

В. В. Митянок

Полесский государственный университет, Беларусь 225710, Пинск, ул. Днепровской флотилии, 23, e-mail: mitsianok@mail.ru

О синтезе некоторых искусственных звуков и слов

Получена 23.01.2017, опубликована 11.02.2017

Описываются результаты численных экспериментов по разложению некоторых звуков и слов речи человека на отдельные моды с медленно дрейфующими амплитудами и обратному их суммированию с целью выявления факторов, как имеющих значение, так и не имеющих значение для автоматического распознавания речи. Приводятся формулы для синтеза некоторых искусственных звуков и слов и числовые значения их параметров.

Ключевые слова: распознавание речи, синтез речи, фазовый анализ звуков.

ВВЕДЕНИЕ

Как хорошо известно, метод преобразований Фурье обладает рядом существенных недостатков [1-4]. В частности, спектры сигналов всегда получаются размытыми. При слишком малой длительности сигнала размытость линий становится настолько большой, что соседние линии спектра могут поглощать друг друга. Наоборот, при слишком большой длительности и при некоторой неустойчивости параметров сигнала на спектре появляется множество узких фальшивых линий, и, чем длиннее сигнал, тем больше этих линий. Фальшивые линии присутствуют на спектрограммах даже в случае идеальных гармонических сигналов, но заданных на *ограниченном* промежутке времени. Все это сказывается и на задачах автоматического распознавания речи человека и верификации и идентификации личности по голосу. Косвенным признаком того, что метод преобразований Фурье не годится для решения этих задач, является то, что, несмотря на многочисленные усилия, серьезные финансовые вложения, данные задачи до сих пор не имеют удовлетворительного решения.

В связи с этим в [3, 5-7] был предложен метод аппроксимации, который предназначен для решения тех же задач, но который не имеет присущих методу преобразований Фурье недостатков. На основе метода аппроксимации был получен ряд принципиальных результатов. Так оказалось, что в спектре отдельных, долго произносимых звуков, присутствуют полуцелые (по отношению к базовой) частоты, действующие «вспышками», имеет место «жесткая» модуляция амплитуд высших мод базовой частотой, причем со срывами. Тем самым было найдено объяснение неудачам метода преобразований Фурье.

В связи с имеющими место определенными успехами метода аппроксимации имеет смысл применить его для создания искусственных звуков и слов. Если искусственные

слова и звуки будут созданы, то тогда станет ясным, на что же именно следует обращать внимание при автоматическом распознавании речи, какие особенности звуковых сигналов позволяют отличать одного диктора от другого, а какие — наоборот, не имеют никакого значения, они случайны, привнесены несовершенством аппарата речеобразования человека, они лишь «путаются под ногами», отвлекая внимание исследователей и заставляя их распылять свои усилия.

Решение задачи синтеза звуков и слов следует начинать именно с выявления *математических* особенностей различных звуковых единиц речи человека. Расшифровка математических особенностей различных звуков человека — это также ключ к расшифровке математических особенностей звуков также и других существ, проживающих на Земле, прежде всего — дельфинов, слонов, китов, и, в более отдаленной перспективе, к пониманию их семантики.

1. МЕТОД АППРОКСИМАЦИИ

Метод основан на функционале [5-7]

$$S = \sum_{i=1}^n [y(t_i) - y_1(t_i)]^2 + \alpha \sum_{k=1}^{n-1} (b_{0,i} - b_{0,i+1})^2 + \alpha \sum_{k=1}^l \sum_{i=1}^{n-1} (a_{k,i} - a_{k,i+1})^2 + \alpha \sum_{k=1}^l \sum_{i=1}^{n-1} (b_{k,i} - b_{k,i+1})^2, \quad (1)$$

где $y(t_i)$ — зависящая от времени аппроксимируемая функция, описывающая сигнал, заданная своими значениями в n последовательных моментах времени от t_1 до t_n ,

$$y_1(t_i) = b_{0,i} + \sum_{k=1}^l [a_{k,i} \sin(\omega_k t_i) + b_{k,i} \cos(\omega_k t_i)], \quad i=1 \dots n \quad (2)$$

— аппроксимирующая функция, $b_{0,i}$ — дрейфующий нуль (начало отсчета), $a_{k,i}$, $b_{k,i}$ — дрейфующие амплитуды синус- и косинус- волн (параметры аппроксимирующей функции), ω_k — их несущие частоты, l — количество волн (мод) в аппроксимирующей функции. В (1) и (2) для простоты можно принять $t_i = i$, хотя это и не обязательно. Функционал (1) сконструирован как сумма слагаемых двух видов: слагаемые, не содержащие множитель α , отвечают за близость между аппроксимируемой и аппроксимирующей функциями, слагаемые, содержащие α , отвечают за сглаживание прыжков дрейфующих амплитуд волн (мод) при переходе по оси времени между соседними моментами дискретизации. Чем бóльшим выбрано значение α , тем более гладкими будут получаться амплитуды волн. Вычисляя частные производные (1) по дрейфующим амплитудам и по дрейфующему началу отсчета и приравнивая результаты нулю, получим систему линейных алгебраических уравнений относительно параметров аппроксимирующей функции. Решив эту систему, найдем эти параметры и тем самым произведем разложение аппроксимируемой функции на сумму волн с медленно меняющимися амплитудами. Найденные таким путем $b_{0,i}$, $a_{k,i}$, $b_{k,i}$ можно подставить в (2). Полученную в результате аппроксимирующую функцию можно назвать восстановленным звуком. Если затем вычесть восстановленный звук из

исходного звука и подвергнуть разность преобразованиям Фурье, то часто выясняется, что существуют еще какие-то несущие частоты, которые не были замечены при первом разложении в ряд (интеграл) Фурье по причине малой интенсивности несомых ими мод. В частности, этим способом в [7] было установлено, что в спектре многих звуков присутствуют полуцелые (по отношению к базовой) несущие частоты.

