

часть 14
Тембр, часть 3
Ирина Алдошина

Тембр и фазовый спектр

Все изложенные в предыдущих двух статьях результаты по установлению связи воспринимаемого тембра с акустическими характеристиками сигнала относились к амплитудному спектру, точнее, к временному изменению спектральной огибающей (в первую очередь смещению энергетического центра амплитудного спектра-центроида) и развертыванию во времени отдельных обертонов.

В этом направлении было проделано наибольшее количество работ и получено много интересных результатов. Как уже было отмечено, на протяжении почти ста лет в психоакустике превалировало мнение Гельмгольца о том, что наша слуховая система не чувствительна к изменениям фазовых соотношений между отдельными обертонами. Однако постепенно были накоплены экспериментальные данные о том, что слуховой аппарат чувствителен к изменениям фаз между различными компонентами сигнала (работы Шредера, Хартмана и др.). В частности, было установлено, что слуховой порог к фазовому сдвигу в двух- и трехкомпонентных сигналах в области низких и средних частот составляет 10...15 градусов.

В 80-х годах это привело к созданию ряда акустических систем с линейно-фазовой характеристикой. Как известно из общей теории систем, для неискаженной передачи сигнала необходимо, чтобы соблюдались постоянство модуля передаточной функции, т.е. амплитудно-частотной характеристики (огибающей амплитудного спектра), и линейная зависимость фазового спектра от частоты, т.е. $\varphi(\omega) = -\omega T$.

Действительно, если амплитудная огибающая спектра сохраняется постоянной, то, как было сказано выше, искажений звукового сигнала при этом не должно происходить. Требования же к сохранению линейности фазы во всем диапазоне частот, как показали исследования Блауерта, оказались избыточными. Было установлено, что слух реагирует в первую очередь на скорость изменения фазы (т.е. ее производную по частоте), которая называется "групповое время задерживания ГВЗ": $\tau = -d\varphi(\omega)/d\omega$.

В результате многочисленных субъективных экспертиз были построены пороги слышимости искажений ГВЗ (т.е. величины отклонения $\Delta\tau$ от ее постоянного значения) для различных речевых, музыкальных и шумовых сигналов. Эти слуховые пороги зависят от частоты, и в области максимальной чувствительности слуха составляют 1...1,5 мс (рисунок 1). Поэтому последние годы при создании акустической аппаратуры Hi-Fi ориентируются, в основном, на приведенные выше слуховые пороги по искажению ГВЗ.

