менее, представляется целесообразным на первом этапе обработки вычислять АКФ для каждого периода ОТ (минимальное время усреднения), а функции повышения достоверности за счёт обработки данных на большом интервале времени выполнять на втором этапе путём нелинейной фильтрации замеров АКФ или РСФ.

Учёт периодичности АКФ и РСФ

Для периодического процесса АКФ и РСФ имеют пики при значениях τ , кратных периоду ОТ. В реальном РС вследствие нестационарности амплитуды пиков по мере увеличения τ обычно уменьшаются, однако, бывают исключения и пик, например, при задержке на $2 \times \tau$ может быть больше пика при задержке на время τ , равное периоду ОТ. Для исключения таких ошибок иногда значения АКФ умножают на некоторую монотонно уменьшающуюся функцию от τ . Это даёт некоторый эффект, хотя вероятность ошибочных выборов пиков на задержках, меньших периода ОТ, увеличивается.

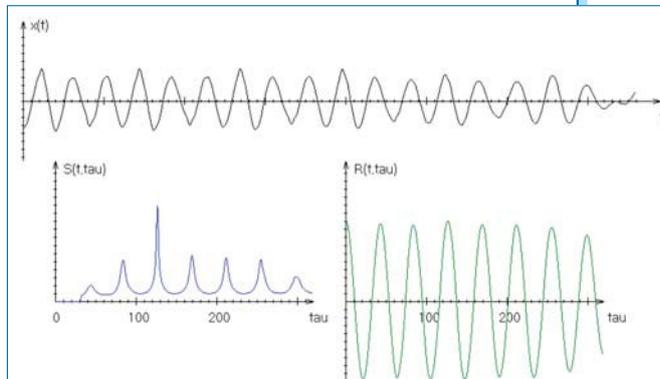


Рис. 5. Пример функции $S(t, \tau)$ и $R(t, \tau)$ сигнала $x(t)$ с подавлением частот ниже 300 Гц

Из теории статистического оценивания следует, что при измерении периода ОТ надо учитывать не только пик при задержке на τ , но также и пики, соответствующие задержкам на $2 \times \tau$, $3 \times \tau$ и т.д. [6, 8]. Действительно, присутствие пика при задержке $2 \times \tau$ можно рассматривать как подтверждение достоверности замера по пику при задержке τ (если такой имеется). Для проверки эффективности одного из возможных алгоритмов учёта периодичности РСФ проведён следующий эксперимент.

Использовались тестовые файлы с речью дикторов типа S [11]. В исходных файлах частота дискретизации 8 кГц. Путём интерполяции получены отсчёты с частотой следования 16 кГц. Приведённые в работе [11] данные о периодах ОТ представляют усреднённые значения, следующие через 11,25 мс. На участках с быстрыми изменениями ОТ реальные значения периодов могут существенно (больше порогового значения $\xi=0,1$) отличаться от приведённых усреднённых значений, поэтому использованы откорректированные данные о периодах, следующие через 4 мс. Увеличена доля вокализованных участков в исследуемом речевом материале. В разметке [11] эта доля составляет около 52%, в используемой нами разметке вокализованные участки занимают 60% общей длительности, увеличение достигнуто за счёт учёта вокализованных звуков с малой амплитудой.

При измерении периодов ОТ по положению максимума РСФ в 6,44% случаев наблюдались грубые ошибки (отличие измеренного периода от истинного более 10%). Если же выбирать два самых больших пика РСФ $m_1(\tau_1)$ и $m_2(\tau_2)$ ($\tau_2 > \tau_1$) и при соблюдении условия $|\tau_1 - \tau_2/2| < \delta$ увеличивать амплитуду пика $m_1(\tau_1)$ на $\nu \times m_2(\tau_2)$, после чего определять период ОТ по положению пика с максимальной амплитудой, то при таком алгоритме количество грубых ошибок измерения ОТ уменьшается до 3,69% (если τ — количество тактов частоты 16 кГц, то оптимальное значение $\delta=6$. Оптимальное значение $\nu=0,31$).

