

Часть 14.1

Тембр, часть 1.

Ирина Алдошина

Определения тембра

Как уже было отмечено в предыдущих статьях, одной из главных задач психоакустики является установление соответствия между объективными параметрами звука (интенсивность, длительность, периодичность, расположение в пространстве и др.) и его субъективно воспринимаемыми характеристикам (высота, громкость, маскировка, тембр и др.). Как известно, связь между ними неоднозначна и нелинейна. Однако можно сказать, что субъективно ощущаемая высота тона связана в первую очередь с частотой (периодичностью), громкость - с интенсивностью и т. д. Высота тона позволяет классифицировать звуки по линейной шкале (выше-ниже), и служит в музыке основой мелодии, гармонии, интонации и пр. В свою очередь громкость определяет музыкальную динамику (ff...pp), баланс инструментов в ансамбле и является объемной характеристикой звука (больше-меньше).

Самым сложным субъективно ощущаемым параметром является тембр. С определением этого термина возникают сложности, сопоставимые с определением понятия "жизнь": все понимают, что это такое, однако над научным определением наука бьется уже несколько столетий. Аналогично с термином "тембр": всем ясно, о чем идет речь, когда говорят "красивый тембр голоса", "глухой тембр инструмента" и т. д., но... О тембре нельзя сказать "больше-меньше", "выше-ниже", для его описания используются десятки слов: сухой, звонкий, мягкий, резкий, яркий и т. д. (О терминах для описания тембра поговорим отдельно).

Уже более двухсот лет многие выдающиеся ученые пытаются дать научное определение этого параметра, которое, естественно, меняется с расширением наших представлений о механизмах работы слуховой системы. Определение тембра дается в трудах таких всемирно известных ученых, как Гельмгольц (1877), Флетчер (1938), Ликлайде (1951), Плом (1976), Наутсм (1989), Россин (1990), Ханде (1995).

Тембр (timbre-фр.) означает "качество тона", "окраску тона" (tone quality).

Американский стандарт ANSI-60-дает такое определение: "Тембр - атрибут слухового восприятия, который позволяет слушателю судить, что два звука, имеющие одинаковую высоту и громкость, различаются друг от друга".

В трудах Гельмгольца содержится следующее заключение: "разница в музыкальном качестве тона (тембре) зависит только от присутствия и силы парциальных тонов (обертонов), и не зависит от разности фаз, с которой эти парциальные тоны вступают в композицию". Это определение почти на сто лет определило направление исследований в области восприятия тембров, и претерпело существенные изменения и уточнения только в последние десятилетия. В трудах Гельмгольца был сделан еще целый ряд тонких наблюдений, которые подтверждаются современными результатами. В частности, им было установлено, что восприятие тембра зависит и от того, с какой скоростью парциальные тоны вступают в начале звука и умирают в его конце, а также, что наличие некоторых шумов и нерегулярностей помогает в распознавании тембров отдельных инструментов.

В 1938 г. Флетчер заметил, что тембр зависит от обертоновой структуры звука, но также изменяется при изменении громкости и высоты тона, хотя обертоновая структура может при

этом сохраняться. В 1951 г. известный специалист Ликлайдер добавил, что тембр является многомерным объектом восприятия - он зависит от общей обертоновой структуры звука, которая также может меняться с изменением громкости и высоты тона.

В 1973 г. к определению тембра, данному в вышеприведенном стандарте ANSI, было сделано следующее добавление: "тембр зависит от спектра сигнала, но он также зависит от формы волны, звукового давления, расположения частот в спектре и временных характеристик звука".

