

ИНТЕРАКТИВНАЯ МУЗЫКА В ИГРАХ: РАЗРАБОТКА СИСТЕМ ДИНАМИЧЕСКОЙ МУЗЫКИ, КОТОРАЯ АДАПТИРУЕТСЯ К ДЕЙСТВИЯМ ИГРОКА И ИЗМЕНЯЕТСЯ В ЗАВИСИМОСТИ ОТ ИГРОВОГО ПРОЦЕССА.

А. А. Бондарук Я. Р. Сахацкая

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Рябычина О. П. – канд. тех. наук

В данной статье рассматриваются алгоритмы для разработки систем динамической музыки.

Современные видеоигры представляют собой не только развлекательные продукты, но и сложные мультимедийные системы, в которых музыка играет критическую роль в создании иммерсивного опыта. Динамическая музыка, адаптирующаяся к игровому процессу, стала ключевым элементом взаимодействия игрока с виртуальной средой. Она используется в таких популярных играх как Hades, Red Dead Redemption 2, No Man's Sky". Несмотря на это, аудиодизайн нередко включается на поздних стадиях разработки, что снижает потенциал интеграции звука с геймплеем и визуальной составляющей.

Для работы подобных систем используются различные алгоритмы адаптации музыкального сопровождения к действиям игрока и изменению игрового процесса. Алгоритмы можно разделить на две группы: адаптивные и генеративные

Адаптивная музыка является наиболее распространенной техникой для создания интерактивных музыкальных систем. Это техника, которая динамически изменяет музыку, используя заранее подготовленные аудиофрагменты. Она может добавлять или убирать инструменты, менять темп, тон и эффекты, создавая ощущение интерактивности. Расширяя рамки адаптивной музыки, важно изучить концепции горизонтального и вертикального наслоения, алгоритмы с использованием эмоциональных моделей и модульную композицию.

Горизонтальное перетекание реализуется через смену музыкальных фрагментов в зависимости от событий игры. Вертикальная структура работает с музыкальными слоями, которые включаются и выключаются динамически. Вектор состояния можно описать как

$$S = (s_1, s_2, \dots, s_n), s_i \in [0, 1] \quad (1)$$

Он отражает громкость или интенсивность слоя. Изменения происходят по формуле сглаживания:

$$s_i(t) = \alpha \cdot s_i(t-1) + (1 - \alpha) \cdot s_i^{\text{target}}, \alpha \in (0, 1) \quad (2)$$

Алгоритмы с использованием эмоциональных модулей представляют собой модели, описывающие эмоции в виде координат в многомерном пространстве. Чаще всего — это двумерная модель Валентности и Активации (Valence-Arousal), иногда с добавлением напряжения (Tension) как третьей оси. Музыкальные переходы происходят на основе изменений в эмоциональном состоянии игрока.

Алгоритм модульной композиции — это алгоритм при которых музыка разбивается на небольшие фрагменты (модули), а затем эти модули комбинируются в реальном времени в зависимости от происходящего в игре.

Генеративные алгоритмы создают музыку в реальном времени при помощи стохастических процессов, создают музыку в реальном времени на основе параметров игрового процесса. К таким алгоритмам относятся Марковские цепи (Markov Chains), Эволюционные алгоритмы, Фрактальные алгоритмы и Нейросети (особенно RNN, LSTM)

Марковские цепи — это генерация последовательностей нот, основанных на вероятности перехода от одной ноты к другой. Музыкальная последовательность в марковской модели определяется множеством состояний и матрицей вероятностей переходов:

$$M = \{m_1, \dots, m_n\}, P(m_i \rightarrow m_j) \quad (3)$$

Такие подходы позволяют создавать уникальные вариации, однако страдают от трудностей в передаче нарратива и высокой вычислительной сложности.

