

РАЗРАБОТКА КОНЦЕПТУАЛЬНОЙ МОДЕЛИ ГАРМОНИЗАЦИИ ЗВУКОВОГО И ВИЗУАЛЬНОГО РЯДА

Гарифуллин Ф.А.¹, Шкляр А.В.²

¹ТПУ ИШИТР ОАР, гр.8ИМ42, e-mail: fag6@tpu.ru

²ТПУ ИШИТР ОАР., к.т.н. доцент ОАР ИШИТР, e-mail: shklyarav@tpu.ru

Аннотация

В данной статье рассматривается вывод концептуальной модели гармонизации на основе кроссмодальных соответствий, а именно измеримых параметров аудио и видео составляющих. Исследование базируется на основе экспериментальных данных о соответствии звучания и визуального образа.

Ключевые слова: кроссмодальные соответствия, аудиовизуальное восприятие, коэффициент гармоничности.

Введение

Современные мультимедийные проекты объединяют различные модальности – визуальные, звуковые, иногда тактильные и др. От того, насколько согласованно и гармонично сочетаются между собой эти элементы, зависит целостность пользовательского восприятия. Дизайн, учитывающий многосенсорный опыт, должен обеспечивать оптимальное согласование стимулов разных модальностей. Иными словами, необходима **кросс-модальная гармонизация** – приведение компонентов разных чувств (например, изображения и звука) к единому выразительному замыслу. Исследования в области восприятия показывают, что соответствие между стимулами разных модальностей может улучшать когнитивную обработку: эффект кроссмодального соответствия проявляется в более быстром и точном распознавании при согласованном сочетании стимулов. В контексте цифрового дизайна это означает, что продуманная комбинация визуальных и звуковых решений способна сделать взаимодействие пользователя более интуитивным и приятным.

Понятие гармонии [1] исторически применялось в музыке и изобразительном искусстве, а позднее стало использоваться и в межсенсорном контексте, так, еще в 1940 г. мультфильм «*Fantasia*» студии Disney продемонстрировал возможность синхронного сочетания анимации и музыки, которое было воспринято как единое гармоничное целое. В современную эпоху мультимедиа задача одновременного воздействия на несколько чувств актуальна там, где важно единообразие аудиовизуального стиля. Однако достижение такой гармонии на практике сопряжено с рядом проблем, так как проектирование мультимедийных продуктов часто разделено по модальностям: визуальные элементы создаются отдельно от звуковых, что может приводить к разобщенности восприятия. Отсутствие общей «метрики» или правил для согласования цвета, формы, звука и других чувств затрудняет работу дизайнера: гармония между изображением и музыкой носит во многом интуитивный характер и опирается на субъективное мнение. Это создает риск, что при совмещении проработанных по отдельности компонентов итоговое впечатление зрителя окажется несвязным. Проблема усугубляется тем, что восприятие каждой модальности имеет свою специфику: человеческий глаз и ухо обрабатывают информацию по-разному, и прямого соответствия, подобного музыкальной гармонии (созвучие тонов), между цветами и звуками не существует в явном виде. Тем не менее, психологические исследования выявили целый ряд **кроссмодальных соответствий**. За последние десятилетия накоплено множество научных данных о таких соответствиях, но до сих пор они слабо формализованы именно в прикладных инструментах дизайна. Таким образом, существует разрыв между знаниями о восприятии и практикой мультимедийного проектирования. И выстраивается гипотеза о том, что измеримые параметры музыки (темп, громкость, тональность, спектр) могут быть сопоставлены с параметрами визуала (динамика

монтажа, колористика, форма и композиция) так, что при их согласованности возникает ощущение целостной гармонии, а при намеренном, но осмысленном контрасте – создаётся необходимый художественный эффект. В рамках данной работы будет рассмотрен процесс реализации прототипа модели на основе известных кроссмодальных соответствий. Результатом работы является визуализация модели в качестве таблицы соответствий и формулы вычисления коэффициента гармоничности кадра на основе измеримых аудио и видео параметров.