Только к 1976 г. в работах Пломпа было доказано, что ухо не страдает "фазовой глухотой", и восприятие тембра зависит как от амплитудного спектра (в первую очередь, от формы спектральной огибающей), так и от фазового спектра. В 1990 году Россинг добавил, что тембр зависит от временной огибающей звука и его длительности. В работах 1993-1995 гг. отмечено, что тембр является субъективным атрибутом того или иного источника (например, голоса, музыкального инструмента), то есть он позволяет выделить этот источник из различных звуковых потоков в различных условиях. Тембр обладает достаточной инвариантностью (стабильностью), что позволяет сохранить его в памяти, а также служит для сравнения ранее записанной и вновь поступившей в слуховую систему информации об источнике звука. Это предполагает определенный процесс обучения - если человек никогда не слышал звучание инструмента данного тембра, то он его и не узнает.

За последние годы исследованию восприятия тембра и установлению его зависимостей от физических характеристик звука было посвящено сотни статей и книг. Большие научные коллективы ведущих университетов и институтов (Стенфордский университет, Гарвардский университет, университет Беркли, ИРКАМ и др.) ведут интенсивные исследования, поскольку результаты, полученные в этом направлении, являются принципиально важными для решения общей проблемы "расшифровки слухового образа", что необходимо в целях дальнейшего развития аудиотехники, систем мультимедиа и других направлений. Рассказать об огромном количестве полученных за последние годы результатов по изучению восприятия тембра в коротких статьях невозможно, поэтому постараюсь остановиться только на некоторых из них, поскольку они чрезвычайно важны для работы звукорежиссера со звуком.

Тембр и акустические характеристики звука

Современные компьютерные технологии позволяют выполнить детальный анализ временной структуры любого музыкального сигнала - это может сделать практически любой музыкальный редактор, например, Sound Forge, Wave Lab, SpectroLab и др. Примеры временной структуры (осциллограмм) звуков одной высоты (нота "до" первой октавы), создаваемых различными инструментами (орган, скрипка), показаны на рисунке 1. Как видно из представленных волновых форм (т. е. зависимости изменения звукового давления от времени), в каждом из этих звуков можно выделить три фазы: атаку звука (процесс установления), стационарную часть, процесс спада. В различных инструментах, в зависимости от используемых в них способов звукообразования, временные интервалы этих фаз разные - это видно на рисунке 1.

У ударных и щипковых инструментов, например гитары, короткий временной отрезок стационарной фазы и атаки и длинный по времени - фазы затухания. В звуке органной трубы можно видеть достаточно длинный отрезок стационарной фазы и короткий период затухания и т. д. Если представить отрезок

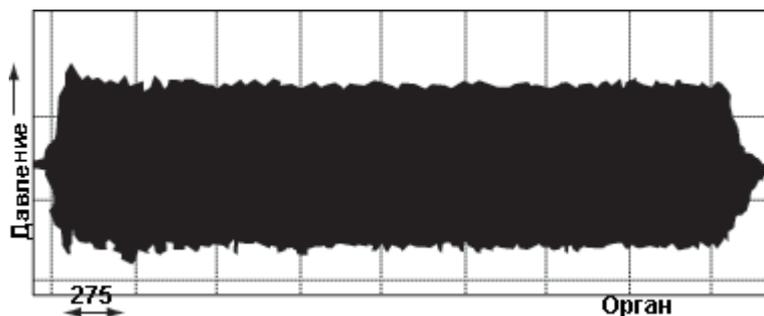
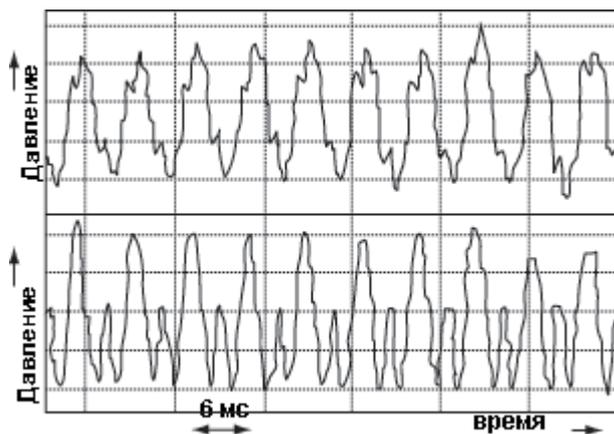


Рис. 1 Осциллограммы (волновая) форма звуков



стационарной части звучания более растянутым во времени (Рис. 2), то можно отчетливо видеть периодическую структуру звука. Как уже было сказано в предыдущих статьях, эта периодичность является принципиально важной для определения музыкальной высоты тона, поскольку слуховая система только для периодических сигналов может определить высоту, а непериодические сигналы воспринимаются ею как шумовые.

