

**A VARIAÇÃO DO TIMBRE DA CLARINETA
NA INTERPRETAÇÃO DE FRASES MUSICAIS:
ANÁLISE DAS VARIAÇÕES DAS AMPLITUDES E FREQUÊNCIAS
DOS COMPONENTES HARMÔNICOS**

por Maurício Alves Loureiro

Introdução

O conceito de timbre se refere à cor ou à qualidade do som e está normalmente dissociado dos conceitos de intensidade e altura. Timbre é definido pela ASA (American Standard Association) como “aquele atributo do sentido auditivo em termos do qual o ouvinte pode julgar que dois sons similarmente apresentados e tendo a mesma intensidade e altura, são dissimilares” [Risset e Wessel, 1982]. Pesquisas na percepção do timbre demonstram que a distribuição da energia espectral e a variação temporal desta distribuição fornecem os determinantes acústicos de nossa percepção de qualidade sonora [Wessel, 1979]. A psicoacústica utiliza o termo *timbre* para especificar de uma maneira não muito precisa, os mecanismos perceptivos que classificam os sons em determinadas categorias, como por exemplo em família de instrumentos musicais ou vozes de indivíduos. Estes mecanismos de alguma maneira agrupam sons tocados por um mesmo instrumento ou falados por uma mesma pessoa, mas que podem ser bem distintos entre si, seja em altura, intensidade ou duração. De fato o conceito de timbre tem sido sempre relacionado com sons de instrumentos musicais ou de voz e é neste âmbito que a maioria das pesquisas em timbre têm se desenvolvido [Luce, 1963; Risset, 1991; McAdams e Bregman, 1979; McAdams, 1987; Gordon e Grey, 1977, 1978; Grey, 1975, 1978; Barrière, 1991].

Na década de 60, inúmeros trabalhos de pesquisa realizados com métodos de análise e síntese de sons de instrumentos musicais mostraram ser bastante adequada a representação destes sons através de curvas de variações temporais das amplitudes e das frequências dos componentes harmônicos obtidas por diferentes técnicas [Luce, 1963; Luce e Clark, 1965 e 1967; Risset, 1965; Strong e Clark, 1967a e 1967b; Fredman, 1965 e 1968]. Strong e Clark sintetizaram sons de instrumentos de sopro por síntese aditiva, somando os harmônicos com suas amplitudes controladas por um envelope espectral. Foi utilizado, num primeiro momento, o mesmo envelope para todas as frequências, e em seguida vários envelopes que variavam de acordo com as frequências dos harmônicos [Strong e Clark, 1967a e 1967b]. Risset analisou sons de trompete e obteve curvas individuais de evolução temporal das amplitudes e frequências para cada harmônico. Através de testes auditivos, concluiu que certas características timbrísticas do som do trompete obedecem mais a leis de variação do espectro do que a características fixas da estrutura espectral [Risset, 1965; Risset e Mathews, 1969; Risset e Wessel, 1982].

Objetivos

O conceito de timbre acima descrito, se torna ainda mais impreciso quando consideramos que um mesmo instrumento musical pode produzir timbres diferentes, de acordo com a maneira em que é tocado. Apesar da existência de inúmeros estudos sobre o timbre de instrumentos musicais, tal como o trabalho exaustivo de J. M. Grey [Grey, 1975], estes trabalhos têm sido feito em notas isoladas, abordando comparações entre diferentes instrumentos musicais e fora de qualquer contexto musical, focalizando apenas o mecanismo perceptivo que discrimina um instrumento musical de outro. Pouco foi realizado em relação ao controle e à percepção da variação de timbre pelo intérprete,

num mesmo instrumento. As variações dos parâmetros do som, que se dão dentro do âmbito de uma mesma nota, e que podem ser percebidas como variações de timbre, são menos explícitas, se comparadas com as variações de duração ou intensidades, dificultando a medição e a classificação de seus estados de evolução ao longo do tempo e por conseguinte sua especificação no sistema tradicional de notação musical. Assim, os instrumentos acústicos fornecem para o timbre pouco ou nenhum controle manipulável de escalonamentos gradativos, tal como as alturas, intensidades e durações possibilitam, para que um compositor possa utilizar este parâmetro com precisão quantitativa em uma composição musical.