Основная часть

Модель кроссмодальной гармонизации представляет собой трёхкомпонентную структуру. На первом этапе создаётся список количественно измеримых критериев (аудио- и видеохарактеристик), а также соответствующих им эмоциональных состояний [2]. Для каждого параметра определяется диапазон значений (например, для яркости, громкости, темпа и т. д.) с привязкой к описанным настроениям (таблица 1). Это формализует эмоциональную составляющую, делая её более измеримой. В таблице представлены несколько основных значений из основной модели для визуализации параметров.

Таблица 1. Параметры и диапазоны значений

Параметр	Диапазон / Условие	Категория восприятия	Ассоциируемое настроение
Темп (BPM)	< 80	Медленный	Спокойствие, меланхолия
Темп (BPM)	> 130	Быстрый	Бодрость, возбуждение
Частота звука (Гц)	< 500	Низкая	Массивность, основательность
Частота звука (Гц)	> 2000	Высокая	Лёгкость, тревожность
Яркость (L*, V)	> 80%	Яркое	Радость, открытость
Яркость (L*, V)	< 20%	Тёмное	Уныние, напряжённость

Второй частью исследования является сопоставление критериев на основе устойчивых кроссмодальных корреляций [3]. В рамках данной структуры каждой паре соответствий были заданы коэффициенты гармоничности на основе статистических, культурных и гипотетических доказательств (таблица 2). Значения коэффициентов задаются в интервале от 0 до 1, отражая силу и достоверность взаимосвязи, при этом учитываются как статистически подтверждённые, так и реже встречающиеся корреляции, возникающие из-за культурных или индивидуальных особенностей.

Таблица 2. Кроссмодальные соответствия

Параметры	Диапазоны значений	Настроение	Тип связи	Коэффициент достоверности
Частота звука – Яркость	500–2000 Гц : 0.2– 0.8 яркости	Светлое / активное	Статистическая	0.9
Громкость – Яркость	-20 до -10 LUFS ↔ яркость > 0.5	Интенсивность / сила	Структурная	0.8
Тональность – Цветовой тон	Мажор ↔ теплые, Минор ↔ холодные	Радость / грусть	Семантическая	0.75
Темп – Яркость	60–80 BPM ↔ низкая, 120+ ↔ высокая	Медитативность / возбуждение	Гипотетическая	0.55

Дополнительно в рамках данной части модели была разработана визуализация таблицы связей в формате трехмерного графа. Для визуализации были использованы библиотеки python. Результат визуализации представлен на рисунке 1.

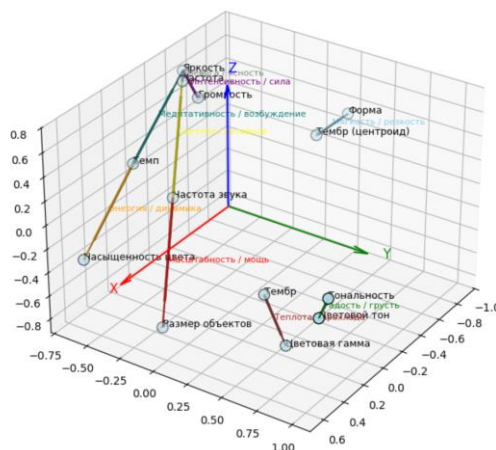


Рис. 1. Визуализация кроссмодальных связей

Третьей частью модели является общая формула для выявления итогового коэффициента гармоничности мультимедийного кадра, учитывающая соответствия, выстроенные выше. Формула опирается на взвешенный учёт параметров и коэффициентов достоверности [4]. В упрощённом виде её можно представить как сумму произведений нормированных значений параметров на соответствующие веса и коэффициенты, однако в практической реализации допустимы более сложные функции, учитывающие нелинейные эффекты или перекрёстные взаимодействия.

$$H = \frac{\sum_{i=1}^n w_i * m_i}{\sum_{i=1}^n w_i} \quad (1)$$

где H – итоговый коэффициент гармоничности (0 до 1), m_i – значение соответствия между параметрами (0 до 1), w_i – вес достоверности соответствия.