Как утверждает классическая теория, развиваемая, начиная с Гельмгольца почти все последующие сто лет, восприятие тембра зависит от спектральной структуры звука, то есть от состава обертонов и соотношения их амплитуд. Позволю себе напомнить, что обертоны - это все составляющие спектра выше фундаментальной частоты, а обертоны, частоты которых находятся в целочисленных соотношениях с основным тоном, называются гармониками.

Как известно, для того, чтобы получить амплитудный и фазовый спектр, необходимо выполнить преобразование Фурье от временной функции (t), т. е. зависимости звукового давления p от времени t .

С помощью преобразования Фурье любой временной сигнал можно представить в виде суммы (или интеграла) составляющих его простых гармонических (синусоидальных) сигналов, а амплитуды и фазы этих составляющих образуют соответственно амплитудный и фазовый спектры.

С помощью созданных за последние десятилетия цифровых алгоритмов быстрого преобразования Фурье (БПФ или FFT), выполнить операцию по определению спектров можно также практически в любой программе обработки звука. Например, программа SpectroLab вообще является цифровым анализатором, позволяющим построить амплитудный и фазовый спектр музыкального сигнала в различной форме. Формы представления спектра могут быть различными, хотя представляют они одни и те же результаты расчетов. На рисунке 3 представлены в виде АЧХ амплитудные спектры различных музыкальных инструментов (осциллограммы которых были показаны на рисунке

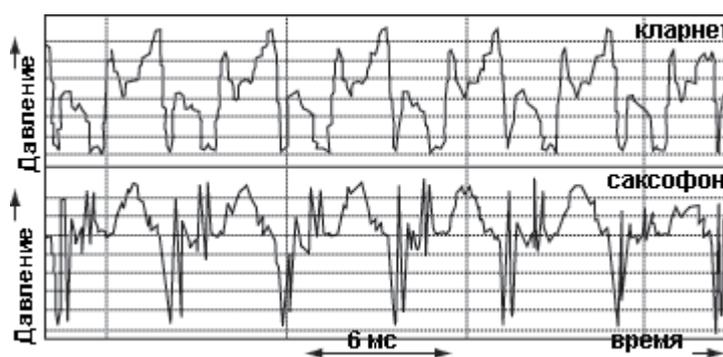
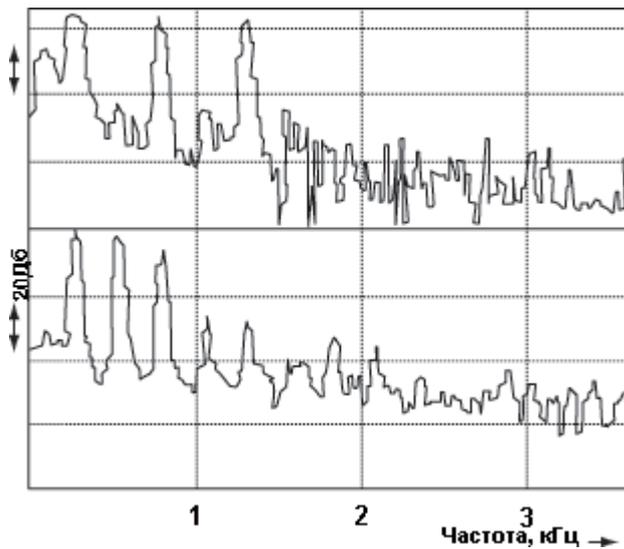


Рис. 2 Периодическая структура звуков.
а-орган; б-кларнет, саксофон



2). АЧХ представляет здесь зависимость амплитуд обертонов в виде уровня звукового давления в дБ, от частот.