Каждую из мод, входящих в (2), можно переписать в физически более информативном виде:

$$a_{k,i} \sin(\omega_k t_i) + b_{k,i} \cos(\omega_k t_i) = c_{k,i} \sin(\omega_k t_i + \varphi_{k,i}), \quad k=1\dots l, i=1\dots n \quad (3)$$

и тогда аппроксимирующая функция выглядит так:

$$y_1(t_i) = b_{0,i} + \sum_{k=1}^l c_{k,i} \sin(\omega_k t_i + \varphi_{k,i}), \quad i=1\dots n. \quad (4)$$

Здесь $c_{k,i}$ — дрейфующая *общая* амплитуда волны (моды), $\varphi_{k,i}$ — дрейфующая фаза. Всюду ниже под термином *амплитуда* будет пониматься именно *общая амплитуда*, численные значения фаз даются в радианах, а значения дрейфующих амплитуд — в компьютерных битах. Запись звуков производилась при частоте дискретизации 44100 Гц, поэтому в качестве единицы времени принята 1/44100 доля секунды.

2. СИНТЕЗ ИСКУССТВЕННЫХ МОНОЗВУКОВ

Исследовались те гласные звуки, которые можно было произносить долго — это звуки (монозвуки) «А», «О», «У», «Э», «Ы», «И», полученные от нескольких респондентов, женщин и мужчин. Звуки раскладывались на моды пропорциональной ловящей сетью [5] и затем восстанавливались. Во всех случаях восстановленный звук звучал также как и исходный.

Для того чтобы ответить на вопрос, что именно делает звук «А» звуком «А», звук «О» звуком «О» и т.д., перед суммированием (4) были проведены математические эксперименты по сознательному искажению амплитуд и (или) фаз.

Во-первых, фазы всех целых мод, кроме базовой, заменялись на искусственно вычисленные, связанные с фазой базовой моды формулой

$$\varphi_{k,i} = k\varphi_{1,i} + r_k, \quad k=1\dots l, i=1\dots n, \quad (5)$$

где k — номер моды, $\varphi_{1,i}$ — дрейфующая фаза базовой моды, r_k — массив произвольных чисел. Фаза базовой моды не менялась. Не менялись и все дрейфующие амплитуды. Как оказалось, звучание звуков от такой замены не изменяется.

Во-вторых, обратим внимание на поведение дрейфующих амплитуд одного из образцов звука «А» (рис. 1).

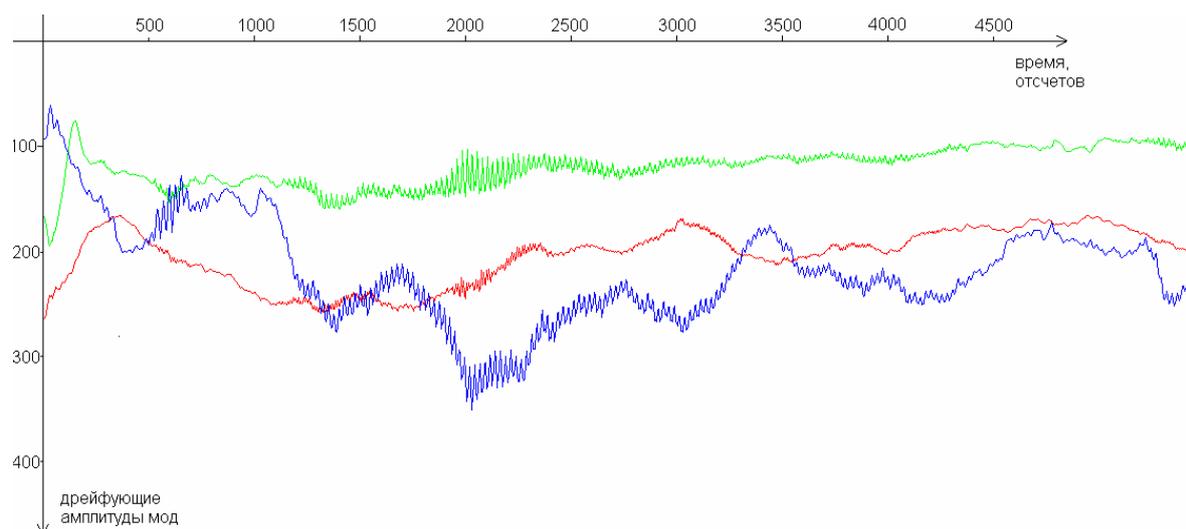


Рисунок 1. Поведение 3-х нижних амплитуд целых мод звука «А». Красная линия — амплитуда моды № 1, зеленая линия — моды № 2, синяя линия — моды № 3

Как видно из рис. 1, дрейфующие амплитуды как бы хаотически вибрируют вокруг их средних значений. Сам факт хаотичного поведения общих амплитуд наводит на мысль о том, что хаотичность — это нечто привнесенное, не имеющее отношения к индивидуальности звуков. Так и оказалось. Дрейфующие амплитуды можно усреднить по отрезку звучания достаточной длины, и затем заменить фактические дрейфующие амплитуды их усредненными значениями. Звук, полученный после такого искажения амплитуд, звучит так же, как и исходный.

В третьих, выяснилось, что при суммировании мод (4) можно опустить дрейфующий нуль и полуцелые моды. И от такого отбрасывания звук не меняется. А вот если фазу каждой из мод, в том числе и базовой, на всем отрезке звучания заменить на постоянное, но случайное число, то качество звука ухудшается. Вместо четкого звука слышится то, что скорее можно назвать звучанием зуммера.

