

SÍNTESE DE SONS MUSICAIS BASEADA NA TRANSFORMADA DE KARHUNEN-LOÈVE

Carlos Alberto Ynoguti & Marcelo Basílio Joaquim

ynoguti@inatel.br
DTE-INATEL

Av. João de Camargo, 510
CEP37540-000 Santa Rita do Sapucaí, MG – Brasil

marcelo@sel.eesc.sc.usp.br
SEL-EESC-USP

Av. Trabalhador São-carlense, 400
CEP. 13.566-590 São Carlos, SP - Brasil

Resumo - Tradicionalmente, a técnica de síntese que tem obtido os sons sintetizados mais próximos dos originais é a técnica aditiva com funções base de Fourier. Entretanto, a carga computacional imposta por esta técnica torna a sua utilização para aplicações em tempo real extremamente onerosa, devido ao grande número de senóides necessárias para uma síntese de alta qualidade. Substituindo este conjunto de funções por outro menor, consegue-se uma redução na carga computacional imposta por esta técnica, minimizando o custo para a sua implementação em tempo real. Neste artigo é discutido um modelo de síntese aditiva com funções base derivadas das técnicas estatísticas de Karhunen-Loève. Também são apresentados resultados práticos de tons musicais sintetizados com esta técnica.

Abstract – Traditionally, the Fourier based additive synthesis is the method that have achieved the best results concerning to the quality of the generated sounds. However, the computational load imposed by this technique is extremely high, making its real time implementation difficult. Substituting the Fourier functions by another set of more complex functions, derived from the Karhunen-Loève statistical techniques, one achieves a reduction on the amount of the necessary operations so as to permit its real time implementation. In this article we discuss a additive synthesis model with base functions derived from Karhunen-Loève statistical techniques. Also, practical results of musical tones synthesized with this technique are presented.

I. INTRODUÇÃO

Com o grande avanço experimentado pela tecnologia digital nos últimos anos, os custos de implementação diminuíram bastante e a confiabilidade de sistemas nela baseados aumentou extraordinariamente. Isto faz com que seja muito mais viável a sua utilização em aplicações musicais do que a alguns anos atrás.

Existem várias razões para que cada vez mais músicos se utilizem desta tecnologia para o seu trabalho. Uma delas é a maior variedade de sons gerados pelo computador. Estes não são limitados àqueles disponíveis pelos instrumentos já existentes: pode-se gerar uma infinidade de novos sons através de processamentos digitais. Outra razão é a precisão com que o computador trata os sinais: ele é capaz de analisar e processar a forma de onda do som até ao nível da amostra individual [5].

Outra vantagem da tecnologia digital na produção de sons musicais seria a seguinte: apesar da grande qualidade sonora de um piano Stenway, por exemplo, este é um instrumento difícilimo de transportar, e que exige uma manutenção periódica quanto a sua afinação [9]. Os instrumentos de orquestra em geral são bastante sensíveis a mudanças de temperatura e umidade, e desafinam facilmente com variações destes fatores. Pode-se imaginar o trabalho que dá para afinar as 88 notas de um piano... Com a tecnologia digital, poderia-se construir sintetizadores que soassem como um Stenway, com a vantagem de serem portáteis e não desafinarem nunca. É lógico que a construção de um sintetizador que reproduza todas as características e nuances de um instrumento real é bastante difícil e complicada, mas este é um objetivo a ser alcançado, e vários pesquisadores estão se empenhando firmemente neste propósito.

Neste trabalho foi estudado e implementado um método de síntese baseado em análise utilizando a transformada de Karhunen-Loève, método também conhecido como Análise de Componente Principal [4]. Testes subjetivos de audição mostraram que é possível obter sons sintetizados bastante similares aos originais para sons de guitarra, violão, flauta e violino usando no máximo 8 componentes, o que representa uma grande economia de memória e tempo de processamento quando comparada à técnica baseada em Fourier, onde são necessárias aproximadamente 30 senóides para uma síntese de boa qualidade.