Иногда спектр представляют в виде дискретного набора обертонов с разными амплитудами. Спектры могут быть представлены в виде спектрограмм, где по вертикальной оси отложена частота, по горизонтальной - время, а амплитуда представлена интенсивностью цвета (Рис. 4). Кроме того, существует форма представления в виде трехмерного (кумулятивного) спектра, о котором будет сказано далее.

Для построения указанных на рисунке 3 спектров, в стационарной части осциллограммы выделяется некоторый временной отрезок, и проводится расчет усредненного спектра по данному отрезку. Чем больше этот отрезок, тем точнее получается разрешающая способность по частоте, но при этом могут теряться (сглаживаться) отдельные детали временной структуры сигнала. Такие стационарные спектры обладают индивидуальными чертами, характерными для каждого музыкального инструмента, и зависят от механизма звукообразования в нем.

Например, флейта использует в качестве резонатора открытую с двух концов трубу, и поэтому содержит в спектре все четные и нечетные гармоники. При этом уровень (амплитуда) гармоник быстро уменьшается с частотой. У кларнета используется в качестве резонатора труба, закрытая с одного конца, поэтому в спектре, в основном, содержатся нечетные гармоники. У трубы в спектре много высокочастотных гармоник. Соответственно, тембры звучания у всех этих инструментов совершенно разные: у флейты - мягкий, нежный, у кларнета - матовый, глуховатый, у трубы - яркий, резкий.

Исследованию влияния спектрального состава обертонов на тембр посвящены сотни работ, поскольку эта проблема чрезвычайно важна как для проектирования музыкальных инструментов и высококачественной акустической аппаратуры, особенно в связи с развитием аппаратуры Hi-Fi и High-End, так и для слуховой оценки фонограмм и др. задач, встающих перед звукорежиссером. Накопленный огромный слуховой опыт наших замечательных звукорежиссеров - П.К. Кондрашина, В.Г. Динова, Е.В. Никульского, С.Г. Шугаля и др. - мог бы представить бесценные сведения по этой проблеме (особенно если бы они написали о нем в своих книгах, чего хотелось бы им пожелать).

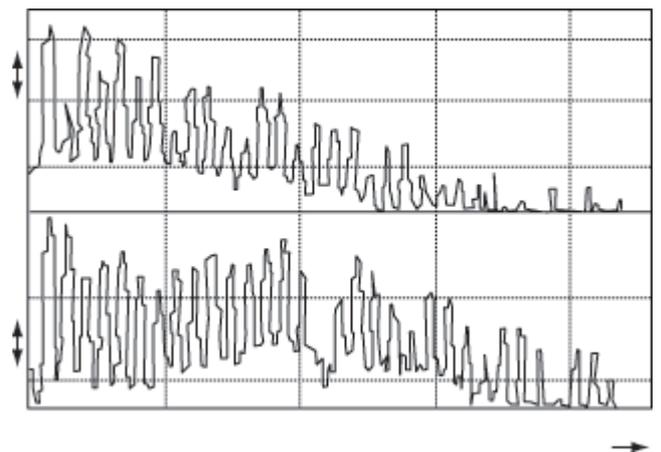


Рис. 3 Амплитудные спектры а-органа, б - кларнета, саксофона

Поскольку этих сведений чрезвычайно много и они часто противоречивы, приведем только некоторые из них.

