

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

“Todos nós fazemos o 'do, re, mi,' mas você tem que achar as outras notas por você mesmo.”

— **Louis Armstrong**

“É claro que existem muitas maneiras de se tratar o blues, mas continuará sendo o blues.”

— **Count Basie**

2 Relatório de Atividades

2.1 Resumo do plano de trabalho

Estava programado na proposta original do projeto a criação de programas para a criação de música automática, além de um período para documentação e divulgação do trabalho. O trabalho acabou tomando um outro rumo, entretanto, e preferiu-se dar uma atenção mais profunda ao funcionamento do efeito de *overdrive*. As principais atividades empreendidas foram portanto:

- Realização de um experimento psicoacústico relativo à capacidade de discriminação de parâmetros de *overdrive*.
- Análise da relação entre cadeias de Markov e ruídos $1/f^\alpha$

2.2 Resumo das atividades empreendidas

A principal atividade empreendida neste último período da bolsa foi um experimento psicoacústico relativo à capacidade de ouvintes em discriminar gravações processadas por diferentes parâmetros de *overdrive*. Também foram feitas análises mais aprofundadas na questão da natureza da reverberação como distorção de fase no domínio da frequência, e também foi feita uma averiguação da similaridade do uso de ruídos $1/f^\alpha$ e cadeias de Markov em composição automática.

3 Relatório Técnico-científico - 3.2 Objetivos

Overdrive

3.1 Experimentos psicoacústicos — Introdução e Fundamentação

“O conhecimento fala, mas a sabedoria ouve.”
— Jimi Hendrix

A distorção de *overdrive*, ou sobrecarga, está presente em caixas acústicas, amplificadores e pedais de efeito, e é frequentemente utilizada no meio musical contemporâneo como ferramenta de interesse artístico. Esta distorção pode ser controlada por diversos parâmetros, e oferece uma gama muito grande de novos timbres quando aplicada a qualquer instrumento musical.

Existe muito pouca pesquisa de caráter científico a respeito da percepção deste tipo de distorção na intensidade em que é utilizada pelos músicos. Geralmente o interesse pelo *overdrive* é apenas em sua percepção em baixa intensidade. Estes estudos estão relacionados com a avaliação de qualidade de amplificadores e caixas acústicas de alta fidelidade, que não devem exibir este efeito. Em nosso contexto, pelo contrário, o efeito é desejável.

Por ser um efeito não-linear, o funcionamento do *overdrive* é de difícil análise e interpretação. Como veremos a seguir, é possível entretanto realizar experimentos coerentes relativos a este efeito musical.

O estudo deste efeito possui um bom potencial para estudos científicos relevantes. Em primeiro lugar, porque já existe muito conhecimento informal sendo divulgado a respeito do efeito. Também é importante que o estudo seja desenvolvido de forma independente das empresas que fabricam o efeito. A ação de publicidade das empresas hoje em dia faz com que discussões sobre o tema envolvam mais enunciados publicitários do que deliberações científicas. O efeito também traz uma ferramenta interessante para o estudo de timbre, permitindo variar o timbre de um instrumento ou gravação de forma contínua, controlada e reprodutível, e dentro de um contexto musical existente e relevante.

O objetivo do experimento realizado era determinar a sensibilidade de um ouvinte a variações nos parâmetros de um efeito de *overdrive* sendo aplicado a diferentes gravações. Esta medida seria similar a medidas do tipo “just noticeable difference”, ou JND, para os parâmetros de *overdrive* do modelo utilizado aqui.

3.3 Metodologia

O modelo de *overdrive* utilizado foi o mesmo já desenvolvido anteriormente nesta pesquisa. Um parâmetro diz a partir de onde o sinal deve ser “cortado”, e o outro controla o ângulo do corte, sendo que a função de distorção é normalizada para que amostras de valor igual a ± 1.0 sejam mapeadas em amostras do mesmo valor.

Para a realização dos testes foi criado um programa na linguagem C++. Os testes foram realizados no laboratório do CEFALA, com o sinal sendo apresentado aos ouvintes por fones de ouvido, reproduzidos a partir da placa de som dos computadores. Todos sinais reproduzidos foram normalizados durante a reprodução para ficarem com a mesma potência média.

