

УДК 534.3

СПЕКТРАЛЬНОЕ СОПОСТАВЛЕНИЕ МУЗЫКАЛЬНЫХ ПРОИЗВЕДЕНИЙ

А. Б. АНАНЬЕВ, Д. В. ПРОСВИРОВ

Национальный технический университет Украины “КПИ”

Получено 21.10.2004

Приведена методика расчета расстояния между спектрами музыкальных произведений в оркестровом исполнении, усредненными на длительном временном интервале. На примере выборки музыкальных произведений разных жанров показано, что такие характеристики отражают степень тембрового подобия исполняемой музыки. Полученные результаты могут служить объективной опорой при обсуждении сходства обобщенного звучания различных музыкальных произведений.

Наведено методику розрахунку відстані між спектрами музичних творів в оркестровому виконанні, осередненими на тривалому часовому інтервалі. На прикладі вибірки музичних творів різних жанрів показано, що такі характеристики відображають ступінь тембрової подібності музики, що виконується. Одержані результати можуть служити об'єктивною опорою при обговоренні подібності узагальненого звучання різних музичних творів.

A technique for calculation of distance between long-time averaged spectra of musical compositions performed by orchestra is presented. As it was shown on an example of a sample list formed by musical compositions belonging to different genres, such characteristics reflect the timbre similarity degree of the performed music. Obtained results may serve the objective support when arguing on generalized sound resemblance of different musical compositions.

ВВЕДЕНИЕ

Настоящая работа посвящена исследованию спектральных свойств таких весьма длительных и широкополосных звуковых сигналов как музыкальные произведения в оркестровом (ансамблевом) исполнении.

Несмотря на то, что вычисление спектра $F(\omega)$ сигнала $s(t)$ ($0 \leq t \leq T$) в принципе предполагает преобразование сигнала целиком,

$$F(\omega) = \frac{1}{T} \int_0^T s(t) e^{-j\omega t} dt, \quad (1)$$

обычная практика спектральной обработки музыкальных произведений состоит в вычислении так называемого кратковременного (short-time) или, по Харкевичу [1], “мгновенного” спектра

$$F(\omega; t) = \frac{1}{\tau} \int_t^{t+\tau} s(\nu) e^{-j\omega \nu} d\nu, \quad (2)$$

который отражает характер изменения во времени “частотного состава” сигнала.

Такие сменяющие друг друга во времени спектральные “срезы” присутствуют на пульте звукорежиссера. Графически их также принято отображать в виде трехмерного набора сечений рельефной поверхности или плоской спектрограммы, где уровень сигнала задается цветовой шкалой. Такая частотно-временная развертка позволяет визуально следить за развитием мелодии и

ритмикой музыкального произведения, одновременно оценивая тембровый характер звука в каждый отдельный момент.

Впрочем, иногда в литературе встречается понятие “спектр оркестра” [2, 3], под которым понимают long-term-average spectrum – усредненный на длительном временном интервале спектр

$$F(\omega) = \frac{1}{N} \sum_{k=1}^N \frac{1}{\tau} \int_{t_k}^{t_k+\tau} s(\nu) e^{-j\omega \nu} d\nu. \quad (3)$$

Интерес к исследованию спектров вида (3) сформировался в 20-е годы XX века, когда развитие электроакустической аппаратуры позволило строить измерительные тракты для оценки звуковой энергии в заданной полосе частот. Одна из первых фундаментальных работ в этом направлении [2] основывалась именно на применении третьоктавного аппаратного анализатора звука. С его помощью были проведены измерения распределения энергии по частоте для речевых сигналов, а затем выполнена большая серия аналогичных измерений как для отдельных инструментов симфонического оркестра, так и для музыкальных произведений, исполняемых оркестром в целом. При этом музыканты, участвующие в измерительных экспериментах с индивидуальными инструментами, исполняли на них произвольные характерные музыкальные фрагменты, а оркестры в процессе измерений даже могли использовать отрывки разных музыкальных пьес для различ-

ных анализируемых участков частотного диапазона.

