

МЕТОД КОЛИЧЕСТВЕННОЙ ОЦЕНКИ СЛУХОВОГО ВОСПРИЯТИЯ НЕЛИНЕЙНЫХ ИСКАЖЕНИЙ. ЧАСТЬ 2. ВЫБОР МЕРЫ ТОНАЛЬНОСТИ СИГНАЛА И УЧЕТ ЕЕ ВЛИЯНИЯ НА СУБЪЕКТИВНОЕ ВОСПРИЯТИЕ ИСКАЖЕНИЙ

М.О. Пахомов^{1*}, В.С. Рожнов^{1, 2}

¹Санкт-Петербургский филиал корпорации LG Electronics inc.,
Санкт-Петербург, 191123, Российская Федерация

²Санкт-Петербургский государственный университет телекоммуникаций им. проф. М.А. Бонч-Бруевича,
Санкт-Петербург, 193232, Российская Федерация

*Адрес для переписки: mpakhomov39@gmail.com

Информация о статье

УДК 621.391

Язык статьи – русский

Ссылка для цитирования: Пахомов М.О., Рожнов В.С. Метод количественной оценки слухового восприятия нелинейных искажений. Часть 2. Выбор меры тональности сигнала и учет ее влияния на субъективное восприятие искажений // Труды учебных заведений связи. 2018. Т. 4. № 3. С. 98-105.

Аннотация: Во второй части работы более полно раскрывается тема учета спектральных характеристик тестового сигнала при оценке субъективного восприятия нелинейных искажений. Выбрана наиболее значимая характеристика – тональность и предложена количественная мера этой характеристики – коэффициент тональности. Проведены субъективные тесты прослушивания для определения зависимости слухового восприятия нелинейных искажений от тональной структуры сигнала и спектрального распределения отношения «Шум-Маска».

Ключевые слова: нелинейные искажения, одновременная маскировка, отношение «Шум-Маска», коэффициент тональности, линейная модель множественной регрессии.

Введение

Приступая к описанию задачи, решаемой в представленной работе, прежде всего, следует отметить тот факт, что влияние тональной структуры референсных музыкальных сигналов на слышимость нелинейных искажений был отмечен в середине XX века и подробно описан в работах основоположников психоакустики – Э. Цвикера, Б. Мура и пр. Учет меры предсказуемости спектра сигнала при вычислении допустимого отношения «сигнал-шум» широко используется в современных алгоритмах кодирования музыкальных сигналов, например AAC (от англ. Advanced Audio Coding). Эти исследования помогают более точно вычислять пороги маскировки сигнала нелинейных искажений полезным сигналом, но не обеспечивают необходимой точности для предсказания субъективной оценки в случае, когда искажения заметно превышают порог маскировки. Другими словами, можно ли провести линейную (или нелинейную) зависимость между количеством слышимых шумов и субъективной оценкой. Тесты про-

слушивания показывают, что для набора музыкальных сигналов с разными спектральными характеристиками зависимости между отношением «Шум-Маска» (NMR, от англ. Noise-to-Mask Ratio) и оценкой предпочтения достаточно слабые [6, 7].

В первой части работы был описан способ точной оценки отношения «Шум-Маска» для многотональных тестовых сигналов. Была выдвинута гипотеза, что влияние слышимых шумов на субъективное предпочтение зависит от частотного диапазона, в котором эти шумы превышают порог маскировки. Данная гипотеза была проверена на одном музыкальном сигнале. В выбранный музыкальный сигнал искусственно вносились нелинейные искажения разных уровней в различные частотные диапазоны и слушателям предлагалось оценить качество звучания с точки зрения предпочтения. Проверка результатов показала различную скорость деградации качества звучания в зависимости от величины «Шум-Маска» для разных частотных диапазонов. Было отмечено, что спектр сигнала различался по своей тональной структуре

в зависимости от частоты. Для объяснения полученных результатов было принято решение провести серию дополнительных тестов прослушивания с набором сигналов, имеющих различную тональную структуру в соответствующих диапазонах частот.

В представленной работе описывается и обосновывается выбор количественной меры тональности сигнала, референсных музыкальных сигналов и способы внесения нелинейных искажений для получения тестовых сигналов. Представлены выходные переменные модели и результаты тестов прослушивания.

1. Выбор меры тональности сигнала

Известно, что тональная структура сигнала играет существенную роль в вычислении порога одновременной маскировки. Рассмотрим известные способы определения тональных и шумоподобных компонент в музыкальном сигнале. Тональными называют те отсчеты быстрого преобразования Фурье, которые обусловлены наличием во входном сигнале продолжительных во времени гармонических составляющих (синусоид). Маскирующий эффект тональных и не тональных компонент различен, и поэтому с целью выявления тональных компонент выполняется специальный анализ спектра.