В поисках объяснений этому явлению были проделаны следующие математические эксперименты. Усредненный амплитудный спектр каждого из изучаемых звуков соединялся в формуле (4) с дрейфующими фазами от любого другого из этих же звуков и от любого из других респондентов. После такой операции звук не менялся, звучал четко и соответствовал именно амплитудам. Так чем же объясняется ухудшение качества звука при замене дрейфующих фаз на константы? Оказалось, что во всех случаях нормированные фазы (то есть фазы, деленные на номер моды [8]) не являются строгими константами, а дрейфуют вокруг неких средних значений с неустойчивым периодом от 1.5 до 2.5 Гц и с неустойчивой амплитудой 0.2-2 радиан. Они как бы «испорчены». В связи с этим возникло предположение, что именно так и должно быть. Что мозг слушателя уже готов к тому, что диктор будет производить сигнал с испорченной фазой, а звук с неиспорченной фазой мозгом слушателя за звук не воспринимается. Это предположение оправдалось. Когда в качестве фазы принималась хаотически меняющаяся (в определенных рамках) величина, то звук вновь звучал четко и распознаваемо.

Подытоживая все вышесказанное, получаем, что для синтеза вышеуказанных звуков, вместо (4), как один из вариантов, можно принять формулу

$$y_1(i) = \sum_{k=1}^l c_k \sin(\omega_k i + kp \sin(i/3300) + r_k), \quad i=1\dots n, \quad (6)$$

где усредненные значения амплитуд c_k приведены в нижеследующей таблице 1, ω_k — несущие частоты, кратные базовой, значение которой приведено в последней строчке таблицы 1, r_j — массив произвольных чисел, n — длина отрезка звучания (в отсчетах дискретизации). Множитель p в (6) может принимать любое значение в интервале $[1\dots 10]$, но лучшее звучание наблюдается при $p=2$ для звуков «Э», «Ы», и $p=4$ для звуков «А», «О», «У», «И». За основу получения усредненных общих амплитуд в таблице 1 был взят голос автора. В (6) также принято $t_i = i$. Внутренний синус в (6) обеспечивает порчу фазы. (Возможны и другие варианты порчи фаз.)

3. СИНТЕЗ ИСКУССТВЕННЫХ СЛОВ

При изучении слов возникает проблема краевых эффектов. В случае долгих гласных звуков эта проблема решалась просто. Начало и конец записи попросту обрезались по 10-30 процентов общей длины записи, после обрезания оставался достаточный для изучения отрезок звуковой кривой. При изучении же слов так поступать нельзя, так как при обрезании можно было случайно отрезать и звуки, входящие в слово и существенные для его распознавания. Поэтому было принято следующее решение: вместо некоторого слова, на одном дыхании, произносилась последовательность 4-х этих же слов, образующих в сумме мультислово. Так, для изучения слова МАЛИНА записывалось мультислово МАЛИНАМАЛИНАМАЛИНАМАЛИНА (удобнее обозначать как 4МАЛИНА) по возможности так, чтобы оно для диктора представлялось одним словом. Затем определялась базовая частота мультислова и производилось разложение звуковой кривой по пропорциональной ловящей сети [5]. После этого визуализировались дрейфующие амплитуды и производился поиск повторяющихся характерных участков (зарубок), хотя бы для одной из мод. В некоторых случаях оказывалось, что для поиска зарубок имеет смысл предварительно выполнить некоторые математические преобразования. Так, например, для мультислова 4МАЛИНА оказалось, что если провести численное дифференцирование амплитуд по времени, то зарубки четко просматриваются на модах № 3, 5, 6 (рис. 2).

Таблица 1. Значения амплитуд различных мод простых гласных звуков

Номер моды	А	О	У	Э	Ы	И
1	637	613	1060	566	1757	914
2	375	714	814	540	354	112
3	674	836	303	1007	65	22
4	794	495	0	61	0	0
5	753	51	0	114	25	0
6	180	0	0	123	51	0
7	49	0	0	90	140	0
8	19	0	0	97	32	0
9	15	0	0	183	54	16
10	17	0	28	93	111	49
11	17	10	0	114	14	30
12	21	17	0	120	10	35
13	8	0	0	44	22	71
14	16	0	34	31	92	135
15	16	0	7	42	26	147
16	16	15	17	54	26	110
17	30	30	8	79	8	35
18	34	12	0	45	0	6
19	13	0	0	37	0	8
20	0	0	0	18	7	5
21	0	0	0	45	0	14
22	0	0	0	25	9	21
23	0	12	0	0	6	14
24	0	0	0	0	12	5
25	0	0	10	0	10	9
26	0	0	10	0	8	22
27	0	0	5	0	7	16
28	0	0	3	0	0	24
29	0	11	4	0	0	42
30	0	15	8	0	0	18
31	0	18	13	0	0	14
32	0	18	15	0	0	13
Базовая частота	0.0269	0.0262	0.0305	0.0268	0.0302	0.0291

Примечание: допустимы небольшие (в пределах 10-30 процентов) изменения амплитуд, заметно не влияющие на звук. Возможно также одновременное пропорциональное изменение всех амплитуд некоторого звука — этому соответствует изменение громкости. Данные получены усреднением по 20 образцам длительностью по 2-3 сек каждый.

