

Метод генерации музыкального фрагмента, соответствующего эмоциональному состоянию человека, с использованием марковских моделей

Студент: Фролова Елена Владимировна ИУ7-81Б
Научный руководитель: Гаврилова Юлия Михайловна

Москва, 2023

Рост спроса на фоновую музыку

Согласно отчету **IFPI** о мировом музыкальном рынке за 2022 год:

- доходы от разнообразной музыки в цифровом формате выросли на 11,5% относительно 2021 года;
- доходы от создания фоновой музыки за 2022 год составили 67% от общих доходов от музыки.

Фоновая музыка активно используется для:

- поддержания определенного ритма:
 - спорт, например бег, йога, велоспорт;
 - концентрация внимания;
 - создание рабочей атмосферы;
- удержание и управление вниманием пользователей:
 - музыка в видеоиграх;
 - музыка в фильмах.

Цель и задачи работы

Цель: разработать метод генерации музыкального фрагмента, соответствующего эмоциональному состоянию человека, с использованием марковских моделей.

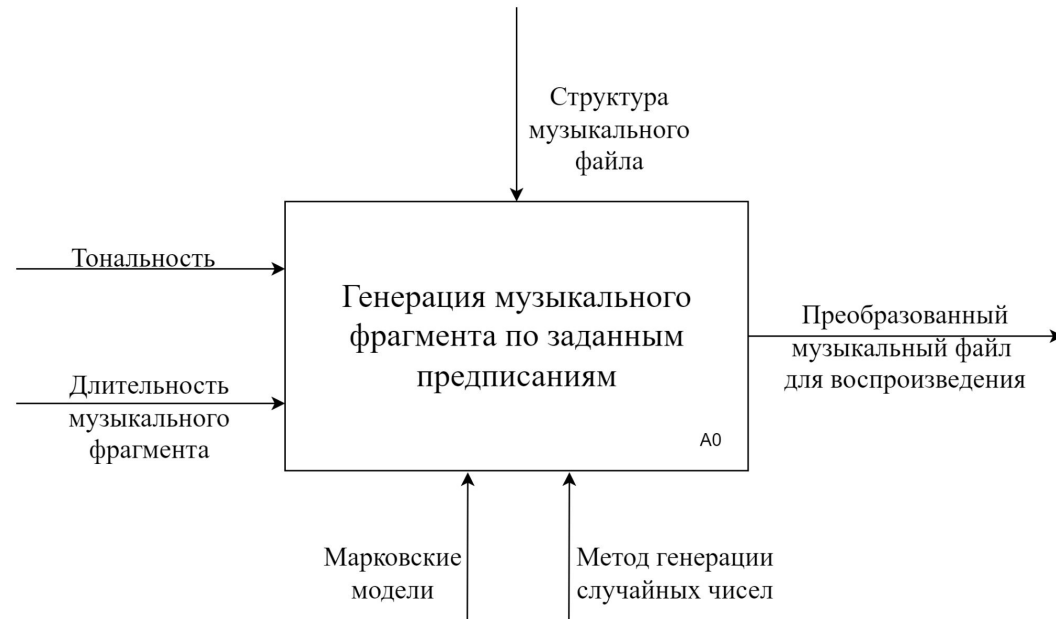
Задачи:

- провести анализ предметной области генерации музыкальной композиции и существующих методов машинной генерации музыки;
- разработать метод генерации музыкального фрагмента, соответствующего эмоциональному состоянию человека, с использованием марковских моделей;
- разработать программное обеспечение, реализующее спроектированный метод;
- провести исследование соответствия сгенерированных музыкальных фрагментов выбранной эмоции.

