

ИЗМЕРИТЕЛЬНАЯ ТЕХНИКА

УДК 621.391

Д. Ю. Ваурун, магистрант;
В. Б. Гитлин, доктор технических наук, профессор
Ижевский государственный технический университет имени М. Т. Калашникова

ВЫДЕЛЕНИЕ ОСНОВНОГО ТОНА МЕТОДОМ SWIPE ИЗ СИГНАЛА, ПРОШЕДШЕГО ТЕЛЕФОННЫЙ КАНАЛ

Рассмотрена работа алгоритмов SWIPE и SWIPE' при выделении основного тона речевого сигнала, ограниченного полосой телефонного канала. Показано, что для реальных сигналов, ограниченных полосой частот телефонного канала, указанные алгоритмы склонны к переходу на гармоники основного тона. Предложен алгоритм, позволяющий повысить надежность выделения основного тона алгоритмами SWIPE и SWIPE' для сигналов указанного типа.

Ключевые слова: основной тон, гармоники, скалярное произведение, пилообразный сигнал

Алгоритм SWIPE (Sawtooth Inspired Pitch Estimator – оценка основного тона по пилообразному сигналу) [1] оценивает частоту основного тона (ОТ) речи и музыки путем поиска максимума скалярного произведения между спектром анализируемого сигнала и спектром пилообразного сигнала, период которого известен. Вычисления выполняются во всем диапазоне частот, вплоть до частоты Найквиста. Решение ТОН/НЕ ТОН принимают по превышению максимумом скалярного произведения определенного порога.

Алгоритм SWIPE и его модификация SWIPE' относятся к интегральным алгоритмам [2]. Испытания этих алгоритмов показали [3], что они хорошо работают при низких соотношениях сигнал/шум. Однако количество больших ошибок выделения ОТ для сигнала, прошедшего телефонный тракт, существенно возрастало. Под большими ошибками в этих испытаниях понимали отклонения измеренных значений частоты ОТ от эталонных, превышающие 15 %.

Таблица 1. Большие ошибки выделения ОТ методами GS2 и SWIPE'

Алгоритмы	Большие ошибки (%)			
	Чистый сигнал	С/Ш 5дБ	Телефонный	Клиппированный
GS2	1.9	15.24	6.19	0.55
SWIPE'	1.31	6.66	47.54	0.03

Было также проведено сравнение надежности выделения ОТ алгоритмом SWIPE' с надежностью выделения ОТ алгоритмом GS2 [3]. Алгоритм GS2 относится к локальным алгоритмам [4]. Сравнение работы этих алгоритмов выполняли только по количеству больших ошибок. Результаты этого сравнения показаны в табл. 1 [3]. Отметим, что алгоритм SWIPE' требовал тщательного подбора порога принятия решения ТОН/НЕ ТОН, в то время как алгоритм GS2 работает в полностью автоматическом режиме [4].

Отмеченные особенности алгоритмов SWIPE и SWIPE' потребовали анализа неудовлетворительной их работы с сигналом, прошедшим телефонный канал, и разработки мер по уменьшению количества больших ошибок для этого типа сигнала.

В методах SWIPE и SWIPE' скалярное произведение вычисляется в модифицированной шкале частот ERBs (Equivalent-rectangular-bandwidth scale) [1]. Шкала ERBs моделирует свойства слуховой системы человека. Пересчет из частотной шкалы в Гц в шкалу ERBs выполняется по формуле [1]:

$$ERBs = 21,41 \log_{10}(1 + f / 229) . \quad (1)$$

Шкала ERBs растягивает область низких частот спектра, придавая больший вес основной частоте спектра, и тем самым снижает вероятность перехода на гармоники или субгармоники.

В качестве тестового образца (template), период которого известен и с которым производится сравнение, в методах SWIPE и SWIPE' выбран пилообразный сигнал [1]. Амплитуда гармоник пилообразного сигнала обратно пропорциональна номеру гармоники, что соответствует типовому наклону спектра вокализованного сигнала в 6 дБ/окт [2]. Вероятность получения максимума скалярного произведения на частоте ОТ, совпадающей с частотой ОТ анализируемого сигнала, повышается.