II. BREVE INTRODUÇÃO À SÍNTESE DE SONS MUSICAIS [7]

O objetivo dos métodos de síntese digital é produzir um sinal digital que quando convertido em uma tensão

analogica e tocado através de um sistema de som de alta fidelidade produza o som musical desejado. O problema da síntese resume-se então ao de gerar este sinal. Isto é feito utilizando várias técnicas fundamentais de síntese diferentes.

Uma técnica de síntese pode ser entendida como uma fórmula para produzir as amostras de um sinal amostrado. Esta expressão tem geralmente muitas variáveis, tais como frequência, duração e volume. Para que haja um maior realismo na reprodução, é necessário que estas variáveis mudem com o tempo durante a duração do som. Isto significa que estas variáveis precisam ser funções do tempo.

O grau em que o som sintetizado se iguala ao som original é uma medida da validade do modelo e da técnica de estimação. Se o modelo não é válido para o som em questão, ou se a técnica de estimação não é acurada, não se pode esperar duplicar o som original.

Existem três tipos principais de síntese: direta, *musique concrète* e baseada em análise. Na síntese direta, o compositor gera diretamente a forma de onda do som desejada, por exemplo em um computador. Na *musique concrète* digitaliza-se sons já existentes e fazem-se processamentos matemáticos sobre estes para gerar novos sons. Na síntese baseada em análise, tons naturais existentes são digitalizados e analisados. Os resultados da análise são usados como dados para dirigir a síntese. Estas funções são formas de onda amostradas processadas a partir dos sons dos instrumentos musicais existentes. Estas podem então ser modificadas pelo compositor. A expressão de síntese é um modelo matemático da forma de onda. No modelo de síntese baseada em análise o procedimento de análise é uma tentativa de estimar os parâmetros do modelo para um dado som.

Podem ser também identificadas três técnicas de síntese principais: aditivas, subtrativas e não lineares. No modelo aditivo a forma de onda final é gerada a partir da soma de componentes mais simples. Para a técnica subtrativa tem-se o oposto: o som é gerado a partir de uma forma de onda espectralmente rica introduzida em filtros variantes no tempo. Dentro da classe das técnicas não lineares, a principal é aquela baseada na equação de modulação em frequência (FM), explorada inicialmente por Chowning em Stanford [3], e que é a base da maioria dos sintetizadores comerciais utilizados atualmente.

Existem vantagens e desvantagens para cada uma destas técnicas. A técnica de FM é a mais simples computacionalmente, e por isso é bastante fácil a sua implementação em tempo real. Entretanto o som gerado é relativamente pobre e sem graça, e não existe a possibilidade da análise de sons já existentes (é uma técnica que só funciona no modo direto).

Para a técnica subtrativa, os problemas são os mesmos da FM, sendo que existe a vantagem de permitir a análise, o que melhora a reprodução de sons já existentes.

A síntese aditiva com funções base senoidais é a que tem alcançado os melhores resultados na reprodução de sons de instrumentos musicais [5], apresentando também a possibilidade da análise. A grande desvantagem desta técnica é a enorme quantidade de dados e conseqüentemente de cálculos envolvida durante o processo de síntese. Diminuindo-se o conjunto de senóides por outro menor é possível diminuir a quantidade de operações envolvidas, diminuindo assim o tempo total de processamento.

Neste artigo procura-se alcançar este objetivo com a substituição das funções base senoidais por outro conjunto menor, baseado nas técnicas estatísticas de Karhunen-Loève.

III A TRANSFORMADA DE KARHUNEN-LOÈVE

A transformada de Karhunen-Loève é ótima em relação ao erro quadrático médio de truncamento [1]. Com isto, consegue-se uma representação de sinais com menos funções base do que qualquer outra conjunto de funções ortogonais. Esta propriedade é bastante útil para a síntese de sons musicais pelo método aditivo, pois reduz o número de operações necessárias.

