

Relatório Parcial de Atividades de Iniciação Científica

Relativo ao período de Agosto de 2003 a Janeiro de 2004

Nicolau Leal Werneck

20 de Fevereiro de 2004

1 Identificação

Título do projeto:

Análise da influência de características acústicas de instrumentos em diferentes contextos musicais

Bolsista: Nicolau Leal Werneck – 9916164

Curso: Engenharia Elétrica

Instituição Financiadora: CNPq (pelo PIBIC)

Ingresso no programa: Agosto de 2002. Renovação em Agosto de 2003.

Orientador: Hani Camille Yehia, Dr.

Departamento: Engenharia Eletrônica

Unidade: Escola de Engenharia da UFMG

Endereço para correspondência:

Departamento de Engenharia Eletrônica

Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627,

31270-010 Belo Horizonte, Fax. 3499-4850, Fone: 3499-4868,

E-mail: hani@cpdee.ufmg.br

“Nem um pouco de inteligência nem imaginação
nem os dois juntos são necessários para ser genial.
Amor, amor, amor, esta é a alma da genialidade.”

— **Wolfgan Amadeus Mozart**

“Em duas ocasiões eu já fui questionado por membros do Parlamento:
'Diga, sr. Babbage, se você colocar números errados na sua máquina,
ela irá dizer as respostas certas?' Eu não sou capaz de compreender
exatamente que espécie de confusão de idéias provocaria tal questão”

— **Charles Babbage**

2 Introdução

2.1 Resumo da proposta de pesquisa

Este é um projeto de pesquisa da área de análise musical. Como outras pesquisas da área, tenta-se encontrar modelos que expliquem como características subjetivas que ouvintes e compositores atribuem a músicas se manifestam na concretização física da mesma.

O objetivo do trabalho a longo prazo é compreender os mecanismos da criação e apreciação artística em termos científicos e matemáticos. Existem duas áreas da engenharia que possuem uma estreita relação com estes problemas: a inteligência artificial, e a área de desenvolvimento de ferramentas para produção artística.

A Engenharia Elétrica possui diversas ferramentas matemáticas que podem ser empregadas neste tipo de estudo aliadas a experimentos psicofísicos. Este trabalho tem por objetivo não só estudar a arte de “alto nível”, relacionada à inteligência artificial, como também o funcionamento de aparelhos de síntese e processamento musical.

2.2 Resumo do plano de trabalho

As primeiras atividades sugeridas na proposta da bolsa fora:

- Realização de uma pesquisa bibliográfica.
- Estudo de procedimentos utilizados em testes psicométricos.
- Estudo de algoritmos de composição automática.
- Criação e validação de modelos computacionais de composição automática.

2.3 Participação em eventos

No período abrangido por este relatório o bolsista participou da Semana de Iniciação Científica da UFMG, e também do SIBGRAPI, Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens, que aconteceu no começo de Outubro.

2.4 Resumo das atividades empreendidas

Optou-se por adiar o estudo de composição automática. Ao invés disso foi decidido que seria realizada uma experiência envolvendo somente a capacidade de percepção de distorção musical, como descrito mais à frente.

A pesquisa focou-se no funcionamento de duas distorções musicas utilizadas largamente em guitarras elétricas hoje em dia: o *overdrive* ou saturação, e a reverberação. Modelos matemáticos foram criados para estes efeitos, e também foram feitas implementações computacionais destes.

3 Pesquisa Bibliográfica

No final do ano de 2003 nosso grupo de pesquisa aceitou sugestões de compras de livros. Os seguintes livros foram sugeridos e adquiridos. Eles possuem ligação direta ou indireta com o tema da pesquisa:

- Virtual Music: Computer Synthesis of Musical Style, David Cope

- The Cognition of Basic Musical Structures, David Temperley
- The Computational Beauty of Nature: Computer Explorations of Fractals, Chaos, Complex Systems, and Adaptation, Gary William Flake
- The Mind's I: Fantasies and Reflections on Self & Soul, editado por Douglas R. Hofstadter
- Nexus: Small Worlds and the Groundbreaking Theory of Networks, Mark Buchanan
- Self-Organized Criticality : Emergent Complex Behavior in Physical and Biological Systems, Henrik Jeldtoft Jensen
- Digital Filters, Richard Hamming
- Numerical Methods for Scientists and Engineers, Richard Hamming

4 Pesquisa de ferramentas

4.1 Síntese musical

Inicialmente imaginou-se que seria possível já realizar o experimento em cima de músicas sintetizadas por computador a partir de arquivos MIDI de performances de músicos. Isto facilitaria o futuro desenvolvimento de uma experiência envolvendo composição automática de arquivos deste tipo.

O programa *Csound* foi o primeiro candidato a sintetizador para ser utilizado. Entretanto descobriu-se apenas tarde demais que este programa ainda não suporta MIDI em tempo real em nossa plataforma de desenvolvimento.

