

ings 2017 (Lisbon, Portugal, July 20-22, 2017): part of the Multi Conference on Computer Science and Information Systems 2017: Proceedings / ed. by Piet Kommers; IADIS (International Association for Development of the Information Society). – [Lisbon, Portugal], 2017. – P. 35-42.

## ПРОГРАММНАЯ ГЕНЕРАЦИЯ ЗВУКОВ ПО ЦВЕТОВОЙ ГАММЕ ИЗОБРАЖЕНИЙ С ИСПОЛЬЗОВАНИЕМ РЕКУРРЕНТНОЙ НЕЙРОННОЙ СЕТИ

*Никитин Н.А., Розалиев В.Л., Орлова Ю.А.*

*(г. Волгоград, Волгоградский Государственный Технический Университет)  
nikitin.nikita@outlook.com, vladimir.rozaliiev@gmail.com, yulia.orlova@gmail.com*

## PROGRAM FOR SOUND GENERATION BASED ON IMAGE COLOR SPECTRUM WITH USING THE RECURRENT NEURAL NETWORK

*Nikitin N.A., Rozaliev V.L., Orlova Yu.A.*

*(Volgograd, Volgograd State Technical University)*

**Abstract.** This work is devoted to development and approbation of the program for sound generation based on image color spectrum with using the recurrent neural network. The work contains a description of the transition between color and music characteristics, the rationale for choosing and the description of a recurrent neural network. The choices of the neural network implementation technology as well as the results of the experiment are described.

**Keywords:** artificial neural networks, recurrent neural network, long-short term memory, Python, Keras, Newton correlation scheme, sampling.

**Введение.** С тех пор как музыку стали записывать на бумаге в виде нотных знаков, стали появляться оригинальные «способы» ее сочинения. Одним из самых первых методов алгоритмической композиции стал способ сочинения музыки, придуманный Моцартом – «Музыкальная игра в кости» [1]. Первое компьютерное музыкальное произведение – «Iliac Suite for String Quartet» – было создано в 1956 году пионерами применения компьютеров в музыке – Лежарен Хиллер и Леонард Айзексон [2]. В этом произведении использованы почти все главные методы алгоритмической музыкальной композиции: теория вероятностей, марковские цепи и генеративная грамматика.

Развитие компьютерной музыки, в том числе и генерации звуков по изображению, в прошлом веке было сильно ограничено вычислительными ресурсами – покупать и содержать мощные ЭВМ могли позволить себе лишь крупные университеты и лаборатории, а первым персональным компьютерам не хватало вычислительной мощности. Однако в XXI веке, изучением компьютерной музыки может заниматься практически каждый человек. В настоящее время компьютерная музыка может применяться во многих отраслях: создание музыки для компьютерных игр, рекламы и фильмов. Сейчас, для создания фоновых музыкальных композиций в компьютерных играх и рекламе, компании нанимают профессиональных композиторов или покупают права на уже написанные музыкальные произведения. Однако в таком жанре, требования к музыкальной композиции не велики, а значит, данный процесс можно автоматизировать, что позволит компаниям снизить расходы на сочинение композиций. Также, генерацию звуков по изображению можно применить в образовательном процессе [3]. Взаимодействие музыки и изобразительного искусства в процессе интегрированной образовательной деятельности с детьми дошкольного возраста может осуществляться в форме сочетания восприятия произведений музыкального и изобразительного искусства на основе общности их настроения, стиля, жанра, что способствует развитию музыкального восприятия у дошкольников [4].

Наибольших успехов автоматизация процесса написания и создания музыки достигла сравнительно недавно (в последние десятилетия), однако по большей части связана с изуче-

нием и повторением различных музыкальных стилей [5]. Поскольку процесс создания музыки сложно формализуем, то для программного (автоматизированного) создания композиций лучше всего подходят искусственные нейронные сети, так как они позволяют выявить связи, которые не видит человек [6]. Помимо этого, для снижения роли пользователя-композитора в генерации музыкальных произведений, было принято решение брать часть музыкальных характеристик с изображения. В связи с этим, целью данной работы является увеличение гармоничности и мелодичности программной генерации звуков по цветовой гамме изображений посредством использования нейронных сетей.

**От цветовых характеристик к музыкальным.** Для снижения роли пользователя-композитора в генерации звуков, часть характеристик музыкального произведения получается путём анализа цветовой гаммы изображения. Таким образом, характер полученной музыкальной композиции будет соответствовать входному изображению. Данная особенность делает возможным применение данного подхода для создания фоновых музыкальных произведений в компьютерных играх, рекламе и фильмах.