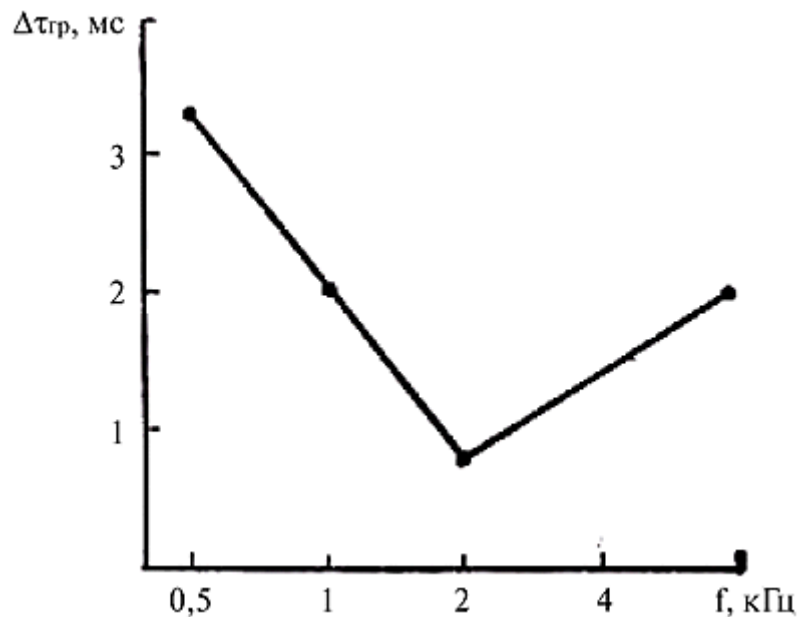


Рис. 1
Зависимость порогов слышимости искажений ГВЗ
от частоты

Что касается влияния фазовых искажений на оценку тембра различных музыкальных инструментов, то были выполнены исследования на синтезированных звуках различных музыкальных инструментов по субъективным оценкам изменения тембра при введении различных фазовых искажений. В реальных музыкальных инструментах генерируются достаточно сложные сигналы с большим количеством обертонов, определенной негармоничностью между ними, сложным динамичным развитием во времени спектральной и временной огибающих, а следовательно определенной динамикой развития во времени фазовых соотношений.

В одной из последних фундаментальных работ, посвященных этому вопросу (Галембо, Аскенфельд, Кадди - 2001 г.), были выполнены исследования на синтезированных звуках фортепиано. Для анализа был выбран низкочастотный диапазон (основные тона ниже 100 Гц), поскольку, как известно (см. "Звукорежиссер", 6/1999), при анализе сигнала в слуховой системе в области низких частот превалируют временные процессы.

Как уже было отмечено в той же статье, посвященной определению высоты тона, если какой-то из обертонов необычно ведет себя по амплитуде, то слуховая система выделяет его из общего ряда, и назначает ему отдельную высоту тона. Оказалось, что подобное поведение в фазовых спектрах также приводит к аналогичным результатам.

Кроме того, при оценке высоты тона "неразвернутых" гармоник, определение высоты происходит по их огибающей, частота которой равна фундаментальной частоте. Если все обертоны музыкального тона находятся в фазе, то периодичность огибающей становится четко выраженной, и возрастает точность определения высоты тона ("сила высоты тона"). Если фазовые соотношения между гармониками становятся различными, то волновая структура суммарного звука претерпевает существенные изменения (пример показан на рисунке 2), и высота тона становится менее определенной (уменьшается сила высоты звука).

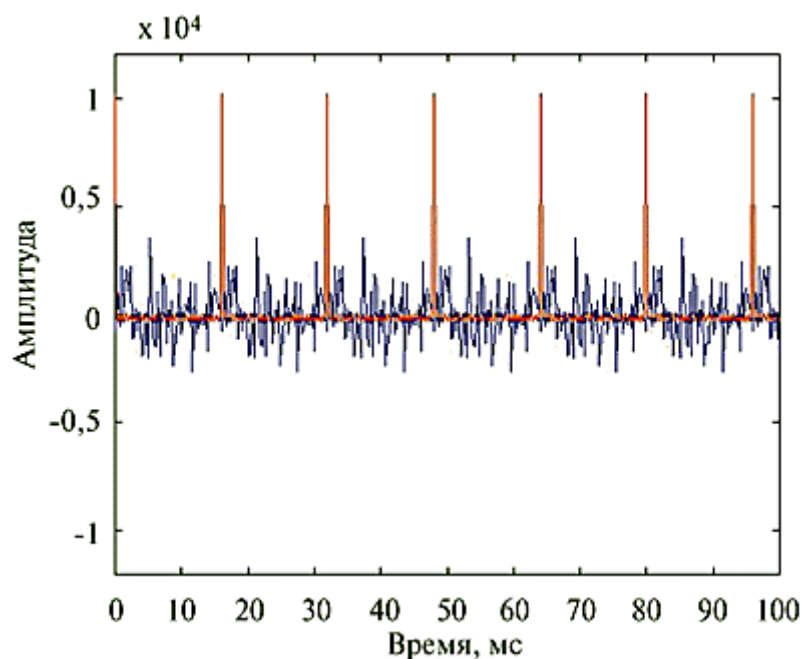


Рис. 2

Вид волновой формы при разных соотношениях фаз обертонов; красная - все обертоны имеют одинаковые начальные фазы, синяя - фазы распределены случайно

Таким образом, если фазовые соотношения оказывают слышимое влияние на определение высоты тона, то можно ожидать, что они окажут существенное влияние и на распознавание тембра.