Амплитудный спектр сигнала всегда неотрицателен. В работе [1] вместо спектра пилообразного сигнала в ядре интегрального преобразования, вычисляющего скалярное произведение, используют косинус. Амплитуды лепестков косинусов уменьшаются с повышением номера лепестка. Но, в отличие от спектра пилообразного сигнала, знаки лепестков косинуса меняются с положительных на отрицательные и наоборот. Наличие отрицательных лепестков подавляет шум между гармониками анализируемого сигнала и повышает надежность правильного решения о частоте ОТ.

Максимум скалярного произведения может быть достигнут в том случае, когда форма спектральных лепестков сигнала близка к форме лепестков косинуса ядра. С целью получения спектральных лепестков сигнала необходимой формы перед выполнением преобразования Фурье сигнал умножают на окно Ханна, имеющего ширину, определяемую предполагаемым значением частоты ОТ. Повышая надежность выделения ОТ, изменение ширины окна требует перевычисления спектра, что существенно увеличивает количество вычислений при непосредственной реализации алгоритмов SWIPE и SWIPE'. До вычисления скалярного произведения, кроме того, берут корень квадратный из огибающей спектра сигнала. Эта операция способствует приближению спектра сигнала к спектру ядра преобразования и повышает выделенность максимума скалярного произведения. В работе [1] изложены методы, позволяющие существенно снизить количество вычислений при реализации алгоритмов SWIPE и SWIPE'. Однако при этом несколько увеличивается погрешность вычисления скалярного произведения.

С целью снижения вероятности перехода на субгармоники в алгоритме SWIPE' в отличие от алгоритма SWIPE скалярное произведение вычисляют только для гармоник с номерами, соответствующими простым числам.

Отмеченные выше особенности спектра речевого сигнала выполняются не всегда. Огибающая спектра сигнала сильно зависит от формантной структуры и от типа источника звука. Например, фонема /и/ имеет сильный спад огибающей спектра в полосе 500...1 500 Гц и подъем в области 1700...2 200 Гц [5, 6]. Звонкие согласные, со смешанным источником возбуждения, также характеризуются подъемом спектра в высокочастотной области. В системах обработки речи часто искусственно выполняют подъем спектра в области высоких частот. Гармоническая структура спектра в большинстве случаев сохраняется лишь в диапазоне частот 2...2,5 кГц, причем с повышением частоты форма гармоник размывается [4–6]. Указанные особенности спектров реальных звуков могут служить причиной ошибок при выделении ОТ методами SWIPE и SWIPE'. Пример среднего спектра слога /не/ (Фраза «Не видали мы такого невода», диктор vbg), иллюстрирующий отмеченные выше особенности спектра речи, показан на рис. 1. Показаны средние спектры чистого сигнала (без аддитивного шума и без какого-либо ограничения спектра) и сигнала, ограниченного полосой телефонного канала. По оси абсцисс на рис. 1 отложена частота в Гц, по оси ординат – логарифм амплитуды. Спектр получен с помощью программы SIS.6x [7].

На рис. 2 показана осциллограмма и контур ОТ фразы «Не видали мы такого невода» этого же диктора, полученный алгоритмом SWIPE. Предельное верхнее значение частоты ОТ, задаваемое как один

из входных параметров алгоритма, равно $F_H = 500$ Гц. За исключением оценки интервалов наличия и отсутствия вокализации, этот контур достаточно близко соответствует эталонному контуру ОТ. Частота ОТ на графике рис. 2 изменяется в диапазоне от 80 до 160 Гц.