Такой подход подразумевает наличие некоторого усредненного слухового впечатления, ассоциированного с представлением об оркестре. Однако мы понимаем, что в природе нет обобщенного оркестра, а существуют симфонические, камерные, джазовые оркестры, рок-ансамбли и студийные смешанные электронно-инструментальные ансамбли, исполняющие весьма различающиеся между собой музыкальные программы. Ситуация еще более усложняется, когда речь идет о музыке, записанной на разных физических носителях, имеющих специфические характеристики. И с каждым из перечисленных вариантов оркестрового звучания у опытного слушателя связано некоторое тембровое ощущение – нечто вроде музыкального “послевкусия”, которое позволяет высказываться о сходстве или различии звучания тех или иных пар музыкальных произведений в конкретных исполнениях.

Следует отметить, что подобные методики используются и для оценки качества звучания вокалистов в различных студиях звукозаписи [4]. Усредненные на больших временных интервалах спектры достаточно успешно применяются для анализа разборчивости речевых сигналов [5, 6] и выделения значимых диагностических признаков при электронной аускультации легких [7].

Поставленная нами задача состоит в том, чтобы определить, в какой мере результаты построения спектра в соответствии с формулой (3) могут служить объективной опорой при сопоставлении характера звучания различных записей музыкальных произведений.

1. МЕТОДИКА

Фактическому анализу мы подвергаем усредненный спектр амплитуд записи $\hat{F}(\omega) = |F(\omega)|$, который трактуем как вектор в n -мерном метрическом пространстве (n – количество дискрет для $\omega \geq 0$ в спектре, вычисляемом методом быстрого преобразования Фурье). Расстояние между спектрами зададим с помощью функционала

$$d(\hat{F}_x(\omega), \hat{F}_y(\omega)) = \sum_{i=1}^n |\hat{F}_x(\omega_i) - \hat{F}_y(\omega_i)|, \quad (4)$$

который, по нашему мнению, наиболее отвечает существу поставленной задачи.

Заметим, что вычисление функционала (4) предвзвешивается определенным “выравниванием” спектральных функций. Прежде всего мы обнуляем

все малозначимые компоненты спектра, уровень которых лежит ниже -60 дБ от пикового значения, а затем нормируем оставшиеся компоненты по величине площади под спектральной кривой.

Расстояние $d(\hat{F}_x(\omega), \hat{F}_y(\omega))$, вычисленное с применением указанных процедур, используется в качестве объективного критерия сравнения музыкальных записей.

2. ИССЛЕДУЕМЫЙ МАТЕРИАЛ

Для нахождения расстояния между спектрами были отобраны десять записей музыкальных произведений, как субъективно сходных по характеру звучания, так и взаимно контрастных (см. список).

- #1. И. С. Бах, Сюита номер 2 си минор, BWV 1067, длительность 20 мин. 41 сек., исполняет Штуттгартский камерный оркестр – широко известная сюита из семи произведений с характерным звучанием струнного камерного оркестра.
- #2. И. С. Бах, Overture из сюиты #1, длительность 6 мин. 45 сек. – отдельно подвергнутая анализу первая пьеса Сюиты.
- #3. И. С. Бах, Badinerie из сюиты #1, длительность 1 мин. 27 сек. – отдельно подвергнутая анализу весьма популярная седьмая пьеса сюиты с выразительным соло флейты на протяжении всего номера.
- #4. И. С. Бах, Концерт для двух скрипок с оркестром ре минор, BWV 1043, длительность 16 мин. 29 сек., исполняет оркестр Camerata Romana – близкое к #1 по камерному составу инструментов исполнение крупного произведения другим оркестром.
- #5. Р. Вагнер, вступление к опере “Нюрнбергские мейстерзингеры”, длительность 9 мин. 54 сек., исполняет Берлинский филармонический оркестр – произведение иного стиля, исполненное классическим составом симфонического оркестра.
- #6. Д. Гарланд, “In the mood”, длительность 3 мин. 17 сек., исполняет биг-бэнд Глена Миллера, записана в декабре 1939 в радиостудии Нью-Йорка, перезаписана со старого носителя в 1997. Пьеса широко известна по фильму “Серенада Солнечной Долины”, принадлежит к раритетным образцам звучания джазового биг-бэнда.

#7. К. Портер, “Begin the beguin”, длительность 3 мин. 9 сек., исполняет вокальное трио “The Andrews Sisters” в сопровождении оркестра Гленна Миллера, записана в радиостудии в 1940 г., перезаписана в 1997. Пьеса представляет собой тембровую смесь биг-бенда и женского вокального трио.