Для экспериментов были выбраны звуки с основным тоном 27,5 и 55 Гц и со ста обертонами, с равномерным соотношением амплитуд, характерным для звуков фортепиано. При этом исследовались и тоны со строго гармоничными обертонами, и с определенной характерной для звуков фортепиано негармоничностью, которая возникает из-за конечной жесткости струн, их неоднородности, наличия продольных и крутильных колебаний и др.

Исследуемый звук синтезировался как сумма его обертонов: $X(t) = \sum A(n) \sin[2\pi f(n)t + \varphi(n, 0)]$

Для фазы, начальная фаза была принята равной нулю для всех обертонов $\varphi(n, 0) = 0$;

- Б - альтернативная слуховых экспериментов было выбраны следующие соотношения начальных фаз для всех обертонов:

- А - синусоидальная фаза (синусоидальная для четных и косинусоидальная для нечетных), начальная фаза $\varphi(n, 0) = \pi/4[(-1)^n + 1]$;

- С - случайное распределение фаз; начальные фазы при этом изменялись случайным образом в интервале от 0 до 2π .

В первой серии экспериментов все сто обертонов имели одинаковые амплитуды, различались только их фазы (основной тон 55 Гц). При этом прослушиваемые тембры получились различными:

- в первом случае (А), прослушивалась отчетливая периодичность;

- во- втором (Б), тембр был ярче и прослушивалась еще одна высота тона на октаву выше первой (правда высота не была четкой);

- в третьем (С) - тембр получился более равномерный.

Необходимо заметить - вторая высота прослушивалась только в наушниках, при прослушивании через громкоговорители все три сигнала отличались только тембром (сказывалась реверберация).

Это явление - изменение высоты тона при изменении фазы некоторых составляющих спектра - можно объяснить тем, что при аналитическом представлении преобразования Фурье сигнала типа Б, его можно представить как сумму двух комбинаций обертонов: сто обертонов с фазой типа А, и пятьдесят обертонов с фазой, отличающейся на $3\pi/4$, и амплитудой больше в $\sqrt{2}$. Этой группе обертонов слух назначает отдельную высоту тона. Кроме того, при переходе от соотношения фаз А к фазам типа В смещается центроид спектра (максимум энергии) в сторону высоких частот, поэтому тембр кажется ярче.

Аналогичные эксперименты со сдвигом фаз отдельных групп обертонов также приводят к появлению дополнительной (менее ясной) виртуальной высоты тона. Это свойство слуха связано с тем, что слух сравнивает звук с определенным имеющимся у него образцом музыкального тона, и если какие-то гармоники выпадают из типичного для данного образца ряда, то слух выделяет их отдельно, и назначает им отдельную высоту.

Таким образом, результаты исследований Галембо, Аскенфельда и др. показали, что фазовые изменения в соотношениях отдельных обертонов достаточно отчетливо слышны как изменения тембра, и в некоторых случаях - высоты тона.

Особенно это проявляется при прослушивании реальных музыкальных тонов фортепиано, в которых амплитуды обертонов убывают с увеличением их номера, имеют место особая форма огибающей спектра (формантной структуры), и отчетливо выраженная негармоничность спектра (т.е. сдвиг частот отдельных обертонов по отношению к гармоническому ряду).

Во временной области наличие негармоничности приводит к дисперсии, то есть высокочастотные компоненты распространяются по струне с большей скоростью, чем низкочастотные, и волновая форма сигнала изменяется. Наличие небольшой негармоничности в звуке (0,35%) добавляет некоторую теплоту, жизненность звучания, однако, если эта негармоничность становится большой, в звучании становятся слышны биения и другие искажения.