3.3.1 Funcionamento do Programa

Os parâmetros de entrada do programa são o nome de um arquivo com a amostra que será distorcida e apresentada, os valores dos parâmetros de referência e quantas apresentações devem ser feitas. O arquivo de entrada não deve possuir formatação — deve ser apenas uma seqüência de amostras de 16 bits, little-endian, de um único canal de áudio amostrado a $44100 \cdot s^{-1}$.

A cada iteração do laço principal do programa, um novo par de parâmetros é sorteado, e o ouvinte pode escolher entre ouvir o som distorcido com o par de parâmetros de referência, ou com o par de parâmetros sorteado. O sorteio foi feito de forma adaptativa, como explicado a seguir. Após decidir se a distorção causada pelos parâmetros sorteados

é similar ou não à causada pelos parâmetros de referência, o ouvinte pressiona o botão adequado, e uma linha contendo os parâmetros de referência, os sorteados e a resposta é emitida na saída padrão do programa. A interface gráfica criada pode ser vista na figura 1.

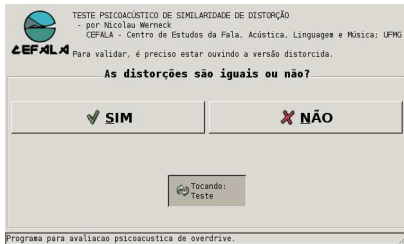


Figura 1: Interface do programa de testes.

3.3.2 Sorteio das amostras

A cada iteração do laço principal do programa, um novo par de parâmetros é sorteado. O sorteio é baseado em uma função de densidade de probabilidade — PDF. Foi criado um sistema para que a função se adaptasse às respostas dos testes, sorteando parâmetros coerentes com as respostas dos ouvintes. Não faz muito sentido sortear amostras com muita variância, por exemplo, se as respostas positivas se concentrarem muito próximas à referência. Foram feitas as seguintes considerações para o projeto deste sistema adaptativo:

- É desejável que um ouvinte ouça as amostras dando as duas respostas em iguais proporções. Isto evita que o sujeito fique “viciado” em alguma resposta. Isto também dá ao sujeito uma melhor percepção de que ele está sendo testado.
- É desejável que se façam muitas amostragens no limiar da região entre respostas positivas e negativas.

- Entretanto, também é desejável que eventualmente sejam sorteados parâmetros muito próximos e muito distantes dos de referência, para que se possa medir a largura da faixa de transição entre as regiões de similaridade e diferença.

A PDF utilizada para os sorteios precisava ser uma função paramétrica para que a distribuição pudesse ser adaptada às respostas. Foi projetada uma função controlada por dois parâmetros, α e β . A função da PDF do raio está apresentada na figura 2. Esta função causa uma distribuição no espaço em que um histograma levantado a partir do raio teria um resultado como o apresentado na figura 3. Desta forma, a distribuição é uniforme na superfície quando o raio é menor do que α , e hiperbólica quando maior. O limite da PDF é quando o raio é igual a β .

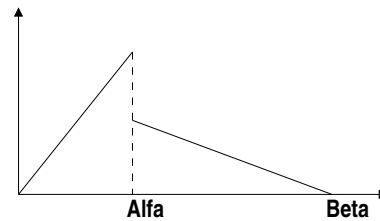


Figura 2: Função de densidade de probabilidade do raio da distribuição.

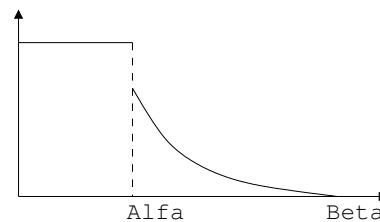


Figura 3: Aspecto do histograma radial da distribuição.

O parâmetro α determina o raio a partir de onde a distribuição deixa de ser uniforme. A partir daí, a PDF do raio é linear e decrescente até o ponto β . A curva é ajustada para que a probabilidade de $r < \alpha$ seja igual a 0.5. Em nosso experimento o valor de β foi ajustado hora em $5 \times \alpha$, e hora em $3 \times \alpha$.