#8. Ч. Вил, “Bad for me”, длительность 5 мин. 35 сек., исполняет Ди Ди Бриджутер в сопровождении инструментального ансамбля – пример студийной записи высококачественной танцевальной пьесы, которую исполняет известная джазовая вокалистка на звуковой основе, созданной электронным синтезатором, бэк-вокалом, квартетом духовых инструментов и ритм-группой.

#9. Р. Суржа, “Пароль”, длительность 4 мин. 30 сек., исполняет Самая-Т – насыщенная эффектами высококачественная отечественная студийная композиция.

#10. Д. Гарланд, “In the mood”, длительность 3 мин. 18 сек., исполняет оркестр Хуго Штрайсера – римейк пьесы #6, практически совпадающий с ней по аранжировке и длительности, исполненный современным биг-бендом и записанный в современной студии.

Все записи представлены для анализа на компакт-аудио носителях в формате 44100 Гц, 16 бит.

3. ТЕХНОЛОГИЯ

Для всех перечисленных музыкальных произведений спектр (3) вычислялся методом быстрого преобразования Фурье при 32768 отсчетах, т.е. с разрешением по частоте 1.3 Гц. Характерный вид спектральной функции показан на рис. 1 на примере пьесы Баха #1. Ясно, что получаемым при этом результатам присущ естественный разброс значений спектральной функции в соседних частотных каналах, который может заметно влиять на вычисляемую величину расстояния $d(\hat{F}_x(\omega), \hat{F}_y(\omega))$. Поэтому для каждой спектральной функции $F(\omega)$ выполнялось сглаживание скользящим полуоктавным усредняющим фильтром. Это приводило к существенному выравниванию близлежащих спектральных данных (см. рис. 1), делая их пригодными для реализации методики, описанной в предыдущем разделе.

На рис. 2 представлены две спектральные функции, которые дают исходные данные для вычисления расстояния по формуле (4). Очевидна суще-

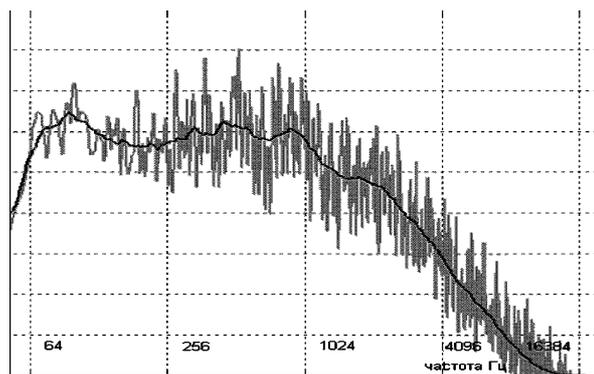


Рис. 1. Характерный вид спектральной функции для произведения #1 без сглаживания и с полуоктавным сглаживанием

ственная близость между спектрами двадцатиминутной сюиты, #1 (кривая 1), состоящей из семи разнохарактерных пьес, и полтора минутной пьесы из этой сюиты, #3 (кривая 2), казалось бы, абсолютно заполненной звуком солирующей флейты. При сравнении спектров следует иметь в виду, что на рисунках и частота, и амплитудный уровень отображены в логарифмической шкале. Мы не маркируем ось амплитуд, поскольку сетка на ней соответствует относительным значениям уровней амплитуды (в дБ), полученным после их нормировки по площади под кривой. Важно, что все используемые спектральные данные получены при одних и тех же параметрах обработки. И поскольку в пространстве спектральных векторов не задана норма (отсутствует единичный вектор), то существенным является лишь отношение “дальше/ближе” между парами спектральных кривых.

На рис. 3 показаны формы спектров одной и той же оркестровой пьесы, доступной по архивной записи оркестра Глена Миллера, #6 (кривая 1), и в записи современного оркестра, #10 (кривая 2). Очевидна существенная разница между этими кривыми даже в привычных по технической литературе логарифмических шкалах, однако суммирование в формуле (4) оперирует линейно изменяющейся частотой. Поэтому приводим иллюстративные аналоги рис. 2 и 3 на рис. 4 и 5 соответственно при линейной шкале частоты, усиливающей это впечатление.