Исследовалась зависимость вероятности грубых ошибок от частоты среза ФНЧ, ограничивающего полосу частот, поступающих на вход измерителя. При измерении периодов ОТ по максимуму РСФ при работе с РС, содержащими низкочастотные составляющие (тестовые сигналы типа S [11]), оказалось целесообразным установить частоту среза ФНЧ равной максимальной частоте ОТ (500 Гц). При увеличении полосы анализируемых частот до 1000 Гц количество грубых ошибок увеличивается в 1,07 раза. В дальнейших экспериментах частота среза ФНЧ выбрана равной 500 Гц. Очевидно, что при работе с РС с подавленными низкочастотными составляющими частоту среза ФНЧ необходимо увеличить.



Приведённые результаты относятся к случаю измерения периодов ОТ по функции РСФ $S(t, \tau)$ для фиксированных значений t без учёта результатов измерений на соседних интервалах. Далее рассматриваются возможности повышения достоверности за счёт учёта зависимостей между периодами ОТ.

Учёт зависимостей между периодами ОТ

В качестве кандидатов на результат замера периода ОТ функции $S(t, \tau)$ для момента времени t рассматриваются восемь значений $\tau_1(t), \tau_2(t), \dots, \tau_8(t)$, соответствующих восьми наибольшим пикам функции $S(t, \tau)$. Каждому замеру $\tau_i(t)$ присваивается вес, пропорциональный величине функции $S(t, \tau_i(t))$. Простейший способ учёта результатов замеров на соседних сегментах состоит в следующем. Для каждого $\tau_i(t)$ строится траектория движения ОТ по прошлым и будущим значениям и вес этих траекторий прибавляется к весу замера $\tau_i(t)$, после чего выбирается замер с максимальным суммарным весом. Траектории можно строить по правилу выбора ближайших (минимум абсолютной величины разности) замеров ОТ при движении вперёд и назад. Вес траектории можно вычислять в виде суммы весов выбранных замеров, умноженной на весовой коэффициент W_k . Этот алгоритм назовём Алг. 1.

В Таблице 1 приведены результаты экспериментов. Длина траектории L определяет количество соседних замеров, учитываемых при движении в каждую из сторон. $L=0$ означает, что соседние замеры не используются. Коэффициенты W_k определяют веса, с которыми складываются замеры $S(t-k, \tau)$ и $S(t+k, \tau)$ при формировании суммарного веса замеров для момента t . P_e — количество грубых сбоев в процентах.

Таблица 1

Процент ошибок P_e для Алг. 1

L	0	1	2	3	5
$P_e, \%$	3,69	2,01	1,74	1,41	1,30
W_k	—	0,60	0,38	0,32	0,12

На рис. 6 представлен пример работы Алг.1. По оси ординат отложены значения задержки τ при измерении периодов ОТ. Значения τ задаются количеством тактов частоты 16 кГц. По оси абсцисс отложены моменты времени t , следующие через 4 мс (64 такта частоты 16 кГц). Для каждого t выбираются восемь наибольших пиков РСФ, значения τ для которых отображены цветными кружками. Цвета назначены в порядке уменьшения амплитуд пиков следующим образом: красный, жёлтый, зелёный, синий, чёрный, далее белый. Правильные значения периодов представлены красной траекторией. Выбор результата измерения ОТ производится для $t=10$. Результаты предыдущих выборов отображены двойной зелёной траекторией. Рис. 6 показывает, что полученная траектория измерений ОТ совпадает с траекторией правильных значений, грубые ошибки измерения периодов ОТ, которые произошли бы при $t=3$ и $t=8$ в случае измерения путём выбора максимального пика РСФ без учёта замеров на соседних интервалах, исправляются.

На рис. 7 представлен пример, когда обработка замеров согласно Алг. 1 не исправляет, а даже увеличивает ошибки. Однако такие случаи встречаются редко, в среднем обработка замеров по Алг. 1 позволяет уменьшить количество грубых ошибок примерно в три раза.

Ошибки определения ОТ чаще всего случаются на начальных и конечных участках вокализованных сегментов. Здесь обычно наблюдаются значительные вариации периодов ОТ. Кроме того, при формировании траектории одна из её частей попадает на паузу или на невокализованный сегмент речи. Функция (1) инвариантна к уровню РС, поэтому на паузах иногда встречаются случайные выбросы, искажающие вес траектории. Для устранения этого недостатка при формировании веса траектории учитываются уровни сигнала на соответствующих участках траектории путём умножения значений пиков функции $S(t, \tau)$ на корень квадратный из средней амплитуды сигнала для времени t .