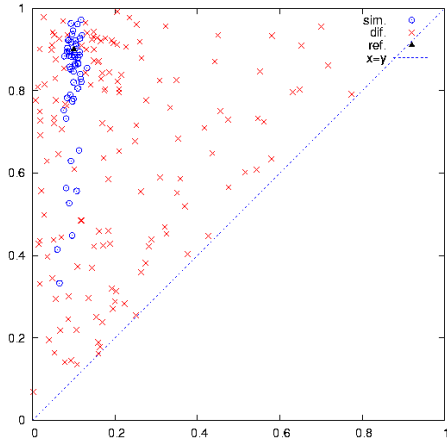


Figura 4: Amostra “byrds” com referência em 0.1,0.9.

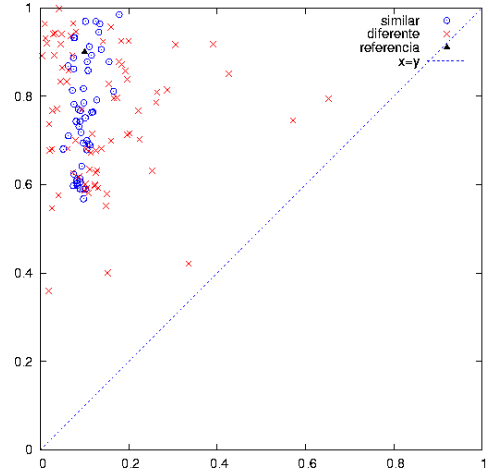


Figura 6: Amostra “tomzé” com referência em 0.1,0.9.

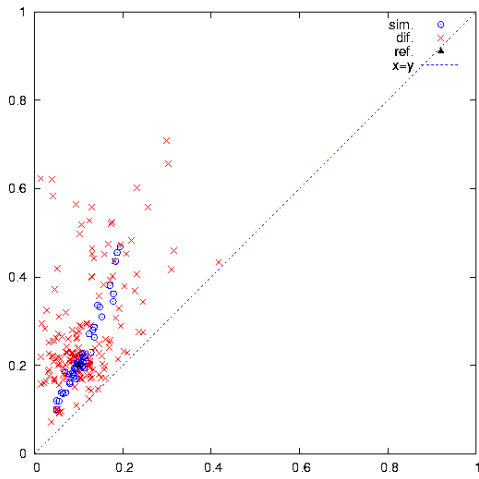


Figura 5: Amostra “byrds” com referência em 0.1,0.2.

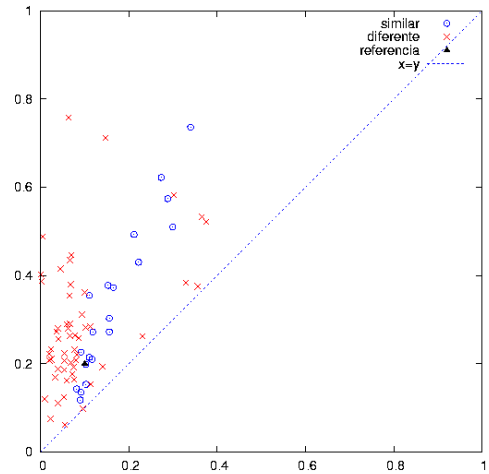


Figura 7: Amostra “tomzé” com referência em 0.1,0.2.

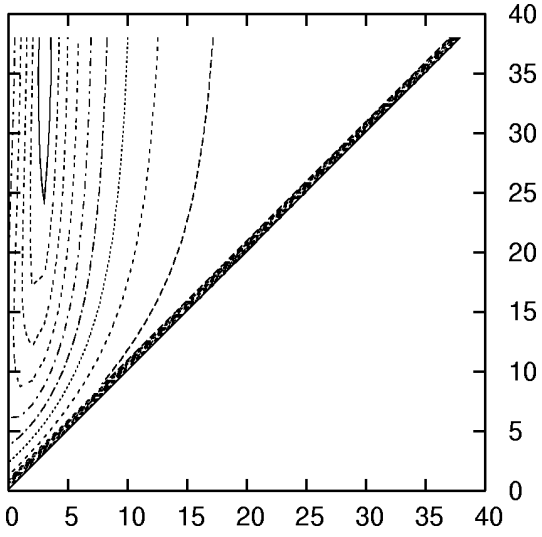


Figura 8: SNR da amostra "byrds" com referência em 0.1, 0.9.