Спектральный анализ проводится в критических полосах слуха. В первой психоакустической модели (MPEG Layer 1) отыскиваются локальные максимумы энергетического спектра. Если мощность компоненты на определенную величину (7 дБ) превышает соседние локальные минимумы, данная компонента заносится в список тональных.

Во второй психоакустической модели (MPEG Layer 3) не классифицируются компоненты на тональные и шумоподобные. Вместо этого вычисляется индекс тональности, характеризующий близость данной компоненты к тональным или шумоподобным. Индекс тональности характеризует степень предсказуемости спектральной компоненты по значениям, вычисленным на предыдущих кадрах. Пусть $r_\omega = r_\omega(n)$, $\varphi_\omega = \varphi_\omega(n)$ обозначают соответственно амплитуду и фазу компоненты с номером ω в преобразовании Фурье для выборки n . Предсказанные значения амплитуды и фазы вычисляются по формулам (1а), (1б):

$$\hat{r}_\omega = 2r_\omega(t-1) - r_\omega(t-2), \quad (1a)$$

$$\hat{\varphi}_\omega = 2\varphi_\omega(t-1) - \varphi_\omega(t-2). \quad (1b)$$

Мера непредсказуемости определяется следующим образом:

$$c_\omega = \frac{\sqrt{(r_\omega \cos \varphi_\omega - \hat{r}_\omega \cos \hat{\varphi}_\omega)^2 + (r_\omega \sin \varphi_\omega - \hat{r}_\omega \sin \hat{\varphi}_\omega)^2}}{r_\omega + |\hat{r}_\omega|}. \quad (2)$$

Всего вычисляются меры непредсказуемости для 57 поддиапазонов. Для полосы с номером b через $\omega_l(b)$ обозначим нижнюю ее границу, а через $\omega_h(b)$ – верхнюю. Энергия поддиапазона равна:

$$e_b = \sum_{\omega_l(b)}^{\omega_h(b)} r_\omega^2, \quad (3)$$

а взвешенная мера непредсказуемости:

$$c_b = \sum_{\omega_l(b)}^{\omega_h(b)} r_\omega^2 c_\omega. \quad (4)$$

Эта пара величин используется как мера тональности данного поддиапазона.

К числу недостатков данного метода можно отнести высокую вычислительную сложность алгоритмов. Поскольку в будущем планируется предсказывать влияние вносимых искажений на субъективное восприятие музыкальных сигналов при потоковой цифровой обработке (выбор параметров обработки зависит от допустимого значения NMR), решено использовать более простой с точки зрения вычислительных затрат алгоритм, основанный на расчете равномерности спектра в полосе частот.

Количественной мерой тональности предлагается считать коэффициент тональности (КТ) [3] (также известный как коэффициент энтропии Винера), который вычисляется на основе вычисления меры пологости спектра в полосе частот (5):

$$\text{Tonality} = \min\left(\frac{\text{SFM}_{db}}{-60}, 1\right), \quad (5)$$

где SFM (*от англ.* Spectral Flatness Measure) – мера спектральной пологости: $\text{SFM}_{db} = 10 \cdot \log_{10}(\text{SFM})$, вычисляемая по формуле (6):

$$\text{SFM}(\text{num_band}) = \frac{(\prod_{k \in \text{num_band}} a(k))^{1/K}}{\frac{1}{K} \sum_{k \in \text{num_band}} a(k)}, \quad (6)$$

где $a(k)$ – амплитуды k -ой компоненты спектра.

Коэффициент тональности вычисляется для 4 диапазонов частот от 250 Гц до 8 кГц: басовый (< 500 Гц), низкий средний (500–1000 Гц), высокий средний (1–4 кГц) и высокий (> 4 кГц).

Коэффициент тональности, вычисленный по формуле (5), для белого шума близок к 0, для однотонального сигнала – к единице.

В сверхвысоком диапазоне частот обычно наблюдается понижение энергии компонент с увеличением частоты. Для устранения недостатка данного метода – высокий коэффициент тональности при значительном наклоне огибающей спектра в широком диапазоне, принято решение разбить диапазон сверхвысоких частот на четыре поддиапазона и рассчитывать КТ в этом диапазоне как усредненное значение КТ для поддиапазонов.