Não é o objetivo deste trabalho o estudo teórico da transformada, e sim a sua utilização para a síntese de sons musicais. Assim será dada apenas uma breve explicação de como é aplicada no presente trabalho. Uma demonstração teórica da transformada pode ser vista em [1].

Com a transformada de Karhunen-Loève, um conjunto de sinais $\{x_\xi\}$ é representado por:

$$x_\xi \equiv \sum_{i=1}^N y_{\xi i} \phi_i \quad (1)$$

onde os ϕ_i são os autovetores da matriz de covariância \mathbf{R}_x dos sinais sob análise, a qual é dada por

$$\mathbf{R}_x = \frac{1}{M} \sum_{\xi} x_\xi x_\xi^t \quad (2)$$

onde os x_ξ são as funções amostra armazenadas em vetores coluna, e M é o número de funções amostra.

O problema de encontrar os autovalores λ_i e os autovetores ϕ_i de uma matriz simétrica é bem conhecido, e existem vários algoritmos propostos para este fim. Neste trabalho foi utilizado o algoritmo QL em conjunção com o método de Householder para tridiagonalização de matrizes reais e simétricas [6]. Posteriormente, os autovetores foram ordenados de acordo com os seus respectivos autovalores, do maior para o menor.

Os coeficientes y_i são obtidos a partir de

$$y_{\xi_i} = x_{\xi}^t \phi_i \quad (3)$$

Com isto, os sinais podem ser representados através dos autovetores de \mathbf{R}_x .

O erro quadrático médio de truncamento é dado pela soma dos autovalores correspondentes ao autovetores desprezados. Se forem selecionados L de um total de N autovetores ($L < N$), o erro de truncamento é dado por:

$$\varepsilon_{\min}(L) = \sum_{i=L+1}^N \lambda_i \quad (4)$$

A eliminação de autovetores que não contribuem significativamente na reconstrução do sinal implica na diminuição do número de funções base. Com isso o sinal pode ser aproximado por

$$\hat{x}_{\xi} = \sum_{i=1}^L y_{\xi_i} \phi_i \quad (5)$$

onde \hat{x}_{ξ} é o sinal x_{ξ} reconstruído com L funções base.

IV ADAPTAÇÃO DA TRANSFORMADA DE KL PARA SÍNTESE DE SONS MUSICAIS

A aplicação direta da transformada de KL em um conjunto de tons musicais não é prática. Na Equação (2), pode-se ver que o comprimento de uma função amostra determina as dimensões da matriz de covariância. 1 s de tom musical consiste de 44100 amostras a uma taxa de amostragem comum de 44.1 kHz. Ao invés de tentar derivar as funções base de KL para conseguir tais comprimentos, é mais viável obter um conjunto de funções base periódicas de KL derivadas de segmentos de tons ao invés de tons completos. Com estes dados, as funções base são afetadas por coeficientes de amplitude e fase de modo a reproduzir as variações temporais do som original.

A. Normalização em Frequência

Como visto acima, os segmentos coletados dos sons musicais precisam ser periódicos e normalizados em frequência, de modo que a parte analisada do sinal contenha todas as informações necessárias à síntese.

Poderia-se tomar de cada instrumento sempre uma amostra com a mesma frequência fundamental para garantir as restrições acima. Entretanto, sabe-se que o timbre de um instrumento muda sensivelmente dependendo da oitava em que se toca: o som de um saxofone nas notas graves é bem diferente nas notas mais agudas por exemplo. Além disso, estes tons são em geral não periódicos. O valor $x_{\xi}(k + Td)$ pode diferir apreciavelmente de $x_{\xi}(k)$ devido a variações naturais no envelope de amplitude e frequência do tom original através de seu tempo de vida. Diante disso, torna-se

necessário fazer algum tipo de normalização em frequência nas funções amostras. Neste trabalho foram utilizadas as técnicas de interpolação e decimação [8]. Para prevenir o aliasing, pode ser necessária uma redução na largura de banda se o número de amostras em um segmento precisar ser reduzido durante a normalização do período. Como pode-se ver da Equação (2), o comprimento de uma função amostra normalizada (Td) determina as dimensões da matriz de covariância. É necessário então estabelecer um compromisso entre o desejo por uma matriz pequena (o que diminui o tempo de cálculo) com a necessidade de preservar a largura de banda nos dados originais tanto quanto possível. No presente trabalho foi escolhido um valor $Td = 103$ amostras.

