

بسم الله الرحمن الرحيم



HOSSAM MAGHRABY



شبكة المعلومات الجامعية التوثيق الالكتروني والميكرو فيلم



HOSSAM MAGHRABY

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها
على هذه الأقراص المدمجة قد أعدت دون أية تغيرات



يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار



HOSSAM MAGHRABY



**COMPUTER SCIENCE DEPARTMENT
FACULTY OF COMPUTER AND INFORMATION
SCIENCES AIN SHAMS UNIVERSITY**

**Computational Intelligence Techniques in Music
Composition**

A thesis submitted to the Department of Computer Science,
Faculty of Computer and Information Sciences,
Ain Shams University, Cairo, Egypt.

In partial fulfillment of the requirements for the degree of Doctor of
Philosophy, Ph.D. in Computer Science

By

Nermin Naguib Jean Siphocly

Assistant Lecturer in the Department of Computer Science,
Faculty of Computer and Information Sciences,
Ain Shams University, Cairo, Egypt.

Under the Supervision of

Prof. Dr. Abdel-Badeeh M. Salem

Professor of Computer Science,
Faculty of Computer and Information Sciences,
Ain Shams University, Egypt.

Prof. El-Sayed M. El-Horbaty

Professor of Computer Science
Faculty of Computer and Information Sciences,
Ain Shams University, Egypt.

May 2021
Cairo, Egypt

Acknowledgments

“Thus far the LORD has helped us.”

My deepest gratitude goes to God who helped me to get this far in my Ph.D. journey, feeling His presence in the hard as well as the good times.

I wish to express my sincere appreciation for my supervisors: Prof. Abdel-Badeeh M. Salem and Prof. El-Sayed M. El-Horbaty for their continuous efforts, support, encouragement, and guidance.

Last but not least, profound thanks go to my family; my parents and my husband, for always being in my back and for their continuous and unconditional love.

Nermin Siphocly

May 2021

Abstract

Engaging computers in composing musical pieces is a challenging and trending field of research. The musical tasks that can be performed or aided by computers' computational powers, are numerous. This thesis is concerned with computational intelligence techniques in music composition. Its main objective is to introduce various intelligent techniques for performing miscellaneous music composition tasks. To achieve this objective, the thesis first provides a thorough survey on the most famous artificial intelligence algorithms used in computer music composition discussing their applications, strengths, and weaknesses. The thesis then proposes multiple applications adopting some of the studied artificial intelligence and machine learning algorithms; including rule-based, case-based reasoning, artificial neural networks, and the relatively new: “generative adversarial networks”.

The contributions of this thesis include: First, providing a comprehensive survey on the field of computer music generation highlighting the most famous adopted algorithms, their most recent applications, their weaknesses, and strengths. Second, proposing an intelligent algorithm for major/minor melody conversion, comparing between rule-based and case-based reasoning in performing the task. This application also introduces a smart method for musical scale detection.

Third, developing an intelligent secondary melody generator with two techniques: artificial neural networks and case-based reasoning. Fourth, comparing between both techniques in performing the task of secondary melody generation. The comparison results show that case-based reasoning secondary melody generator outperformed the artificial neural networks generator by a success percentage of 50%.

Fifth, presenting a novel approach for accompaniment generation using pix2pix generative adversarial networks which is considered from the state-of-the-art in the machine learning area. Sixth, presenting a novel musical data representation which enhanced the pix2pix network training and the overall results. Specifically, this work suggests using color encodings to represent music notes inside images. Experimental results show that the proposed music representation achieved better results on pix2pix GANs over the traditional representations, reaching a loss function value of 0.001.

Seventh, studying the effectiveness of the proposed color encoded data representation showing that it outperformed the previously known representations. Eighth, proposing a post-processing technique on the generated images for enhancing the quality of the generated music, based on erosion. Ninth, to introducing two music evaluation metrics to automate the assessment of the generated music based on harmony and dissimilarity. Experimental results show that the proposed post-processing technique enhanced the musical harmony and the dissimilarity of the generated music by 51.26% and 81.98% respectively. Finally, listing multiple promising ideas for future work and research related to the field of computer music generation.