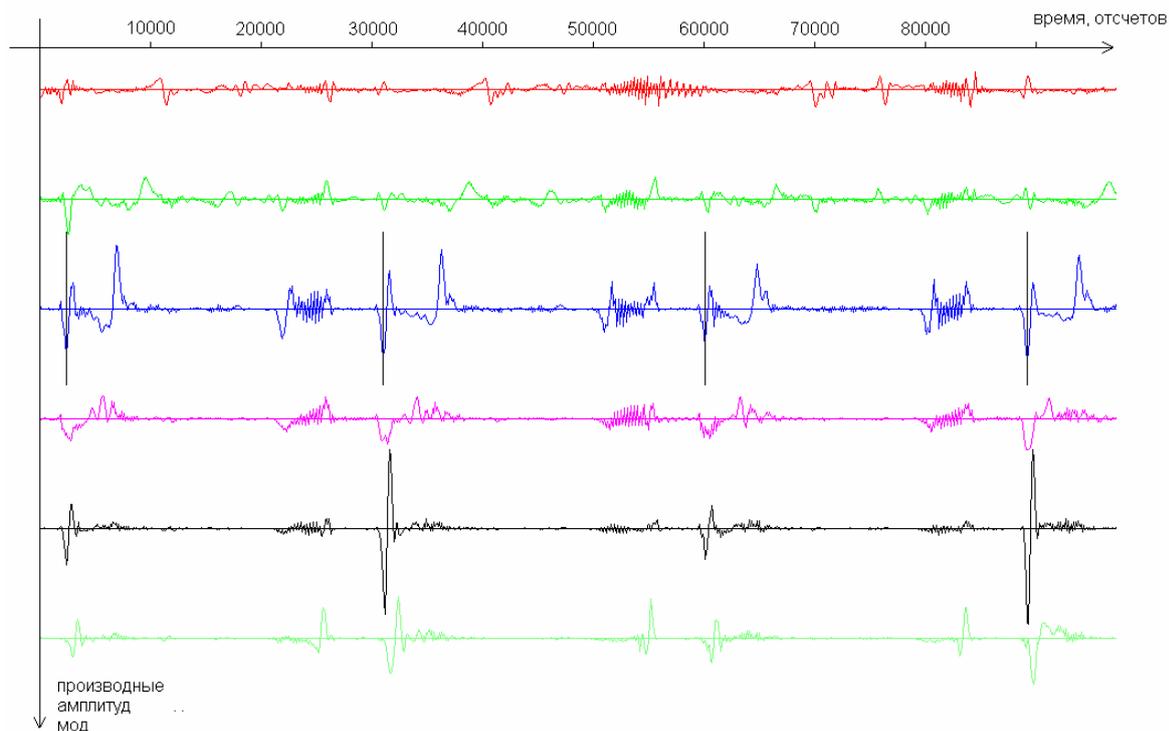


Рисунок 2. Графики производных амплитуд первых 6 целых мод, сверху вниз по порядку номеров, одного из образцов мультислова 4МАЛИНА. Зарубки отмечены на графике моды № 3 (синяя линия) отрезками вертикальных черных прямых, пересекающих синюю

Очевидно, что на интервале между соседними зарубками, каждый из звуков слова (фонем) встречается ровно по одному разу, *хотя и не обязательно слово начинается именно с зарубки*. Таким образом, все то, что находится между соседними зарубками, следует рассматривать не как слово, а как *представителя слова*. Слово отличается от *представителя слова* тем, что какая-то часть слова отрезается от его начала и переносится в конец, в результате чего создается *представитель слова*. Для анализа удобнее использовать именно *представителей слов*, что и сделано в настоящей статье.

Так как во всех образцах мультислово состояло из 4-х состыкованных слов, то в каждом образце разыскивались по 4 зарубки и, соответственно, выделялись по 3 *представителя слова*. После разложения *представителя слова* на моды производилось восстановление *представителя слова*, несколько одинаковых копий которого последовательно состыковывались между собой. Как и ожидалось, восстановленное мультислово звучало также как и исходное мультислово.

Для создания же тех слов, которые можно назвать искусственными, следует найти некую математическую формулу, наподобие формулы (6), применение которой и позволит генерировать слово. Ниже будут представлены математические формулы, полученные на основе одного из образцов *представителя слова* МАЛИНА.

Для нахождения этих формул производилось разложение звуковой кривой мультислова 4МАЛИНА по пропорциональной ловящей сети с количеством мод,

равным 24 и при базовой частоте $\omega_1=0.025$. Затем визуально находились зарубки и выделялись представители слов, которые и подвергались дальнейшему изучению.

Так как в состав слова входят разные звуки (фонемы), и так как соответствующие им амплитуды мод, в соответствии с табл. 1 различны, то аппроксимировать амплитуды мод представителя слова постоянными числами не получится. Поэтому был избран другой путь. Как оказалось, для всех мод дрейфующие амплитуды выглядят как суммы колоколообразных функций, частично наплывающих друг на друга (см. рис. 3 ниже). Следовательно, амплитуды мод могут быть аппроксимированы суммой нескольких функций Гаусса с различными параметрами.

$$c_{k,i} = \sum_{l=1}^m A_{kl} \exp\left[-\frac{(i-\mu_{k,l})^2}{\sigma_{k,l}^2}\right], \quad i=1\dots n, \quad k=1\dots 12. \quad (7)$$

Здесь i — время (номер отсчета записи), $c_{k,i}$ — дрейфующая амплитуда моды номер k , m — количество функций Гаусса, аппроксимирующих амплитуду моды номер k (в настоящем исследовании всюду $m=6$), $A_{kl}, \mu_{k,l}, \sigma_{kl}$ — параметры функций Гаусса, n — длина слова (в отсчетах дискретизации, в настоящем исследовании $n=29299$).

Эти параметры могут быть подобраны одним из известных методов, например, методом покоординатного приближения, методом скорейшего спуска или любым иным из методов нелинейной аппроксимации. Приведем результат такой аппроксимации на примере амплитуд первых 8 мод представителей слова МАЛИНА (рис. 3).