Анализ общей структуры спектров различных инструментов, показанных на рисунке 5,

позволяет сделать следующие выводы:

- при отсутствии или недостатке обертонов, особенно в нижнем регистре, тембр звука становится скучным, пустым - примером может служить синусоидальный сигнал от генератора;
- присутствие в спектре первых пяти-семи гармоник с достаточно большой амплитудой придает тембру полноту и сочность;
- ослабление первых гармоник и усиление высших гармоник (от шестой-седьмой и выше) придает тембру резкость, скрипучесть;

Анализ огибающей амплитудного спектра для различных музыкальных инструментов позволил установить (Кузнецов "Акустика музыкальных инструментов"):

- плавный подъем огибающей (увеличение амплитуд определенной группы обертонов) в области 200...700 Гц позволяет получить оттенки сочности, глубины;
- подъем в области 2,5...3 кГц придает тембру полетность, звонкость;
- подъем в области 3...4,5 кГц придает тембру резкость, пронзительность и др.

Одна из многочисленных попыток классифицировать тембровые качества в зависимости от спектрального состава звука приведена в вышеуказанной книге (Рис. 6).

Многочисленные эксперименты с оценкой качества звучания (а, следовательно, тембра) акустических систем позволили установить влияние различных пиков-провалов АЧХ на заметность изменения тембра. В частности, показано, что заметность зависит от амплитуды, расположения по частотной шкале и добротности пиков-провалов на огибающей спектра (т. е. на АЧХ). В средней области частот пороги заметности пиков, т. е. отклонения от среднего уровня, составляют 2...3 дБ, причем заметность изменения тембра на пиках больше, чем на провалах. Узкие по ширине провалы (менее 1/3 октавы) почти не заметны на слух - по-видимому, это объясняется тем, что именно такие узкие провалы вносит помещение в АЧХ различных звуковых источников, и слух к ним привык.

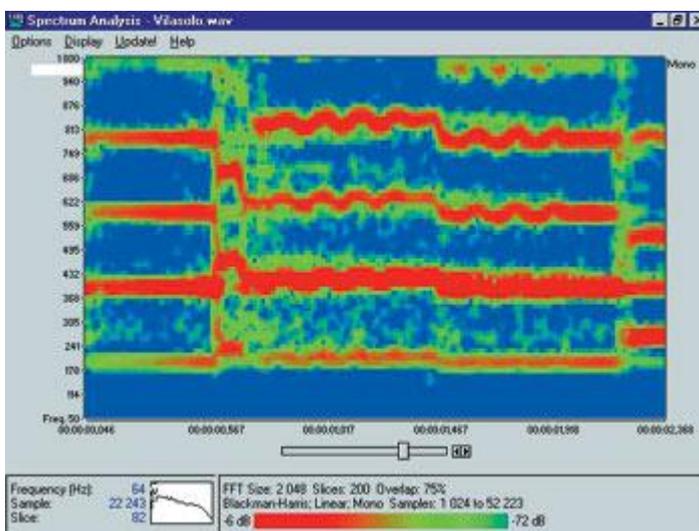


Рис. 4 Спектрограмма звука скрипки

Существенное влияние оказывает группировка обертонов в формантные группы, особенно в области максимальной чувствительности слуха. Поскольку именно расположение формантных областей служит главным критерием различимости звуков речи, наличие формантных частотных диапазонов (т. е. подчеркнутых обертонов) значительно влияет на восприятие тембра музыкальных инструментов и певческого голоса: например, формантная группа в области 2...3 кГц придает полетность, звонкость певческому голосу и звукам скрипки. Эта третья форманта особенно выражена в спектрах скрипок Страдивари.