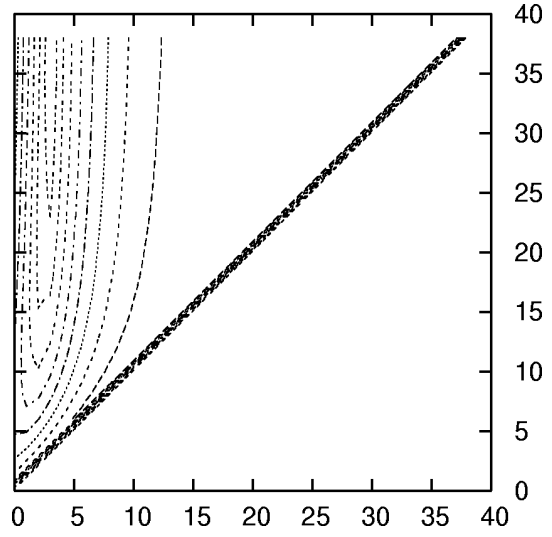


Figura 10: SNR da amostra "tomzé" com referência em 0.1, 0.9.

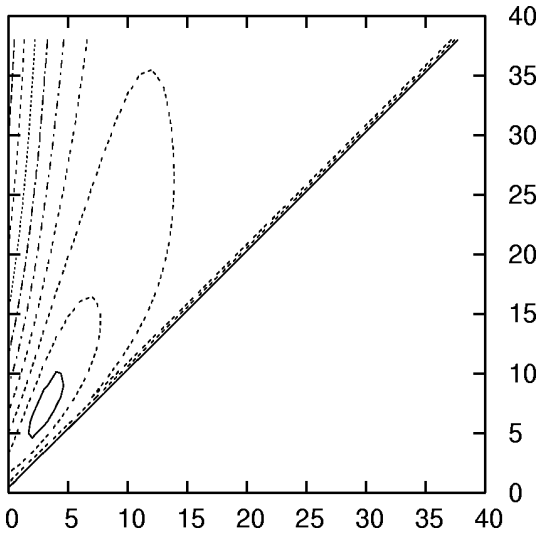


Figura 9: SNR da amostra "byrds" com referência em 0.1, 0.2.

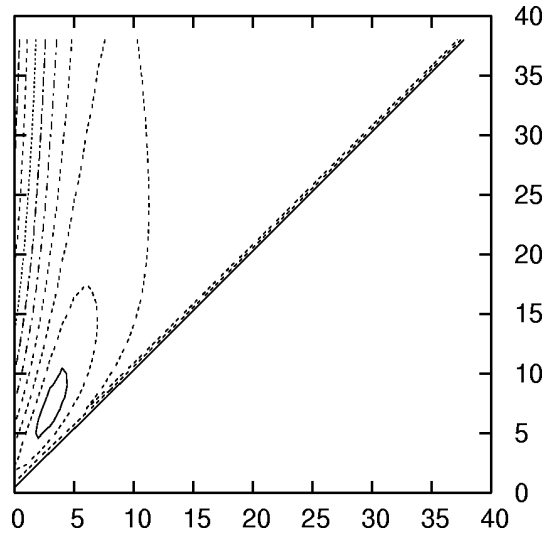


Figura 11: SNR da amostra "tomzé" com referência em 0.1, 0.2.

3.4 Resultados e Análises

A partir dos pontos testados pelos ouvintes foram feitos gráficos em que cada ponto possui um símbolo no caso da comparação ser positiva, e outro para uma comparação negativa. Alguns destes gráficos estão apresentados à frente.

É possível observar nos gráficos que os pontos de identificação positiva se situaram em regiões bastante lineares. A espessura destas regiões variou de acordo com a amostra utilizada, e o ângulo e comprimento das regiões variou com o ponto de referência.

Para buscarmos análises físicas que possam explicar o processo de percepção destas amostras, é preciso comparar as superfícies geradas pela funções a partir dos parâmetros com as regiões de similaridade encontradas. As regiões amostradas devem se situar em regiões em que ocorre uma variação com uma certa magnitude máxima nas funções de análise, que seria o limiar de diferenciação. Quanto mais larga a região amostrada, menor deve ser a magnitude do gradiente da superfície naquela região. A regiões também devem se situar sobre curvas de nível da superfície.