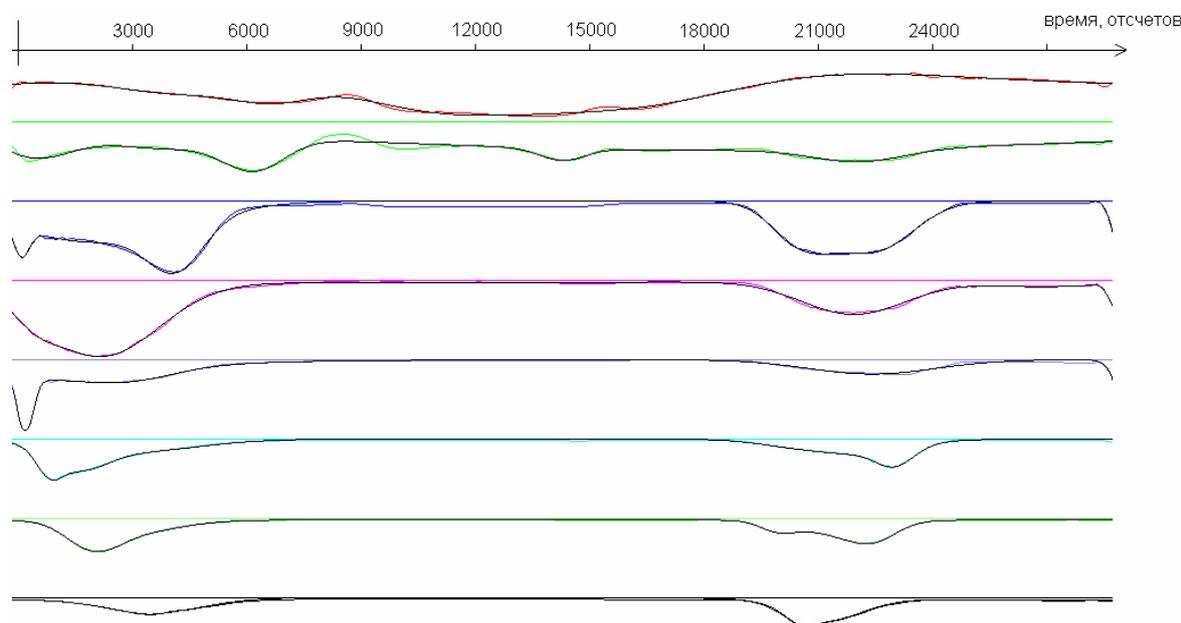


Рисунок. 3 Амплитуды первых 8 мод представителя слова МАЛИНА, сверху вниз по порядку номеров. Цветная кривая линия — амплитуда дрейфующей моды, ровная горизонтальная линия того же цвета — ее нуль отсчета. Линия черного цвета, налагаемая на цветную и частично закрывающая ее — результат аппроксимации амплитуды суммой функций Гаусса. Длина слова равна 29299 отсчетам

Как видно из рис. 3, имеется некоторое различие между аппроксимируемыми и аппроксимирующими функциями. Однако при синтезе мультислова это различие на слух незаметно.

Таким образом, дрейфующие амплитуды представителей слов могут быть представлены в виде (7). Необходимые значения параметров функций Гаусса для амплитуд мод слова МАЛИНА приведены в таблицах 2-4.

Таблица 2. Параметры функций Гаусса (7), аппроксимирующих амплитуды мод 1-8.

		k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8
l=1	μ_{kl}	16730	17893	4323	2354	2774	2145	22530	21836
	σ_{kl}	9721	13579	1530	1707	2384	1639	350	1305
	A_{kl}	424	349	1174	943	337	168	205	293
l=2	μ_{kl}	16849	6563	22228	22501	380	23359	2596	4208
	σ_{kl}	815	1432	1743	2002	250	591	1435	1965
	A_{kl}	157	627	911	319	425	208	420	134
l=3	μ_{kl}	13573	21594	1483	827	22480	1107	22082	20307
	σ_{kl}	3011	1231	1279	684	1870	410	1854	213
	A_{kl}	384	405	329	355	214	490	218	168
l=4	μ_{kl}	6006	14569	4854	23225	24021	23927	23285	5312
	σ_{kl}	6638	847	493	212	433	284	440	340
	A_{kl}	681	255	713	155	81	536	504	84
l=5	μ_{kl}	24635	647	279	4650	1030	21712	5313	3885
	σ_{kl}	198	471	286	1464	116	1733	953	326
	A_{kl}	135	270	469	211	56	162	120	69
l=6	μ_{kl}	-378	23894	23649	23884	622	3490	19962	21669
	σ_{kl}	1913	446	688	224	83	2242	233	661

Таблица 3. Параметры функций Гаусса (7), аппроксимирующих амплитуды мод 9-16.

		k=9	k=10	k=11	k=12	k=13	k=14	k=15	k=16
l=1	μ_{kl}	4954	15200	5119	13537	21521	14556	10661	11294
	σ_{kl}	1093	322	922	1386	4111	586	572	856
	A_{kl}	215	98	121	104	59	115	164	82
l=2	μ_{kl}	21888	6992	6721	5376	5298	11574	13438	13950
	σ_{kl}	1400	5854	9238	759	1723	1333	1106	1157
	A_{kl}	163	43	17	127	117	45	56	21
l=3	μ_{kl}	7889	22192	9970	22117	13747	23722	5219	21801
	σ_{kl}	1079	1783	294	1547	2486	10699	2266	1931
	A_{kl}	94	52	99	61	61	23	25	14
l=4	μ_{kl}	5338	5452	22145	13225	9907	9927	22018	6065
	σ_{kl}	468	654	2018	287	452	321	4521	9833
	A_{kl}	451	48	48	39	59	50	18	7
l=5	μ_{kl}	15522	10650	11319	14229	21835	13606	11831	4696
	σ_{kl}	383	781	961	658	1086	208	391	1292
	A_{kl}	68	30	91	156	115	37	61	10
l=6	μ_{kl}	6301	21104	14450	5738	23694	5702	13329	21273
	σ_{kl}	299	693	1095	11539	146	690	112	351