Таким образом, безусловно справедливо утверждение классической теории, что воспринимаемый тембр звука зависит от его спектрального состава, то есть расположения обертонов на частотной шкале и соотношения их амплитуд. Это подтверждается многочисленной практикой работы со звуком в разных областях. Современные музыкальные программы позволяют легко проверить это на простых примерах. Например, можно в Sound Forge синтезировать с помощью встроенного генератора варианты звуков с различным

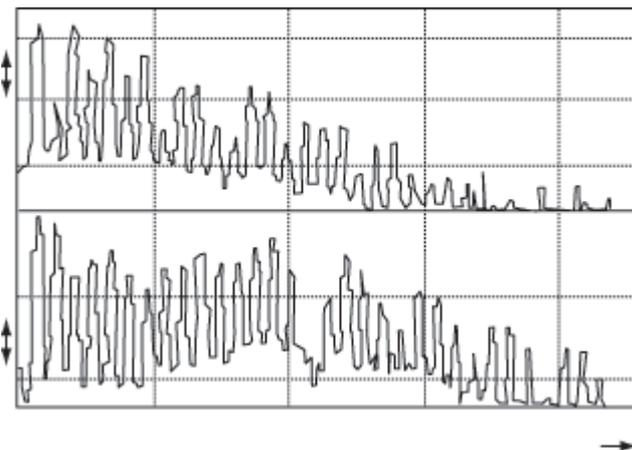
спектральным составом, и послушать, как изменяется тембр их звучания.

Из этого следуют еще два очень важных вывода:

- тембр звучания музыки и речи изменяется в зависимости от изменения громкости и от транспонирования по высоте.

При изменении громкости меняется восприятие тембра. Во-первых, при увеличении амплитуды колебаний вибраторов различных музыкальных инструментов (струн, мембран, дек и др.) в них начинают проявляться нелинейные эффекты, и это приводит к обогащению спектра дополнительными обертонами. На рисунке 7 показан спектр фортепиано при разной силе удара, где штрихом отмечена шумовая часть спектра.

Во-вторых, с увеличением уровня громкости изменяется чувствительность слуховой системы к восприятию низких и высоких частот (о кривых равной громкости было написано в предыдущих статьях). Поэтому при повышении громкости (до разумного предела 90...92 дБ) тембр становится полнее, богаче, чем при тихих звуках. При дальнейшем увеличении громкости начинают сказываться сильные искажения в источниках звука и слуховой системе, что приводит к ухудшению тембра.



Транспонирование мелодии по высоте также меняет воспринимаемый тембр. Во-первых, обедняется спектр, поскольку часть обертонов попадает в неслышимый диапазон выше 15...20 кГц; во-вторых, в области высоких частот пороги слуха значительно выше, и высокочастотные обертоны становятся не слышны. В звуках низкого регистра (например, в органе) обертоны усиливаются из-за повышения чувствительности слуха к средним частотам, поэтому звуки низкого регистра звучат сочнее, чем звуки среднего регистра, где такого усиления обертонов нет. Следует отметить, что поскольку кривые равной громкости, как и потери чувствительности слуха к высоким частотам, в значительной степени индивидуальны, то и изменение восприятия тембра при изменении громкости и высоты также очень различаются у разных людей.

Однако, накопленные к настоящему времени экспериментальные данные позволили выявить определенную инвариантность (стабильность) тембра при целом ряде условий. Например, при транспонировании мелодии по частотной шкале оттенки тембра, конечно, меняются, но в целом тембр инструмента или голоса легко опознается: при прослушивании, например, саксофона или другого инструмента через транзисторный радиоприемник можно опознать его тембр, хотя спектр его был значительно искажен. При прослушивании одного и того же инструмента в разных точках зала его тембр так же меняется, но принципиальные свойства тембра, присущие данному инструменту, остаются.

Некоторые из этих противоречий удалось частично объяснить в рамках классической спектральной теории тембра. Например, было показано, что для сохранения основных признаков тембра при транспонировании (переносе по частотной шкале) принципиально важным является сохранение формы огибающей амплитудного спектра (т. е. его формантной структуры). Например, на рисунке 8 показано, что при переносе спектра на октаву в том случае, когда структура огибающей сохраняется (вариант "а"), вариации тембра менее значительны, чем при переносе спектра с сохранением соотношения амплитуд (вариант "б"). Этим объясняется то, что звуки речи (гласные, согласные) можно распознать независимо от того, с какой высотой (частотой фундаментального тона) они произнесены, если при этом сохраняется расположение их формантных областей относительно друг друга.



