

Aplicação de redes neurais para auxílio nas composições musicais utilizando compassos como primitivas e inspiração em relevos naturais

Débora Cristina Corrêa¹, José Hiroki Saito¹

¹Departamento de Computação – Universidade Federal de São Carlos (UFSCAR)
Caixa Postal 676 – 13.565-905 – São Carlos – SP – Brasil

{debora_correa, saito}@dc.ufscar.br

Abstract. *This paper describes a neural network application for helping musical composition using as inspiration the landscapes contour. During the training phase, the neural network learns certain aspects of musical structure by means of measure examples taken from melodies of the training set. The chosen melodies for training are folk melodies, once they are simple and monophonic. During composition, the system uses the measures learned to compose new melodies using as input the extracted data of the landscapes contour. A preliminary version of this system was implemented and the results obtained using one hundred measures.*

Resumo. *Esse artigo descreve uma aplicação de redes neurais para auxílio nas composições musicais utilizando como inspiração as ondulações dos acidentes geográficos. Durante a fase de treinamento, a rede neural aprende certos aspectos da estrutura musical através de exemplos de compassos extraídos de melodias do conjunto de treinamento. As melodias escolhidas para o treinamento são melodias folclóricas, por serem simples e monofônicas. Durante a composição, o sistema utiliza o repertório de compassos usados no treinamento para compor novas melodias, usando como entrada dados extraídos de relevos naturais. Uma versão preliminar desse sistema foi implementada e os resultados foram obtidos para 100 compassos.*

1. Introdução

As Redes Neurais Artificiais (RNAs), também conhecidas como sistemas conexionistas, constituem uma forma de computação não-algorítmica inspirada na estrutura e processamento do cérebro humano. A computação é realizada por várias unidades de processamento simples, os neurônios, conectados em rede e atuando em conjunto. As RNAs passam por um processo de aprendizagem que consiste em receber um conjunto de treinamento, para a extração de características necessárias para representar a solução desejada. Nesse procedimento, as redes neurais artificiais adaptam suas interconexões até que os padrões de excitação desejados estejam próximos do comportamento desejado. [HAYKIN, 1999] [BRAGA, LUDERMIR, CARVALHO, 2000]

Para a composição musical, os modelos conexionistas, assim como os outros modelos que envolvem aprendizado de máquina, são capazes de aprender padrões e características presentes nas melodias do conjunto de treinamento e obter

generalizações dessas características para a composição de novas melodias. [TODD, 1989] [MOZER, 1994] [CHEN, MIKKULAINEN, 2001] [ECK, SCHMIDHUBER, 2002]

O presente trabalho consiste no uso de RNAs para o aprendizado de compassos de melodias, usando como entradas, dados extraídos de relevos naturais, e como saídas as amostras de compassos de melodias. Para a composição, novos dados de relevos são apresentados à rede para a geração de novas melodias.

O artigo está organizado como segue: a seção 2 consiste na descrição do projeto proposto e a seção 3 consiste nos resultados obtidos e conclusão do trabalho.

2. Descrição do Trabalho

Além do conhecimento musical, emoções e intenções, o compositor geralmente conta com uma inspiração para o desenvolvimento de uma nova melodia. Um dos objetivos desse trabalho é adicionar ao processo de composição uma forma de inspiração, proveniente da Natureza. Propõe-se que as novas composições sejam inspiradas na Natureza através da utilização das ondulações dos acidentes geográficos durante a fase de treinamento e composição. Na fase de treinamento, a rede neural deve aprender os compassos de melodias escolhidos pelo usuário e, na fase de composição, deve ser capaz de gerar novas melodias com compassos previamente treinados.

O trabalho considera, em princípio, que um compasso consiste numa primitiva melódica que possui condições de se encadear com outros compassos para a formação de uma melodia. Para isso, a restrição no treinamento seria a escolha dos compassos de um mesmo estilo musical.

Na implementação preliminar relativa ao presente trabalho, as notas musicais são representadas por suas alturas e durações. As notações rítmicas estão limitadas em colcheia (♩), semínima (♪) e mínima (♫). As pausas não são usadas. Os compassos são representados pelas notas e pelos seus atributos de duração e formados de modo que possuam quatro tempos cada. Para isso, a colcheia é a unidade de tempo. Assim, a semínima e mínima valem dois e quatro tempos, respectivamente.

As alturas das notas são representadas pela combinação de números inteiros e intervalos musicais. A cada nota é atribuído um número inteiro (Figura 1).

Cada intervalo determina uma distância de frequência de uma nota para a outra, em semitons (utilizado nesse trabalho) ou em frequências logarítmicas. Mesmo com um número fixo de neurônios, vários intervalos de notas podem ser atingidos com essa representação. Além disso, essa representação é menos dependente de uma determinada tonalidade e pode descobrir estruturas comuns entre melodias.

O intervalo entre duas notas s e t pode ser determinado pela equação 1:

$$\text{intervalo}(s, t) = t - s \quad (1)$$

em que s e t são os valores inteiros que representam as notas musicais.