Таблица 4. Параметры функций Гаусса (7), аппроксимирующих амплитуды мод 17-24.

		k=17	k=18	k=19	k=20	k=21	k=22	k=23	k=24
l=1	μ_{kl}	11001	10789	11399	22447	2957	7025	6663	5063
	σ_{kl}	895	551	1552	1584	2224	14510	13426	721
	A_{kl}	57	41	99	58	22	4	3	23
l=2	μ_{kl}	23261	21952	22397	10229	11011	4883	9687	5555
	σ_{kl}	853	1969	1744	5447	6237	275	213	211
	A_{kl}	18	31	40	24	9	7	8	19
l=3	μ_{kl}	13741	4245	3968	5305	22491	9632	5377	3792
	σ_{kl}	1698	3442	3829	569	2104	226	279	412
	A_{kl}	20	20	35	57	10	12	3	13
l=4	μ_{kl}	4561	11603	21250	12022	3705	3806	12670	8576
	σ_{kl}	3661	412	530	726	209	1759	251	13198
	A_{kl}	16	41	73	35	16	7	4	4
l=5	μ_{kl}	20353	12337	23905	1647	1398	8543	14156	1789
	σ_{kl}	256	3179	414	1451	283	296	358	243
	A_{kl}	19	25	37	48	13	6	3	6
l=6	μ_{kl}	21393	20473	22888	3798	10113	7368	11834	22493
	σ_{kl}	726	200	204	461	389	381	227	901
	A_{kl}	30	17	42	46	11	7	3	4

Число k в табл. 2-4 означает номер моды, число l — номер функции Гаусса, входящей в аппроксимирующую сумму (7). Что касается фаз мод, то их типичное поведение представлено ниже (рис. 4).

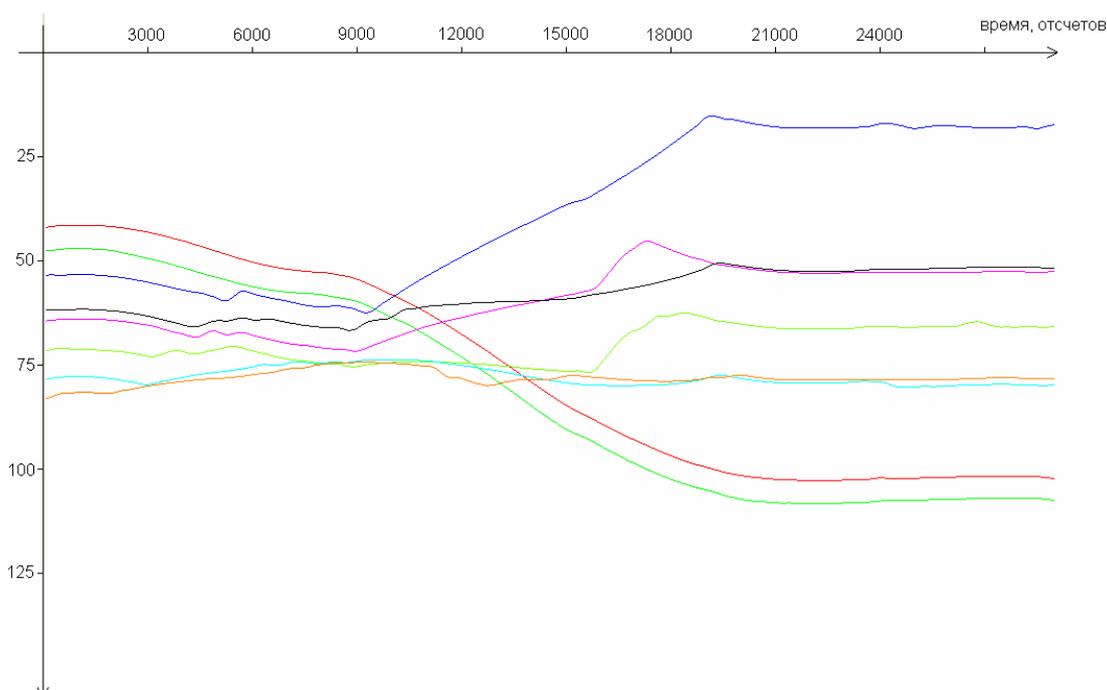


Рисунок 4. Нормированные фазы первых 8 мод представителя слова МАЛИНА. По порядку номеров — красный цвет, зеленый, синий, фиолетовый, черный, бирюзовый, голубой, кирпичный

Как видно из рис. 4, в некоторых случаях фазы испытывают резкий излом. Такие функции неудобны для аппроксимации степенными функциями, поэтому было принято решение разбить всю длину представителя слова на участки, а в качестве границ между этими участками выбрать те моменты времени, где хотя бы одна из фаз испытывает излом. Между границами участков фазы можно линейно интерполировать. Границы участков линейной интерполяции и значения фаз в граничных точках для мод 1-24 приведены в таблицах 5-7.