Метод генерации музыкального фрагмента

Ограничения метода:

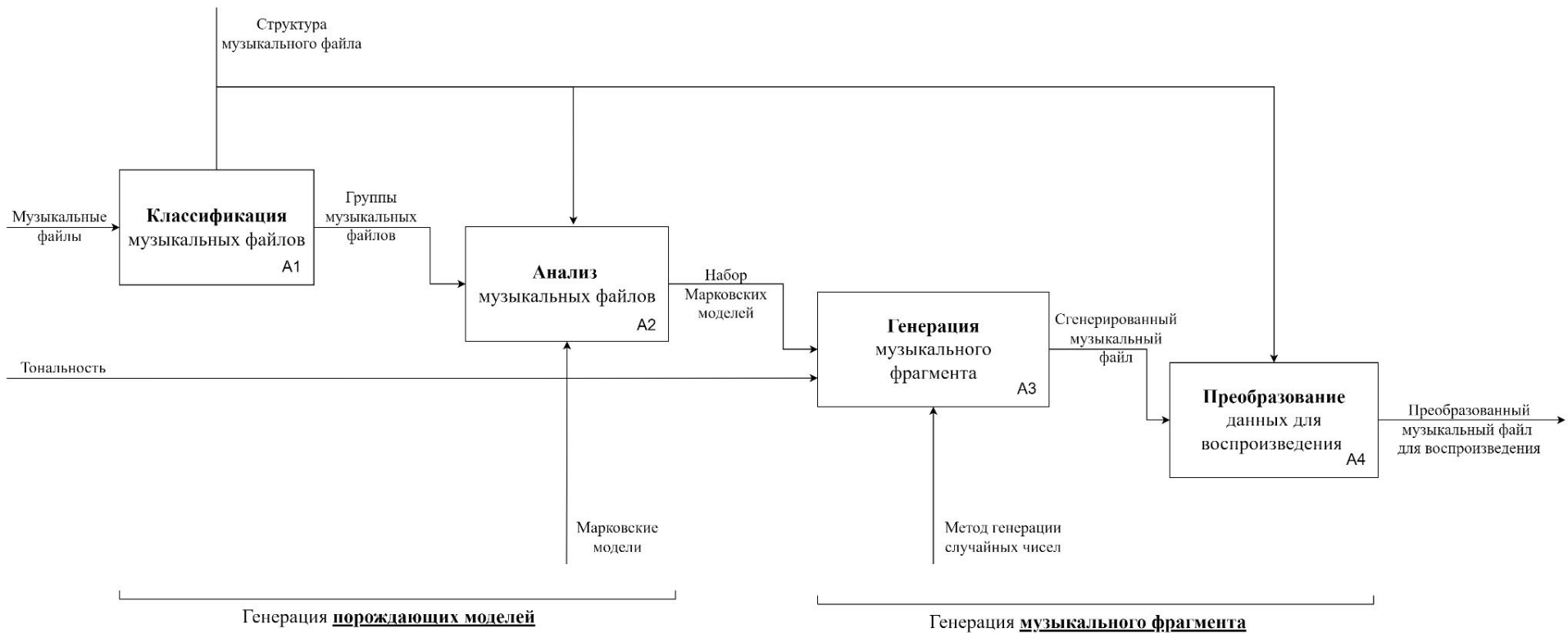
- музыкальный фрагмент генерируется для одного инструмента – фортепиано;
- сгенерированный музыкальный фрагмент не имеет жанровых особенностей.



Классификация существующих методов машинной генерации музыки

Критерий сравнения	Л-системы	Искусственные Нейронные сети	Цепи Маркова	Порождающие грамматики
Степень достаточности малой обучающей выборки	высокая	низкая	высокая	высокая
Необходимость преобразования данных и вывода правил для обучения модели	средняя	средняя	средняя	низкая
Степень вариативности структуры музыкальной композиции	средний	низкая	средняя	высокая
Степень динамичности условий генерации музыки	средняя	высокая	высокая	средняя