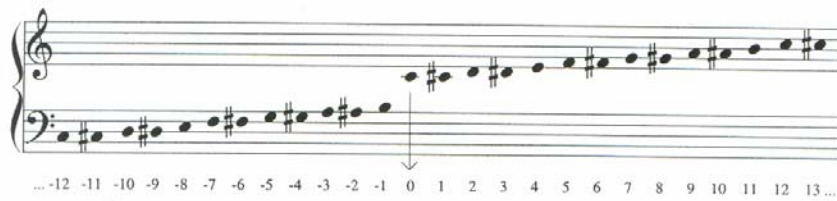
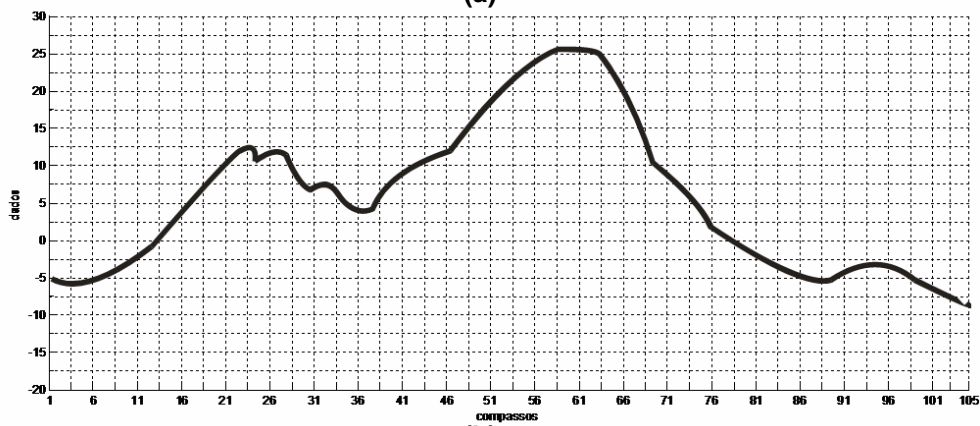


Figura 1. Representação das notas musicais por valores inteiros

Os dados que representam os padrões de entrada são codificados a partir do contorno de ondulações geográficas. A Figura 2 apresenta um exemplo dessa codificação.



(a)



(b)

Figura 2. Representação da entrada da rede

O contorno das ondulações é extraído de uma imagem e convertido para valores numéricos inteiros como mostra o gráfico da Figura 2 (b).

A rede utilizada na versão preliminar é uma MLP (*Multi-Layer Perceptron*), fortemente conectada, com 4 entradas de codificações de ondulações geográficas; 4 saídas de altura das notas e 4 saídas de atributos de duração, treinada com o algoritmo *Back-Propagation*. [HAYKIN, 1999]

3. Resultados e Conclusão

Como exemplo de aplicação, foram usados 100 compassos extraídos de 11 músicas folclóricas brasileiras como, por exemplo, “Samba lêlê”, “Escravos de Jó”, “Marcha Soldado”, para o treinamento da rede neural MLP.

Usando taxa de aprendizagem, $\alpha = 0.001$, no algoritmo *back-propagation* padrão, um número médio de 500 mil iterações foi suficiente para o treinamento de todos os compassos. Os neurônios da camada escondida e das quatro últimas saídas (neurônios de atributos de duração) usaram funções de ativação sigmóide, enquanto que os quatros primeiros neurônios de saída (notas), funções lineares.

Durante a composição, quando o padrão de entrada coincidiu com um padrão usado para treinamento a saída resultou no compasso treinado, conforme esperado. Para os padrões de entrada com os valores próximos aos padrões de entrada de treinamento, as saídas são próximas aos padrões desejados, revelando assim, a generalização da rede para composições de novas melodias.

O sistema assim constituído pode ser aperfeiçoado com a inclusão de outros atributos para as saídas, como dinâmicas de intensidade e andamento e pausas. Além disso, as notas podem ser polifônicas, a unidade de tempo diferente da colcheia (para utilização de outras notações rítmicas) e os compassos podem ter uma quantidade diferente de tempos.

O sistema é útil para a obtenção automática de melodias para que o usuário possa fazer sua análise posterior.

Como trabalhos futuros, além da implementação dos aperfeiçoamentos já descritos, deve-se incluir os algoritmos de análise dos resultados da composição, melhorando a capacidade de generalização da rede.

4. Referências Bibliográficas

BRAGA, A. P.; LUDERMIR, T. B.; CARVALHO, A. P. “Redes Neurais Artificiais – Teoria e Aplicações.”, Rio de Janeiro - RJ, LTC, 2000.

CHEN, C. C. J.; MIIKKULAINEN, R. (2001) “Creating Melodies with Evolving Recurrent Neural Network.” In Proceedings of the International Joint Conference on Neural Networks, IJCNN’01, 2001, p. 2241 – 2246, Washington - DC.

ECK, D.; SCHMIDHUBER, J. (2002) “A First Look at Music Composition using LSTM Recurrent Neural Networks.” Technical Report: IDSIA-07-02.

HAYKIN, S. “Neural Networks – A Comprehensive Foundation.” Prentice Hall, USA, 1999.

MOZER, M. C. (1994) “Neural network music composition by prediction: Exploring the benefits of psychoacoustic constraints and multiscale processing.” Connection Science, 6(2-3), p. 247 – 280.

TODD, P. M. A (1989) “Connectionist Approach to Algorithmic Composition.” Computer Music Journal: Vol.13, No. 4.