2. Сигналы

В предыдущей работе [1] были введены и подробно описаны условные обозначения для акустических сигналов, используемых в исследовании. Здесь приведем только обозначения:

- референсный музыкальный сигнал (МСреф) – неискаженный музыкальный сигнал;
- референсный многотональный сигнал (МТСреф) – многотональный сигнал, синтезированный из референсного музыкального сигнала;
- тестовый музыкальный сигнал (МСтест) – референсный музыкальный сигнал с внесенными нелинейными искажениями;
- тестовый многотональный сигнал (МТСтест) – референсный многотональный сигнал с внесенными нелинейными искажениями;
- очищенный тестовый многотональный сигнал (МТСтест_очищ), имеющий в своем составе только спектральные компоненты тестового многотонального сигнала, содержащиеся в референсном многотональном сигнале;
- сигнал нелинейных искажений (СНИ) – сигнал, содержащий только новые спектральные компоненты тестового многотонального сигнала после его прохождения через звуковой тракт испытуемого устройства, т.е. изначально отсутствующие в референсном многотональном сигнале.

В качестве референсных музыкальных сигналов были отобраны 5 фрагментов музыкальных треков с разными уровнями коэффициента тональности (таблица 1).

ТАБЛИЦА 1. Коэффициенты тональности для МСреф

№ п\п	Фрагмент	Тональность в диапазоне			
		Бас < 400 Гц	Низкие Средние 0,4–1 кГц	Высокие средние 1–4 кГц	Высокие 4–8 кГц
1	Adele, Someone Like You Part 1 (piano)	0,28	0,23	0,2	0,13
2	Adele, Someone Like You Part 2 (voice)	0,23	0,27	0,27	0,8
3	Jamiroquai – Music of the mind	0,25	0,33	0,15	0,09
4	Muse – Follow me	0,28	0,11	0,08	0,06
5	Toto – Simon solo	0,20	0,07	0,06	0,09

Каждому МСреф был создан МТСреф. В данном исследовании мы использовали метод создания многотонального сигнала, основанный на прореживании тональных компонент референсного музыкального сигнала [1]. В основе этого алгоритма создания МТСреф лежит последовательный анализ спектров выборок МСреф, выделение компонент, удовлетворяющих некоторым критериям отбора и обнулением прочих. В случае отсутствия явно выраженных максимумов сигнала в некоторой полосе имеющийся шумовой сигнал заменяется тональной компонентой, расположенной в цен-

тре этой полосы и обладающей такой же энергией. Информация о типе компонент в полосе (тональная или шумоподобная) хранится в специальном файле-паспорте сигнала. Выбор этого типа многотонального сигнала обусловлен тем, что его тональная структура наиболее близка к структуре музыкального сигнала.

После внесения искажений в МСреф и МТСреф получены МСтест и МТСтест для тестов прослушивания и анализа искажений соответственно.

3. Искажения

В процессе исследований в референсные музыкальные и многотональные сигналы вносились нелинейные искажения, как в отдельных узких полосах частот, так и одновременно во всем диапазоне слышимых частот. Уровень сигнала нелинейных искажений был привязан к порогу его маскировки полезным сигналом в каждом диапазоне частот. Для моделирования нелинейных искажений использовались способы, описанные в [1]. Основным алгоритмом создания сигнала нелинейных искажений является жесткое клиппирование исходного музыкального сигнала.

Приведем поэтапное описание процедуры создания сигнала нелинейных искажений.

На первом этапе МСреф обрабатывается фильтром низкой частоты с частотой пропускания 200-400 Гц.

На втором этапе полученный низкочастотный сигнал подвергается процедуре жесткого клиппирования. Уровень ограничителя выбирается с учетом количества искаженных участков сигнала: 1 %, 2 %, 5 % и 10 %. Влияние временного распределения искаженных фрагментов на субъективное предпочтение решено исследовать в заключительной части работы. В данном исследовании используется такой уровень ограничителя, чтобы выше него оказалось 5 % отсчетов низкочастотного сигнала.

На третьем этапе полученный сигнал обрабатывается фильтром высокой частоты с частотой пропускания 200–400 Гц (частота пропускания зависит от верхней частоты исходного низкочастотного сигнала). В результате получается широкополосный сигнал нелинейных искажений.

На четвертом этапе полученный сигнал нелинейных искажений обрабатывается полосовыми фильтрами с полосами пропускания < 600 Гц; 600–1000 Гц, 1–2 кГц, 2–4 кГц, > 4 кГц. В результате получаются сигналы нелинейных искажений в пяти частотных диапазонах.

На пятом этапе происходит корректировка уровней сигналов шумов. Для этого полученные сигналы смешиваются с МТСреф и вычисляются величины NMR. Обычно уровни NMR выше 4 дБ

отмечаются слушателями как критические с точки зрения восприятия.