Рис. 6 Классификация тембров

Таким образом, подводя итоги, полученные классической теорией тембра с учетом результатов последних лет, можно сказать, что тембр, безусловно, существенно зависит от усредненного спектрального состава звука: количества обертонов, их относительного расположения на частотной шкале, от соотношения их амплитуд, то есть формы спектральной огибающей (АЧХ), а точнее, от спектрального распределения энергии по частоте.

Однако, когда в 60-х годах начались первые опыты синтеза звуков музыкальных инструментов, попытки воссоздать звучание, в частности, трубы по известному составу ее усредненного спектра оказались неудачными - тембр был совершенно не похож на звук медных духовых инструментов. То же относится и к первым попыткам синтеза голоса. Именно в это период, опираясь на возможности, который предоставили компьютерные технологии, началось развитие другого направления - установление связи восприятия тембра с временной структурой сигнала.

Прежде, чем переходить к результатам, полученным в этом направлении, надо сказать следующее.

Первое. Довольно широко распространено мнение, что при работе со звуковыми сигналами достаточно получить информацию об их спектральном составе, поскольку перейти к их временной форме всегда можно с помощью преобразования Фурье, и наоборот. Однако, однозначная связь между временным и спектральным представлениями сигнала существует только в линейных системах, а слуховая система является принципиально нелинейной системой, как при больших, так и при малых уровнях сигнала. Поэтому обработка информации в слуховой системе происходит параллельно как в спектральной, так и во временной области (см. "Звукорежиссер" 6/1999 г.).

Разработчики высококачественной акустической аппаратуры сталкиваются с этой проблемой постоянно, когда искажения АЧХ акустической системы (то есть неравномерность спектральной огибающей) доведены почти до слуховых порогов (неравномерность 2 дБ, ширина полосы 20 Гц...20 кГц и т. д.), а эксперты или звукорежиссеры говорят: "скрипка звучит холодно" или "голос с металлом" и т.п. Таким образом, информации, полученной из

спектральной области, для слуховой системы недостаточно, нужна информация о временной структуре. Неудивительно, что методы измерений и оценки акустической аппаратуры существенно изменились за последние годы - появилась новая цифровая метрология, позволяющая определить до 30 параметров, как во временной, так и в спектральной областях.

Следовательно, информацию о тембре музыкального и речевого сигнала слуховая система должна получать как из временной, так и из спектральной структуры сигнала.

Второе. Все полученные выше результаты в классической теории тембра (теории Гельмгольца) базируются на анализе стационарных спектров, полученных из стационарной части сигнала с определенным усреднением, однако принципиально важным является то обстоятельство, что в реальных музыкальных и речевых сигналах практически нет постоянных, стационарных частей. Живая музыка - это непрерывная динамика, постоянное изменение, и это связано с глубинными свойствами слуховой системы.

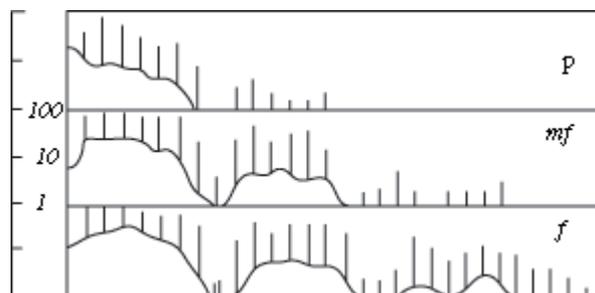


Рис. 7 Зависимость состава спектра от силы удара

Исследования физиологии слуха позволили установить, что в слуховой системе, особенно в ее высших разделах, имеется множество так называемых нейронов "новизны" или "опознавания", т. е. нейронов, которые включаются и начинают проводить электрические разряды, только если есть изменения в сигнале (включение, выключение, изменение уровня громкости, высоты и т. д.). Если же сигнал стационарный, то эти нейроны не включаются, и контроль за сигналом осуществляет ограниченное количество нейронов. Это явление широко известно из повседневной жизни: если сигнал не меняется, то часто его просто перестают замечать.