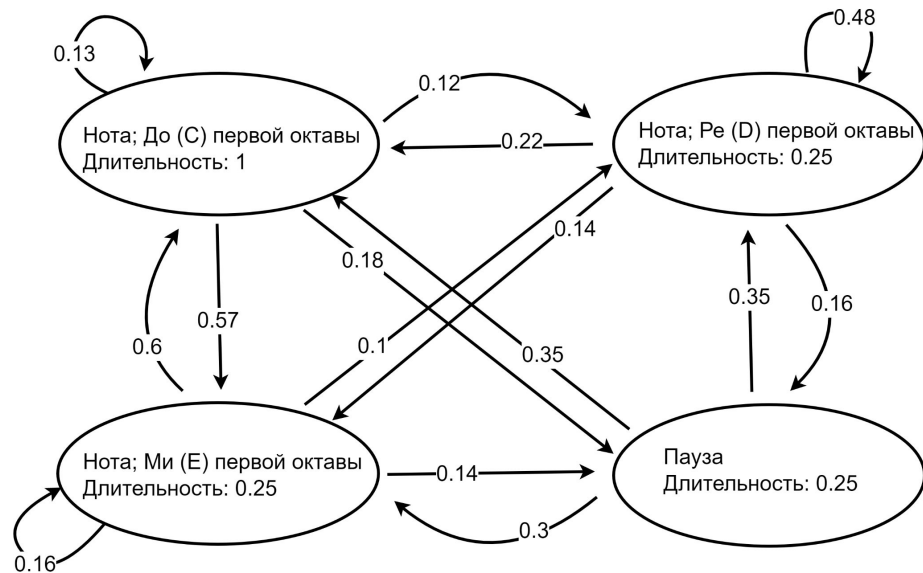
Функциональная модель разрабатываемого метода генерации музыкального фрагмента



Представление музыкального файла в виде марковской модели

Марковская модель состоит из:

- множества всех возможных состояний, где состояние это:
 - либо 1 звук:
 - нота + высота звучания;
 - длительность.
 - либо комбинация из нескольких звуков:
 - аккорд (совокупность нот);
 - длительность.
- матрицы переходных вероятностей;
- вектора начальных вероятностей, или тоники.

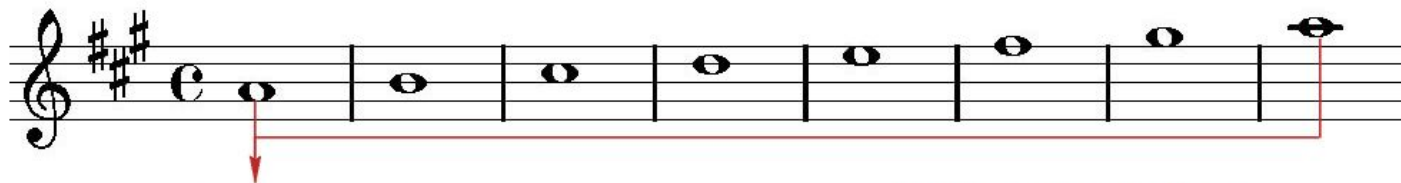


Пример Марковской модели 1-го порядка

Вектор начальных вероятностей (тоники)

Тональность музыкального произведения определяется как совокупность наклонения и тоники:

- Тональность – звукоряд, состоящий из восьми идущих подряд звуков (октава), имеющих высотную периодичность.
- Наклонение – общий окрас звукоряда, придающий ему грустное (минорное) или веселое (мажорное) настроение.
- Тоника – первая нота в тональности. Тоника является устойчивым звуком, который явным или неявным образом начинает и завершает произведение.



Тоника: нота Ля

Наклонение: мажорное

Тональность (Ля-мажор) = Тоника (нота Ля) + Наклонение (мажор)

Определение эмоциональной окраски музыкального текста

Связь между тоникой и цветом может быть объяснена через длину волны и частоту:

- каждая тоника имеет частоту, которая определяет ее высоту;
- каждый цвет определяется длиной волны и тоже имеет частоту.

$$\text{Частота (} \nu \text{)} = \frac{\text{Скорость волны (} V \text{)}}{\text{Длина волны (} \lambda \text{)}}$$

При увеличении высотности звука на 41 октаву частота попадает в спектр видимого излучения. Посредством такого сдвига тоника сопоставляется определённому цвету.