Decidiu-se então abandonar este programa em nome de uma opção que se mostrou mais moderna ainda mais flexível, o programa *sfront*. Este programa é uma das primeiras aplicações que estão surgindo da codificação de áudio do padrão MPEG4. Esta codificação trata o sinal de forma estruturada, como acontece com as imagens nesse padrão.

Este programa já é capaz de criar um sintetizador de MIDI em tempo real, entretando funciona apenas para dispositivos reais. Será necessário acrescentar

no `sfront` a funcionalidade de ler um dispositivo virtual em linux.

4.2 Experimentos psicoacústicos

No semestre anterior foi constatado que a linguagem `tcl` sendo usada para controlar programas com interfaces simples é bastante eficiente.

Foi estudado nesse semestre a possibilidade de se criar uma interface em `tcl` para controlar programas sendo executados pelo interpretador `Octave`, similar ao `Matlab`, porém de código aberto.

5 Relatório Técnico-científico

5.1 Introdução e fundamentação

As análises musicais feitas neste projeto se separam em análises de timbre e de estrutura musical. Existem variáveis que já são utilizadas hoje como classificadores de timbre extremamente eficientes, como o centróide e fluxo espectrais, e tempo de ataque.

Os primeiros trabalhos em análise de timbre se focaram todos no relacionamento entre instrumentos individuais. Nosso trabalho segue uma tendência moderna em tentar enxergar como as variáveis de análise de timbre se comportam dentro de um mesmo instrumento quando este é utilizado de diferentes maneiras.

No semestre anterior foram feitas medidas quantitativas simples da influência da distorção no centróide espectral e na potência do sinal. Pretendia-se com estas medidas da potência fazer alguma análise da mudança do comportamento do som dos instrumentos no tempo, mas ainda não foi encontrado nada muito explícito.

Discute-se aqui também algumas conclusões tiradas acerca de estudos de reverberação.

5.2 Objetivos

Desejávamos realizar ao menos um experimento psicoacústico relacionado à pesquisa até março deste ano. Foi determinado que o melhor experimento que poderia ser feito seria a medida de uma espécie de limiar de percepção de diferença de distorção musical. Este tipo de estudo é o início do casamento da

análise física do instrumento com sua percepção por ouvintes.

5.3 Metodologia

5.3.1 Programas computacionais

Inicialmente foi criado um programa para fazer a saturação com nosso modelo matemático, porém em tempo real. O programa lê um arquivo de áudio e escreve o sinal distorcido na saída. Na prática a entrada é um *fifo* em que outro programa joga um sinal de áudio. Durante o funcionamento é possível passar para o programa os valores desejados para os parâmetros do modelo. Isto é feito a partir de *threads*.

Foram criados também um programa simples que faz a execução em laço de um sinal de áudio, e também um programa de compressão dinâmica. A compressão dinâmica busca manter a potência média do sinal constante. O programa simplesmente calcula a potência média, ou valor quadrático médio do sinal no último intervalo de tempo, e então multiplica o valor da amostra atual da seguinte forma:

$$y[n] = x[n] \times \frac{RMS_{desejado}}{RMS_{janela}}$$

O tamanho da janela em que se calcula o valor RMS é único parâmetro que pode ser ajustado desta maneira. Um jeito melhor de fazer esse tipo de processamento é ponderando-se o vetor das últimas amostras por alguma janela, implementando-se assim um filtro FIR. Isto reduziria o pequeno ruído introduzido pelo compressor.

Todo o sistema é controlado então por um ouvinte, que varia a intensidade de distorção com um mouse eu uma interface gráfica. O parâmetro que o ouvinte controlou neste experimento foi o limiar de distorção. O outro parâmetro, que é o valor a partir de onde o sinal é colocado, foi fixado no valor máximo de reprodução do sinal. A posição do controle do ouvinte era elevada à quarta potência antes de ser transmitida ao programa de distorção. Isto permitiu uma maior resolução de controle para o sujeito.

5.3.2 Medições psicométricas

Inicialmente imaginou-se que um experimento interessante seria ver a capacidade de igualar a distorção de um sinal com a de um outro apresentado anteriormente. Entretanto trabalhando-se com o programa de controle da distorção, percebeu-se que uma medida mais interessante poderia ser feita.

Partindo-se do sinal puro, conforme a distorção é aumentada, inicialmente existe uma região em que a distorção é perceptível, mas ainda não possui interesse musical. A partir de um certo ponto, a distorção passa a ficar interessante, e então ela começa a se tornar marcante demais, criando um ruído que se se destaca do instrumento.

Pedi-se para que ouvintes ouvissem continuamente à gravação de uma música tocada numa guitarra, e procurassem estes dois pontos de mudança. A gravação foi processada por um filtro passa-baixas para que o efeito provocasse mudanças mais bruscas.

5.4 Resultados alcançados e discussão

As medições foram realizadas com somente 3 ouvintes por enquanto. É preciso ainda conseguir gravações de melhor qualidade e sem muita variação de intensidade entre as notas.