Ключевыми характеристиками музыкального произведения является его тональность и темп. Именно эти параметры определяются путём анализа цветовой гаммы изображения. Для начала определим соотношение цветовых и музыкальных характеристик [7] (таблица 1).

Таблица 1 – Соотношение цветовых и музыкальных характеристик

Цветовые характеристики	Музыкальные характеристики
Отенок (красный, синий, жёлтый...)	Нота (до, до-диез, ре, ре-диез, ми, ми, фа, фа-диез, соль, соль-диез, ля, ля-диез, си)
Цветовая группа (тёплый/холодный)	Музыкальный лад (мажор/минор)
Яркость	Октава ноты
Насыщенность	Длительность ноты

Затем, необходимо определить схему соотнесения названия цвета и ноты. На данный момент существует большое количество подобных схем, однако в данной работе была выбрана схема Ньютона.

Как видно из таблицы 1, тональность произведения определяется двумя цветовыми характеристиками – отенок и цветовая группа, а темп – яркость и насыщенность. Алгоритм определения тональности опирается на анализ изображения и таблицу 1, состоит из 3 шагов и описан ниже.

Шаг 1. Преобразуем 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).

Шаг 2. Анализируя в целом изображение, определяем преимущественный цвет.

Шаг 3. Определяем название и цветовую группу преимущественного цвета.

Шаг 4. Согласно таблице 1 и схеме Ньютона определяем тональность произведения (нота и музыкальный лад).

Для определения темпа произведения, необходимо получить яркость и насыщенность (по параметрам saturation и brightness) преимущественного цвета, и рассчитать темп, согласно данным параметрам.

**Выбор нейронной сети для генерации музыкальных композиций.** Важной особенностью нейронных сетей прямого распространения (feedforward neural networks) является то, что у данной нейросети есть общее ограничение: и входные и выходные данные имеют фиксированный, заранее обозначенный размер, например, картинка 100×100 пикселей или последовательность из 256 бит. Нейросеть с математической точки зрения ведет себя как обыч-

ная функция, хоть и очень сложно устроенная: у нее есть заранее обозначенное число аргументов, а также обозначенный формат, в котором она выдает ответ.

Вышеперечисленные особенности не представляет больших трудностей, если речь идет о тех же картинках или заранее определенных последовательностях символов. Но для обработки любой условно бесконечной последовательности, в которой важно не только содержание, но и порядок, в котором следует информация, например, текст или музыка необходимо использовать нейронные сети с обратными связями – рекуррентные нейронные сети (RNN). В рекуррентных нейросетях нейроны обмениваются информацией между собой: например, вдобавок к новому кусочку входящих данных нейрон также получает некоторую информацию о предыдущем состоянии сети. Таким образом в сети реализуется «память», что принципиально меняет характер ее работы и позволяет анализировать любые последовательности данных, в которых важно, в каком порядке идут значения [8].

Однако большой сложностью сетей RNN является проблема исчезающего (или взрывного) градиента, которая заключается в быстрой потере информации с течением времени. Конечно, это влияет лишь на веса, а не состояния нейронов, но ведь именно в них накапливается информация. Сети с долгой краткосрочной памятью (long short term memory, LSTM) стараются решить вышеупомянутую проблему потери информации, используя фильтры и явно заданную клетку памяти. У каждого нейрона есть клетка памяти и три фильтра: входной, выходной и забывающий. Целью этих фильтров является защита информации. Входной фильтр определяет, сколько информации из предыдущего слоя будет храниться в клетке. Выходной фильтр определяет, сколько информации получают следующие слои. Такие сети способны научиться создавать сложные структуры, например, сочинять тексты в стиле определённого автора или сочинять простую музыку, однако при этом потребляют большое количество ресурсов [9].

Таким образом, для реализации программы автоматизированной генерации музыкальных композиций по цветовой гамме изображений необходимо использовать именно рекуррентные нейронные сети с долгой краткосрочной памятью – RNN LSTM (долгая краткосрочная память – разновидность архитектуры рекуррентных нейронных сетей). Именно данный вид нейронных сетей используется для генерации музыкальных композиций в программе Magenta – это музыкальный проект с открытым исходным кодом от Google, также RNN LSTM используется в программе сочинения композиций в стиле И.С. Баха – BachBot, а также в DeepJaz – система позволяет генерировать джазовые композиции на основе анализа midi файлов [10].