Foi plotada a relação sinal ruído em sinais distorcidos com diferentes parâmetros, com relação à referência, que seria o “sinal” da medida. É possível observar que as curvas possuem alguma relação com as regiões de identificação de similaridade, mas a análise que pretendemos fazer pretende ser ainda mais rigorosa do que isso. Estes gráficos estão aqui apenas a título de demonstração de com seria uma função que tentasse explicar as regiões de similaridade.

Algumas alterações podem ser feitas na realização do teste de forma a modificar a aquisição das regiões de similaridade. Podem ser introduzidos ruídos nas amostras, o que deve alargar as regiões. Também deveria ser feita uma normalização das amostras de forma que elas possuíssem a mesma magnitude de *loudness* ao invés de simplesmente o mesmo valor RMS. Outra modificação importante a ser feita no teste é modificar o algoritmo de sorteio com PDF adaptativa para que ela acompanhe formatos comuns como os encontrados neste experimento.

3.5 Conclusões

O experimento apresentou resultados coerentes, que poderão ser mais significativos se as medidas forem repetidas de forma mais precisa.

Já foi possível observar, em primeiro lugar, que é possível questionar-se a precisão necessária a uma curva de *overdrive* para que ela reproduza com sucesso um certo efeito. Isto é importante, porque existe muita discussão acerca de necessidade ou não das curvas de *overdrive* serem suaves ou poderem ser abruptas, como é o caso em circuitos eletrônicos mais modernos.

Também é possível observar que um novo par de parâmetros mais eficiente do que o utilizado pode ser criado para se trabalhar com *overdrive* de forma mais próxima da sua percepção. Um dos parâmetros estaria relacionado com a direção ortogonal das regiões medidas, e o outro parâmetro com a direção paralela. Desta forma o novo par de parâmetros conseguirá formar uma espécie de espaço com “uniformidade perceptual”. O parâmetro ortogonal seria mais sensível do que o outro, e portanto poderia variar em um intervalo maior.

4 Relatório Técnico-científico - Outros Estudos

4.1 Objetivos

No período compreendido por este relatório também foram feitos estudos sobre a similaridade entre ruídos gerados por cadeias de Markov e ruídos do tipo $1/f^\alpha$, no contexto de pesquisas a serem feitas futuramente com composição automática de músicas.

Também foram feitas algumas análises mais profundas sobre a reverberação vista como distorção de fase.

4.2 Metodologia

4.2.1 Cadeias de Markov

Existem vários jeitos de se fazer composição musical automática com computadores. Basicamente, é pre-

ciso escolher uma classe de funções e mapeá-las a parâmetros sonoros de modelos matemáticos de instrumentos musicais. Cadeias de Markov já foram utilizadas em composições computacionais, como as de Iannis Xenakis [1]. Outro tipo de sinal utilizado em composições automáticas são ruídos $1/f^\alpha$ [2][3].

No caso de cadeias de Markov, cada estado é associado a uma altura de nota, ou outra característica, como duração ou volume. Probabilidades de transição são então atribuídas ao sistema. No caso de ruídos $1/f^\alpha$, um ruído deste tipo é gerado, com valores reais, e então quantizado, ou mapeado de alguma forma na altura de notas.

A pergunta que se fez foi: até que ponto é possível representar-se um sinal $1/f^\alpha$ quantizado com cadeias de Markov? Se uma cadeia pudesse ser criada, capaz de substituir um gerador de ruído $1/f^\alpha$, ficaria provado que ambos sistemas são equivalentes para a criação de músicas computacionais.

Foram sintetizados então vários ruídos $1/f^\alpha$, subsequentemente quantizados. Histogramas foram levantados destes sinais, e com eles foram definidas cadeias de Markov de primeira ordem. As cadeias foram então utilizadas para tentar resintetizar os tipos de sinais.

4.2.2 Reverberação

Foram analisadas funções de transferência geradas anteriormente com ruído browniano complexo.