На заключительном шаге полученные сигналы нелинейных искажений смешиваются с соответствующими МСреф для получения набора МСтест. Полученный набор сигналов используется в тестах прослушивания.

4. Выходные переменные

Для оценки влияния нелинейных искажений на субъективное предпочтение необходимо учитывать эффект одновременной маскировки искажений полезным сигналом и разное восприятие слышимых нелинейных искажений слуховой системы человека в зависимости от характеристик сигнала.

Для учета эффекта маскировки искажений полезным сигналом использовалась упрощенная психоакустическая модель стандарта MPEG-1 ISO/IEC 11172-3.

Анализируемый тестовый многотональный сигнал разбивается на выборки, содержащие 2048 отсчетов сигнала, и для каждой выборки сигнала вычисляется энергетический линейчатый спектр $X(k)$. Для уменьшения краевых эффектов при выполнении БПФ используется окно Ханна $h(n)$. Для полученного энергетического спектра выполняется процедура нормирования, здесь максимальной спектральной компоненте присваивается уровень 96 дБ, а все значения остальных спектральных компонент подвергаются процедуре коррекции относительно этой величины. Для каждой тональной и шумоподобной компоненты (информация о типе компоненты берется из паспорта сигнала) выполняется расчет коэффициентов маскировки и индивидуальных кривых маскировки $M[z(i), z(j)]$, и затем для них вычисляются пороги маскировки $N_{TC}[z(i)]$ и $N_{MC}[z(i)]$. Для аппроксимации порогов маскировки вне критических полос слуха используется выражение [2]:

$$B(\Delta z) = A_0 + 15,81 + \frac{7,5(\Delta z + 0,474) - 17,5\sqrt{1 + (\Delta z + 0,474)^2}}{1}, \quad (7)$$

где $A_0 = 8\min((\Delta z - 0,5)^2 - 2(\Delta z - 0,5), 0)$.

Пример семейства индивидуальных кривых маскировки для различных уровней маскирующего тона приведены на рисунке 1. В завершении, рассчитанные пороги суммируются вместе с абсолютным порогом слышимости для получения глобального порога маскировки $N_{MT}(i)$ в каждой субполосе психоакустического анализа [4–8].

Также следует отметить, что поскольку наблюдаются существенные расхождения между теоретически рассчитанными и измеренными для конкретного слушателя кривыми маскировки, были проведены специальные дополнительные тесты прослушивания для определения индивидуаль-

ных кривых маскировки для выбранных экспертов. Полученные поправки (в основном касающиеся крутизны заднего фланца кривой относительно порога слышимости) использовались для вычисления глобальных порогов маскировки, используемых в работе.

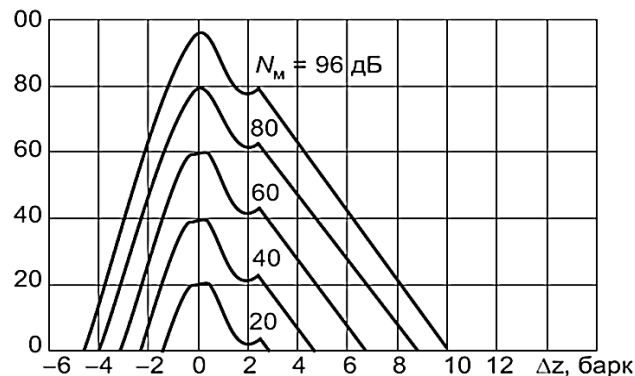


Рис. 1. Аппроксимации кривых маскировки (психоакустическая модель 2 стандарта MPEG ISO/IEC 111722-3)

В качестве количественной характеристики слышимых нелинейных искажений используется часто зависимое отношение «Шум-Маска» [6]:

$$\text{NMR}(n) = 10\log_{10} \left(\frac{1}{Z} \sum_{k=0}^{Z-1} \frac{E_{\text{noise}}[k, n]}{M[k, n]} \right), \quad (8)$$

здесь и ниже $E_{\text{noise}}[k, n]$ – энергия k -ой спектральной компоненты сигнала искажений; $M[k, n]$ – уровень порога глобальной маскировки; n – номер выборки.