Возвратимся к логарифмическим шкалам частоты. На рис. 6 приведено сочетание спектров сюиты Баха #1 (кривая 1) и оркестрового произведения Р. Вагнера #5 (кривая 2), а на рис. 7 – спектры весьма контрастных по звучанию сюиты #1 и современной студийной аранжировки #9 (кривая 2). Что касается спектров для пьес в современ-

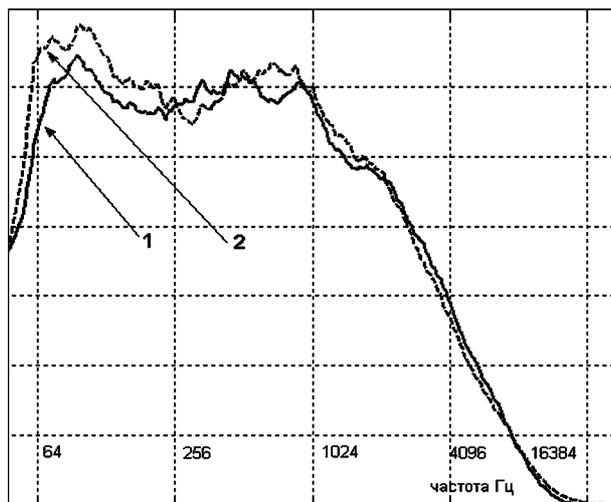


Рис. 2. Спектры произведений #1 (1) и #3 (2), логарифмический масштаб частоты

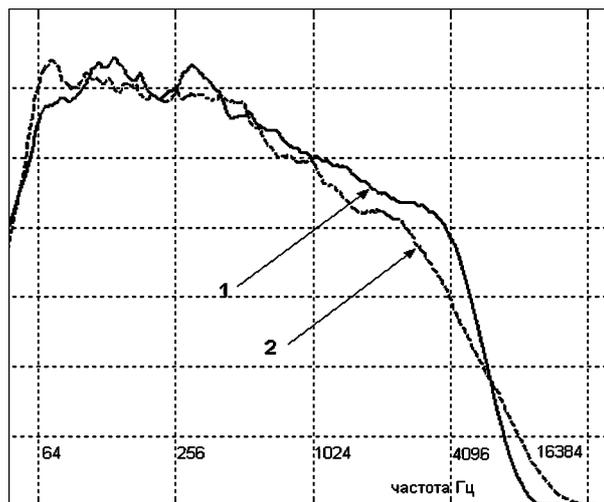


Рис. 3. Спектры произведений #6 (1) и #10 (2), логарифмический масштаб частоты

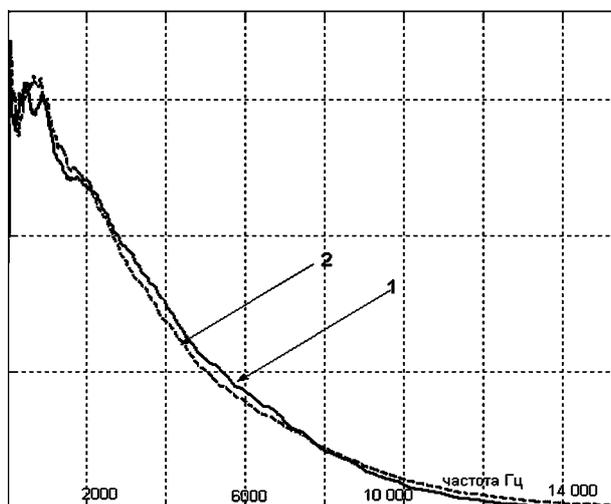


Рис. 4. Спектры произведений #1 (1) и #3 (2), линейный масштаб частоты

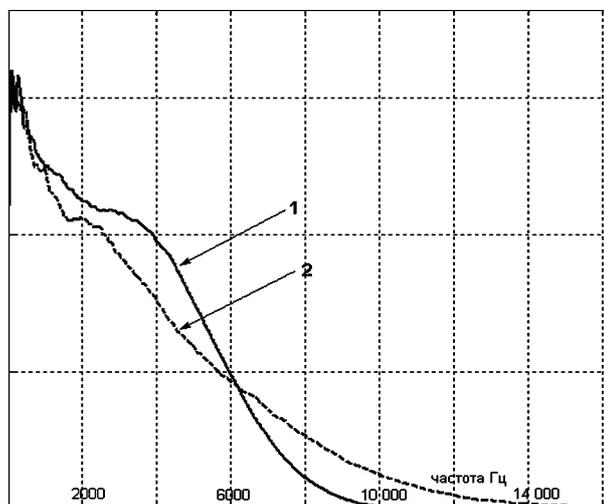


Рис. 5. Спектры произведений #6 (1) и #10 (2), линейный масштаб частоты

ном исполнении и студийной записи, то они также существенно различаются (здесь кривая 1 соответствует биг-бенду #10, а кривая 2 – аранжировке #9, насыщенной электронными инструментами и эффектами, на рис. 8 частота дана в логарифмическом, а на рис. 9 – в линейном масштабе). Указанная разница отражает значимые слуховые тембровые различия за счет перераспределения энергии в пьесе #9 в область крайних низких и крайних высоких частот.