B. Alinhamento em Fase

Um dos problemas com a transformada de KL é que esta é fortemente dependente da fase relativa entre os sinais que constituem a matriz de covariância. O alinhamento em fase das funções amostra tende a concentrar a maior parte da energia nos primeiros autovetores [2]. O alinhamento força características similares das funções amostra a ocorrerem aproximadamente ao mesmo tempo. Por exemplo, se cada função amostra tem um pico o qual contém uma larga fração da energia do sinal, então estas funções amostra são alinhadas de modo que estes picos ocorram ao mesmo tempo.

A idéia básica é ajustar as fases das funções amostra de modo a otimizar o alinhamento, no sentido de uma correlação cruzada, com a primeira função base $\phi_1(t)$. A função de correlação entre dois sinais x e y é dada pela expressão

$$\gamma_{xy}(\tau) = \sum_{n=0}^N x(n)y(n+\tau) \quad (6)$$

onde: γ_{xy} é a correlação cruzada, τ é o deslocamento relativo entre os sinais, e N é o tamanho da janela em que é feita a análise.

Calcula-se inicialmente $\phi_1^{(1)}(t)$, a função base principal de KL para os dados de entrada iniciais. Um modo fácil de calcular $\phi_i(t)$ é utilizar o chamado método das potências [10]. Cada função amostra $x_i(t)$, é deslocada ao longo do eixo de tempo até que seja alcançada a máxima correlação com $\phi_1^{(1)}(t)$. Os dados deslocados tornam-se a entrada para a próxima iteração. Calcula-se então $\phi_1^{(2)}(t)$, e repete-se o processo. O algoritmo para quando não ocorrer mais nenhuma mudança de fase em nenhuma das funções amostra.

Resta ainda o problema de escolher o tamanho N da janela. Para se ter um retrato fiel do comportamento do sinal, N foi escolhido de modo a conter um período do sinal. Dado que os sinais foram normalizados para terem

um período fundamental de 103 amostras, escolheu-se $N = 103$.

Nas Figuras 1 e 2 tem-se uma ilustração de como funciona o alinhamento em fase. Na Figura 1 tem-se três sinais: uma onda quadrada, uma triangular e uma senóide com fases aleatórias. Na Figura 2 são mostrados os mesmos sinais alinhados em fase.

C. Classificação dos Sinais

A concentração de energia nas primeiras funções base depende fortemente do grau de similaridade entre as funções amostra: quanto maior a similaridade, maior a concentração de energia. O agrupamento dos sinais em classes de acordo com a sua similaridade faz com que seja possível reconstruí-los com menos funções base, desde que o erro de truncamento torna-se menor com a maior concentração de energia.

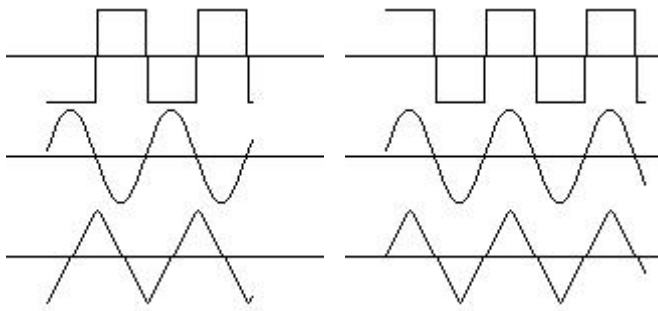


Figura 1: Ondas quadrada, senoidal e triangular com fases aleatórias. **Figura 2: Ondas quadrada, senoidal e triangular alinhadas em fase.**

Uma maneira de se medir a similaridade entre dois sinais é realizar a mesma correlação cruzada da equação (6).