В первой части работы для связи субъективной оценки с количественной характеристикой шумов через взвешивающую функцию G была введена выходная переменная: воспринимаемый уровень искажений (PNL, от англ. Perceive Noise Level). Было отмечено, что эта функция зависит от частоты:

$$\text{PNL}(n) = \sum_{k=0}^{Z-1} \text{NMR}(k, n) * G(k). \quad (9)$$

Как было сказано выше, в процессе анализа результатов тестов прослушивания наблюдается зависимость субъективного предпочтения искаженных сигналов не только от уровня нелинейных искажений, но и от тональной структуры сигнала – $T(f)$. Таким образом, выражение (9) можно переписать следующим образом:

$$\text{PNL}(n) = \sum_{k=1}^K \text{NMR}(k, n) * G(T(k), k). \quad (10)$$

Влияние нелинейных искажений на субъективную оценку качества воспроизводимого испытуемым устройством сигнала определяется средней величиной PNL, полученной для всего сигнала в целом (11):

$$TPNL = \frac{1}{N} \sum_{n=1}^N PNL(n), \quad (11)$$

где N – количество выборок в сигнале; n – порядковый номер выборки.

Основной задачей этой части работы является вычисление взвешивающей функции $G(T(f), f)$.

5. Эксперименты

Тесты прослушивания проводились в звукоизолированной комнате с уровнем фонового шума ~ 20 dB SPL в диапазоне частот 100 Гц–20 кГц. В качестве экспертов привлекались молодые люди в возрасте от 24 лет до 31 года без дефектов слуха, имеющие аудиометрический порог ниже 20 дБ в полосе частот 125 Гц–8 кГц (8 человек). Для оценки слуховой заметности вносимых искажений использовалась семибалльная шкала, при этом наиболее предпочтительное звучание получало оценку 7 баллов, наименее предпочтительное – 1 балл.

Тестовые музыкальные сигналы через внешнюю 24-х битовую звуковую карту MOTU UltraLite-mk3 подавались на усилитель HiFiMan EF-100, после чего воспроизводились с помощью высококачественных головных телефонов Sennheiser HD 800, имеющих крайне низкий коэффициент нелинейных искажений ($< 1\%$).

Для определения индивидуальных порогов маскировки для каждого из экспертов были проведены специальные тесты прослушивания, в которых перед экспертами ставилась задача разделить набор тестовых музыкальных сигналов на две группы: в первую группу вошли сигналы, в которых искажения не слышны, во вторую – сигналы со слышимыми искажениями. Разброс индивидуальных порогов маскировки составил в среднем 1 дБ. В представленном исследовании для расчета отношения «Шум-Маска» использовалось среднее значение порога для всех экспертов.

Перед началом исследований экспертам были предъявлены образцы отобранных МСреф (свободные от искажений, с оценкой 7 баллов). Аналогичным образом были предъявлены также образцы сигналов с максимальным количеством внешних нелинейных искажений (как наихудшее звучание, с оценкой 1 балл). Экспертам предлагалось прослушать весь набор записей в любом произвольном порядке для оценки их качества, затем, прослушивая повторно, поставить оценку в соответствии с собственным предпочтением, используя семибалльную шкалу.

Для определения взвешивающей функции $G(T(f), f)$ были проведены две серии субъективно-статистических экспертиз, для которых был подготовлен набор сигналов с известным количеством нелинейных искажений (см. 3. Искажения),

присутствующих, как в отдельных полосах частот, так и в нескольких частотных диапазонах одновременно.

В первой серии тестов прослушивания экспертам предлагалось оценить с точки зрения предпочтительности звучания набор тестовых музыкальных сигналов, имеющих нелинейные искажения только в одной полосе частот. Всего было предложено для прослушивания 5 групп (по количеству МСреф) по 28–33 тестовых фонограммы в каждой. Экспертам требовалось рассортировать их на 7 групп по степени предпочтительности. С помощью этой серии экспериментов удалось определить зависимости субъективной оценки от величины NMR и коэффициента тональности в полосе частот, содержащей нелинейные искажения (рисунок 2).

Во второй серии тестов прослушиваний экспертам предлагалось ранжировать с точки зрения предпочтительности звучания набор тестовых фонограмм, имеющих слышимые нелинейные искажения одновременно в нескольких полосах частот. Для проведения этой серии прослушиваний были также сформированы пять наборов референсных музыкальных сигналов, каждый из которых содержал около 10 образцов. Каждый образец представлял собой фрагмент МСреф, смикшированный с сигналами нелинейных искажений, содержащих различные комбинации уровня искажений в разных диапазонах частот.

В результате этих экспериментов была получена зависимость оценки предпочтения от величины TPNL (11).

6. Результаты

После обработки результатов были рассчитаны зависимости оценки предпочтения от величины NMR и тональности референсного сигнала в полосе частот. Затем полученные зависимости аппроксимированы прямыми вида (12):

$$\text{Score}(\text{band}) = 7 - k_{\text{band}}(\text{Tonality}) * \text{NMR}(\text{band}) + \varepsilon, \quad (12)$$

где $k_{\text{band}}(\text{Tonality})$ – коэффициент регрессии для диапазона частот band и тональности в полосе частот (tonality); ε – погрешность.