4.3 Resultados e Análises

4.3.1 Cadeias de Markov

Nas figuras 12 e 13 estão apresentados os resultados obtidos para $\alpha = \{0, 0.25, 0.5, 0.75, 1, 1.5, 2, 3\}$. É possível ver claramente que no primeiro caso o ruído gerado será branco, já que para qualquer que for o valor atual da cadeia, a PDF é a mesma. com o aumento de α , existe uma tendência em que a próxima amostra esteja mais próxima. Em $\alpha = 1$, o caso do ruído Browniano, as PDF seguem a diagonal principal, como era de se esperar, já que em cada nível haveria um sorteio de uma PDF de variância reduzida centrada naquele nível.

Os espectros reconstruídos não parecem possuir distorções muito significativas, apesar de que uma análise mais adequada deve ser feita, com mais pontos, e com uma medida numérica do α gerado. O experimento já foi o suficiente, entretanto, para mostrar que cadeias de Markov podem reproduzir muito bem as características dos ruídos $1/f^\alpha$, inclusive do valorizado ruído rosa, com $\alpha = 1/2$.

É muito tentador afirmar que há uma correlação direta entre a inclinação das PDF e o fator α , ao analisar as figuras. Para valores de $\alpha > 1$, vemos ao invés disto uma diminuição violenta na variância das PDF, o que pode ser explicado pelo fato de que as curvas geradas foram normalizadas pelo maior valor antes do levantamento dos histogramas.

Não houve uma medição rigorosa da correlação entre a inclinação nestas cadeias de Markov com o fator α , mas é muito provável que o procedimento realizado aqui seja similar a procedimentos para a determinação da dimensão fractal de funções.

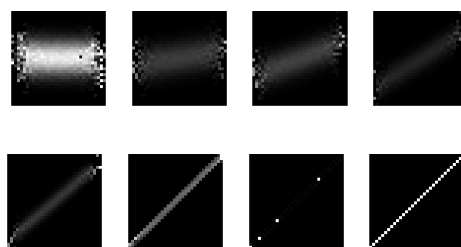


Figura 12: Pesos das cadeias de Markov criadas, apresentados como uma imagem. As matrizes estão ordenadas por α .

4.3.2 Reverberação

Agora já vemos claramente a correlação entre o fato da função de transferência ser um ruído browniano, e seu comportamento no domínio do tempo. A envoltória de uma resposta a impulso da reverberação de uma sala normal atua como um filtro passa-baixas no domínio da frequência, filtrando o ruído branco e o transformando em um ruído próximo do browniano.

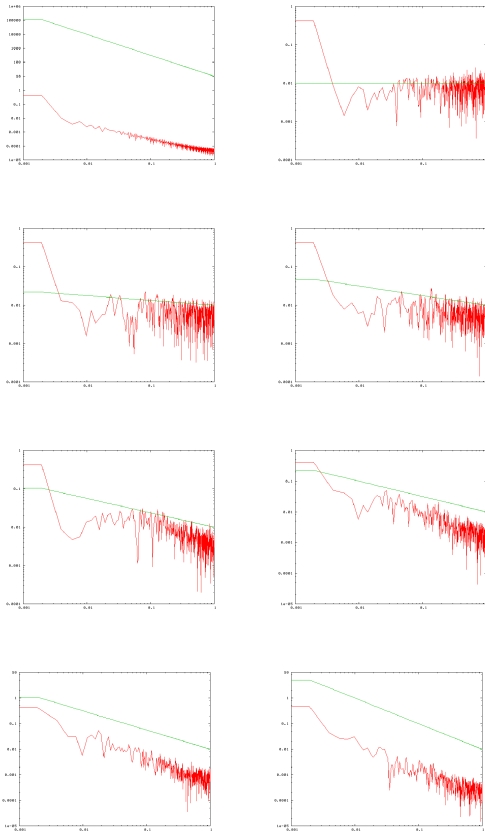


Figura 13: Espectros de frequência originais e criados com as cadeias de Markov.

À medida que a variância do ruído browniano cresce, o limite das variáveis de fase, entre 0 e π , é extrapolado, fazendo com que o ruído vá se tornando branco. No domínio do tempo, isto se traduz em uma função de transferência cada vez mais larga, culminando com um ruído branco.