Como γ_{xy} depende de τ , foi feita uma varredura em τ , e foi escolhido o valor máximo assumido por γ_{xy} como a medida de similaridade entre os sinais.

Com o valor da correlação entre os sinais em mãos, pode-se realizar a separação dos sinais em classes de similaridade. Escolhendo-se um limiar, todos os sinais com correlação igual ou acima deste seriam agrupados em uma mesma classe.

D. Detecção de Envoltória

Os sons musicais são sinais variantes no tempo. São estas variações que determinam o seu realismo e naturalidade. O sistema de análise recupera um período do sinal na região de sustentação do tom através da Equação (1). O objetivo é sintetizar tons inteiros e nestes, tanto a frequência como a amplitude variam com o tempo.

Foram sintetizados com sucesso vários tons com a técnica de síntese aditiva com funções base senoidais (síntese de Fourier), utilizando envoltórias independentes

para cada harmônica, aplicando uma aproximação linearizada por partes em cada uma delas [5]. A idéia é aplicar esta mesma técnica na síntese de sons com funções base de KL, segundo o algoritmo proposto em [10].

Um som musical variante no tempo pode ser aproximado por:

$$x(n) \approx \sum_{i=0}^{N-1} y_i(n) \tilde{\phi}_i(\theta_i(n)) \quad (7)$$

A idéia básica do algoritmo de detecção de envoltória é determinar, em pontos espaçados no tempo, quais valores dos coeficientes y_i e das fases $\theta_i(n)$ aproximam melhor o sinal $x(n)$. Nos pontos intermediários, estes valores são obtidos por interpolação.

A fase $\theta(k_j)$ que satisfaz a condição acima foi determinada como sendo aquela que maximizasse a expressão abaixo:

$$y(k_j) = \sum_{k=k_j}^{k_j+Td} x(k) \tilde{\phi}_0(\theta(k_j) + \hat{\omega}(k - k_j) T_s) \quad (8)$$

onde y é a correlação cruzada entre o sinal e o primeiro autovetor, $\tilde{\phi}_0$ é o primeiro autovetor, k_j é o ponto onde é feita a análise, $\hat{\omega}$ é a frequência fundamental do sinal em rad/s, e T_s é o período de amostragem do sinal.

Os coeficientes y_i são então calculados através da expressão

$$y_i(k_j) = \sum_{k=k_j}^{k_j+Td} x(k) \tilde{\phi}_i(\theta(k_j) + \hat{\omega}(k - k_j) T_s), \quad (9)$$

$$i = 1, 2, \dots, N - 1$$

V. PARTE EXPERIMENTAL

A. Sistema Experimental e Material para Análise.

Com as considerações acima, montou-se um sistema a fim de verificar experimentalmente o desempenho da transformada na síntese de sons musicais. Um diagrama em blocos do sistema desenvolvido é mostrado na Figura 3.

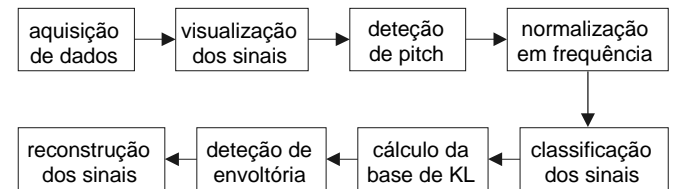


Figura 3. Sistema de análise-síntese de KL.

Para os testes foram selecionados sinais de violão, flauta, guitarra elétrica e violino, sendo este executado de três maneiras diferentes: pizzicato, arco e vibrato. De cada um destes tipos foram selecionadas 15 notas (duas

oitavas em tons naturais) perfazendo um total de 90 sinais de teste.