Para generalizar os resultados, buscamos compreender os valores encontrados em relação ao valor RMS do sinal, e também em relação ao histograma do sinal original. O Histograma está na figura 1, e também é mostrada uma curva gaussina de variância igual. Nota-se que a gaussiana não é uma aproximação perfeita. Outras curvas ainda deveriam ser testadas.

Inicialmente imaginava-se que a região em que a saturação começaria a ser sentida seria em torno do ponto em que metade do histograma do sinal estivesse acima do limiar de distorção. Entretanto, para todos ouvintes o ponto em que a saturação foi considerada excessiva foi apenas em um 1% do valor máximo de reprodução, onde mais de 99% do sinal já estava acima do limiar. Neste ponto um ruído de fundo contínuo começa a ser percebido de forma distinta do sinal.

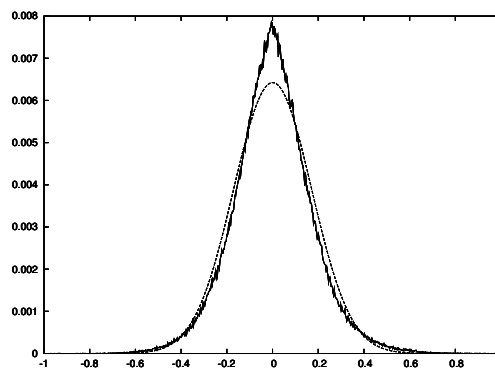


Figura 1: Histograma do sinal original, e aproximação por gaussiana.

O ponto em que os ouvintes julgaram que a distorção já estava funcionando de forma musical foi somente quando 95% do sinal se encontrou além do limiar de distorção, em 7% do valor máximo de reprodução.

Outro resultado importante e que era esperado, é que mesmo com a compressão dinâmica os ouvintes ainda perceberam um aumento na intensidade do som com a quantidade de distorção. Isto pode ser atribuído à questão da medida de *loudness*. Isto deve ser validado por outro experimento mais tarde.

6 Conclusões

6.1 Saturação

É interessante observar que o efeito de saturação não foi julgado excessivo pelos ouvintes, mesmo com quase todo o sinal acima do limiar de distorção. Isto mostra que a saturação sozinha não é um efeito poderoso o suficiente para desfigurar por completo o som de um instrumento musical.

Novos experimentos com saturação ainda deverão ser realizados sob melhores condições para que resultados mais importantes sejam alcançados. As seguintes idéias estão sendo analisadas:

- Uso de filtragem anterior e posterior à saturação.
- Confecção de um pequeno equipamento de controle para que os ouvintes regulem os parâmetros

de forma mais cômoda.

subseção Reverberação Quanto aos estudos feitos sobre a reverberação; em primeiro lugar percebeu-se que não era correto afirmar que a reverberação era um fenômeno muito próximo ao de um filtro IIR. Na realidade, apesar da resposta a impulso da reverberação ser assintótica e infinita, ela não possui a característica de realimentação dos filtros IIR convencionais. Talvez fosse adequado dizer que a reverberação provoca respostas IIR aperiódicas, ou algo similar.

Já havia sido constatado que a função de transferência sintetizada que melhor conseguiu emular um efeito de reverberação foi uma que possuía o valor absoluto constante e um ruído browniano como fase. A interpretação adequada deste fenômeno é a seguinte: A resposta a impulso da uma reverberação perfeita seria um ruído branco ponderado por uma curva assintótica, exponencial ou potencial. Qualquer uma destas curvas atuaria como um filtro passa-baixas atua no domínio da frequência, fazendo assim com que o sinal da função de transferência seja o sinal de um ruído branco filtrado, o que tipicamente se aproxima de um ruído browniano.

Também foi estudada que curva representaria melhor o decaimento da intensidade da resposta a impulso. Existem dois fatores que contribuem para a formação da resposta a impulso. Por um lado a intensidade do impulso cai com o inverso do tempo. Isto se vê facilmente, dado que a energia de um impulso em espaço aberto cai com o inverso da distância, a intensidade da onda varia com a raiz quadrada da energia, e o tempo é proporcional à distância. Por outro lado, os vários reflexos funcionam como se o ouvinte estivesse recebendo a soma de sinais recebidos por diversos reflexos espalhados de forma uniforme e aleatória pelo espaço, no caso de uma reverberação ideal.

Esta distribuição dos reflexos do ouvinte crescem com o quadrado da distância. Isto pode ser visto considerando que há uma densidade volumétrica de ouvintes, e calculando-se a quantidade de ouvintes existentes numa casca esférica. Temos portanto que a energia do impulso tende a se transformar em um sinal de potência constante com o tempo.

Isto, naturalmente é apenas um caso limite de uma reverberação em que não há nenhum tipo de perda de energia sonora, nem mesmo para energia térmica.

Esta análise serve para mostrar que não é óbvio que a envoltória das respostas a impulso de reverberações seja uma exponencial ou uma potencial simples. Uma análise mais aprofundada provavelmente revelará que se trata de uma soma de potenciais.