**REDES NEURAIS APLICADAS À MÚSICA PARA O TREINAMENTO DE UM OUVIDO ABSOLUTO ARTIFICIAL**

Carlos Henrique Alvarenga Dos Santos <carlosho200@hotmail.com>

**1 INTRODUÇÃO**

Para quem faz uso de algum instrumento musical já se deparou com o seguinte problema: Ir a Internet a procura das notações que representam as partituras e não encontrar. Além disso a mesma pode estar incorreta.

Isso provoca uma grande perca de tempo que poderia ser usado para praticar. Muitas das vezes pode causar frustração no que era para ser um lazer ou diversão.

No mundo é possível encontramos pessoas que possuem um tipo de ouvido chamado absoluto onde a percepção de qual nota está sendo tocada no momento é instantânea, o que facilitaria o processo citado acima. Essa técnica pode ser natural ou pode-se aprimorar algo parecido que é o ouvido relativo, porém pode levar anos e gerar esforço exacerbado sendo apropriada para músicos profissionais.

Diante disso proponho a criação de uma aplicação que funcionará como um ouvido absoluto detectando a composição de uma música (melodia, harmonia e ritmo) em tempo real. Ao final este produzirá a partitura. Isso é possível pois um computador é capaz de receber, processar e dar uma saída, ou seja, ele funciona como uma extensão do ser humano.

Segundo Malvino, o uso do computador para a execução de processamento do sinal de som é quase que fundamental. O computador não apenas processa informações matemáticas relativas ao sinal, mas também pode capturar o sinal sonoro transformando-o em informação digital (MALVINO, 1985).

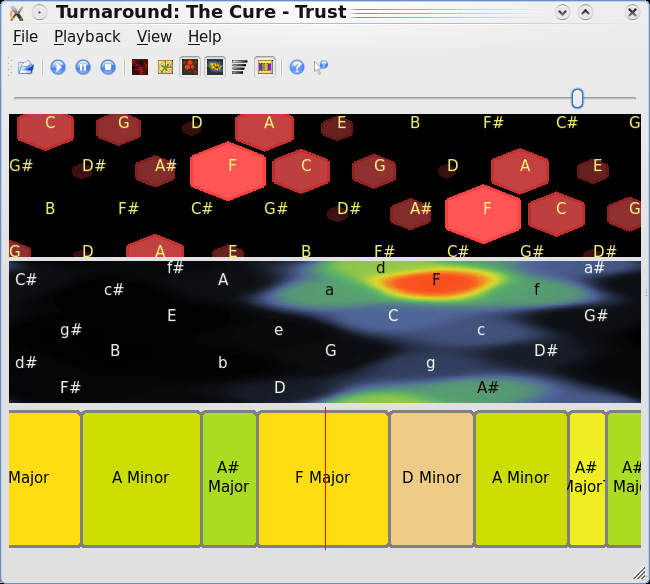
Ao final existirá um ouvido absoluto artificial que produzirá um arquivo com as notações que representam as partituras ou cifras ao qual chamarei de descriptografia neste artigo.

Em relação a softwares relacionados são poucos, seja código aberto ou fechado, que possuem eficiência. Diante disso serão apresentados dois que são: Chordata e Sonic Visualiser ambos para computadores de mesa.

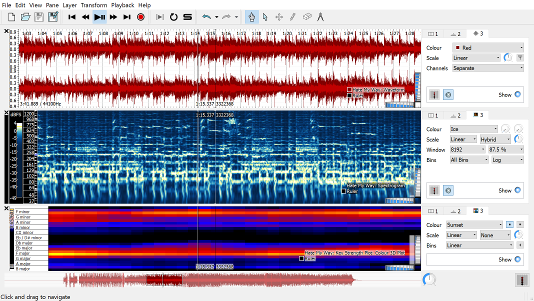
Chordata é uma aplicação desenvolvida por Pawel Bartkiewicz como seu projeto do Google Summer of Code utilizando a biblioteca CLAM escrita C++, onde é possível através de um arquivo de música descobrir os acordes da música seja ela somente instrumental ou com voz. A aplicação existe desde 2008 porém foi em 2010 o lançamento da sua primeira versão estável.

Para Pawel, Chordata é um aplicativo simples mas poderoso que analisa os acordes de qualquer arquivo de música no seu computador. Você pode usá-lo para voltar e avançar a música enquanto assiste a visualizações perspicazes dos recursos tonais da música. As combinações de teclas e as interações do mouse para a navegação de músicas são projetadas pensando em um músico com um instrumento em mãos (BARTKIEWICZ, 2010).