4. РЕЗУЛЬТАТЫ

Таблица содержит результаты расчета расстояний $d(\hat{F}_x(\omega), \hat{F}_y(\omega))$ между спектрами по формуле (4) для всех записей музыкальных произведений

из исследованной выборки. Рассмотрим некоторые из спектральных пар более подробно.

Например, расстояние между спектрами пьес #1 и #3 (см. рис. 2) прочитывается в строке 1, колонке 3 и составляет 0.068. Еще раз напомним, что это число нельзя как-либо интерпретировать само по себе (поскольку в выбранном метрическом пространстве норма нами не определена). Однако видно, что расстояние между спектрами записей #1 и #5 (см. рис. 6) составляет 0.081, в то время как между спектрами #9 и #10 (см. рис. 8) – 0.548. Если к этому добавить, что спектры пьес #9 и #1 (см. рис. 7) “удалены” на 0.636, то можно высказать суждение, что в тембровом отношении симфоническое исполнение Вагнеровской пьесы, в общем, должно создавать впечатление не

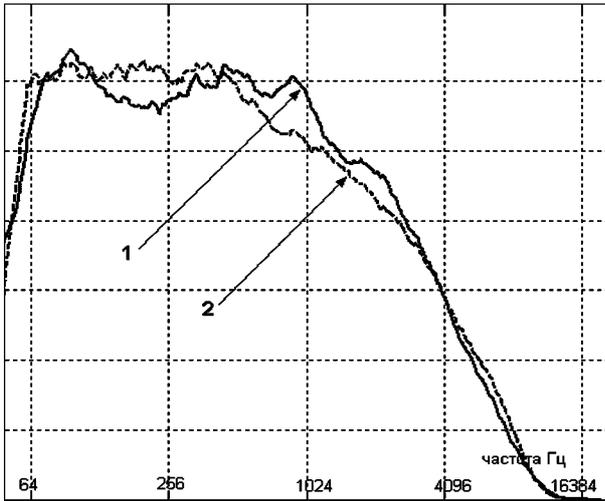


Рис. 6. Спектры произведений #1 (1) и #5 (2), логарифмический масштаб частоты

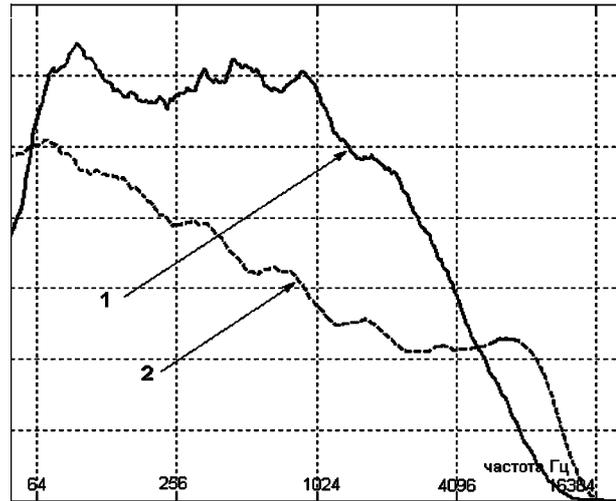


Рис. 7. Спектры произведений #1 (1) и #9 (2), логарифмический масштаб частоты

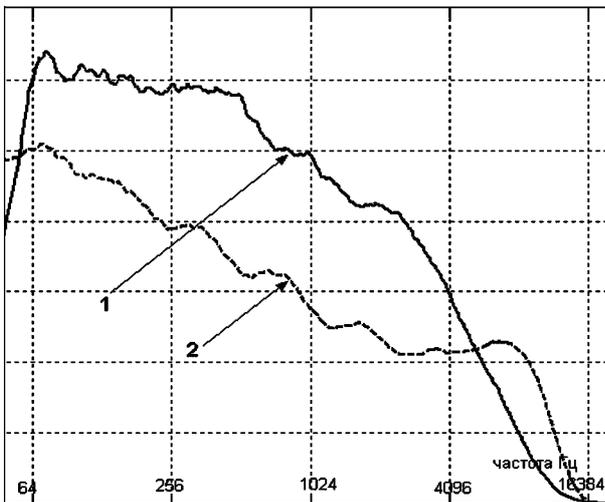


Рис. 8. Спектры произведений #9 (1) и #10 (2), логарифмический масштаб частоты

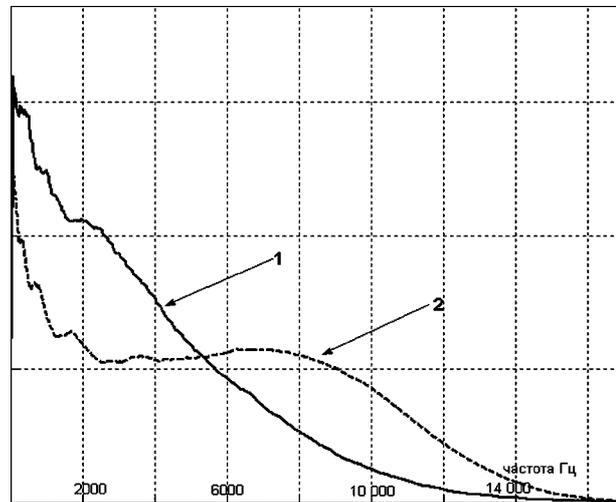


Рис. 9. Спектры произведений #9 (1) и #10 (2), линейный масштаб частоты

очень отличающееся от камерного исполнения Баховской сюиты. В то же время, не только разница впечатлений между камерной сюитой #1 и синтетической композицией #9 ожидается существенной, но даже разница между упомянутой композицией и звучанием биг-бенда #10 примерно столь же велика. Но наибольшего тембрового расхождения можно ожидать при сравнении архивных записей оркестра Гленна Миллера #6, #7 и записи ансамбля с участием Ди Ди Бриджуотер #8.