Estes foram normalizados em frequência, classificados em classes de similaridade e alinhados em fase. O processo de classificação gerou 5 classes:

1. sinais com $\gamma_{xy} \geq 0.8$ (flauta e violino pizzicato)
2. sinais com $\gamma_{xy} \geq 0.8$ (violão e guitarra)
3. sinais com $0.5 \leq \gamma_{xy} < 0.8$ (violão, violino, flauta, pizzicato e vibrato)
4. sinais com $0.3 \leq \gamma_{xy} < 0.5$ (violino, pizzicato, violão e vibrato)
5. sinais com $\gamma_{xy} < 0.3$ (violino, violão, vibrato, pizzicato, guitarra e flauta).

B. Testes de Audição.

Separados em classes, os sinais foram analisados, calculando-se a base de KL para cada classe e fazendo-se a detecção de envoltória para os sinais. Cada sinal foi reconstruído várias vezes, cada uma com um número diferente de autovetores de modo a poder avaliar mais tarde, em testes de audição, qual seria o número mínimo destes para uma síntese de alta qualidade.

Segundo Pollard (1988), pelo menos no estágio atual de conhecimento, o meio mais confiável de verificar a similaridade entre dois sons é realizar testes de julgamento subjetivo de similaridade, em ensaios de audição, com pessoas treinadas.

Os vários sons reconstruídos foram então tocados aos pares, junto com os sons originais, isto é, foi tocado em sequência o som original e o sintetizado, não necessariamente nesta ordem. Também, os sons dos vários tipos de instrumentos foram tocados intercaladamente, de maneira aleatória, com o intuito de evitar influir na avaliação do ouvinte. O número total de pares de sons para a análise foi de 158. Caso desejassem, tinham também a liberdade de repeti-los.

Após ouvir cada par de sons, os sujeitos eram convidados a atribuir as seguintes notas:

1. se os sinais fossem diferentes
2. se fossem pouco parecidos
3. se fossem bem parecidos
4. se fossem iguais

C. Avaliação dos Resultados.

Na apreciação dos testes de audição, considerou-se como sendo uma síntese de boa qualidade os testes que resultaram em uma nota igual ou superior a 3 na escala do item anterior. Foi feita uma ponderação entre as respostas de várias pessoas, e com isto conseguiu-se detectar o número mínimo de autovetores para reconstrução de cada um dos sinais. Estes resultados são mostrados na Tabela 1.

Os resultados obtidos foram muito bons. para os sons de violino tocado com arco e com vibrato. Isto é muito positivo, pois os métodos de síntese aditiva com funções

base de Fourier apresentam uma grande dificuldade de tratar com sons que apresentam vibrato.

TABELA I
Resultados dos Testes de Audição.

Classes	Correlação entre os sinais	n ^o de autovetores p/ síntese
1	$\gamma_{xy} \geq 0.8$	5
2	$\gamma_{xy} \geq 0.8$	5
3	$0.5 \leq \gamma_{xy} \leq 0.8$	8
4	$0.3 \leq \gamma_{xy} \leq 0.5$	6
5	$\gamma_{xy} \leq 0.3$	7

Os sons percussivos não tiveram resultados tão bons quanto os anteriores, apresentando uma espécie de zumbido ao final da região de decaimento, principalmente nas reconstruções feitas com um número elevado de funções base. Nas Figuras 4 e 5, tem-se os espectros de um tom de guitarra original e reconstruído, respectivamente. Analisando-os, pode-se notar que para o sinal reconstruído, há uma presença maior nas componentes de média/alta frequências, comparando-se com o sinal original. Isto explica o zumbido ouvido durante a execução.

Uma das razões apontadas para isto foi que os sons percussivos apresentam uma variação espectral muito grande durante o seu tempo de vida. As harmônicas mais altas são mais amortecidas do que as mais baixas no decorrer do tempo. Como as funções base geradas pela transformada de KL apresentam espectros mais amplos (não são limitados a apenas uma harmônica como no caso da Transformada de Fourier), esta variação harmônica parece não ser reproduzida fielmente. Deste modo, pode-se inferir que nas reconstruções com um número grande de autovetores (mais do que 10, por exemplo) a região de decaimento apresenta estes problemas. entretanto, se a reconstrução for feita com poucos autovetores, o ataque é prejudicado. Uma solução que obteve bons resultados foi reconstruir o ataque com um número maior de autovetores que o decaimento. Isto minimizou bastante o zumbido no final do tom.