Esta é a interface do Chordata que pode ser baixado no link a seguir: <http://clam-project.org/index.html>



Diferentemente o Sonic Visualiser é completo e dispõe diversas funcionalidades a mais do que o Chordata. Ele é distribuído pela GNU General Public License e possui versões para Windows, Mac OS/X e Linux. Foi desenvolvido no Centro de Música Digital da Queen Mary, Universidade de Londres, sendo sua primeira versão lançada em 2011. E possui o funcionamento igual a o anterior onde é carregado um arquivo de música para a aplicação e onde é feito o processamento. Este é a interface do Sonic Visualiser que pode ser baixado no link a seguir: <https://sonicvisualiser.org/index.html>



**2 RESULTADOS E DISCUSSÕES**

É conhecido como música todo som que possui em sua composição a melodia, harmonia e ritmo. Ela está presente na sociedade antes mesmo de Cristo, existem evidências desde a pré-história onde foi encontrada uma flauta feita de osso.

“Segundo Wolfflin e Junior, a música em si, além de ter em sua essência todas as leis físicas da ondulatória sonora, ela é uma forma de arte no que se refere a apresentação estética e do belo” (WOLFFLIN; JUNIOR, 2000, p. 33).

Como mencionado anteriormente o software funcionará baseado nas pessoas que possuem o ouvido absoluto e produzirá ao final a descriptografia. Para Matras, o processo de captura do som é realizado através do uso de placas de som que convertem o sinal analógico em digital, para que seja possível manipulá-lo através de sistemas computacionais. Para converter o som em um sinal digital, é necessário, antes, que o som seja convertido primeiro em um sinal elétrico (MATRAS, 1991).

Para a aplicação do que foi proposto será utilizado o aprendizado de máquina redes neurais que simula o cérebro humano. Isso se faz necessário pois os computadores são burros e não fazem a diferenciação entre voz, acorde ou marcação de tempo, ou seja, não possuem regras e nem conhecimento sobre o assunto.

Segundo Zacan para que as RNA(Redes Neurais Artificiais) possa ser utilizadas na solução de alguma tarefa, é necessário que elas passam inicialmente por uma fase de aprendizagem, onde a rede extrai as informações relevantes dos padrões de informações apresentados para ela, criando assim uma representação própria para o problema (ZANCAN, 2010).

Operacionalmente podemos considerar uma rede neural como uma “caixa de processamento” que pode ser treinada para que, a partir de um conjunto de dados de entrada (inputs), possa gerar uma ou mais saídas (outputs). Por exemplo, pode-se treinar uma rede neural para que a partir do espectro de RMN-H (input) de um composto orgânico ela apresente como saída (output) a estrutura molecular do mesmo, ou ainda os inputs podem ser dados clínicos de um paciente e os outputs o diagnóstico e tipo de evolução. Enfim, as aplicações são extremamente variadas e estão presentes em quase todas as áreas do conhecimento humano (CERQUEIRA EDUARDO O DE ANDRADE AND, 2001, p. 864).

Já para Treleaven, as redes neurais Artificiais são dispositivos não lineares, inspirados na funcionalidade dos neurônios biológicos, aplicados no reconhecimento de padrões na otimização e na previsão de sistemas complexos (P.TRELEAVEN M.VELLASCO, 2000, p. 27).

Uma das áreas de pesquisa mais fascinante presentemente é a simulação de capacidades cognitivas de um ser humano. Projetam-se máquinas capazes de exibir um comportamento inteligente, como se fossem reações humanas. A inteligência do ser humano é a mais avançada dentro do universo das criaturas e o local dessa inteligência dentro do corpo humano é o cérebro. As entidades básicas são os neurônios, interconectados em redes o que permite a troca de informação entre eles, criando a inteligência biológica (RAUBER, 2005, p. 3).

A grande vantagem da utilização deste mecanismo é que a mesma pode ser utilizada para aprendizagem de quase tudo, pois já dizia Mendel:

A aprendizagem realizada pela RNA, pode ser definida como um processo pelo qual seus parâmetros são ajustados através de uma forma continuada de estímulo pelo ambiente no qual a rede está operando, sendo o tipo específico de aprendizagem realizada definido pela maneira particular como ocorrem os ajustes realizados nos parâmetros (MENDEL, 1970, p. 318).

Em contrapartida isso não quer dizer que seja tão eficiente quanto o cérebro biológico humano assim como disse Paula:

Muitas são as dificuldades quando se tenta imitar o funcionamento do sistema nervoso humano, pois a complexidade das estruturas das Redes Neurais Biológicas é muito maior do que os modelos matemáticos usados nas RNA. Sendo assim, uma RNA pode ser representada por um conjunto de várias unidades interconectadas, denominadas de neurônios artificiais, cada um contendo uma porção local de memória (PAULA, 2000, p. 108).

Mas as ferramentas da neurocomputação já foram mais arcaicas e iniciaram com notoriedade próximo ao fim da década de 50 com o Perceptron que foi um avanço mas possuía alguns erros de classificação.

Um ponto marcante na história das redes neurais artificiais foi a apresentação de um modelo de um neurônio artificial por McCulloch and Pitts em 1943. As atividades nessa linha de pesquisa culminaram na concepção do perceptron por Rosenblatt em 1958 (RAUBER, 2005).