Негармоничность приводит также к тому, что если в начальный момент фазы обертонов находились в детерминированных соотношениях, то при ее наличии соотношения фаз со временем становятся случайными, пиковая структура волновой формы сглаживается, и тембр становится более равномерным - это зависит от степени негармоничности. Поэтому мгновенное измерение регулярности соотношения фаз между соседними обертонами может служить индикатором тембра.

Таким образом, эффект фазового перемешивания за счет негармоничности проявляется в некотором изменении восприятия высоты тона и тембра. Необходимо заметить, что эти эффекты слышны при прослушивании на близком расстоянии от деки (в позиции пианиста) и при близком расположении микрофона, причем слуховые эффекты различаются при прослушивании в наушниках и через громкоговорители. В реверберационном окружении сложный звук с высоким пик-фактором (что соответствует высокой степени регуляризации фазовых соотношений) говорит о близости источника звука, поскольку по мере удаления от него фазовые отношения приобретают все более случайный характер за счет отражений в помещении. Этот эффект может служить причиной разных оценок звучания пианистом и слушателем, а также разного тембра звука, записанного микрофоном у деки и у слушателя. Чем ближе, тем выше регуляризация фаз между обертонами и более определенная высота тона, чем дальше, тем более равномерный тембр и менее четкая высота.

Работы по оценке влияния фазовых соотношений на восприятие тембра музыкального звука сейчас активно изучаются в различных центрах (например, в ИРКАМе), и можно ожидать в ближайшее время новых результатов.

Тембр и общие принципы распознавания слуховых образов

Тембр является идентификатором физического механизма образования звука по ряду признаков, он позволяет выделить источник звука (инструмент или группу инструментов), и определить его физическую природу.

Это отражает общие принципы распознавания слуховых образов, в основе которых, как считает современная психоакустика, лежат принципы гештальт-психологии (gestalt, нем. - "образ"), которая утверждает, что для разделения и распознавания различной звуковой информации, приходящей к слуховой системе от разных источников в одно и то же время (игра оркестра, разговор многих собеседников и др.) слуховая система (как и зрительная) использует некоторые общие принципы:

- сегрегация - разделение на звуковые потоки, т.е. субъективное выделение определенной группы звуковых источников, например, при музыкальной полифонии слух может отслеживать развитие мелодии у отдельных инструментов;
- подобие - звуки, похожие по тембру, группируются вместе и приписываются одному источнику, например, звуки речи с близкой высотой основного тона и похожим тембром определяются, как принадлежащие одному собеседнику;
- непрерывность - слуховая система может интерполировать звук из единого потока через маскер, например, если в речевой или музыкальный поток вставить короткий отрезок шума, слуховая система может не заметить его, звуковой поток будет продолжать восприниматься как непрерывный;
- "общая судьба" - звуки, которые стартуют и останавливаются, а также изменяются по амплитуде или частоте в определенных пределах синхронно, приписываются одному источнику.

Таким образом, мозг производит группировку поступившей звуковой информации как последовательную, определяя распределение по времени звуковых компонент в рамках одного звукового потока, так и параллельную, выделяя частотные компоненты присутствующие и изменяющиеся одновременно. Кроме того, мозг все время проводит сравнение поступившей звуковой информации с "записанными" в процессе обучения в памяти звуковыми образами. Сравнивая поступившие сочетания звуковых потоков с имеющимися образами, он или легко их идентифицирует, если они совпадают с этими образами, или, в случае неполного совпадения, приписывает им какие-то особые свойства (например, назначает виртуальную высоту тона, как в звучании колоколов).

Во всех этих процессах распознавание тембра играет принципиальную роль, поскольку тембр является механизмом, с помощью которого экстрактируются из физических свойств признаки, определяющие качество звука: они записываются в памяти, сравниваются с уже записанными, и затем идентифицируются в определенных зонах коры головного мозга (Рис. 3).