Полученные результаты также представлены на рис. 10 в виде графика. На горизонтальной оси отложен номер музыкального произведения n , а на вертикальной – величина расстояния между спе-

ктрами. Каждая из линий на этом рисунке отвечает одному из произведений и проходит на горизонтальной оси через 0 при n , соответствующем номеру произведения, для которого эта линия построена (очевидно, что удаление спектра от самого себя равно нулю). При остальных n эта линия указывает удаление данного спектра от спектров других музыкальных записей из исследованной выборки.

Видно, что общая картина “распадается” на три пучка линий. Первый (штриховые линии) объединяет пьесы #1, #2, #3, #4, #5 и #10. Таким образом, симфоническое и камерное звучание оказывается весьма близким для различных пьес и оркестров, и даже относительно близким к звучанию биг-бенда. Архивные записи #6, #7 (точечные ли-

Таблица. Расстояния $d(\hat{F}_x(\omega), \hat{F}_y(\omega))$ между спектрами музыкальных произведений

#		1	2	3	4	5	6	7	8	9	10
1	Бах, Сюита номер 2	0.000	0.048	0.068	0.134	0.081	0.189	0.188	0.823	0.636	0.106
2	Бах, Overture		0.000	0.082	0.088	0.066	0.194	0.204	0.781	0.594	0.065
3	Бах, Badinerie			0.000	0.149	0.125	0.245	0.236	0.791	0.624	0.124
4	Бах, Концерт для двух скрипок				0.000	0.079	0.237	0.257	0.729	0.522	0.032
5	Вагнер, Вступление к опере					0.000	0.187	0.215	0.796	0.584	0.057
6	Гарланд, "In the mood" (биг-бенд Глена Миллера)						0.000	0.058	0.954	0.752	0.215
7	Портер, "Begin the beguin" (биг-бенд Глена Миллера и вок. трио)							0.000	0.966	0.772	0.243
8	Вил, "Bad for me" (Ди Ди Бриджутер с ансамблем)								0.000	0.432	0.751
9	Суржа, "Пароль" (Самая-Т)									0.000	0.548
10	Гарланд, "In the mood" (биг-бенд Хуго Штрассера)										0.000