Для определения коэффициентов регрессии результаты были сведены в таблицы Excel и построены графики зависимостей и их линейные аппроксимации (см. рисунок 1).

Очевидно, что при дальнейшем увеличении количества слышимых искажений оценка не может снижаться бесконечно, начиная с некоторого уровня зависимости оценок предпочтения от NMR перестают быть линейными. Для определения коэффициентов регрессии мы использовали только линейные участки кривых. В таблице 2 приведены коэффициенты регрессии, полученные для разных сигналов.

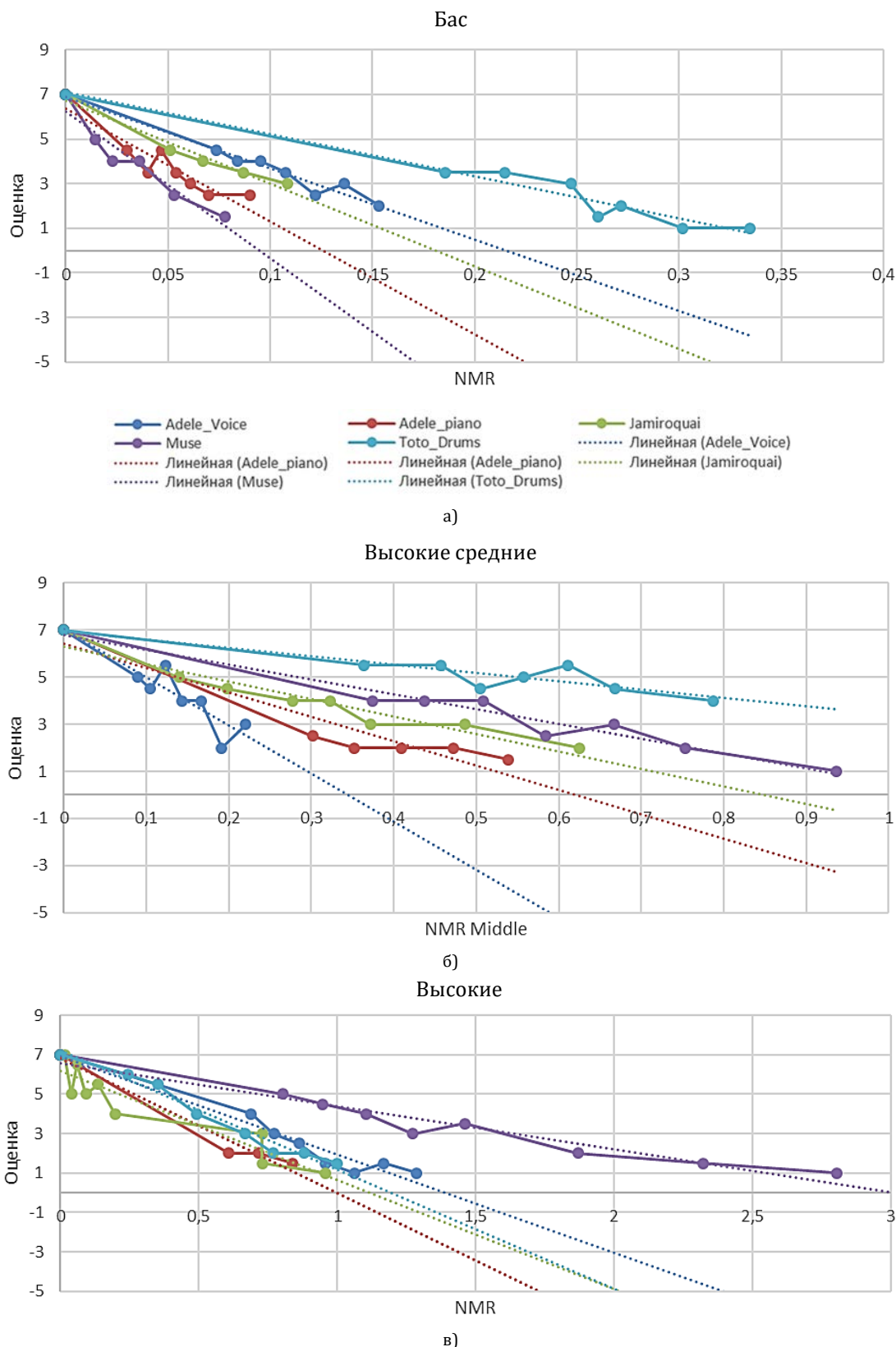


Рис. 2. Кривые изменения предпочтительности звучания от величины отношения «Шум-Маска» при наличии нелинейных искажений в полосе частот: а) 200–400 Гц; б) 1–4 кГц; в) 4–8 кГц

Далее был определен глобальный минимум коэффициентов регрессии и все значения разделены на этот минимум. Полученный набор нормиро-

ванных коэффициентов регрессии задает искомую взвешивающую функцию $G(T(f), f)$ в точках, соот-

ветствующих величинам коэффициентов тональности, полученных в экспериментах (рисунок 3).