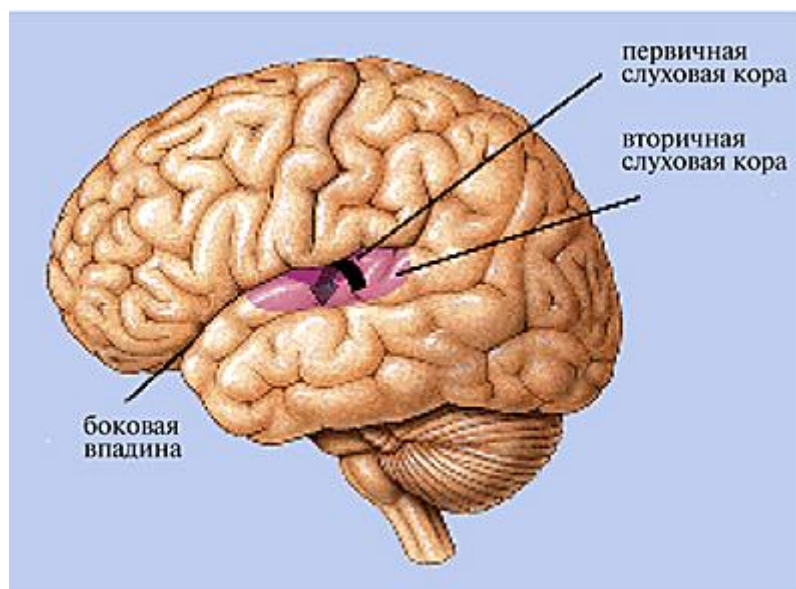


Рис. 3
Слуховые зоны мозга

Тембр - ощущение многомерное, зависящее от многих физических характеристик сигнала и окружающего пространства. Были проведены работы по шкалированию тембра в метрическом пространстве (шкалы - это различные спектрально-временные характеристики сигнала, см. вторую часть статьи в предыдущем номере). В последние годы, однако, появилось понимание, что классификация звуков в субъективно воспринимаемом пространстве не соответствует обычному ортогональному метрическому пространству, там происходит классификация по "субпространствам", связанным с вышеуказанными принципами, которые и не метрические, и не ортогональные.

Разделяя звуки по этим субпространствам, слуховая система определяет "качество звука", то есть тембр, и решает, к какой категории отнести эти звуки. Однако следует отметить, что все множество субпространств в субъективно воспринимаемом звуковом мире строится на основе информации о двух параметрах звука из внешнего мира - интенсивности и времени, а частота определяется временем прихода одинаковых значений интенсивности. Тот факт, что слух разделяет поступившую звуковую информацию сразу по нескольким субъективным субпространствам, повышает вероятность того, что в каком-то из них она может быть распознана. Именно на выделение этих субъективных субпространств, в которых происходит распознавание тембров и других признаков сигналов, и направлены усилия ученых в настоящее время.

Заключение

Подводя некоторые итоги, можно сказать, что основными физическими признаками, по которым определяется тембр инструмента, и его изменение во времени, являются:

- выстраивание амплитуд обертонов в период атаки;
- изменение фазовых соотношений между обертонами от детерминированных к случайным (в частности, за счет негармоничности обертонов реальных инструментов);
- изменение формы спектральной огибающей во времени во все периоды развития звука: атаки, стационарной части и спада;
- наличие нерегулярностей спектральной огибающей и положение спектрального центроида (максимума спектральной энергии, что связано с восприятием формант) и их изменение во времени (Рис. 4);