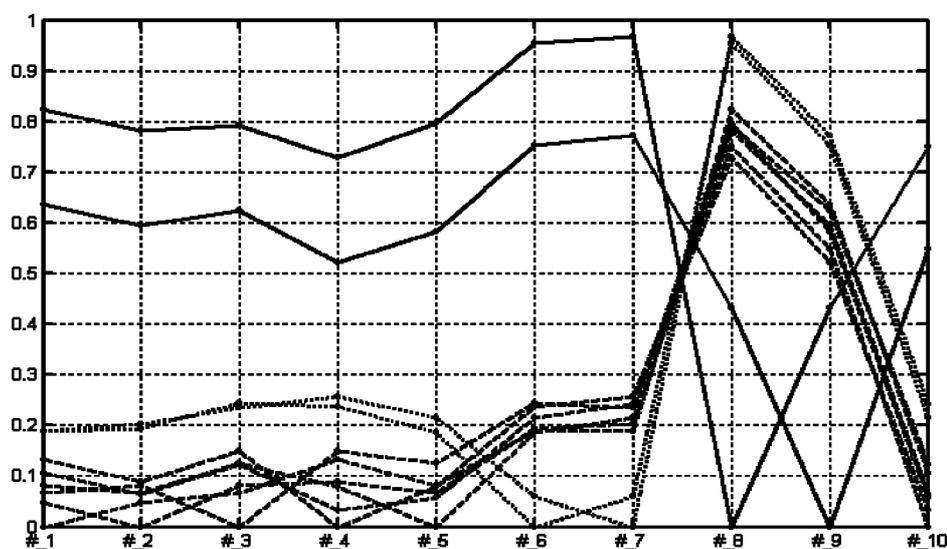


Рис. 10. Графическое представление данных таблицы

нии) в тембровом отношении мало различимы между собой, несмотря на наличие в #7 вокального трио, однако заметно отличаются от всех других образцов. Наконец, пьесы #8 и #9 (сплошные линии), имеющие современную звуковую окраску, не только резко отличаются от всей остальной музыки, но и разнятся между собой, отражая различия в использованных инструментальных средствах.

Нужно сказать, что полученные здесь данные, в общем, подтвердили слуховые впечатления о перечисленных пьесах, которые сложились в кругу исследователей данного вопроса до начала исследования, однако при этом позволили более обоснованно относиться к субъективным оценкам в некоторых спорных случаях.

ВЫВОДЫ

Выбранная методика вычисления расстояний между усредненными спектрами музыкальных произведений дает результаты, отвечающие слуховым впечатлениям при субъективном сравнении этих произведений. Получаемые таким способом данные могут служить объективной опорой при обсуждении относительного сходства или различия в обобщенном звучании произведений.

Следует, однако, указать, что, по нашему мнению, результаты вычислений такого рода, в принципе, не предназначены для оценки качества звучания музыкальных произведений, а призваны лишь вносить ясность при наличии спорных суждений.

1. Харкевич А. А. Спектры и анализ.– М.: ГИ физ.-мат. лит, 1962.– 234 с.
2. Sivian L. J., Dunn H. K., White S. D. Absolute amplitudes and spectra of certain musical instruments and orchestras // J. Acoust. Soc. Amer.– 2, N 3.– 1931.– P. 330–371.
3. Borch D. Z., Sundberg J. Spectral distribution of solo voice and accompaniment in pop music // Speech, Music and Hearing, Quart. Progr. Status Rept (КТН, Stockholm).– 2002.– 43.– P. 31–35.
4. Barnes J. J., Davis P., Oates J. Chapman J. The relationship between professional operatic soprano voice and high range spectral energy // J. Acoust. Soc. Amer.– 2004.– 116, N 1.– P. 530–538.
5. Weinberg B., Horii Y., Smith B. E. Long-time spectral and intensity characteristics of esophageal speech // J. Acoust. Soc. Amer.– 1980.– 67, N 5.– P. 1781–1784.
6. Byrne D. et al. An international comparison of long-term average speech spectra // J. Acoust. Soc. Amer.– 1994.– 96, No. 4.– P. 2108–2120.
7. Вовк И. В., Дахнов С. Л., Крижановский В. В., Олийник В. Н. Возможности и перспективы диагностики легочных патологий с помощью компьютерной регистрации и обработки шумов дыхания // Акуст. вісн.– 1998.– 1, N 2.– С. 21–33.