Таблица 5. Границы участков и значения фаз в граничных точках для мод 1-8

i	k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=8
1	0.95	3.09	-50.42	-79.36	-159.87	20.04	-102.84	-69.90
690	-0.82	-0.32	-53.68	-86.03	-166.07	13.17	-112.42	-69.32
1290	-1.95	-2.64	-57.09	-90.63	-172.00	8.63	-119.83	-82.09
2700	-2.86	-3.44	-59.62	-94.68	-177.02	3.75	-125.56	-89.04
3390	-2.46	-2.29	-58.44	-93.57	-175.31	6.15	-122.18	-85.93
4590	-0.66	1.55	-53.65	-86.50	-166.43	16.94	-108.39	-99.99
5100	0.43	3.82	-51.29	-82.91	-160.86	22.05	-113.91	-105.02
6330	3.37	9.17	-42.68	-88.78	-161.90	13.28	-126.01	-113.96
7200	5.27	12.25	-36.78	-81.00	-157.86	16.27	-124.09	-122.30
8640	8.67	18.36	-25.95	-81.88	-143.67	11.49	-138.23	-132.82
9420	10.15	21.75	-21.34	-78.87	-140.03	22.89	-147.16	-139.02
9840	11.08	23.42	-24.56	-76.15	-137.24	27.44	-148.26	-134.91
10200	12.06	25.03	-22.28	-79.62	-134.69	29.55	-151.03	-136.51
12000	17.33	35.56	-6.47	-108.76	-153.05	11.13	-164.12	-144.46
13050	19.94	40.81	1.70	-120.18	-149.94	11.30	-174.91	-153.02
14100	22.41	45.86	8.73	-119.74	-138.03	5.91	-186.72	-162.71
15600	24.75	50.67	16.90	-109.56	-125.63	21.17	-191.30	-177.79
15990	24.63	49.94	13.26	-108.77	-123.90	22.69	-190.08	-177.05
18600	21.37	43.47	3.62	-120.96	-137.45	34.35	-188.69	-184.76
20550	14.43	30.03	-14.32	-107.43	-154.79	31.90	-189.90	-199.34
21030	12.47	26.16	-12.31	-112.03	-154.90	29.84	-184.95	-196.84
22530	5.56	13.18	-1.39	-108.91	-157.91	22.02	-172.21	-191.61
24180	-2.59	-2.91	7.84	-106.53	-164.44	23.60	-178.08	-203.46
24720	1.36	-9.13	5.98	-106.10	-168.19	21.67	-179.16	-207.24
26070	-4.77	-21.34	13.36	-100.27	-174.73	30.70	-172.06	-211.59
29299	-14.84	-41.38	-15.57	-79.23	-213.37	33.12	-162.31	-227.07

Таблица 6. Границы участков и значения фаз в граничных точках для мод 9-16

i	k=9	k=10	k=11	k=12	k=13	k=14	k=15	k=16
1	-92.32	16.92	39.61	97.73	69.52	-25.98	20.94	66.23
690	-94.15	14.08	38.91	96.60	71.09	-23.75	21.16	55.62
1290	-90.37	16.36	41.01	96.95	70.02	-26.78	32.21	60.45
2700	-95.54	8.80	33.78	99.68	63.88	-25.78	74.95	59.93
3390	-93.66	12.37	30.53	103.55	65.19	-18.26	79.34	61.52
4590	-107.07	12.51	19.12	102.16	56.60	1.04	77.11	58.67
5100	-110.23	9.75	17.81	101.71	57.36	3.60	79.86	61.54
6330	-115.97	6.94	17.68	105.23	62.37	17.89	90.07	63.83
7200	-122.30	2.14	15.88	106.60	62.89	19.71	96.10	72.20
8640	-128.94	-0.82	15.96	107.97	68.99	28.68	100.70	70.22
9420	-130.67	-4.61	13.61	107.98	73.08	32.71	108.13	74.33
9840	-132.28	-6.98	10.53	104.97	73.94	29.25	109.08	78.31
10200	-134.09	-7.00	10.31	106.36	75.45	30.39	106.68	75.77
12000	-136.47	-4.13	21.38	119.18	95.21	51.75	100.23	61.34
13050	-142.32	-7.20	20.59	122.07	100.95	58.69	111.28	61.51
14100	-149.07	-12.20	18.20	122.29	102.97	63.19	117.27	65.99
15600	-145.22	-0.29	7.60	115.94	96.14	54.98	121.72	71.91
15990	-144.82	2.14	10.01	118.42	99.13	53.86	126.75	88.93
18600	-149.07	13.85	12.70	123.59	107.83	47.48	125.86	122.16
20550	-155.03	18.73	15.98	126.84	108.99	56.81	127.40	237.54
21030	-152.85	18.49	13.69	125.05	113.08	50.36	127.90	236.43
22530	-153.01	13.90	2.66	135.48	116.89	47.44	120.85	220.89
24180	-156.85	19.39	-0.02	131.12	117.59	42.29	115.12	211.79
24720	-154.47	21.82	-0.76	133.71	116.95	43.36	114.07	205.35
26070	-154.04	26.22	4.62	134.15	111.33	39.10	114.49	218.35
29299	-164.19	18.14	17.06	149.54	109.52	43.21	113.47	209.18

В табл. 5-7 координаты граничных точек участков разбиения представителя слова МАЛИНА даны в первом столбце. Число k в этих таблицах означает номер моды.

Числа в столбцах 2-9 таблиц 5-7 приведены с точностью 0.01. Эти числа можно домножить на 100, присоединить к ним числа из таблиц 2-4 и первый столбец таблиц 5-7, в результате получится набор из 1082 чисел целого типа, которые в памяти компьютера занимают 4328 байт. Слово МАЛИНА, записанное в формате .wav — около 59600 байт. Таким образом, содержание таблиц 2-7 может рассматриваться как результат сжатия слова МАЛИНА, коэффициент сжатия — около 14. (Причем он может быть увеличен за счет более экономных приемов записи чисел, отбрасывания малоинтенсивных мод и т. д.).