O presente projeto se propõe a investigar a variação do timbre da clarineta ao longo de execuções de frases musicais selecionadas de tal maneira a abranger todos os registros do instrumento, e as relações desta variação voluntária de timbre com a intenção de expressividade do intérprete, através da análise das funções de variação temporal das amplitudes e das frequências dos componentes harmônicos dos sons amostrados destas passagens musicais. Ao invés de tratar notas isoladas e quase-estáticas, este trabalho busca estudar frases musicais inteiras, dentro de contextos musicais representativos do instrumento analisado, visando investigar o significado de parâmetros físicos determinantes da variação do timbre de um instrumento, na realização de um gesto “musical” ou “expressivo” numa interpretação musical. Assim, as amostras analisadas foram tomadas de passagens que caracterizam vários aspectos timbrísticos do instrumento, mas que incluam sobretudo frases com grande teor expressivo. Foram gravadas passagens dos Quintetos para Clarineta e Cordas de Brahms (op. 115), Mozart (Kv. 588) e Weber (op. 34). As gravações foram feitas em estúdio profissional (Escola de Música da UFMG).

Descrição da Metodologia

Medição das curvas de Amplitude e Frequência

As curvas de variação temporal de amplitude e frequência foram calculadas através da *Transformada Discreta de Fourier de Janela Deslizante* (Discrete Short-Time Fourier Transform - DSTFT), cuja implementação foi feita com a utilização do pacote de ferramentas de processamento digital de sinal (DSP toolbox) do software MATLAB. Este método apresentou problemas relacionados ao espalhamento dos valores de amplitude ao longo de bandas largas de frequências, dificultando o mapeamento dos componentes harmônicos. Métodos mais eficientes como o *Filtro Heterodyne*, *Constant-Q*, *Phase Vocoder* e *Wavelet* estão sendo investigados.

Análise dos dados

A partir de uma primeira análise das curvas de variação de amplitude e frequência dos componentes harmônicos, percebeu-se que uma característica comum a dados de síntese aditiva: os envelopes de amplitude apresentam um grande grau de correlação seja no sentido do tempo, como no sentido da frequência. Os métodos de *Análise Estatística de Multivariados*, se mostram adequados para se aplicar a este tipo de dados. Entre estes métodos, foi utilizado neste projeto a *Análise por Componentes Principais* (PCA), que identifica sistematicamente redundâncias nos dados, redistribuindo-os de maneira a proporcionar uma aproximação que represente quase toda a informação contida nos dados originais com um menor volume de dados. Alcançamos assim uma redução de dados que nos facilita detectar a relação de significância perceptiva com os parâmetros medidos. A PCA calcula uma combinação linear das variáveis medidas, que contém a

máxima variância dos dados e pode assim representar os dados originais de acordo com o grau de semelhança identificado. Calcula-se então combinações lineares subsequentes, de maneira que a correlação entre elas seja nula, ou seja, dispostas em bases mutuamente ortogonais (*auto-vetores* dos respectivos *auto-valores* da matriz de Covariância) [Johnson e Wichern, 1998; Dillon e Goldstein, 1984; Rencher, 1995]. A *Análise por Componentes Principais*, tem sido largamente utilizada em aplicações de medidas de percepção de objetos, mostrando excelentes resultados. Neste projeto, o método foi aplicado direto às medições dos parâmetros físicos medidos, com o objetivo de relacioná-los com as características perceptivas. As primeiras experiências com Análise Estatística de Multivariados aplicada a parâmetros físicos de som foram feitas em 1995 [Sandell e Martens, 1995; Charbonneau et al, 1997a e 1997b]. Este método de análise se mostrou efetivo e parece revelar características importantes das variações destes parâmetros, passíveis de especulação de uma “variação de timbre”.

Modelagem da Análise

Primeiramente foram calculados os valores de amplitude de cada harmônico ao longo do tempo, assumindo valores constantes de frequência dos harmônicos. Estes dados estão contidos numa matriz $\mathbf{X}_{[n \times p]}$, onde n é o número medições no tempo e p o número de harmônicos medidos. Assim, cada linha desta matriz representa uma mediação no tempo i ($i = 1, 2, \dots, n$) das amplitudes do harmônico j ($j = 1, 2, \dots, p$) dispostos nas colunas da matriz. A cada tempo i os p harmônicos podem ser considerados como coordenadas de um ponto em um espaço de dimensão p . Calcula-se a matriz de Covariância $\mathbf{S}_{[p \times p]}$ da matriz $\mathbf{X}_{[n \times p]}$, que é uma matriz simétrica de dimensão p .