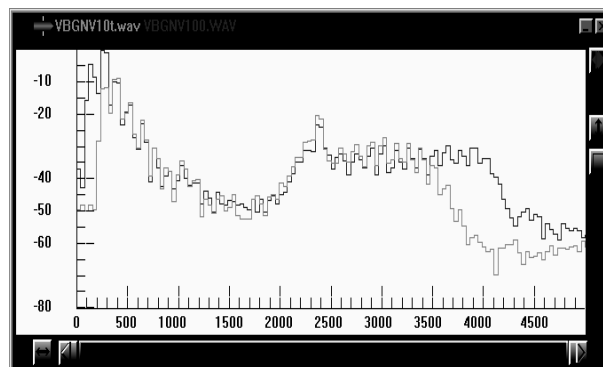


Рис. 1. Средний спектр слога /не/

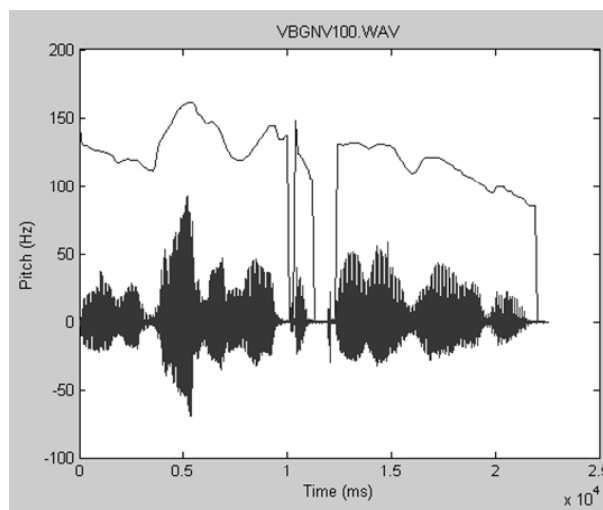


Рис. 2. Контур основного тона, полученный для чистого сигнала алгоритмом SWIPE

На рис. 3 показаны контуры ОТ, полученные методами SWIPE' и SWIPE для сигнала, ограниченного полосой телефонного канала, для этой же фразы и $F_H = 500$ Гц. Видны многочисленные сбои на гармониках ОТ, значение частот которых приближается к F_H . Было сделано предположение, что причиной появления этих сбоев являются меры, предпринятые в методах SWIPE и SWIPE' для борьбы с субгармониками при наличии отмеченных выше особенностей реального речевого сигнала. Было также высказано предположение, что этих ошибок можно избежать, если ограничить диапазон поиска частоты ОТ сверху, уменьшив значение F_H . Анализ рис. 3 также показывает, что информация об истинном значении частоты ОТ сохраняется в контуре ОТ. Она содержится в значениях локальных минимумов траектории ОТ.

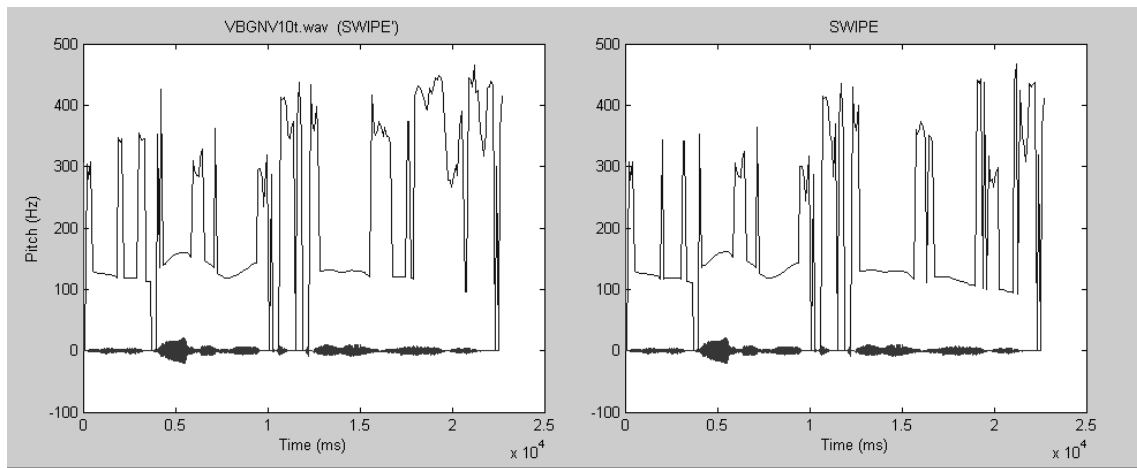


Рис. 3. Контуры основного тона, полученные для сигнала, ограниченного полосой телефонного канала, методами SWIPE' и SWIPE

С целью проверки указанного предположения проведен эксперимент, результаты которого показаны на рис. 4. На этом рисунке представлен контур ОТ, полученный методом SWIPE для сигнала, ограниченного полосой телефонного канала, когда значение F_H снижено до $F_H = 200$ Гц. В этом случае сбои на гармоники отсутствовали. Более того, по сравнению с контуром, показанным на рис. 2, количество ошибочных решений о наличии вокализации даже уменьшилось.

В качестве одного из вариантов автоматического выбора верхней предельной частоты основного тона предложен алгоритм, схема которого показана на рис. 5. Алгоритм использует особенности в полученной алгоритмами SWIPE' и SWIPE траектории ОТ, заключающиеся в сохранении информации об истинных значениях частоты ОТ в минимумах траектории, в возможности коротких интервалов отсутствия вокализации и в наличии больших скачков при сбоях на гармоники ОТ.