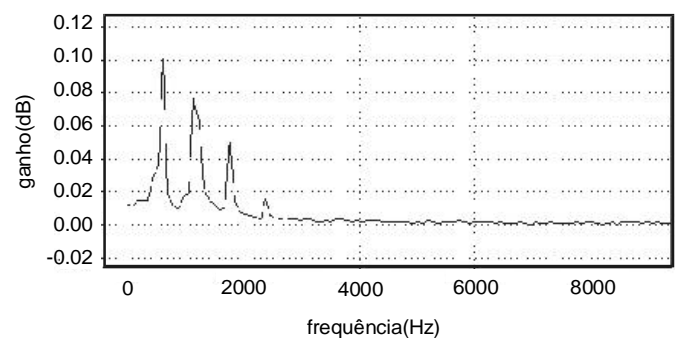


Figura 4: Espectro de um tom de guitarra no final da região de decaimento.

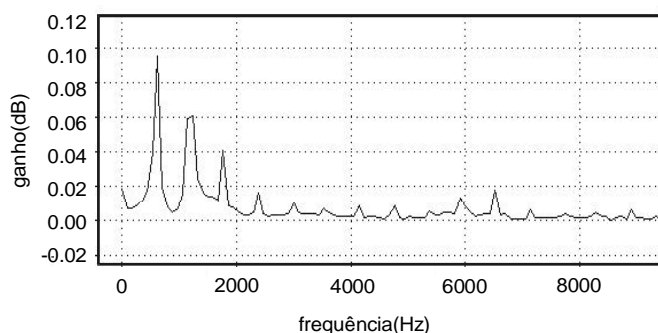


Figura 5: Espectro do tom reconstruído no final da região de decaimento.

VI. CONCLUSÕES

Tradicionalmente considera-se a técnica de síntese aditiva como a que produz sons de mais alta qualidade. O método mais comumente empregado com esta técnica é a síntese aditiva de Fourier, com funções base senoidais. Entretanto, a carga computacional imposta por este método tem limitado o seu uso para aplicações em tempo real.

Neste trabalho foi implementado e testado um modelo de síntese digital de sons musicais usando o modelo de síntese aditiva baseada em análise com funções base de Karhunen-Loève.

Foi demonstrado que um conjunto de funções base de Karhunen-Loève é aquele que apresenta o mínimo erro médio quadrático de truncamento para o conjunto de funções amostra da qual foi derivado. Ainda, a base de Karhunen-Loève apresenta todas as propriedades de qualquer conjunto de funções base ortogonais. Isto permite a representação das funções amostra com menos funções base do que qualquer outro conjunto de funções base ortogonais.

Os testes realizados mostraram que, comparada à síntese aditiva de Fourier, a técnica de Karhunen-Loève apresenta um desempenho melhor no que diz respeito ao compromisso entre a qualidade dos sons sintetizados e o esforço computacional requerido. Com 31 funções base conseguiu-se representar 90 sons musicais, sendo que este número poderia ser bem maior se fossem tomadas mais notas dos mesmos instrumento (por exemplo com sustentados e bemóis).

A transformada foi adaptada para reproduzir não apenas segmentos relativamente curtos, mas tons musicais inteiros, variantes no tempo. Para isto foram formadas bases com um período de cada sinal tomados na sua região de sustentação. Este trecho representaria as características do som sendo sintetizado. De modo a otimizar o desempenho e reduzir ainda mais o número de funções base necessárias à síntese foram aplicadas algumas técnicas de processamento digital de sinais tais

como normalização em frequência e alinhamento em fase.

Os sons foram agrupados em classes antes de serem analisados, o que resultou numa melhor concentração de energia nas funções base, reduzindo assim a quantidade destas para a síntese. Com isto consegue-se alguma economia no hardware necessário para operação em tempo real.