Após decidir-se quantos *componentes principais* serão retidos para a reconstrução dos dados, estes são facilmente calculados multiplicando-se os valores dos componentes pelos *auto-vetores* (bases ortogonais) de cada *auto-valor* de $\mathbf{S}_{[p \times p]}$, dispostos em ordem decrescente. A porcentagem de semelhança da reconstrução dos dados com o número de *componente principais* retidos, é dada pela soma acumulada dos *auto-valores* utilizados na reconstrução. Utilizando-se todos os p *componentes principais* obtém-se uma reconstrução completa (100 %) do sinal.

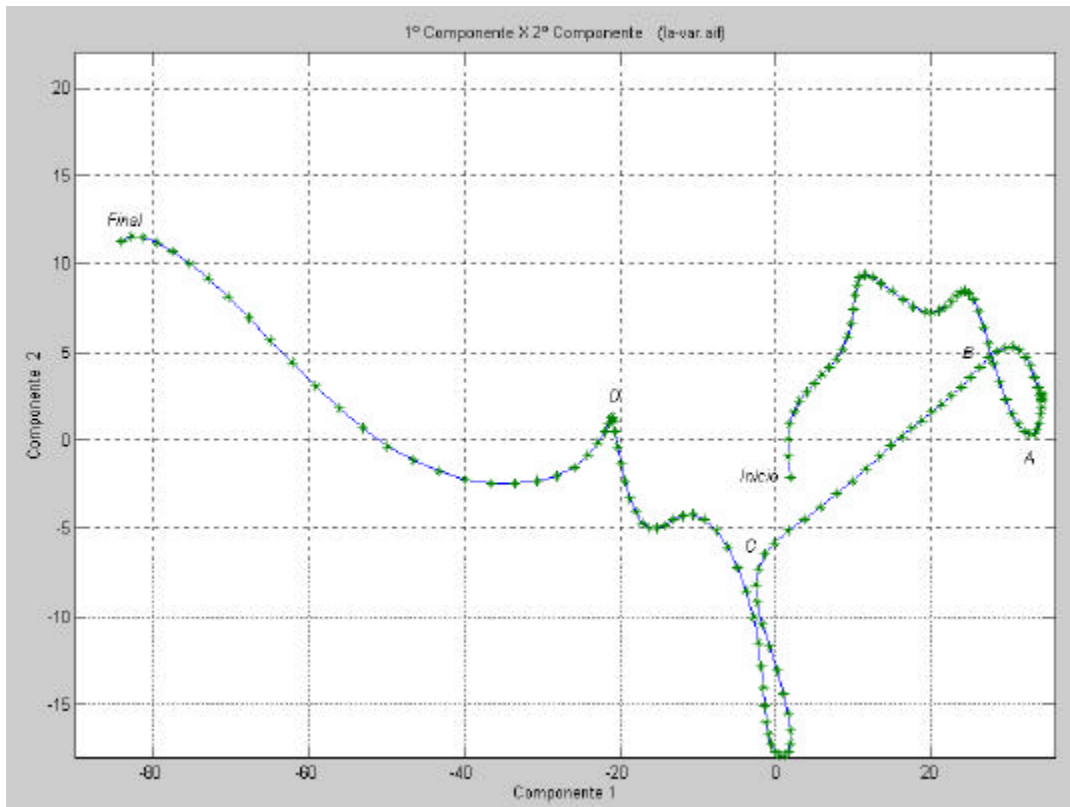
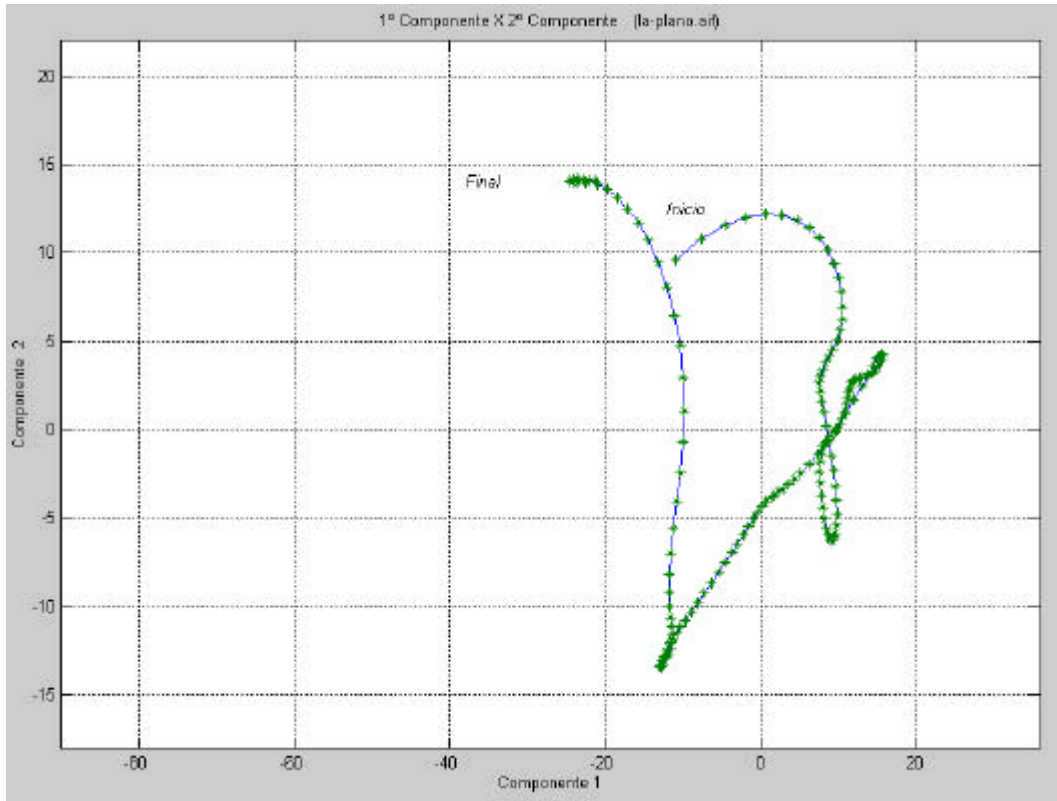
Simplificação de Dados