В Алг. 1 при построении траекторий выбирается ближайший по τ соседний замер, но не учитывается достоверность замера (величина пика РСФ) и абсолютная величина расстояния по τ (вероятность перехода). В более совершенном Алг. 2 при построении траектории выбирается замер с максимальной вероятностью произведения достоверности измерения на вероятность перехода. В логарифмическом масштабе это сумма величины пика РСФ и логарифма вероятности перехода. Движение ОТ аппроксимируется нестационарным Марковским

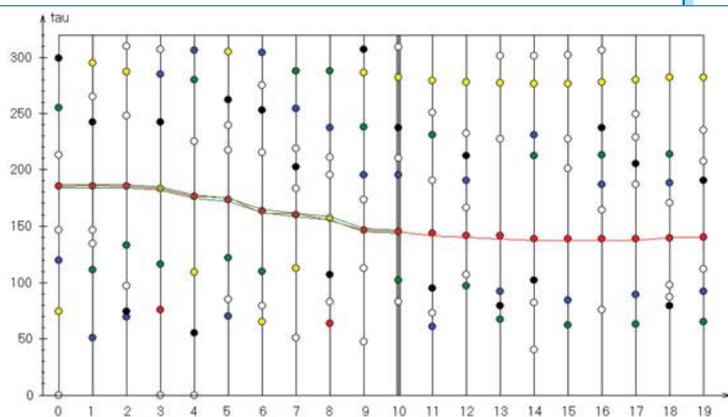


Рис. 6. Замеры ОТ и результаты их обработки. Алгоритм 1. Пример 1

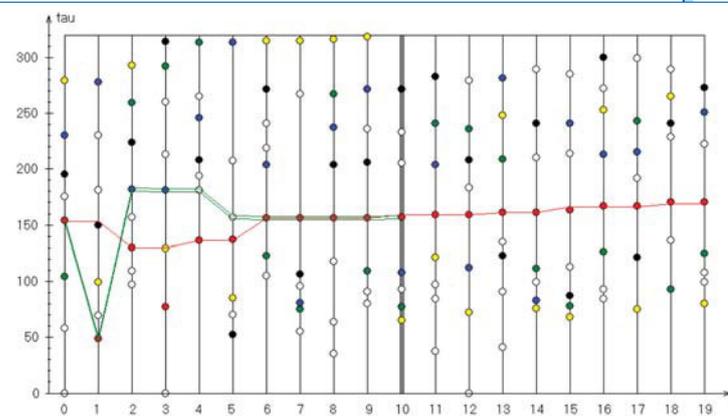


Рис. 7. Замеры ОТ и результаты их обработки. Алгоритм 1. Пример 2

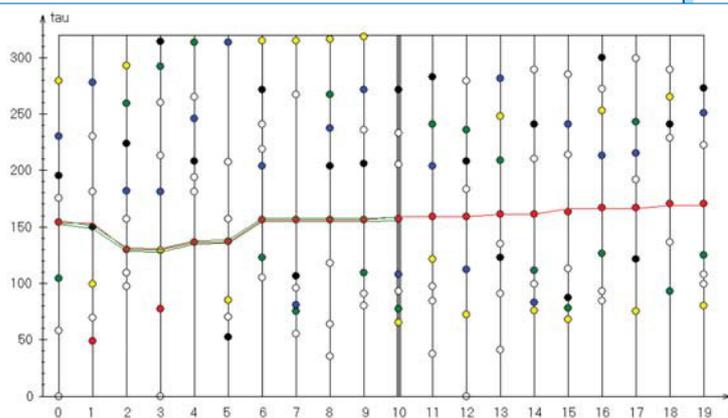


Рис. 8. Замеры ОТ и результаты их обработки. Алгоритм 2. Пример 1

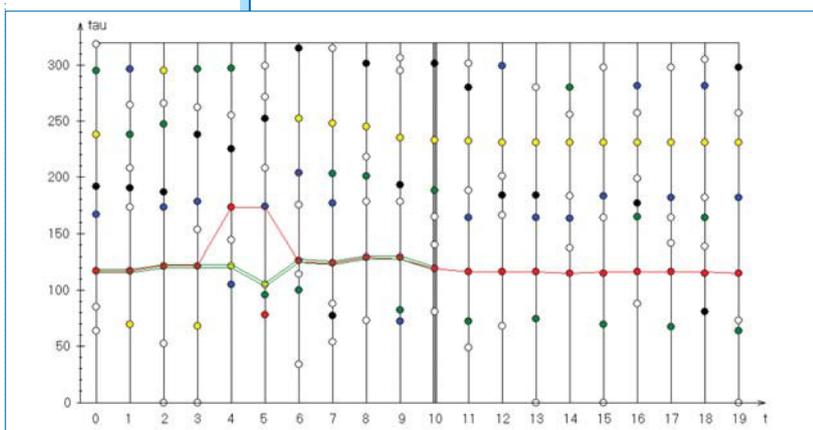


Рис. 9. Замеры ОТ и результаты их обработки. Алгоритм 2. Пример 2

процессом с функцией переходов вида $\exp(-|d\tau|/\sigma(t))$, где $d\tau$ — изменение периода ОТ на интервале между соседними замерами, $\sigma(t)$ — изменяющаяся во времени дисперсия. Дисперсия увеличивается на границах вокализованных сегментов и при резком изменении формы РС. Грубой оценкой $\sigma(t)$ могут служить вариации уровня анализируемого РС. В Алг. 2 вес траектории определяется в виде суммы амплитуд выбранных пиков РСФ и логарифмов вероятностей переходов $-|d\tau|/\sigma(t)$.

Испытания Алг. 2 при длине траектории $L=9$ (дополнительная задержка на $4 \times 9 = 36$ мс) дали 0,44% грубых ошибок. На рис. 8 представлены результаты обработки по Алг. 2 тех же данных, что приведены на рис. 7. Там, где Алг. 1 даёт увеличение количества ошибок, Алг. 2 полностью исправляет ошибки. Но и Алг. 2 не исправляет ошибок во всех случаях, а иногда и увеличивает их количество. На рис. 9 приведён пример, когда резкое изменение периода ОТ при $t=4$ и $t=5$ вызвало две грубые ошибки (при выборе максимального пика РСФ без учёта соседей была бы только одна ошибка при $t=5$). Тем не менее, в среднем Алг. 2 обеспечивает существенное уменьшение количества грубых ошибок. В таблице 2 приведены опубликованные в работе [14] результаты сравнения различных ВОТ с разработанным В.