



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н. Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н. Э. Баумана)

ФАКУЛЬТЕТ «Информатика и системы управления»

КАФЕДРА «Программное обеспечение ЭВМ и информационные технологии»

РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

К КУРСОВОЙ РАБОТЕ

НА ТЕМУ:

*«14_КОТ Метод генерации фрагмента музыкального
произведения по изображению»*

Студент ИУ7-56Б
(Группа)

(Подпись, дата)

Вольняга М. Ю.
(И. О. Фамилия)

Руководитель курсовой работы

(Подпись, дата)

Кострицкий А. С.
(И. О. Фамилия)

2023 г.

СОДЕРЖАНИЕ

ОПРЕДЕЛЕНИЯ	3
ВВЕДЕНИЕ	4
1 Аналитический раздел	5
1.1 Алгоритм анализа изображения	5
1.1.1 Алгоритм определения тональности	7
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	10

ОПРЕДЕЛЕНИЯ

В настоящей расчетно-пояснительной записке применяют следующие термины с соответствующими определениями.

Музыкальный тон — это устойчивый периодический звук. Музыкальный тон характеризуется его длительностью, высотой, интенсивностью (или громкостью) и тембром (или качеством) [1].

Темп (итал. *tempo*) — это скорость движения в музыке, мера времени в музыке [1].

HSB (англ. Hue, Saturation, Brightness — тон, насыщенность, яркость) — цветовая модель, в которой координатами цвета являются: цветовой тон (например, красный, зеленый или сине-голубой), насыщенность (варьируется в пределах 0—100), яркость (варьируется в пределах 0—100) [2].

RGB (англ. Red, Green, Blue — красный, зеленый, синий) — цветовая модель, описывающая способ кодирования цвета для цветовоспроизведения с помощью трех цветов, которые принято называть основными [2]

Кластеризация — это разбиение элементов некоторого множества на группы по принципу схожести. Эти группы принято называть кластерами. [3]

ВВЕДЕНИЕ

Создание музыки творческий процесс, его автоматизация сложна из-за важной роли композитора и труднопонимаемой эмоциональности в музыке [4]. Для автоматической генерации музыкальных композиций с учетом эмоционального состояния пользователя-композитора можно использовать трансляционные модели [5].

Трансляционные модели — это способ генерации музыкальных произведений на основе немusикальных данных, таких как графические образы или текст. Данный процесс может быть случайным или основываться на определенных правилах мелодической структуры, с использованием нейросетевых технологий для преобразования и распознавания исходных данных [6].

Генерация музыки на основе изображения может применяться в компьютерных играх, рекламе и фильмах для создания фоновой музыки. Автоматизация этого процесса позволит сократить расходы компаний, учитывая небольшие требования к фоновой музыке в этих областях [7].

Целью данной работы является изучение (описание) метода генерации фрагмента музыкального произведения по изображению.

Формальная постановка задачи:

- 1) пусть A — множество всех возможных изображений, исходные данные для алгоритма;
- 2) пусть B — множество всех возможных музыкальных композиций, результаты работы алгоритма;
- 3) Тогда $f : A \rightarrow 2^B$ — многозначное отображение, которое каждому элементу (изображению) $a \in A$ сопоставляет множество музыкальных композиций $f(a) \subseteq B$.

1 Аналитический раздел

Генерация фрагмента музыкального произведения по изображению, представляет собой преобразование визуальных данных в последовательности нот с определенным тоном и темпом [8].

Метод генерации фрагмента музыкального произведения по изображению состоит из двух составляющих алгоритмов:

- 1) алгоритм анализа изображения;
- 2) алгоритм генерации музыкального фрагмента с использованием нейронных сетей.

1.1 Алгоритм анализа изображения

Алгоритм анализа изображения позволяет получить музыкальные характеристики изображения — определить тональность и темп получаемой музыкальной композиции [8]. Тональность и темп — являются ключевыми параметрами для трансляции изображения в музыку, поскольку они формируют эмоциональную составляющую произведения, и должны быть определены путем анализа цветовой гаммы изображения [7].