Após as primeiras medições, constatou-se grandes variações dos valores das amplitudes em curtos espaços de tempo, principalmente nos harmônicos de ordem superior. Estas variações impedem que a análise por *componentes principais* forneça dados conclusivos, por absorverem grande parte da detecção da máxima variância, escondendo comportamentos dos parâmetros que possam ser mais significativos para a percepção auditiva. A simplificação do sinal foi feita por eliminação de pontos e por interpolação cúbica (*spline*). A reconstrução do sinal a partir destes dados simplificados mostrou que estas grandes variações de amplitude em curtos espaços de tempo têm pouco significado auditivo (testes de audição foram suficientes para validar esta simplificação). Na análise destes exemplos aqui demonstrada, foram assumidos valores constantes de frequência dos harmônicos. No entanto, foi constatado que esta simplificação, apesar de mostrar pouca significância auditiva para estes exemplos, não pode ser considerada para outros casos, como por exemplo para alturas da região média da clarineta, onde ocorre variações mais acentuadas de frequência.

Resultados Obtidos

Uma análise dos resultados obtidos com dois exemplos de amostras é apresentada a seguir. Estas amostras têm a mesma duração (3,666 segundos) e a mesma frequência (Fá sustenido de última linha da clave de sol = 749 Hz), ambos amostrados a uma taxa de 44.1000 Hz, em 16 bits, em apenas um canal. O primeiro destes exemplos, (*nota expressiva*) corresponde ao Fá sustenido sustentado pela clarineta ao longo do 8º e metade do 9º compassos do quinteto de Brahms (op. 115). A amostra foi retirada da gravação de uma interpretação musical de toda a abertura do quinteto (até o 14º compasso). O outro exemplo (*nota plana*) corresponde à mesma nota, tocada com o mínimo de variação de intensidade e fora do contexto musical. Neste exemplo a matriz $\mathbf{X}_{[n \times p]}$, que contem os dados da transformada de Fourier tem dimensão 156 x 12, ou seja, foram medidos os valores de amplitudes dos 12 primeiros harmônicos em 156 intervalos de tempo, cada um com uma duração de 23 milissegundos (*ms*).