Rosemblatt descreveu o perceptron como tendo três camadas: a primeira que recebe as entradas do exterior e possui conexões fixas (retina); a segunda que recebe impulsos da primeira através de conexões cuja eficiência de transmissão (peso)é ajustável e, por sua vez, envia saídas para a terceira camada (resposta) (ROSEMBLATT, 1958).

Com o passar das décadas a neurocomputação evoluiu e desenvolveu novas técnicas. Surgem as Redes Neurais e seu novo modo de aprendizado de máquina mais eficiente. O aprendizado das RNA pode ser feito através de vários métodos de treinamento, os quais podem ser divididos em dois principais paradigmas: o aprendizado supervisionado e o não supervisionado (PAULA, 2000).

Para Correa, no aprendizado supervisionado, a entrada e a saída são fornecidas por um supervisor externo, que indica explicitamente um comportamento bom ou ruim para a rede, com o objetivo de direcionar o processo de treinamento. Para cada entrada a rede produz uma resposta na saída que é comparada com o sinal de saída desejado e assim, a rede gera um sinal de erro que corresponde à diferença desses sinais. Este sinal de erro obtido é utilizado então, para calcular o ajuste necessário aos pesos sinápticos da rede, que são corrigidos até que a resposta da rede case com a saída desejada (CORREA, 1998).

Já no aprendizado não supervisionado, não há supervisor para acompanhar o processo de aprendizagem. Estas redes ajustam seus pesos através de um monitoramento interno da mesma, o qual analisa as regularidade e tendências dos sinais de entrada, adaptando assim, automaticamente as necessidades da rede (BRAGA, 2000).

Como foi falado o computador não sabe o que é uma música. Por esse motivo uma solução seria o refinamento do sinal.

Para Tafner, o sinal digital precisa passar por um processo de refinamento que o torna apropriado para construir o treinamento de uma RNA. Este processo tem como objetivo identificar os sinais mais significativos, diminuir a quantidade de sinais que irão compor a entrada da RNA e melhorar a representatividade do sinal em relação ao sinal original. Tafner propõe então quatro etapas para o refinamento do sinal: Eliminação do ciclo negativo do sinal amostrado; Redução do sinal amostrado detectando a forma de onda(envoltória); Mediação do sinal reduzido; Normalização do sinal mediado (TAFNER, 1996, p. 102).

Assim a diferenciação da entrada de dados será possível executando o processo e retornando a saída que no caso seria a descriptografia.

**3 CONSIDERAÇÕES FINAIS**

Diante do que foi apresentado o esperado é que este software seja capaz de interpretar música em tempo real, não importando o gênero, e ao final trazer o processamento em forma de partitura ou cifra. Além disso tal software possui a seguinte dificuldade: o fato do computador ter de imitar o ouvido humano, algo muito complicado do ponto de vista computacional. Portanto tal aplicação seria um grande avanço tecnológico.

**4 REFERÊNCIAS**

BARTKIEWICZ, P. Clam CLAM Chordata. 2010. Disponível em: <<http://clam-project.org/>>.

BRAGA, A. P. Redes Neurais Artificiais:Teoria e Aplicação. Rio de Janeiro: LTC, 2000.

CERQUEIRA EDUARDO O DE ANDRADE AND, J. C. d. P. Redes neurais e suas aplicações em calibração multivariada. Química Nova, v. 24, 2001.

CORREA, C. Detecção de ataques em redes de computadores. 1998.

MALVINO, A. P. Microcomputadores e microprocessadores. [S.l.]: McGraw-Hill, 1985.

MATRAS, J. J. O som. [S.l.: s.n.], 1991.

MENDEL, J. M. Adaptive, Learning, and Pattern Recognition Systems; Theory and Applications. New York: Academic Press, 1970. 318 p.

PAULA, M. B. Reconhecimento de palavras faladas utilizando redes neurais. Pelotas: [s.n.], 2000. 108 p.

P.TRELEAVEN M.VELLASCO, M. VLSI Architectures for Neural Networks. [S.l.: s.n.], 2000. v. 9.

RAUBER, T. W. Redes neurais artificiais. Universidade Federal do Espírito Santo, 2005. ROSEMBLATT, F. The perceptron:A probabilistic model for information storage and organization in the brain. New York: Academic Press, 1958.

TAFNER, M. A. Reconhecimento de palavras faladas isoladas usando redes neurais artificiais. RI UFSC, p. 102, 1996.

WOLFFLIN, H.; JUNIOR, J. A. Conceitos fundamentais da história da arte: o problema da evolução dos estilos na arte mais recente. [S.l.]: Martins Fontes, 2000.

ZANCAN, A. L. F. PROTÓTIPO DE UM SISTEMA DE RECONHECIMENTO DE PALAVRAS FALADAS UTILIZANDO REDES NEURAIS ARTIFICIAIS. [S.l.: s.n.], 2010.