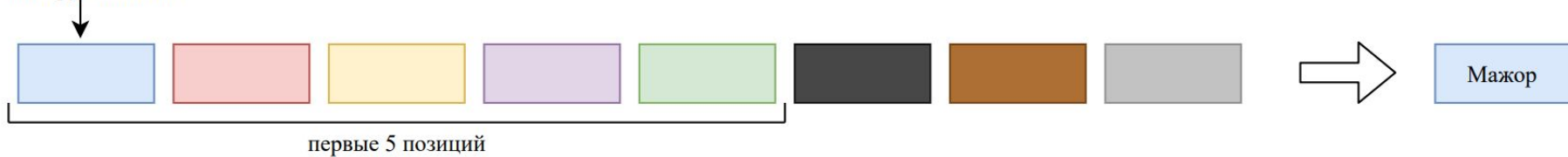
Тоника	Фа Диез (F#)	Соль (G)	Соль Диез (G#)	Ля (A)	Си Бемоль (B#)	Си (B)	До (C)	До Диез (C#)	Ре (D)	Ре Диез (D#)	Ми (E)	Фа (F)
Цвет	Красный	Красный	Желтый	Желтый	Желтый	Зеленый	Зеленый	Зеленый	Синий	Синий	Синий	Красный

Выбор эмоциональной окраски произведения

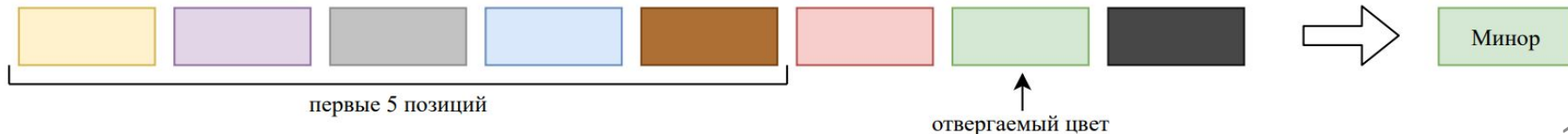
Пользователь проходит тест Люшера, упорядочивая цвета в порядке его предпочтения в данный момент времени:

- Если все основные цвета расположены на первых пяти позициях, фрагмент будет иметь **мажорное наклонение** и тоника определяется по первому, **доминирующему** основному цвету;

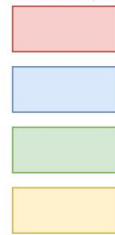
доминирующий цвет



- Иначе фрагмент будет иметь **минорное наклонение** и тоника определяется по последнему в списке основному цвету, **отвергаемому**.



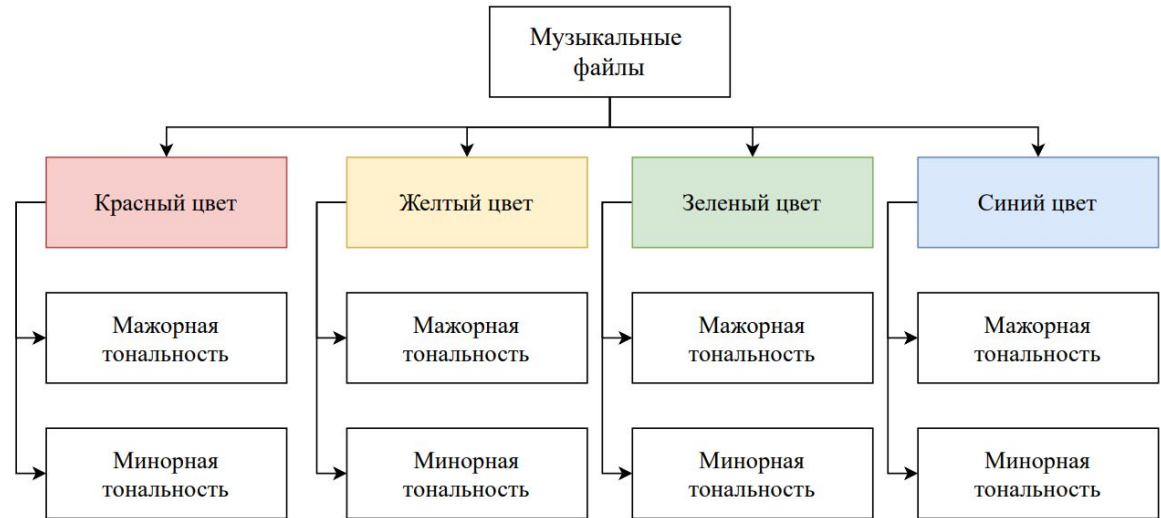
основные цвета





Построение порождающих моделей

- На основе цвета и тональности выделены 8 групп:



- Исходные музыкальные файлы классифицируются в соответствии с выделенными группами;
- В результате, для каждой из 8 групп составляется своя Марковская модель.