A Figura 1 plota o 1º *componente principal* contra o 2º da *nota expressiva*. Este gráfico equivale a projetar os dados medidos em 12 dimensões (medições das amplitudes de cada um dos 12 primeiros harmônicos) em um plano definido por 2 dimensões que representam a maior variação destes dados. Para a *nota expressiva*, estas 2 dimensões resgatam 90% do sinal e para a *nota plana*, 74% (Figura 2). As curvas mostram a evolução da nota no tempo, tendo cada instante de medição representado por um ponto (★ em cor verde). Estes pontos estão ligados em ordem cronológica com o **início** e o **final** do som marcados no gráfico. Como o intervalo de tempo de cada quadro é constante (23 *ms*), a distância entre os pontos no gráfico representa o grau de variação do timbre, ou seja, quanto mais próximos os pontos menor é a variação entre eles.



Análise dos Resultados

- A Figura 1 mostra a extensão da trajetória da variação do timbre da *nota expressiva*. Maiores variações de timbre podem ser observadas no final do som, precisamente nos últimos 25 pontos (do ponto **D** ao **final**), ou seja, nos últimos 575 ms ($575 = 23 \times 25$). Esta grande variação está relacionada com o diminuendo acentuado no final da nota. É interessante notar que o diminuendo da nota começa aproximadamente nos últimos 2 segundos, mas a análise mostra que o timbre começa sofrer maior variação somente a partir do último $\frac{1}{2}$ segundo (é possível notar isso auditivamente, porém testes mais sistemáticos deverão ser feitos).
- Ressalta-se que esta variação é mais acentuada na direção do 1º *componente principal*. Inspecionando-se as curvas individuais dos *componentes* notamos que o 1º *componente* só apresenta valores positivos e que diminui à medida que a intensidade do sinal diminui. Certamente este componente está relacionado com a quantidade física de energia do sinal.
- Observando-se, no gráfico da *nota expressiva*, os 60 pontos anteriores ao ponto **D** (compreendidos entre os pontos **B** e **D**), equivalentes a 1.380 ms e que correspondem ao período de tempo compreendido entre os instantes 1.711 ms e 3.091 ms, nota-se também uma variação de timbre acentuada na direção do primeiro *componente*, porém somente entre os pontos **B** e **C**. Depois a variação ocorre na direção do 2º *componente*, num movimento de ida e volta, por duas vezes, primeiramente verticalmente sobre o ponto **C** e depois sobre o ponto **D**. O 2º *componente* parece estar relacionado com a distribuição de energia ao longo do espectro, mas este fato ainda carece de maior investigação.

- Na região entre o **início** e o ponto **B**, ocorrem variações de timbre, porém bem menos acentuadas do que as acima mencionadas, pois verifica-se que os pontos estão bem mais próximos uns dos outros. É nesta região que ocorre o *crescendo* e portanto conclui-se que a variação de timbre é mais lenta ao longo do *crescendo*, se comparado com o *diminuendo*. Esta variação se dá em primeiro lugar predominantemente no sentido positivo do 1º *componente* (**início** até **A**) e depois também no sentido do 2º *componente* (de **A** até **B**).
- Observando-se a *nota plana* verifica-se claramente uma variação pequena no sentido do 1º *componente*. A variação no sentido do 2º *componente* se dá praticamente em dois movimentos, negativamente até os primeiros 3/4 do sinal e depois positivamente, mostrando um comportamento mais ordenado do 2º *componente* para sons mais “planos”. Este fator deve ser investigado numa comparação entre notas de diferentes alturas do instrumento.
- Verificou-se também que os valores acumulados de semelhança da reconstrução correspondente à *nota plana*, são bem inferiores os valores medidos para a *nota expressiva*, mas crescem bem mais rapidamente, praticamente igualando seus valores quando 6 *componentes principais* são retidos. Este fato parece indicar alguma relação com as flutuações de curta duração, que assumem menor importância nos *componentes principais* de ordem inferior para sinais com maior variação dinâmica. Testes deverão ser feitos com *componentes principais* de ordem superior.