**Проведение эксперимента.** Для подтверждения эффективности предложенных алгоритмов генерации звуков по цветовой гамме изображения, была разработана программа на языке Python, с использованием библиотеки Keras. Для синтеза звуков используется метод сэмплинг (sampling). Программа генерации музыкальных композиций с использованием нейронных сетей была обучена на 29 композициях Людвига ван Бетховена. После обучения был составлен набор из десяти тестовых изображений, имеющих различный тип (абстрактные изображения, пейзажи, города и люди). По всем десяти изображениям были получены и сохранены выходные музыкальные композиции. Данные музыкальные композиции были отправлены на анализ 10 экспертам, которые должны были оценить каждое произведение по следующим критериям: соответствие характеру изображения (по пяти бальной шкале); реалистичность звучания инструмента (фортепьяно или гитара); мелодичность композиции; качество гармонии (аккомпанемента); приятность мелодии для восприятия; цельность композиции; реалистичность/искусственность композиции.

Проанализировав оценки всех экспертов и высчитав средние по каждому критерию, можно сделать вывод о том, что фортепьяно на слух экспертов звучит реалистичнее, чем гитара. Также можно сделать вывод о том, что композиция, сгенерированная по абстрактным изображениям, более приятна на слух, чем генерация по пейзажам. В целом общее впечатление от сгенерированных звуков у экспертов положительное. Среди минусов некоторые экс-

перты выделяют однотипность гармонии, иногда рваность и недостаточную реалистичность произведения, и не достаточную реалистичность гитары.

Делая вывод по каждому критерию можно сказать, что все эксперты оценили на высокий бал соответствие произведения характеру изображения, по второму критерию – инструмент фортепьяно звучит довольно реалистично. Мелодичность композиций разделилась по полам, то есть половина композиций эксперты оценили на высший бал, другую половину на 4, в целом неплохой результат. Качество гармонии также было оценено экспертами на высший бал. Приятность мелодий для восприятия получил 60% высших баллов и 40% четвёрок, что говорит о том, что некоторые произведения звучат не вполне реалистично. Реалистичность и цельность композиций в среднем оценено на 4, что является естественным результатом для компьютерной генерации звуков.

**Заключение.** В ходе выполнения работы была определена схема соотнесения цветовых и музыкальных характеристик, был проведён обзор типов нейронных сетей и выбран наиболее подходящий тип для генерации музыкальных композиций, была детально описана используемая нейронная сеть, была выбрана технология реализации нейронной сети, был выбран метод синтеза звуков, был проведён эксперимент по оценке гармоничности и мелодичности выходных музыкальных композиций.

В ходе анализа различных типов и архитектур ИНС был сделан вывод о том, что наиболее подходящей сетью для обработки музыкальной информации являются рекуррентные нейронные сети (RNN), а именно сети с долгой краткосрочной памятью (long short term memory, LSTM).

В результате проведения эксперимента, была обучена модель (нейронная сеть) на композициях Баха, а также были сгенерированы композиции по 10 изображениям. Данные композиции были отправлены на анализ экспертам. В результате анализа экспертных оценок можно сделать вывод о том, что программа генерирует достаточно мелодичные композиции, однако сказывается, что модель была обучена на небольшом количестве произведений только одного автора.

Работа частично поддержана Российским фондом фундаментальных исследований (проекты 16-47-340320, 17-07-01601).

#### ЛИТЕРАТУРА

1. Фазылова, Э.Ф. Системы генерации музыки или как автоматизировать искусство? // Молодёжный научно-технический вестник. – 2014. URL: <http://sntbul.bmstu.ru/doc/723360.html>. (Дата обращения: 20.03.2017).
2. Ariza, C. Two Pioneering Projects from the Early History of Computer-Aided Algorithmic Composition / C. Ariza // Computer Music Journal. – MIT Press, 2012. – №3. – pp. 40-56
3. Черешнюк, И. Р. Алгоритмическая музыкальная композиция и её место в современном музыкальном образовании / И. Р. Черешнюк // Педагогика искусства. – 2015. – № 3. – С. 65-68.
4. Выготский, Л.С. Воображение и творчество в детском возрасте / Л.С. Выготский. – Москва, 2008 // Мышление и речь : сборник / Л.С. Выготский. – Москва : АСТ, 2008. – С. 497-594.
5. D. Cope, Computer Models of Musical Creativity, MIT Press, Cambridge, Mass., 2005.
6. Mazurowski, L. Computer models for algorithmic music composition / L. Mazurowski // Proceedings of the Federated Conference on Computer Science and Information Systems. – Szczecin, Poland, 2012. – pp. 733–737
7. Caivano, J. L., Colour and sound: Physical and Psychophysical Relations, Colour Research and Application, 12(2), pp. 126-132, 1994
8. Sak, H., Senior, A., Beaufays, F. Long Short-Term Memory Based Recurrent Neural Network Architectures for Large Vocabulary Speech Recognition / H. Sak, A. Senior, F. Beaufays // ArXiv e-prints. – 2014