Генерация марковской модели музыкального фрагмента

Генерация музыкального фрагмента состоит из следующих основных этапов:

- Выбор начального состояния системы, используя вектор начальных вероятностей требуемой тональности;
- Выбор следующего состояния марковской модели на основе:
 - случайно сгенерированного числа;
 - текущего значения ноты;
 - матрицы переходных вероятностей, соответствующей заданной тональности.



Преобразование последовательности состояний к музыкальному файлу

Запись результата генерации производится в MIDI-файл, компонентами которого являются:

- Канальные события:
 - note off – выключение ноты;
 - note on – включение ноты;
- Мета-события:
 - set tempo – текущий темп в микросекундах;
 - time signature – музыкальный размер композиции;
 - key signature – текущая тональность.



Структура программного обеспечения



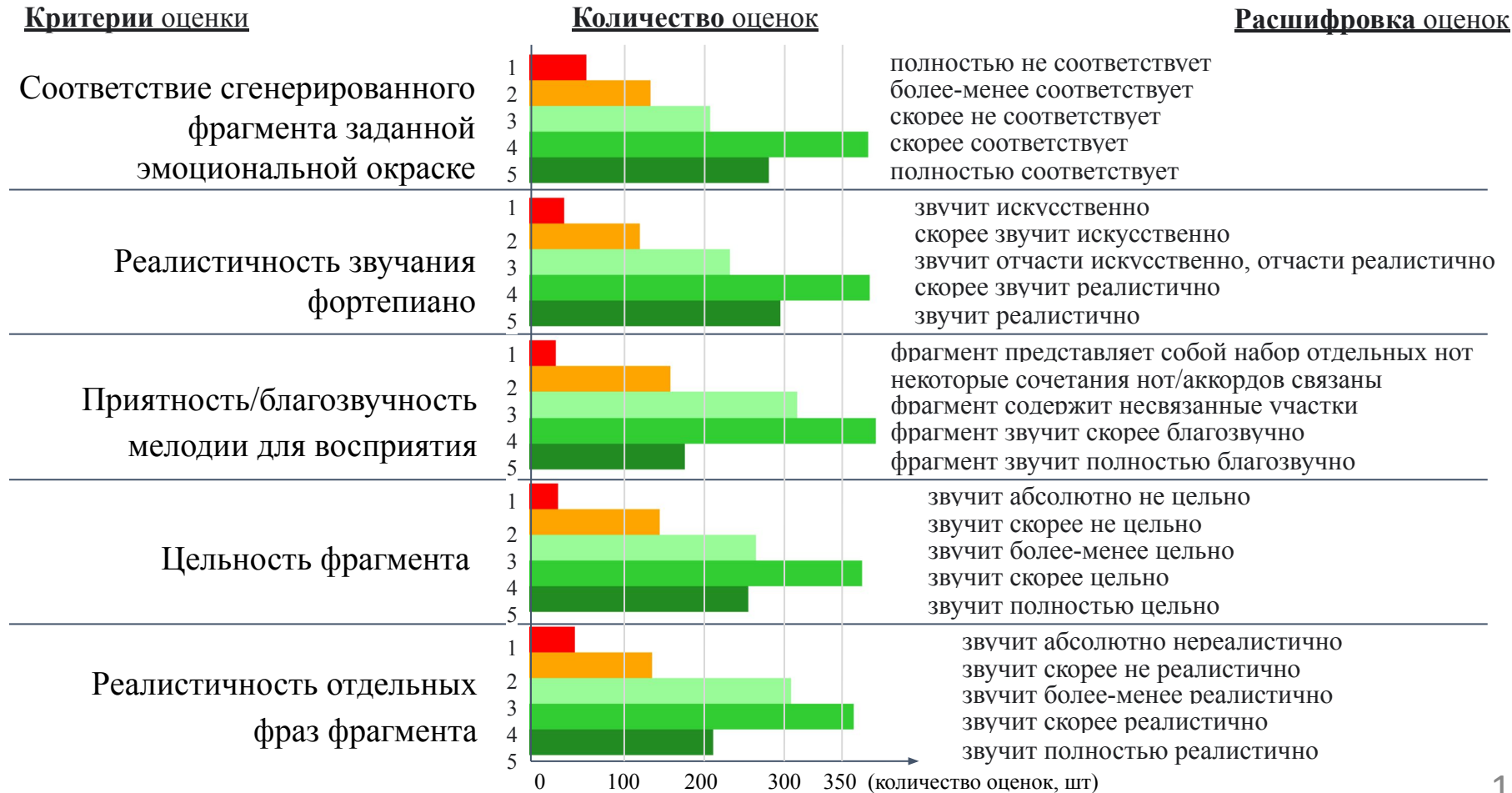
Исследование соответствия сгенерированных музыкальных фрагментов выбранной эмоции