Referências Bibliográficas




- Barrière, J.-B. 1991. Le Timbre, Métaphore pour la Composition. Paris: IRCAM e Cristian Bourgois.
- Chabornneau, G.; Hourdin. C.; Moussa, T. 1997a. "A Multidimensional Scaling Analysis of Musical Instrument's Time-Varying Spectra," Computer Music Journal, vol. 21, no. 2, pp. 40-55.
- Chabornneau, G.; Hourdin. C.; Moussa, T. 1997b. "A Sound Synthesis Technique Based on Multidimensional Scaling of Spectra," Computer Music Journal, vol. 21, no. 2, pp. 56-68.
- Dillon, W. R. & Goldstein, M. 1984. Multivariate Analysis - methods and applications. New York: John Wiley & sons: New York.
- Gordon, J. e Grey, J. M. 1977. "Perception of Spectral Modifications on Orchestral Instrument Tones." Computer Music Journal, vol. 2, no. 1, pp. 24-31.
- Gordon, J. e Grey, J. M. 1978. "Perceptual Effects of Spectral Modifications on Music Timbres," Journal of the Acoustics Society of America. vol. 63, pp. 1493-1500.
- Grey, J. M. 1975. An Exploration of Musical Timbre. Ph.D. Dissertation, Department of Psychology, Stanford University, Department of Music Report STAN-M-2, Palo Alto, CA.
- Grey, J. M. 1978. "Timbre Discrimination in Musical Patterns," Journal of the Acoustics Society of America. vol. 64, pp. 467-472.
- Johnson, R. A. & Wichern, D. W. 1998. Applied Multivariate Statistical Analysis. Upper Saddle, NJ: Prentice-Hall.
- Luce, D. A. 1963. Physical Correlates of Nonpercussive Musical Instrument Tones. Ph. D. Dissertation, Department of Physics, MIT, Cambridge, MA.
- Luce, D. e Clark, M. 1965. "Duration of Attack Transients of Nonpercussive Orchestral Instruments," Journal of the Audio Engineering Society. vol. 13, pp. 194-199.

- Luce, D. e Clark, M. 1967. "Physical Correlates of Brass-Instrument Tones," Journal of the Acoustics Society of America. vol. 42, pp. 1232-1243.
- McAdams, S. e Bregman, A. 1979. "Hearing Musical Streams," Computer Music Journal, vol. 3, no. 4, pp. 26-44. Reimpresso em C. Roads e J. Strawn, eds. Foundations of Computer Music. Cambridge, Massachussets: MIT Press, pp. 658-698.
- McAdams, S. 1987. "Music: A Science of Mind?," Contemporary Music Review, vol. 2. no. 1, pp. 1-61.
- Rencher, A. C. 1995. Methods of Multivariate Analysis. New York: John Wiley & sons: New York.
- Risset, J.-C. 1965. "Computer Study of Trumpet Tones," Journal of the Acoustics Society of America. (Abstracts), vol. 38, pp. 912. 1966. Bell Laboratories Report, Murray Hill, NJ.
- Risset, J.-C e Mathews, M. V. 1969. "Analysis of Musical Instrument Tones, "Physics Today, vol. 22, no. 2, pp. 23-40.
- Risset, J.-C., & D. Wessel. 1982. "Exploration of Timbre by Analysis and Synthesis," in D. Deutsch, ed. Psychology of Music. San Diego, California: Academic Press, pp. 25-58.
- Risset, J.-C. 1991. "Timbre Analysis by Synthesis: Representations, Imitations and Variants for Musical Composition," in G. De Poli, A. Piccialli e C. Roads, eds. Representation of Musical Signals. Cambridge, Massachussets: MIT Press, pp. 7-43.
- Sandell, G. J. e Martens, W. 1995. "Perceptual Evaluation of Principal-Component-Based Synthesis of Musical Timbres," Journal of the Audio Engineering Society. vol. 43, no. 12, pp. 1013-1028.
- Strong, W. e Clark, M. 1967a. "Synthesis of Wind-Instrument Tones," Journal of the Acoustics Society of America. vol. 41, no. 1, pp. 39-52.

Strong, W. e Clark, M. 1967b. "Pertubations of Synthetic Orchestral Wind-Instrument Tones," Journal of the Acoustics Society of America. vol. 41, no. 2, pp. 277-285..

Wessel, David. 1979. "Timbre Space as a Musical Control Structure," Computer Music Journal 3(2): 45-52. Reimpresso in C. Roads e J. Strawn, eds. Foundations of Computer Music. Cambridge: MIT Press, pp. 640-657.

Guia para continuar

-  **Programação da ANPPOM 1999**
-  **Informação dos Participantes**
-  **Saída dos Anais da ANPPOM**