Полученный индекс гармоничности служит ориентиром при анализе эмоционально-эстетической целостности мультимедийного проекта, позволяя увидеть, насколько органично сочетаются звук и визуальный ряд или, напротив, где возникает диссонанс

Таким образом, трёхкомпонентная структура модели (список критериев и диапазонов, система кроссмодальных соответствий и формула расчёта коэффициента) предоставляет базу для более объективного анализа и проектирования гармоничного аудиовизуального контента. В дальнейшем модель может быть доработана путём расширения набора параметров и уточнения весовых коэффициентов на основе дополнительного эмпирического материала, что позволит ещё более точно предсказывать и корректировать степень согласованности различных компонентов мультимедийного произведения.

Заключение

В данной работе предложена модель, позволяющая системно подходить к задаче гармонизации разномодальных компонентов мультимедийных проектов. Модель основана на научных представлениях о кроссмодальных соответствиях и предоставляет практический инструмент для планирования и оценки мультимедийного дизайна. Применение модели дает возможность *предварительного проектирования ощущений* пользователя: еще на стадии разработки можно проверить, насколько согласованно будут работать вместе, например, визуальная анимация и звуковой фон.

Модель кроссмодальной гармонизации интегрирует результаты из психологии и нейроэстетики в понятную форму. Кроме того, требуется дальнейшая проверка модели в экспериментальных условиях: проведение пользовательских тестирований, измерение восприятия гармоничности и коррекция связей графа под конкретные аудитории (учитывая культурные различия в ассоциациях). Тем не менее, уже сейчас модель может найти применение в образовательном процессе (для обучения дизайнеров основам мультисенсорного восприятия) и в практике креативных индустрий. Она предоставляет общий язык для обсуждения решений звукорежиссеров и художников-графиков, облегчая коммуникацию в междисциплинарных командах. В конечном счете, кросс-модальная гармонизация стремится сделать цифровой опыт более человечным, интуитивно цельным. Как образно заметил М. Хаверкамп [5], идеальный синестетический дизайн сочетает общие для всех людей ощущения с индивидуальным творческим видением автора, достигая подлинной глубины воздействия. Предложенная модель является шагом к реализации этой идеальной цели – созданию мультимедийных проектов, где все элементы «звучат» согласованно и резонируют с пользователем на разных уровнях чувств.

Список используемых источников

1. Spence C., Di Stefano N. Crossmodal Harmony: Looking for the Meaning of Harmony Beyond Hearing. *i-Perception*, 13(1), 1–40 (2022). DOI: 10.1177/20416695211073817 (Обзор концепции гармонии в кросс-модальном контексте, примеры «цветомузыки»). [Электронный ресурс]. – URL: ncbi.nlm.nih.gov/pmc/articles/PMC8931439/ (дата обращения: 01.01.2025)
2. Giannos K., Athanasopoulos G., Cambouropoulos E. Cross-Modal Associations Between Harmonic Dissonance and Visual Roughness. *Music & Science*. – 2021. – Vol. – P. 1-15. DOI: 10.1177/20592043211055484 [Электронный ресурс]. – URL: journals.sagepub.com/doi/10.1177/20592043211055484 (дата обращения: 10.02.2025)
3. Полное руководство по распознаванию и работе с эмоциями [Электронный ресурс]. – URL: training-oz.com/polnoe-rukovodstvo-po-raspoznavaniyu-i-rabote-s-emotsiyami-chast-1/ (дата обращения: 11.02.2025)
4. Андриющенко Е.А. На повышенных тонах: роль пространственного познания в кросс-модальном взаимодействии эмоциональной семантики и аудиального восприятия / Е.А. Андриющенко, Е. Н. Блинова, Ю. Ю. Штыров [и др.] // Психология. Журнал Высшей школы экономики. – 2022. – Т. 19, № 4. – С. 736-756.
5. Хаверкамп М. Интервью о кросс-сенсорном дизайне и синестезии восприятия. Роль синестезии в дизайне; принцип сочетания общечеловеческих и индивидуальных восприятий [Электронный ресурс]. – URL: psyjournals.ru/synesthesia (дата обращения: 20.02.2025).