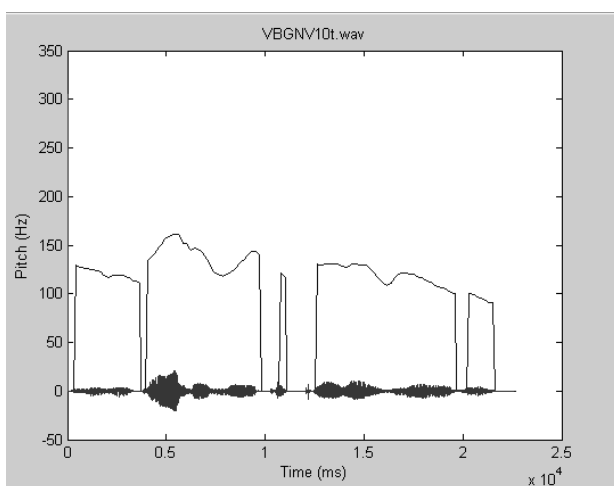


Рис. 4. Контур основного тона, полученный для сигнала, ограниченного полосой телефонного канала, методом SWIPE с верхней предельной частотой основного тона 200 Гц

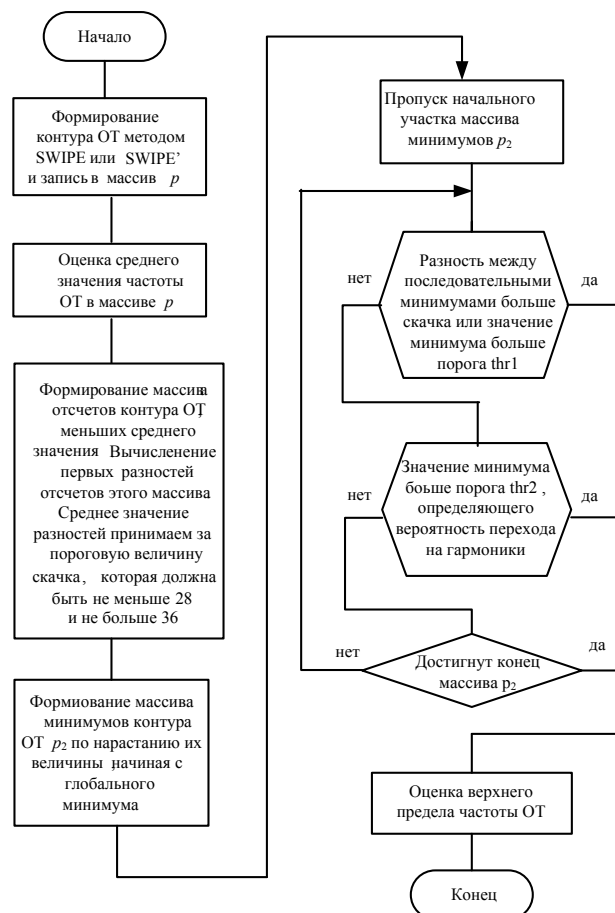


Рис. 5. Структура алгоритма оценки верхнего предела частоты ОТ

Порог $thr1$ учитывает возможные существенные изменения траектории ОТ, связанные с конкретным произнесением. Порог $thr2$ определяет вероятность перехода на гармоники ОТ, используя знание величины глобального минимума траектории ОТ. После оценки верхнего предела траектории ОТ производится повторный запуск алгоритмов SWIPE' или SWIPE. Пример работы корректирующего алгоритма показан на рис. 6.

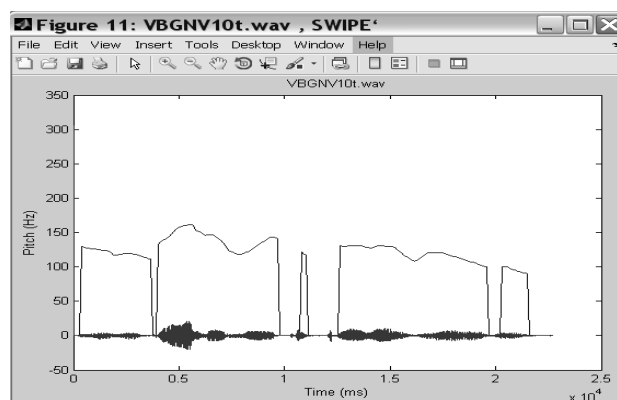


Рис. 6. Пример работы корректирующего алгоритма

Результаты испытаний алгоритмов SWIPE' и SWIPE, совместно с этими же алгоритмами, но с применением корректирующего алгоритма (SWIPE* и SWIPE**), представлены в табл. 2. Можно отметить существенное уменьшение всех типов ошибок в случае применения корректирующего алгоритма, за исключением малых ошибок, когда частота ОТ отличалась от эталонной в диапазоне от 5 до 15 %.