Для музыкального исполнения всякие монотонность и постоянство являются губительными: у слушателя отключаются нейроны новизны и он перестает воспринимать информацию (эстетическую, эмоциональную, смысловую и др), поэтому в живом исполнении всегда есть динамика (музыканты и певцы широко используют различную модуляцию сигнала - вибрато, тремоло и пр.).

Кроме того, каждый музыкальный инструмент, включая голос, обладает особой системой звукообразования, которая диктует свою временную структуру сигнала и его динамику изменения. Сравнение временной структуры звука (Рис. 1) показывает принципиальные различия: в частности, длительности всех трех частей - атаки, стационарной части и спада - у всех инструментов различаются по продолжительности и по форме. У ударных инструментов очень короткая стационарная часть, время атаки 0,5...3 мс и время спада 0,2...1 с; у смычковых время атаки 30...120 мс, время спада 0,15...0,5 с; у органа атака - 50...1000 мс и спад 0,2...2 с. Кроме того, принципиально отличается форма временной огибающей (Рис. 1).

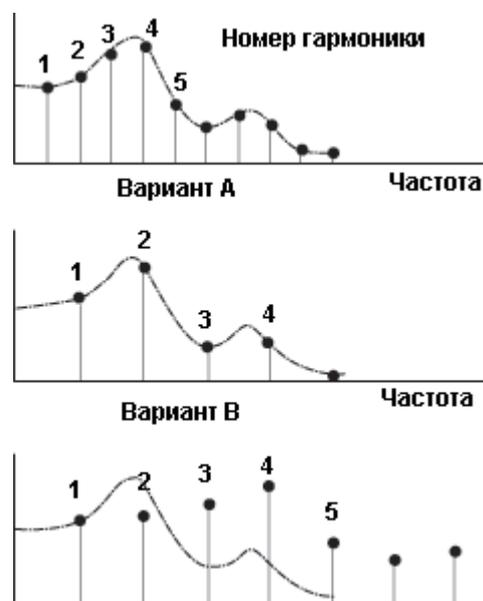


Рис. 8 Изменение огибающей спектра при сдвиге высоты

Эксперименты показали, что, если удалить часть временной структуры, соответствующей атаке звука, или поменять местами атаку и спад (проиграть в обратном направлении), или атаку от одного инструмента заменить атакой от другого, то опознать тембр данного инструмента становится практически невозможным. Следовательно, для распознавания тембра не только стационарная часть (усредненный спектр которой служит основой классической теории тембра), но и период формирования временной структуры, как и период затухания (спада) являются жизненно важными элементами.

Действительно, при прослушивании в любом помещении первые отражения поступают на слуховую систему после того, как атака и начальная часть стационарной части уже была услышана. В то же время на спад звука от инструмента накладывается реверберационный процесс помещения, что значительно маскирует звук, и, естественно, приводит к модификации восприятия его тембра. Слух обладает определенной инерционностью, и короткие звуки воспринимаются как щелчки. Поэтому длительность звука должна быть больше 60 мс, чтобы можно было распознать высоту, и, соответственно, тембр. По-видимому, постоянные должны быть близки.

Тем не менее, времени между началом прихода прямого звука и моментами поступления первых отражений оказывается достаточно, чтобы распознать тембр звучания отдельного инструмента - очевидно, этим обстоятельством и определяется инвариантность (стабильность) распознавания тембров разных инструментов в разных условиях прослушивания. Современные компьютерные технологии позволяют достаточно детально проанализировать процессы установления звука у разных инструментов, и выделить самые существенные акустические признаки, наиболее важные для определения тембра. Детальный анализ этих признаков будет выполнен во второй части статьи.