В. Бабкиным алгоритмом. Таблица дополнена двумя последними столбцами, где приведены данные наших исследований.

Таблица 2

Вероятности грубых ошибок P_e для различных алгоритмов выделения ОТ при отсутствии внешних шумов [14]. Тестовый сигнал типа S [11]

ВОТ	G.729.AB	G.723.1	MELP	LPC10E	HVXC	[14]	Алг. 1	Алг. 2
P_e , %	10,6	10,9	4,6	6,0	5,7	4,5	1,3	0,4

Помехоустойчивость

Работа программы Алг. 2 проверялась при наложении на тестовый сигнал S [11] помех с равномерным спектром. Испытания проводились с помехами с гауссовским распределением (ГП) и с помехами с экспоненциальным распределением амплитуд (ЭП). Отношение сигнал/шум (С/Ш) определялось как отношение средней мощности РС на вокализованных участках к средней мощности помехи. В таблице 3 приведены результаты испытаний.

Таблица 3

Вероятности грубых ошибок P_e для ВОТ Алг. 2 при воздействии помех

С/Ш, дБ	20	15	10	5
ГП, P_e , %	1,6	3,3	6,6	13,3
ЭП, P_e , %	1,0	1,6	2,7	4,8

Полученные данные уступают результатам, опубликованным в работе [14] (для помехи с гауссовским распределением $P_e=6,9\%$ при $C/Ш=0$ дБ). Видимо, это объясняется двумя причинами:

1. В использованной нами разметке сигнала S к вокализованным звукам отнесены сегменты РС с малыми амплитудами, которые в разметке [11] считаются паузой. При воздействии шумов ошибки возникают чаще всего на слабых сигналах.
2. Параметры Алг. 2 оптимизированы на незашумленных РС. При воздействии шумов некоторые параметры целесообразно изменить. Например, надо увеличить время накопления при вычислении РСФ. Последнее исключит возможность отслеживать быстрые изменения периодов ОТ, но средний процент ошибок при воздействии сильных шумов уменьшится.