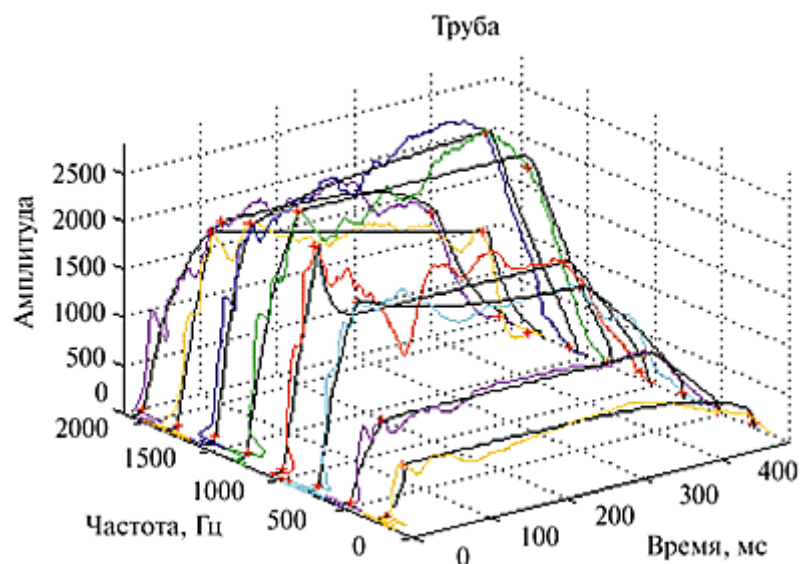


Рис. 4
Общий вид спектральных огибающих
и их изменение во времени

- наличие модуляций - амплитудной (тремоло) и частотной (вibrато);
- изменение формы спектральной огибающей и характера ее изменения во времени;
- изменение интенсивности (громкости) звучания, т.е. характера нелинейности звукового источника;
- наличие дополнительных признаков идентификации инструмента, например, характерный шум смычка, стук клапанов, скрип винтов на рояле и др.

Разумеется, все это не исчерпывает перечень физических признаков сигнала, определяющих его тембр. Поиски в этом направлении продолжаются.

Однако при синтезе музыкальных звуков необходимо учитывать все признаки для создания реалистичного звучания.

Интересная классификация инструментов была предложена в IRCAMе - "бинарное дерево" (рисунок 5). Если выделить пять признаков, и оценить их по 30-бальной шкале (негармоничность, форма атаки, форма спада, vibrато, тремоло, форма огибающей и др.), то все пять исследуемых инструментов можно расположить на "бинарном дереве", что может соответствовать их классификации по тембрам.

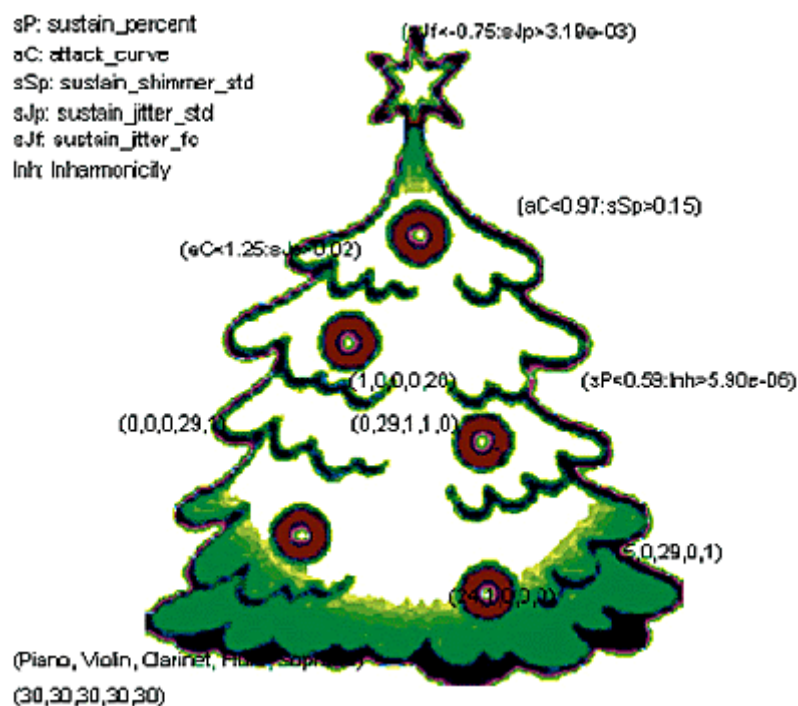


Рис. 5
"Бинарное дерево"