É interessante analisar este fenômeno do ponto de vista da linearidade de sinais. Sinais lineares são aqueles em que a fase pode ser “embaralhada”, e suas propriedades não se alteram. Sinais de música podem ser considerados como “perceptualmente lineares” em uma escala de tempo muito pequena, já que por mais que o sinal físico possua harmônicos bem definidos, sua compreensão pelo ouvido hu-

mano elimina qualquer efeito das distorções de fase. Numa escala maior, entretanto, existem importantes variações de frequência, timbre e amplitude que podem se perder com o embaralhamento de fase.

Um sinal que possui um ruído branco em sua fase é um sinal linear no sentido estrito. A reverberação, atuando como distorção de fase, causaria portanto uma aproximação de um sinal qualquer a um sinal linear.

É interessante ainda observar neste formalismo, que o efeito de overdrive tende a modificar o histograma de sinais gaussianos, para sinais com um histograma de menor entropia, aproximando sinais da não-linearidade, causando um efeito reverso à reverberação.

Toda esta análise é muito intuitiva, mas apresenta questões que podem ser exploradas em estudos futuros.

4.4 Conclusões

Foi possível demonstrar que cadeias de Markov apropriadamente construídas podem emular ruídos $1/f^\alpha$, injustificando o uso destes em detrimento delas.

Quanto à reverberação, esta análise permitiu mostrar, junto do estudo sobre overdrive, que existe muito mais no mundo da análise de timbres a partir de síntese e processamento de sinais, do que o simples controle da amplitude dos harmônicos criados por um instrumento.

A filtragem de frequências específicas é corretamente tida como indesejável no campo do projeto de salas acústicas, e nosso trabalho ajuda a fundamentar este discurso, mostrando quais são as variáveis de uma função de transferência de reverberação que realmente influenciam suas qualidades.

Em geral, entende-se que o overdrive simplesmente amplifica harmônicos do sinal de guitarras elétricas, mas neste trabalho também pudemos ver que existe uma influência enorme na mudança no ataque e decaimento do som do instrumento, além do acréscimo de parciais não harmônicas no som final, mostrando que a questão é mais complexa do que se crê inicialmente.

5 Fatores que interferiram na execução do projeto

Um fator extremamente positivo no andamento do projeto foi o seu desenvolvimento dentro do ambiente do CEFALA, que é um laboratório onde estudantes e pesquisadores interagem constantemente, discutindo diferentes projetos e técnicas aplicadas.

Esta pesquisa teve um grande apoio principalmente do professor orientador Hani Yehia, dos professores Maurício Loureiro e Maurílio Vieira, que souberam enxergar nas propostas do projeto o potencial para estudos avançados em percepção de timbre, e do aluno de doutorado Hugo de Paula, que contribuiu filosoficamente para que o projeto ganhasse mais solidez científica.

Também foi um fator muito positivo o projeto ter sido desenvolvido sem muitas restrições iniciais. Isto permitiu que fossem abordados assuntos que não eram estudados com muita profundidade dentro da UFMG, apesar de inseridos no contexto do grupo de pesquisa. Algumas das questões abordadas nem mesmo possuem muita cobertura da literatura internacional. Este tipo de pesquisa dificilmente poderia ter sido realizada em uma tese de mestrado, por exemplo, e agora que já foi demonstrado o potencial inexplorado dos efeitos estudados aqui, torna-se mais viável que o trabalho possa ter continuidade na pós-graduação.

Alguns destes fatores positivos, entretanto, possuem seu custo. O trabalho ficou muitas vezes estagnado, e muitas vezes não existia uma visão clara do que poderia ou deveria ser feito, o que causou a baixa qualidade e quantidade dos resultados. Mas enquanto o trabalho não gerou um único resultado substancial, ele consolidou uma base para um amplo grupo de pesquisas que podem ser realizadas, além de ter permitido o domínio de várias técnicas que foram aplicadas no projeto, mas podem ser úteis em muitas outras pesquisas relacionadas.

Referências

- [1] I. Xenakis, *Formalized Music*. Indiana University Press, 1971.
- [2] R. F. Voss and J. Clarke, ““1/f noise” in music: Music from 1/f noise,” *JASA*, vol. 63, pp. 258–263, Jan 1978.
- [3] M. Gardner, “White and brown music, fractal curves and one-over-f fluctuations,” *Scientific American*, 1978.

22 de Dezembro de 2004,

Nicolau Leal Werneck

Hani Camille Yehia