Для упрощения расчетов принято решение использовать полиномиальную аппроксимацию усредненных значений коэффициентов регрессии для соответствующих значений тональности – $G(T)$:

$$G(T) = 5349 * T^3 - 2363 * T^2 + 358 * T - 13, \quad (13)$$

где T – рассчитанный коэффициент тональности.

ТАБЛИЦА 2. Коэффициенты регрессии k_{band} (Tonality)

Tonality	k низкие	k средние	k высокие средние	k высокие
0,06	–	–	3,5	2,2
0,08	–	5	6	5
0,09	–	8	6,5	5,4
0,13	–	15	7	6,8
0,15	–	16	–	–
0,2	–	18	–	–
0,23	–	20	–	–
0,25	–	30	–	–
0,27	–	40	–	–
0,28	–	45	–	–
0,3	–	60	–	–
0,33	–	90	–	–

Полученная взвешивающая функция была применена к вычисленным значениям NMR для отдельных выборок тестовых сигналов по формуле (10), значение было усреднено по количеству выборок по формуле (11). Значение субъективной оценки выбиралось по результатам тестов прослушивания, которые приведены на рисунке 3.

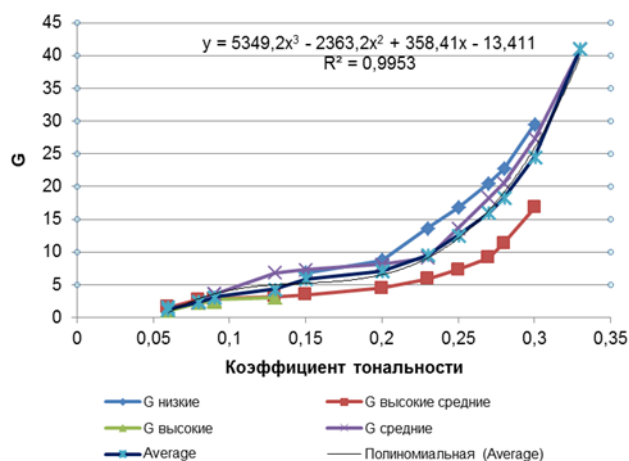


Рис. 3. Зависимости взвешивающей функции от тональности и частоты

Из рисунка 4 видно, что зависимость субъективного предпочтения практически линейно зависит от величины TPNL (корреляция составляет около 0,96 %), в то время как зависимость от величины отношения «Шум-Маска» гораздо менее строгая (корреляция составляет около 27 %).

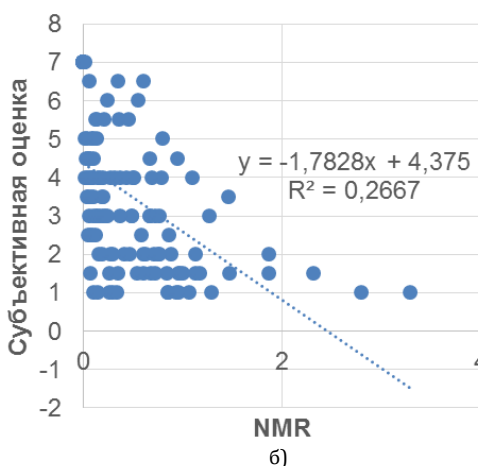
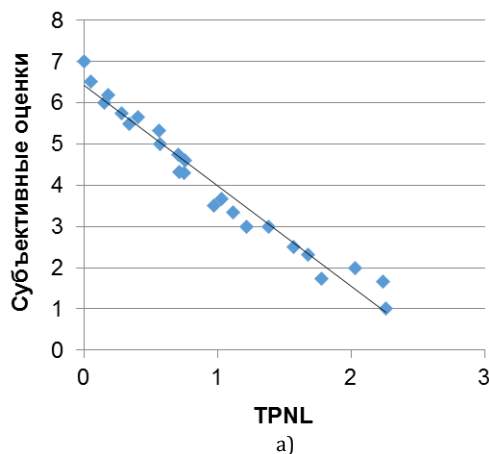