Алгоритм анализа изображения состоит из следующих шагов [8].

- 1) Преобразовать входное изображение из цветовой модели RGB в HSB. Цветовая модель HSB более удобна, так как содержит необходимые характеристики.
- 2) Определить преобладающий цвет изображения.
- 3) Определить тональность произведения.

Определение преобладающего цвета изображения

Для этой задачи традиционно выбирают алгоритм кластеризации К-средних [5; 8]. Данный алгоритм имеет следующие особенности: высокое качество кластеризации, возможность распараллеливания, существует множество модификаций [9].

Алгоритм кластеризации К-средних — это неконтролируемый метод обучения. В этом методе данные разбиваются на кластеры на основе их сходства без использования предварительных обозначений [8].

Алгоритм состоит из следующих шагов [3]

- 1) Выбирается число k — количество кластеров.
- 2) Далее случайным образом из заданного изображения выбирается k точек. На первом шаге эти точки будут считаться «центрами» кластеров. Каждому кластеру соответствует один центр.
- 3) Все точки изображения распределяются по кластерам. Вычисляется расстояние от точки до каждого центра кластера, и точку относят к тому кластеру, расстояние до центра которого будет наименьшим.
- 4) Когда все точки изображения распределены по кластерам, происходит пересчет центров кластеров. В качестве нового центра кластера берется среднее арифметическое всех точек, принадлежащих кластеру.

Пункты 3 и 4 повторяются до тех пор, пока не будет выполнено условие в соответствии с некоторым критерием остановки:

- кластерные центры стабилизировались, то есть все наблюдения принадлежат кластеру, которому принадлежали до текущей итерации;
- число итераций равно максимальному числу итераций.

В методе k -средних ставится цель минимизировать полную внутриклассовую дисперсию:

$$V = \sum_{i=1}^k \sum_{X_j \in C_i} (X_j - \mu_i)^2 \quad (1.1)$$

где X_j — векторы характеристик, k — количество кластеров, C_i — кластеры, μ_i — центры кластеров.

В качестве метрики для данной задачи традиционно берут евклидово расстояние [3; 10; 11].

Евклидово расстояние:

$$\rho(x_i, y_i) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1.2)$$

Результаты алгоритма кластеризации К-средних [12]:

- Центроиды кластеров K , которые можно использовать для маркировки новых данных;
- Метки для обучающих данных (каждая точка данных назначается одному кластеру).

Определить тональность произведения

Тональность и темп — являются ключевыми параметрами для трансляции изображения в музыку, поскольку они формируют эмоциональную составляющую произведения, и должны быть определены путем анализа цветовой гаммы изображения [7]. Для этого нужно установить соответствие между цветовыми и музыкальными характеристиками [5] (таблица 1.1). Затем следует определить схему соотнесения цвета и ноты [5]. В статье [13] описывается множество подобных схем, например соотнесение цветов и нот по И. Ньютону, он искал связь между солнечным спектром и музыкальной октавой, сопоставляя длины разноцветных участков спектра и частоту колебаний звуков гаммы, таблица 1.2.

Таблица 1.1 – Соотношение цветовых и музыкальных характеристик

Цветовые характеристики	Музыкальные характеристики
Оттенок (красный, синий, желтый. . .)	Нота (до, до-диез, ре, ре-диез ...)
Цветовая группа (теплый/холодный)	Музыкальный лад (мажор/минор)
Яркость	Октава ноты
Насыщенность	Длительность ноты

Таблица 1.2 – Соотнесение цветов и нот по И. Ньютону

Цвет	Нота
Красный	До
Фиолетовый	Ре
Синий	Ми
Голубой	Фа
Зеленый	Соль
Желтый	Ля
Оранжевый	Си

1.1.1 Алгоритм определения тональности

Определение тональности основывается на проведении анализа изображения и использовании таблицы 1.1, состоит из следующих шагов, описанных в [5]:

- 1) преобразовать входное изображение из цветового пространства RGB в HSV. Данный шаг позволяет преобразовать изображение к более удобному виду, поскольку HSV пространство уже содержит необходимые характеристики – название цвета (определяется по параметру hue), насыщенность (параметр saturation) и яркость (параметр brightness);
- 2) определить преобладающий цвет изображения;
- 3) определить название и цветовую группу преимущественного цвета;
- 4) согласно выбранной схеме соотнесения цветов и нот, а также результатах, полученных на предыдущих шагах, определяем тональность произведения.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. *Миланич Е. Э., Лойко О. Б.* Учебно-методический комплекс по учебной дисциплине «Элементарная теория музыки». — 2017.
2. Сравнительный анализ использования цветовых пространств RGB и HSB в задачах распознавания объектов при определении позиции предметов в роботизированных системах / Кузьмина Е. К. [и др.] // Системный анализ в науке и образовании. — 2022. — С. 55—63. — URL: <https://sanse.ru/index.php/sanse/article/view/520> ; дата обращения: 09.11.2023.
3. *Котелина Н. О., Матвийчук Б. Р.* Кластеризация изображения методом К-средних // Вестник Сыктывкарского университета. Серия 1: Математика. Механика. Информатика. — 2019. — Т. Выпуск 3 (32). — URL: <file:///D:/downloads/klasterizatsiya-izobrazheniya-metodom-k-srednih.pdf> ; дата обращения: 10.11.2023.
4. *Никитин Н. А.* Модели, методы и средства компьютерного синтезирования музыки по цветовому изображению. — 2022.
5. *Никитин Н. А., Розалиев В. Л., Орлова Ю. А.* Разработка веб-сервиса для генерации музыкальной последовательности по изображению // Молодой ученый. — 2019. — Т. 51, № 289. — С. 27—30. — URL: <https://moluch.ru/archive/289/65648/> ; дата обращения: 06.10.2023.
6. *Малахов В. Е., Rogozинский Г. Г.* Применение декомпозиции для изучения процессов при создании генеративной музыки. — 2022.
7. *Никитин Н. А., Орлова Ю. А., Розалиев В. Л.* Программная генерация звуков по цветовой гамме изображений с использованием рекуррентной нейронной сети // Первая Всероссийская научно-практическая конференция. — 2017. — С. 35—39. — URL: http://fti.ulstu.ru/wp-content/uploads/2022/01/Proceedings_FTI_2017_final.pdf#page=105/ ; дата обращения: 09.11.2023.
8. *Никитин Н. А., Орлова Ю. А., Розалиев В. Л.* Алгоритм генерации музыкальных композиций с использованием интуитивного и эмоционального подходов // Молодой ученый. — 2021. — Т. 24, № 366. — С. 35—

39. — URL: <https://moluch.ru/archive/366/82308/> ; дата обращения: 06.10.2023.
9. *Jain A., Murty M., Flynn P.* Data clustering: A review // ACM Computing Surveys. — 1999. — Т. 31.
10. *Ооржак О. Е.* Алгоритм кластеризации К-средних и его реализация в среде MATLAB. — 2015.
11. *Coates A., Ng A. Y.* Learning Feature Representations with K-means. — 2012. — URL: https://web.archive.org/web/20150621073741/http://www.cs.stanford.edu/~acoates/papers/coatesng_nntot2012.pdf ; дата обращения: 10.11.2023.
12. *Oyelade O., Oladipupo O., Obagbuwa I.* Application of k-Means Clustering Algorithm for Prediction of Students Academic Performance // International Journal of Computer Science and Information Security. — 2010. — Т. 7, № 1. — С. 292—295.
13. *Abdullayev E.* An Experimental Model Of Color And Sound Correlation // Eurasian music science journal. — 2020. — URL: <https://core.ac.uk/download/pdf/336866842.pdf> ; дата обращения: 07.10.2023.