Table of Contents

| | |
|---|-------------|
| Acknowledgments | i |
| Abstract | ii |
| List of Tables | vii |
| List of Figures | viii |
| List of Abbreviations | xi |
| List of Algorithms | xii |
| List of Publications | xiii |
| Chapter 1 Introduction | 1 |
| 1.1 Classification of Computer Music Applications | 2 |
| 1.2 Thesis Problem Statement | 4 |
| 1.3 Thesis Objective | 5 |
| 1.4 Research Methodology | 6 |
| 1.5 Thesis Contributions | 6 |
| 1.6 Thesis Organization | 8 |
| Chapter 2 Background and Related Work | 10 |
| 2.1 Music Improvisation Systems | 10 |
| 2.2 Music Expressiveness Systems | 13 |
| 2.3 Music Composition Systems | 15 |
| 2.3.1 Music Composition Tasks | 15 |
| 2.3.2 Music Composition Applications | 16 |
| 2.4 Brief Musical Background | 18 |
| Chapter 3 Artificial Intelligence Algorithms in Computer Music Composition | 20 |
| 3.1 Rule-Based Systems | 21 |
| 3.1.1 Overview and Description | 21 |
| 3.1.2 Rule-Based Systems in Algorithmic Composition | 24 |
| 3.2 Case-Based Reasoning | 24 |
| 3.2.1 Overview and Description | 25 |
| 3.2.2 CBR in Algorithmic Composition | 26 |
| 3.3 Markov Chains | 27 |
| 3.3.1 Overview and Description | 27 |
| 3.3.2 Markov Chains in Algorithmic Composition | 31 |
| 3.4 Generative Grammars | 33 |
| 3.4.1 Overview and Description | 33 |
| 3.4.2 Generative Grammar in Algorithmic Composition | 36 |
| 3.5 Linear Programming | 38 |
| 3.5.1 Overview and Description | 38 |
| 3.5.2 Linear Programming in Algorithmic Composition | 40 |
| 3.6 Genetic Algorithms | 41 |

| | |
|--|------------|
| 3.6.1 Overview and Description | 42 |
| 3.6.2 GAs in Algorithmic Composition | 44 |
| 3.7 Artificial Immune Systems | 47 |
| 3.7.1 Overview and Description | 48 |
| 3.7.2 AIS in Algorithmic Composition | 50 |
| 3.8 Artificial Neural Networks | 52 |
| 3.8.1 Overview and Description | 52 |
| 3.8.2 ANNs in Algorithmic Composition | 57 |
| 3.9 Deep Neural Networks..... | 58 |
| 3.9.1 Overview and Description | 58 |
| 3.9.2 DNNs in Algorithmic Composition | 60 |
| 3.10 Generative Adversarial Networks | 63 |
| 3.10.1 Overview and Description | 63 |
| 3.10.2 GANs in Algorithmic Composition | 65 |
| 3.11 Comparison and Discussion of AI Algorithms in Music Composition..... | 69 |
| 3.12 Summary | 76 |
| Chapter 4 Proposed Intelligent Techniques for Automating the Conversion between Major and Minor Melodies..... | 78 |
| 4.1 Methodology | 79 |
| 4.1.1 rule-based system in major/minor conversion | 79 |
| 4.1.2 CBR in Major/Minor Conversion | 83 |
| 4.2 Implementation..... | 88 |
| 4.3 Experimental Results | 89 |
| Chapter 5 Proposed Intelligent techniques for Automating Secondary Melody Music Generation..... | 93 |
| 5.1 Methodology | 94 |
| 5.1.1 Secondary Melody Generation using Artificial Neural Networks and Adam Algorithm. | 94 |
| 5.1.2 CBR in Secondary Melody Generation..... | 95 |
| 5.2 Implementation..... | 98 |
| 5.3 Experimental Tests and Results | 98 |
| 5.3.1 Experimentation with ANN parameters..... | 98 |
| 5.3.2 ANN versus CBR Models Experimental Results | 100 |
| 5.4 Discussion..... | 102 |
| Chapter 6 Proposed Intelligent Technique Based on Deep Learning for Accompaniment Music Generation | 105 |
| 6.1 Methodology | 106 |
| 6.1.1 Data Representation | 106 |
| 6.1.2 Network Architecture | 108 |
| 6.1.3 Post-Processing | 110 |
| 6.1.4 Music Evaluation Metrics..... | 112 |
| 6.1.5 Data Representation Variation | 114 |

| | |
|---|------------|
| 6.2 Implementation..... | 116 |
| 6.3 Experimental Tests and Results | 117 |
| 6.3.1 Proposed Color Encoded Versus B/W Data Representation | 117 |
| 6.3.2 Original Versus Dilated Data Representation | 119 |
| 6.4 Discussion..... | 121 |
| Chapter 7 Conclusions and Future Works | 123 |
| References | 128 |

List of Tables

| <i>Number</i> | <i>Page</i> |
|--|-------------|
| Table 2.1 - Summary of AI techniques used in improvisation..... | 11 |
| Table 2.2 - Summary of AI techniques used in expressiveness..... | 14 |
| Table 3.1 - Comparison of generative grammars | 35 |
| Table 3.2 - Comparison between absolute and relative pitch..... | 45 |
| Table 3.3 - Comparison between the top 10 AI algorithms in computer music | 67 |
| Table 5.1 - Minimum error for each network set-up | 99 |
| Table 5.2 - ANN and CBR generators test results..... | 101 |
| Table 6.1 - Comparison between original and dilated representations without post- processing..... | 120 |
| Table 6.2 - Comparison between original and dilated representations with post- processing..... | 120 |