Таблица 2. Сравнительные испытания алгоритмов выделения основного тона

Выделитель ОТ	Ошибка Т/НТ, %	Ошибка НТ/Т, %	Большие ошибки, %	Малые ошибки, %
SWIPE	2,77	18,41	32,92	4,45
SWIPE'	6,16	10,36	47,54	3,79
SWIPE*	3,75	7,80	2,80	8,05
SWIPE**	4,32	8,99	13,70	7,48

D. Yu. Vashurin, Master's Degree student, Kalashnikov Izhevsk State Technical University
V. B. Gitlin, DSc in Engineering, Professor, Kalashnikov Izhevsk State Technical University

SWIPE application algorithm of pitch measuring for telephone bandwidth limited signal

Functioning of SWIPE and SWIPE' algorithms for telephone bandwidth limited signal is examined. These algorithms tend to switch to harmonics within processing of these signals. Correction algorithm to suppress these errors is proposed.

Keywords: pitch, harmonics, inner product, saw-tooth signal

Получено: 08.11.2013

УДК 681.518.3+623.546

А. Ю. Вдовин, кандидат технических наук, доцент;

В. А. Афанасьев, старший преподаватель

Ижевский государственный технический университет имени М. Т. Калашникова

АНАЛИЗ ВАРИАНТОВ УТОЧНЕНИЯ ТРАЕКТОРИИ ПУЛИ В ИНФОРМАЦИОННО-ИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ ДЛЯ ОПРЕДЕЛЕНИЯ ВНЕШНЕБАЛЛИСТИЧЕСКИХ ПАРАМЕТРОВ

Уточнены аналитическим способом величины провисания траектории пули без использования численного интегрирования уравнений внешней баллистики.

Ключевые слова: траектория пули, внешняя баллистика

При определении внешнебаллистических параметров с помощью информационно-измерительных сис-

тем в производственных условиях необходимо обеспечить работу таких систем в реальном масштабе

Библиографические ссылки

1. Camacho, A.; Harris, J. G. A sawtooth waveform inspired pitch estimator for speech and music // Journal of the Acoustical Society of America. – 2008. – Vol. 124. – Pp. 1638–1652.
2. Гитлин В. Б. Основной тон речевого сигнала. – Деп. в ВИНТИ 1998, № 1206-B98.
3. Ваишурин Д. Ю., Гитлин В. Б., Лузин Д. А. Сравнение алгоритмов выделения основного тона GS2 и SWIPE // Информационные технологии в науке, промышленности и образовании : сб. тр. регион. науч.-техн. очно-заочной конф. (27 апр. 2012 г., Ижевск) / М-во образования и науки Российской Федерации, Федеральное гос. бюджетное учреждение высш. проф. образования «Ижевский гос. техн. ун-т им. М. Т. Калашникова» ; [редкол.: В. А. Куликов (науч. ред.) и др.]. – Ижевск : Изд-во ИжГТУ, 2012. – С. 30–34.
4. Гитлин В. Б., Лузин Д. А. Совместный алгоритм выделения основного тона речи по методам GS и автокорреляционной функции спектра // Речевые технологии. – 2008. – № 3. – С. 39–49. – URL: <http://speechtechnology.ru/files/3-2008.pdf> (дата обращения: 11.11.2013).
5. Сапожков М. А. Речевой сигнал в кибернетике и связи. – М. : Связьиздат, 1963. – 452 с.
6. Кодзасов С. В., Кривнова О. Ф. Общая фонетика : учебник / Ин-т «Открытое о-во». – М. : Рос. гос. гуманитар. ун-т, 2001. – 592 с.
7. Система редактирования, анализа и шумоочистки речевых сигналов SIS 6.x : руководство пользователя / Центр речевых технологий. СПб.