Эволюционная модель (или эволюционные алгоритмы) — это класс алгоритмов машинного обучения и оптимизации, которые имитируют процесс естественного отбора. В контексте генеративной

музыки — это значит, что алгоритм "создает" множество музыкальных вариантов, а затем отбирает и улучшает лучшие, исходя из заданных критериев.

Нейросети, особенно рекуррентные (RNN) и их усовершенствованные варианты, такие как LSTM (Long Short-Term Memory), играют важную роль в разработке систем или генеративной музыки. Модели обучаются на нотных данных и могут генерировать музыку с учётом временной зависимости.

Системы на базе машинного обучения могут учитывать игровые параметры и сопоставлять их с эмоциональным откликом. В работе Dorsey (2023) представлена уникальная архитектура, где «дирижёр» (человек-наблюдатель) в реальном времени задаёт эмоциональные метки (интенсивность и чувство победы) в процессе игрового тестирования. Игровые параметры — такие как количество объектов, скорость, позиции и счёт — фиксируются синхронно и затем объединяются по временным меткам. Обучающая модель (включая Random Forest и линейную регрессию) аппроксимирует функцию

$$f: X \rightarrow Y \quad (4),$$

где $X \in \mathbb{R}^n$ — вектор состояния игры; $Y \in [0,1]^2$ — предсказанные эмоциональные реакции.

Лучший результат (RMSE = 0.199 по интенсивности) достигнут при использовании Random Forest с задержкой временной синхронизации в 250 мс.

RMSE показывает, насколько сильно предсказанные значения отклоняются от реальных. Это делается следующим образом: вычисляется разница между предсказанным и реальным значением, эта разница возводится в квадрат (чтобы избежать отрицательных значений), полученные квадраты усредняются.

Практическая реализация адаптивной музыки возможна через middleware-решения — специальные аудиодвижки. К числу наиболее популярных относятся FMOD Studio, Audiokinetic Wwise и iMUSE (Interactive Music Streaming Engine), последний исторически использовался в играх LucasArts. FMOD предлагает композиторам гибкий DAW-подобный интерфейс, позволяющий работать с событиями, параметрическими переходами, логикой случайного выбора слоёв и условиями активации. Wwise предоставляет возможности по интеграции звука с физикой, пространственным размещением и объёмной обработкой. iMUSE, хотя и менее распространён сейчас, был пионером в области плавных музыкальных переходов на основе игровых сценариев.

Современные методы адаптивной и генеративной аудиосистем позволяют динамически подстраивать музыкальное сопровождение под игровой процесс, используя алгоритмы управления слоями, эмоциональные модели и нейросетевые подходы. Несмотря на наличие технологических решений и middleware-инструментов, на практике аудиодизайн нередко интегрируется слишком поздно, что снижает его потенциал. Для максимального эффекта требуется более тесная связка аудиосистем с игровой логикой на ранних этапах разработки. Интеграция машинного обучения и мультимодальных моделей открывает перспективы для более глубокой адаптации музыки к состоянию игрока и контексту сцены, что делает звук полноправным элементом интерактивного повествования. Метрики, такие как RMSE, позволяют оценивать точность соответствия музыки эмоциональному состоянию игрока, а middleware-системы (FMOD, Wwise) обеспечивают гибкую и эффективную интеграцию звука в игровой процесс.

Список использованных источников:

1. Varela, D. *Developing an Adaptive Music System for Video Games: Enhancing Immersion and Engagement Through Interactive Music* / D. Varela // *Master's Thesis in Multimedia. Oporto University, 2024.* — 97 p.
2. Smith, L.F. *The Effect of Dynamic Music in Video Games: An Overview of Current Research* / L.F. Smith // *Bachelor's Thesis in Game Design. Uppsala University, 2020.* — 30 p.
3. Dorsey, A., Neller, T.W., Tran, H.G., Yilmaz, V. *Learning Adaptive Game Soundtrack Control* / A. Dorsey [et al.] // *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence (AAAI-23), 2023.* — P. 16070–16074