Com a técnica de síntese aditiva baseada em análise elimina-se o procedimento cansativo de tentativa-e-erro para a derivação dos parâmetros de síntese. O método utilizado para a detecção das envoltórias de fase e amplitude das funções base funcionou bastante bem, preservando adequadamente os transientes iniciais da porção de ataque dos sons que, como visto anteriormente, desempenham um papel essencial na caracterização destes sinais.

Existem duas desvantagens inerentes à utilização das funções de Karhunen-Loève:

- A síntese com KL é limitada a uma síntese harmônica, exceto para efeitos especiais, o que faz imaginar que a síntese de sons inarmônicos seja dificultada;
- O aliasing é melhor tratado com a síntese de Fourier, onde as harmônicas problemáticas são simplesmente eliminadas, enquanto que com KL é necessário limitar os sinais em banda.

As vantagens inerentes à técnica aditiva, que são a flexibilidade e o controle sobre as funções base utilizadas na síntese são preservadas com a técnica de KL.

O sistema não funcionou bem para sons percussivos (sons onde a parte de sustentação está ausente). Uma das causas apontadas para este baixo desempenho foi que o sistema de análise não consegue reproduzir de maneira fiel a grande variação harmônica sofrida por estes sinais durante o seu tempo de vida.

Um procedimento que ajudou bastante nestes casos foi reconstruir a porção de ataque com um número maior de funções base do que a porção de decaimento. Desta forma, minimiza-se o zumbido no decaimento, e preserva-se o ataque.

Uma versão mais detalhada deste trabalho pode ser encontrada em [12].

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AHMED, N. & RAO, K. R. *Orthogonal transforms for digital signal processing*. pp 180-224. Springer-Verlag. 1975
- [2] CHRISTENSEN, Ronald A. & HIRSCHMAN, Alan, D. Automatic phase alignment for the Karhunen-Loève expansion. *IEEE Transactions on*

- Biomedical Engineering*. v. BME-26, n. 2. pp. 94-99. February. 1979.
- [3] CHOWNING, J. M. The synthesis of complex audio spectra by means of frequency modulation. *Journal of Audio Engineering Society*. v. 21, n 7, pp. 526-534. September. 1973.
- [4] JOHNSON, Richard & WICHERN, D. W. *Applied Multivariate Statistical Analysis*. Prentice Hall. 1998.
- [5] MOORER, James A. Signal processing aspects of computer music - a survey. *Proceedings of the IEEE*. v. 8, n. 8. pp. 1108-1136. August. 1977.
- [6] PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. and VETTERLING, W. T. *Numerical recipes - the art of scientific computing*. Cambridge University Press, Cambridge, 1987.
- [7] RUSCHIONI, Ruggero Andrea. Música digital: histórico e perspectivas. In: WORKSHOP SOBRE COMPUTAÇÃO DE ALTO DESEMPENHO PARA PROCESSAMENTO DE SINAIS. São Carlos, 1993. *Anais*. São Carlos. 1993. pp. 178-191.
- [8] SCHAFER, Ronald W., RABINER, Lawrence R. A. Digital signal processing approach to interpolation. *Proceedings of the IEEE*, v. 61, n. 6. pp. 692-702. June. 1973.
- [9] STANFORD UNIVERSITY. *Music synthesis approaches sound quality of real instruments*. Newsgroups: su.news-service.press-releases. 1994.
- [10] STAPLETON, John C. & BASS, Steven Craig. Synthesis of musical tones based on the Karhunen-Loève transform. *IEEE Transactions on Acoustic, Speech and Signal Processing*, v. 36, n. 3. pp. March. 1986.
- [12] YNOGUTI, Carlos Alberto. *Síntese de sons musicais baseada na transformada de Karhunen-Loève*. São Carlos. Dissertação (Mestrado) - Escola de Engenharia de São Carlos, Universidade de São Paulo. 1995.