9. Doornbusch, P. Gerhard Nierhaus: Algorithmic Composition: Paradigms of Automated Music Generation / P. Doornbusch // Computer Music Journal. - Volume: 34, Issue: 3. – 2014.

10. Brinkkemper, F. Analyzing Six Deep Learning Tools for Music Generation [Электронный ресурс]. – 2015. - Режим доступа: <http://www.asimovinstitute.org/analyzing-deep-learning-tools-music/> (Дата обращения: 03.07.2017).

## **АНАЛИЗ ПЕРСПЕКТИВ РАЗРАБОТКИ ИНТЕЛЛЕКТУАЛЬНЫХ СИТУАЦИОННЫХ ЦЕНТРОВ В КАЧЕСТВЕ ИНФОРМАЦИОННЫХ СИСТЕМ ПОДДЕРЖКИ УПРАВЛЕНЧЕСКИХ РЕШЕНИЙ ПРИ МОНИТОРИНГЕ РЕЧНЫХ ПАВОДКОВ**

*А.А. Скугарев, М.Ю. Катаев*

*(г. Томск, Томский университет систем управления и радиоэлектроники)  
e-mail: skugarev@inbox.ru, kataev.m@sibmail.com*

## **ANALYSIS OF THE PROSPECTS OF INTELLECTUAL SITUATION CENTERS DEVELOPING AS INFORMATION SYSTEMS FOR SUPPORT OF MANAGEMENT DECISIONS IN MONITORING RIVER FLOODS**

*A.A. Skugarev, M.Yu. Kataev*

*(Tomsk, Tomsk University of Control Systems and Radioelectronics)*

**Abstract.** The report describes the analysis of the prospects for the development of intellectual situation centers as information systems for support of management decisions in the monitoring of river floods. The problem of the current situation is that the measurements are multiple times, but point-wise in space. This circumstance seriously complicates the assessment of emerging situations, the forecast of the development of the situation associated with the movement of the flood down the river. The structure of data preparation for processing in the Intelligent Situation Center and the scheme for processing multilevel data for searching for changes and assessing the situation are given.

**Key words:** information technologies, management solutions, support decision, intellectual situation center, river floods

**Введение.** Современные информационные технологии достаточно широко используются в различных областях экономики, управления территориями и социальной сферы. Однако, далеко не всегда использование информационных технологий является эффективным и оправдывает затраченные средства. Так, например, использование ситуационных центров (СЦ) для выработки управленческих решений в условиях возникновения и развития неблагоприятных природных процессов и явлений (весенних паводков на реках, лесных пожаров и др.) не всегда приводит к выработке своевременных и эффективных управленческих решений. Для повышения оперативности и эффективности выработки управленческих решений и прогнозирования авторами предлагается использование нового типа информационных систем – интеллектуальных ситуационных центров (ИСЦ) [1].

**Постановка задачи.** Центральной задачей интеллектуального ситуационного центра является мониторинг состояния контролируемой территории и выявление изменений, которые требуется зафиксировать, принять соответствующие решения и выработать прогнозы. В случае с паводками к таким изменениям относятся резкое увеличение количества осадков, резкое увеличение уровня воды, подтопление территорий, разрушение гидротехнических сооружений и др. Процессы принятия решений и тем более прогнозирования развития ситуаций в ИСЦ являются достаточно сложными для автоматизации, так как связаны с многокритериальным оцениванием с учетом множества данных, условий и требований. Выбор того или иного решения или прогноза выполняется на основе данных, получаемых из всевозможных источников (априорные данные, данные наземных измерений и наблюдений, данные космической съемки, данные съемки с БПЛА и др.). Очевидно, различные типы данных требуют специфического подхода к их обработке и анализу. Алгоритм поиска решения