Приведенные в этих трех статьях сведения являются только несколькими первыми шагами в этом направлении, и далеко не исчерпывают проблемы. Надеюсь, что у нас будет возможность возвращаться к этим проблемам в дальнейшем, а также, надеюсь, что поставленные здесь вопросы заинтересуют наших читателей, и подтолкнут их к проведению научных и практических исследований в направлении анализа и восприятия тембра.

Приложение

Вербальное (словесное) описание тембра

Если для оценки высоты звуков имеются соответствующие единицы измерения: психофизические (мелы), музыкальные (октавы, тоны, полутоны, центы); есть единицы для громкости (соны, фоны), то для тембров такие шкалы построить невозможно, поскольку это понятие многомерное. Поэтому, наряду с описанными выше поисками корреляции восприятия тембра с объективными параметрами звука, для характеристики тембров музыкальных инструментов пользуются словесными описаниями, подобранными по признакам противоположности: яркий - тусклый, резкий - мягкий и др.

В научной литературе имеется большое количество понятий, связанных с оценкой тембров звука. Например, анализ терминов, принятых в современной технической литературе, позволил выявить наиболее часто встречающиеся термины, показанные в таблице. Были сделаны попытки выявить самые значимые среди них, и провести шкалирование тембра по противоположным признакам, а также связать словесное описание тембров с некоторыми акустическими параметрами (см. таблицу 6 в первой части этой статьи, "Звукорежиссер" 2/2001).

Таблица

Основные субъективные термины для описания тембра, используемые в современной международной технической литературе (статистический анализ 30 книг и журналов) Acidlike - кислый

forceful -
усиленный

muffled -
заглушенный

sober - трезвый
(рассудительный)

antique - старинный	frosty - морозный	mushy - пористый	soft - мягкий
arching - выпуклый	full - полный	mysterious - загадочный	solemn - торжественный
articulate - разборчивый	fuzzy - пушистый	nasal - носовой	solid - твердый
austere - суровый	gauzy - тонкий	neat - аккуратный	somber - мрачный
bite, biting - кусачий	gentle - нежный	neutral - нейтральный	sonorous - звучный
bland - вкрадчивый	ghostlike - призрачный	noble - благородный	steely - стальной
blaring - ревуший	glassy - стеклянный	nondescript - неопиcуемый	strained - натянутый
bleating - блеющий	glittering - блестящий	nostalgic - ностальгический	strident - скрипучий
breathy - дыхательный	gloomy - унылый	ominous - зловещий	stringent - стесненный
bright - яркий	grainy - зернистый	ordinary - ординарный	strong - сильный
brilliant - блестящий	grating - скрипучий	pale - бледный	stuffy - душный
brittle - подвижный	grave - серьезный	passionate - страстный	subdued - смягченный
buzzy - жужжащий	growly - рычащий	penetrating - проникающий	sultry - знойный
calm - спокойный	hard - жесткий	piercing - пронзительный	sweet - сладкий
carrying - полетный	harsh - грубый	pinched - ограниченный	tangy - запутанный
centered - концентрированный	haunting - преследующий	placid - безмятежный	tart - кислый
clangorous - звенящий	hazy - смутный	plaintive - заунывный	tearing - неистовый
clear, clarity - ясный	heartly - искренний	ponderous - увесистый	tender - нежный
cloudy - туманный	heavy - тяжелый	powerful - мощный	tense - напряженный
coarse - грубый	heroic - героический	prominent - выдающийся	thick - толстый
cold - холодный	hoarse - хриплый	pungent - едкий	thin - тонкий
colorful - красочный	hollow - пустой	pure - чистый	threatening - угрожающий
colorless - бесцветный	honking - гудящий (автомобильный гудок)	radiant - сияющий	throaty - хриплый
cool - прохладный	hooty - гудящий	raspy - дребезжащий	tragic - трагичный
crackling - трескучий	husky - сиплый	rattling - грохочущий	tranquil - успокаивающий
crashing - ломаный	incandescence -	reedy -	transparent -