При воздействии помех типа ЭП помехоустойчивость ВОТ существенно выше. Видимо, это объясняется тем, что алгоритм вычисления РСФ ориентирован на экспоненциальное распределение помех. В связи с этим возникло предположение, что при помехах с гауссовским распределением измерение периодов ОТ по АКФ даст существенно лучшие результаты, чем измерение по РСФ. Эксперименты не подтвердили это предположение. При измерении периодов ОТ путём выбора максимума АКФ или РСФ (без учёта замеров на соседних сегментах) при воздействии гауссовских помех количество грубых ошибок оказалось примерно одинаковым.

Заключение

Автокорреляционный метод выделения ОТ речи, предложенный проф. А.А. Пироговым 50 лет назад, в настоящее время широко используется при исследованиях речи и в аппаратуре связи и даёт хорошие результаты. Перспективы дальнейшего совершенствования метода связаны с более полным учётом особенностей речеобразования и с использованием аппарата теории статистических решений для синтеза алгоритмов обработки речевого сигнала, приближающихся к оптимальным. Точность измерений определяется уровнем внутренних шумов, которые учитывают отклонения используемой математической модели от реального РС. Нормализация по уровню (коэффициент $gain$ в формуле (1)) позволяет уменьшить уровень шумов, обусловленных изменениями амплитуд РС на анализируемом интервале, где, согласно модели, сигнал представляется периодическим процессом с постоянными амплитудами гармоник. Следующим шагом может быть нормализация по форме РС. Резкие изменения формы речевой волны (рис. 4), обусловленные перестройками артикуляционного аппарата, затрудняют измерения периодов ОТ. Нормализация по форме, например, путём выравнивания спектра также может способствовать решению трудной проблемы выделения основного тона речи.



Литература

1. Пирогов А.А. Устройство для автоматического определения частоты основного тона. Реестр изобретений СССР. Авторское свидетельство №129739 с приоритетом от 08.06.1958 г. Бюллетень изобретений и товарных знаков. 1960. № 13. С. 38.
2. Баронин С.П. Автокорреляционный метод выделения основного тона речи: Сб. трудов Гос. НИИ Мин. связи СССР. Вып. 3(24). М., 1961. С. 93–102.
3. Coding of Speech at 8 kbit/s. ITU-T Recommendation G.729. 1996.
4. Dual Rate Speech Coder. ITU-T Recommendation G.723.1. 1996.
5. Speech Service Option Standard for Wideband Spread Spectrum Systems. ANSI/TIA/EIA-96-C. 1998.
6. Баронин С.П. Статистические методы анализа речевых сигналов // Электросвязь, 1966. № 5. С. 50–56.
7. Баронин С.П. О построении многоканальных схем выделения основного тона речи: Сб. тр. Гос. НИИ мс. 1965. Вып. 3(39). С. 17–24.
8. Баронин С.П. Статистические методы анализа речевых сигналов: Канд. дисс. М.: НИИР, 1968.
9. Баронин С.П., Куштуев А.И. О построении схем адаптации анализаторов частоты основного тона речи: Тез. докл. 7-й Всесоюзной акустической конференции. Л., 1971.
10. Вокодерная телефония / Под ред. А.А. Пирогова. М.: Связь, 1974.
11. Бабкин В.В. Тестовые файлы для оценки помехоустойчивости выделителей ОТ. Файл S. СПб.: Центр ЦОС СПб ГУТ, 2005.
12. Соболев В.Н., Баронин С.П. Исследование сдвигового метода выделения основного тона речи // Электросвязь. 1968. № 12. С. 30–36.
13. Соболев В.Н. Информационные технологии в синтетической телефонии. М.: ИРИАС, 2007.
14. Бабкин В.В. Помехоустойчивый выделитель основного тона речи. 7-я Международная конференция и выставка «Цифровая обработка сигналов и её применение (DSPA-2005)». М., 2005.

Баронин Сергей Павлович —

кандидат технических наук, ведущий научный сотрудник Государственного научно-производственного центра «ВИГСТАР». Занимается исследованиями и разработкой современных систем связи, в том числе систем эффективного кодирования и автоматического распознавания речи. Опубликовал более 100 научных работ, получил авторские свидетельства на 30 изобретений. Работы, посвящённые автокорреляционному методу выделения основного тона речи и статистическим методам анализа речи, в том числе анализу авторегрессионной модели речи (метод линейного предсказания), очистки речи от шумов путём обработки спектров, векторному квантованию дельта-кепстров (фонетической функции Пирогова) с целью сжатия информации, опубликованы ещё в 60-е годы прошлого века.