- Порождающие марковские модели были сгенерированы на **150** произведениях для фортепиано, сгруппированных в **8** классов.
- Был сгенерирован тестовый набор из **8** музыкальных произведений по **25** тактов.

Экспертная группа состояла из **125** человек:

- **86** любителей, не имеющих музыкального образования;
- **28** любителей, имеющих музыкальное образование;
- **11** профессионалов в сфере музыки.

Распределение экспертных оценок по критериям



Метод оценки согласия экспертов

Для оценки согласия экспертов вычислен коэффициент вариации:

$$CV = \frac{\sigma}{\bar{x}} \cdot 100\%$$

σ - стандартное отклонение:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

\bar{x} - среднее значение экспертной оценки:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

n - количество выставленных оценок;

x_i – i -я выставленная оценка.

Критерий	CV
Соответствие сгенерированного фрагмента заданной эмоциональной окраске	38%
Реалистичность звучания фортепиано	32%
Приятность/благозвучность мелодии для восприятия	30%
Цельность фрагмента	32%
Реалистичность отдельных фраз фрагмента	34%

Коэффициент вариации показывает, насколько велик **разброс** экспертных оценок относительно их среднего значения. Чем меньше значение коэффициента, тем более **точными** и **согласованными** являются экспертные оценки.

Экспертная оценка сгенерированных музыкальных фрагментов и отрывков из настоящих сонат

Критерий	Расшифровка ближайших оценок	Сгенерированные музыкальные фрагменты	Отрывки из сонат композитора Альбана Берга
Соответствие сгенерированного фрагмента заданной эмоциональной окраске	3 - более-менее соответствует 4 - скорее соответствует	3,62	3,47
Реалистичность звучания фортепиано	3 - звучит отчасти искусственно, отчасти реалистично 4 - звучит скорее реалистично	3,7	3,53
Приятность/благозвучность мелодии для восприятия	3 - звучит более-менее связано, но содержит много несвязанных участков 4 - звучит скорее благозвучно	3,49	3,12
Цельность фрагмента	3 - попадают несвязанные ноты 4 - звучит полностью цельно	3,61	3,39
Реалистичность отдельных фраз фрагмента	3 - более-менее реалистично 4 - скорее реалистично	3,50	3,23

Заключение

Цель работы достигнута: разработан метод генерации музыкального фрагмента, соответствующего эмоциональному состоянию человека, с использованием марковских моделей.

Решены все поставленные задачи:

- проведен анализ предметной области генерации музыкальной композиции и существующих методов машинной генерации музыки;
- разработан метод генерации музыкального фрагмента, соответствующего эмоциональному состоянию человека, с использованием марковских моделей;
- разработано программное обеспечение, реализующее спроектированный метод;
- проведено исследование соответствия сгенерированных музыкальных фрагментов выбранной эмоции.

Направление дальнейшего развития

- изменение динамики музыкального фрагмента за счет добавления сильных долей, а не только уплотнения музыкального текста;
- улучшение качества сгенерированных музыкальных фрагментов за счет использования марковских моделей более высокого порядка или скрытых марковских моделей;
- использование вспомогательного способа определения эмоционального состояния человека за счет использования датчиков биологических показателей.