Таблица 7. Границы участков и значения фаз в граничных точках для мод 17-24

i	k=17	k=18	k=19	k=20	k=21	k=22	k=23	k=24
1	311.71	169.99	63.09	227.08	526.45	768.11	2296.48	2576.68
690	314.96	170.97	70.72	221.12	522.02	833.38	2408.22	2630.22
1290	313.03	174.25	73.20	222.69	515.66	848.96	2460.27	2647.50
2700	310.35	177.35	74.92	223.61	510.13	841.25	2618.07	2683.95
3390	316.31	178.84	77.35	226.18	519.47	837.06	2666.77	2717.94
4590	317.54	179.83	79.75	229.42	519.95	841.05	2879.09	2734.03
5100	321.17	185.15	85.74	225.66	514.90	837.70	2973.62	2739.71
6330	322.84	182.18	76.59	219.42	530.76	862.17	3068.99	2744.20
7200	337.69	192.95	72.04	209.48	550.89	881.98	3096.14	2741.58
8640	344.20	184.58	61.81	203.30	564.57	893.97	3115.68	2786.30
9420	351.35	179.09	58.09	196.83	564.57	891.53	3122.48	2835.03
9840	350.60	180.05	60.11	196.94	563.20	889.42	3119.50	2833.54
10200	347.37	179.22	59.97	197.40	563.21	909.04	3142.71	2834.43
12000	337.43	174.29	60.18	203.58	577.18	1048.50	3272.57	3064.90
13050	349.22	162.46	52.16	198.51	589.21	1119.77	3284.78	3211.02
14100	360.48	164.60	43.24	190.17	590.02	1152.99	3391.66	3292.31
15600	362.49	169.46	48.20	198.29	604.64	1217.41	3533.45	3523.92
15990	374.23	167.94	78.91	213.47	607.16	1309.90	3545.69	3526.30
18600	446.67	166.39	107.26	276.56	845.99	1377.37	3798.80	3770.87
20550	516.49	181.20	157.22	349.10	985.36	1720.73	4068.01	3976.55
21030	517.00	178.78	160.18	348.40	1001.69	1767.36	4115.29	4034.23
22530	524.84	181.85	156.67	338.35	1028.70	1864.55	4308.88	4169.39
24180	528.21	185.49	156.04	330.36	1016.32	1946.68	4486.44	4257.09
24720	528.74	203.07	167.02	333.34	1012.97	1976.87	4534.50	4376.86
26070	563.06	213.81	171.87	356.10	1021.57	2163.41	4688.76	4581.78
29299	752.59	310.47	253.59	460.31	1177.22	2282.86	4856.29	4771.62

ЗАКЛЮЧЕНИЕ

Амплитуды каждой из мод мультислова 4МАЛИНА можно аппроксимировать суммой функций Гаусса, фазы можно аппроксимировать кусочно-линейно. После этого можно, по формуле (4), воссоздать представителя слова, последовательно состыковать любое число одинаковых представителей слова, и получить тем самым восстановленное мультислово, которое можно преобразовать в любой из звуковых форматов и прослушать. Полученное мультислово звучит также как и исходное, его можно назвать искусственным. Незначительные ошибки аппроксимации на звучании практически не сказываются. Заинтересованный (и даже скептически настроенный) читатель может все это проделать, используя табл. 2-7, формулы (7), (4) и тривиальные формулы линейной интерполяции. При необходимости автор может по запросу читателя выслать файлы, содержащие таблицы настоящей статьи.

ЛИТЕРАТУРА

1. Васильева Л. Г., Жилейкин Я. М., Осипик Ю. И. Преобразования Фурье и вейвлет-преобразования. Их свойства и применение. // Вычислительные методы и программирование: в 3 т. – М., - Т 3, - Вып 1, - С 172-175, 2002.
2. Максимчук И. В., Гергель Л. Г., Осадчий О. В. Сравнительный анализ Фурье и вейвлет-преобразования для анализа сигнала фотоплетизмограммы. [Электронный ресурс] // Современные научные исследования и инновации – М., 2013. - № 6 – Режим доступа: <http://web.snauka.ru/issues>. - Дата доступа: 25 06 2013.
3. Митянок В. В. Определение числовых характеристик высокочастотных звуков речи на основе аппроксимации гармоническими функциями // Известия НАН Беларуси, сер ф.-м.н., - Минск, - № 2, С 111-118. 2009.
4. Митянок В. В. О физической структуре простых гласных звуков речи человека // Открытые семантические технологии проектирования интеллектуальных систем: материалы VI международной научно-технической конференции OSTIS-2016, Минск, 18-20 февраля 2016, -Минск: БГУИР, 2016, С. 404-410.
5. Митянок В. В., Коновалова Н. В. Применение фазового анализа звуков речи для распознавания человека по его голосу. [Электронный ресурс] // Техническая акустика. – Электрон. журн.- СПб., - 2013. № 4.- Режим доступа: <http://www.ejta.org>, свободный.
6. Митянок В. В. О числовых характеристиках некоторых низкочастотных звуков человеческой речи. [Электронный ресурс] // Техническая акустика. – Электрон. журн.- СПб., 2008.- № 15.- Режим доступа: <http://www.ejta.org>, свободный.
7. Митянок В. В. О физической структуре звуков З, Зь, Ж, Жь. [Электронный ресурс] // Техническая акустика. – Электрон. журн.- СПб., - 2014.- № 9.- Режим доступа: <http://www.ejta.org>, свободный.
8. Митянок В. В. К проблеме идентификации и верификации личности по фазовым характеристикам звуков речи [Электронный ресурс] // Техническая акустика. – Электрон. журн.- СПб., - 2015.- № 7.- Режим доступа: <http://www.ejta.org>, свободный.