	накаленный	пронзительный	прозрачный
creamy - сливочный	incisive - резкий	refined - рафинированный	triumphant - торжествующий
crystalline - кристаллический	inexpressive - невыразительный	remote - удаленный	tubby - бочкообразный
cutting - резкий	intense - интенсивный	rich - богатый	turbid - мутный
dark - темный	introspective - углубленный	ringing - звенящий	turgid - высокопарный
deep - глубокий	joyous - радостный	robust - грубый	unfocussed - несфокусированный
delicate - деликатный	languishing - печальный	rough - терпкий	unobtrusive - скромный
dense - плотный	light - светлый	rounded - круглый	veiled - завуалированный
diffuse - рассеянный	limpid - прозрачный	sandy - песочный	velvety - бархатистый
dismal - отдаленный	liquid - водянистый	savage - дикий	vibrant - вибрирующий
distant - отчетливый	loud - громкий	screamy - кричащий	vital - жизненный
dreamy - мечтательный	luminous - блестящий	sere - сухой	voluptuous - пышный(роскошный)
dry - сухой	lush (luscious) - сочный	serene, serenity - спокойный	wan - тусклый
dull - скучный	lyrical - лирический	shadowy - затененный	warm - теплый
earnest - серьезный	massive - массивный	sharp - резкий	watery - водянистый
ecstatic - экстатический	meditative - созерцательный	shimmer - дрожащий	weak - слабый
ethereal - эфирный	melancholy - меланхолический	shouting - кричащий	weighty - тяжеловесный
exotic - экзотический	mellow - мягкий	shrill - пронзительный	white - белый
expressive - выразительный	melodious - мелодичный	silky - шелковистый	windy - ветряный
fat - жирный	menacing - угрожающий	silvery - серебристый	wispy - тонкий
fierce - жесткий	metallic - металлический	singing - певучий	woody - деревянный
flabby - дряблый	misty - неясный	sinister - зловещий	yearning - тоскливый
focussed - сфокусированный	mournful - траурный	slack - расхлябанный	
forboding - отталкивающий	muddy - грязный	smooth - гладкий	

Однако, главная проблема состоит в том, что нет однозначного понимания различных субъективных

терминов, описывающих тембр. Приведенный в таблице перевод далеко не всегда соответствует тому техническому смыслу, которое вкладывается в каждое слово при описании различных аспектов оценки тембра.

В нашей литературе раньше был стандарт на основные термины, но сейчас дела обстоят совсем печально, поскольку не ведется работа по созданию соответствующей русскоязычной терминологии, и употребляется много терминов в разных, иногда прямо противоположных, значениях.

В связи с этим AES при разработке серии стандартов по субъективным оценкам качества аудиоаппаратуры, систем звукозаписи и др. начал приводить определения субъективных терминов в приложениях к стандартам, а так как стандарты создаются в рабочих группах, включающих ведущих специалистов разных стран, то эта очень важная процедура приводит к согласованному пониманию основных терминов для описания тембров.

В качестве примера приведу стандарт AES-20-96 - "Рекомендации для субъективной оценки громкоговорителей", - где дано согласованное определение таких терминов, как "открытость", "прозрачность", "ясность", "напряженность", "резкость" и др.

Если эта работа будет систематически продолжаться, то, возможно, основные термины для словесного описания тембров звуков различных инструментов и других звуковых источников будут иметь согласованные определения, и будут однозначно или достаточно близко пониматься специалистами разных стран. Мы постараемся информировать об этом наших читателей.