Рис. 4. Зависимость оценки: а) от средней величины воспринимаемых искажений (TPNL) (взвешенных величин NMR); б) от величины NMR

7. Заключение

В работе представлены результаты исследования влияния тональной структуры сигнала на субъективное восприятие музыкальных композиций с заметными нелинейными искажениями. Предложена вычислительно эффективная метрика тональности спектра выборки музыкального сигнала, позволяющая количественно оценивать тональность спектра в режиме обработки потокового аудиосигнала. Показаны зависимости влияния отношения «Шум-Маска» на субъективную оценку от коэффициента тональности. Вычислена взвешивающая функция $G(T)$, позволяющая установить связь между количеством слышимых искажений и субъективным предпочтением. Результаты экспериментов показали высокую точность предсказания субъективной оценки предпочтения сигналов, содержащих широкополосные нелинейные искажения.

Данная работа пока не может претендовать на полноту, так как в ней отсутствуют исследования влияния распределения звуковых искажений во времени и учета пиковых значений звуковых искажений в одном блоке искаженных кадров (параметр, аналогичный выходной переменной мо-

дели RDF в инструменте PEAQ). В заключительной части исследования будут представлены результаты исследования влияния временного распределения искаженных фрагментов на оценку субъек-

тивного предпочтения искаженных музыкальных сигналов, а также влияние сильно искаженных фрагментов.

Список используемых источников

1. Пахомов М.О., Рожнов В.С. Метод количественной оценки слухового восприятия нелинейных искажений. Часть 1. Испытательные сигналы и меры количественной оценки нелинейных искажений // Труды учебных заведений связи. 2018. Т. 4. № 1. С. 75–84. DOI:10.31854/1813-324x-2018-1-75-84.
2. Ковалгин Ю.А., Вологдин Э.И. Цифровое кодирование звуковых сигналов: Учебное пособие. СПб.: Корона Принт, 2004. 240 с.
3. Dubnov S. Generalization of spectral flatness measure for non-Gaussian linear processes // IEEE Signal Processing Letters. 2004. Vol. 11. Iss. 8. PP. 698–701. DOI: 10.1109/LSP.2004.831663.
4. Moore B.C. An Introduction to the Psychology of Hearing. London: Academic Press, 1989.
5. Moore B.C.J. A Model for the Prediction of Thresholds, Loudness, and Partial Loudness // Journal of the Audio Engineering Society. 1997. Vol 45. №. 4. PP. 224–240.
6. Tan C.T., Moore B.C.J., Zacharov N., Mattila V.V. Predicting the Perceived Quality of Nonlinearly Distorted Music and Speech Signals // Journal of the Audio Engineering Society. 2004. Vol. 52. № 7/8. PP. 699–711.
7. Temme S., Olive S., Tatarunis S., Welti T., McMullin E. The Correlation between Distortion Audibility and Listener Preference in Headphones // 137th AES Convention, Los Angeles, USA. 2014.
8. Zwicker E., Fastl H. Psychoacoustics: Facts and Models. Berlin, Heidelberg: Springer-Verlag, 1990.

* * *

METHOD FOR QUANTITATIVE EVALUATION OF AUDITORY PERCEPTION OF NONLINEAR DISTORTION. PART 2. METRIC FOR MUSIC SIGNAL TONALITY AND ITS IMPACT ON SUBJECTIVE PERCEPTION OF DISTORTIONS

M. Pakhomov¹, V. Rozhnov^{1, 2}

¹St. Petersburg Branch of the corporation LG Electronics inc.
St. Petersburg, 191123, Russian Federation

²The Bonch-Bruевич State University of Telecommunications,
St. Petersburg, 193232, Russian Federation

Article info

Article in Russian

For citation: Pakhomov M., Rozhnov V. Method for Quantitative Evaluation of Auditory Perception of Nonlinear Distortion. Part 2. Metric for Music Signal Tonality and its Impact on Subjective Perception of Distortions // Proceedings of Telecommunication Universities. 2018. Vol. 4. Iss. 3. PP. 98–105.

Annotation: In the work presented the topic of taking into account the spectral characteristics of the test signal when evaluating the subjective perception of nonlinear distortions is disclosed. The most significant characteristic – tonality is chosen and a quantitative measure of this characteristic is proposed – the tonality coefficient. Subjective listening tests were conducted to determine the dependence of the auditory perception of nonlinear distortions on the tone structure of the signal and the spectral distribution of the noise-mask relationship.

Keywords: multitone signal, nonlinear distortions, simultaneous masking, noise-to-mask ratio, tonality coefficient, linear multiregression model.