List of Figures

| <i>Number</i> | <i>Page</i> |
|---|-------------|
| Figure 1.1 - Classification of computer music generation systems | 3 |
| Figure 2.1 - Piano | 19 |
| Figure 2.2 - (a) C Major scale notes, (b) C Minor scale notes | 19 |
| Figure 3.1 - Forward chaining algorithm..... | 22 |
| Figure 3.2. Backward chaining algorithm | 23 |
| Figure 3.3. CBR cycle diagram..... | 25 |
| Figure 3.4. (a) Markov state space graph example - (b) Markov transition matrix example..... | 28 |
| Figure 3.5. Forward algorithm (adapted from [59])..... | 30 |
| Figure 3.6. Viterbi algorithm (adapted from [59]) | 30 |
| Figure 3.7. Baum-Welch algorithm (adapted from [59]) | 31 |
| Figure 3.8 - Derivation of a generative grammar | 35 |
| Figure 3.9 - Chomsky hierarchy | 35 |
| Figure 3.10. Graph of linear programming model for a two decision variables problem | 39 |
| Figure 3.11. Simplex algorithm (basic version)..... | 40 |
| Figure 3.12. The genetic algorithm | 42 |
| Figure 3.13. Examples on mutation and crossover | 43 |
| Figure 3.14. Negative selection algorithm (reproduced from [76]) | 49 |

| | |
|--|-----|
| Figure 3.15. Clonal selection algorithm (reproduced from [76]) | 50 |
| Figure 3.16. Artificial neuron (adapted from [80]) | 52 |
| Figure 3.17. Backpropagation algorithm (adapted from [82]) | 56 |
| Figure 3.18. Convolutional neural network architecture (adapted from [86]) | 59 |
| Figure 3.19. Convolution pseudocode (adapted from [87])..... | 60 |
| Figure 3.20. Generative adversarial network architecture (adapted from [93])..... | 63 |
| Figure 4.1 - Rule-based major/minor conversion algorithm..... | 82 |
| Figure 4.2 - Case adaptation visual example | 85 |
| Figure 4.3 - Case-based major/minor conversion algorithm. | 85 |
| Figure 4.4 - (a) Rule-based versus case-based major/minor conversion (b) Rule-based versus case-based conversion error rate per scale | 90 |
| Figure 4.5 – Major to minor melody conversion sample with CBR..... | 91 |
| Figure 5.1 - CBR secondary melody generation pseudocode. | 97 |
| Figure 5.2 - Training loss graphs relevant to the network parameters changes..... | 100 |
| Figure 5.3 - Average error of mismatches..... | 102 |
| Figure 5.4 - Secondary melody generation sample with CBR..... | 103 |
| Figure 6.1 - Sample of main melody represented as an image..... | 107 |
| Figure 6.2 - A magnified sample of the color encoded accompaniment music..... | 108 |
| Figure 6.4 - Architecture of the generator network..... | 109 |
| Figure 6.3 - Architecture of the discriminator network..... | 110 |
| Figure 6.5 - Sample piano-roll of the generated accompaniment music | 111 |
| Figure 6.6 - The 3X3 structuring element used in the proposed post-processing technique..... | 112 |

Figure 6.7 - The result of applying the proposed post-processing technique112

Figure 6.8 - Snapshot of the original proposed data representation (left) and its dilated variation (right).115

Figure 6.9 - The loss curves for pix2pix GANs trained with b/w images (left column) versus the proposed coloring schemed images (right column)117

Figure 6.10 - Samples from the output of the pix2pix training with b/w representation (upper row) versus the proposed color encoded representation (lower row).119

List of Abbreviations

| | |
|------|--|
| AI | Artificial Intelligence |
| AIS | Artificial Immune System |
| AN | Artificial Neuron |
| ANN | Artificial Neural Network |
| BM | Basis function Model |
| BN | Bayesian Network |
| CBR | Case-Based Reasoning |
| CNN | Convolutional Neural Network |
| CRBM | Conditional Restricted Boltzmann Machine |
| CRBM | Convolutional Restricted Boltzmann Machine |
| DFA | Deterministic Finite state Automaton |
| DNN | Deep Neural Network |
| FO | Factor Oracle |
| GA | Genetic Algorithm |
| GAN | Generative Adversarial Network |
| HMM | Hidden Markov Model |
| IE | Inference Engine |
| KB | Knowledge Base |
| LSTM | Long Short-Term Memory |
| NFA | Non-deterministic Finite state Automaton |
| PTGG | Probabilistic Temporal Graph Grammar |
| RNN | Recurrent Neural Network |
| TIS | Tonal Interval Space |
| WCL | Weighted Centroid Localization |
| WM | Working Memory |