

Timbre de um instrumento musical: caracterização e representação

Maurício Alves Loureiro (UFMG, Belo Horizonte)

mauricio@musica.ufmg.br

Hugo Bastos de Paula (PUC Minas, Belo Horizonte)

hugo@pucminas.br

Resumo: A representação do timbre de um instrumento musical envolve problemas de grande complexidade. Apesar da conhecida correlação entre o timbre e o conteúdo espectral do som, o mapeamento das características espectrais dos mais variados tipos de sons produzidos por um instrumento e sua utilização de forma semanticamente relevante exige uma metodologia de análise de dados específica. Este trabalho apresenta uma abordagem para este problema através do mapeamento das curvas de variação temporal das amplitudes dos componentes harmônicos, extraídos através da Transformada Discreta de Fourier, utilizando técnicas de Análise por Componentes Principais (PCA). As bases ortogonais definidas pela PCA possibilitaram grande redução de dados e a criação de subespaços timbrísticos capazes de representar os sons do instrumento em várias alturas e níveis de intensidade. Técnicas de classificação de dados permitiram uma análise semântica destes espaços timbrísticos, possibilitando a classificação de grupos de timbres semelhantes neste espaço e ratificando a PCA como uma forma eficiente da representação da dinâmica de timbres de instrumentos musicais.

Palavras-Chave: timbre de instrumentos musicais; espaço timbrístico; classificação de timbres; acústica musical.

Timbre of a musical instrument: characterization and representation

Abstract: The representation of the timbre of a musical instrument involves problems of great complexity. Despite the already known correlation between timbre and the spectral content of the sound, the mapping the spectral characteristics of the great variety of sounds produced by an instrument demands specific methodologies of data analysis. This study presents an approach for this problem through the mapping of the amplitude time-varying curves of the harmonic components, extracted through the Discret Fourier Transform, using techniques of Analysis of Principal Components (PCA). The orthogonal bases defined by the PCA made possible a significant data reduction and the creation of timbral sub-spaces capable of representing the sounds of the instrument in several pitch and intensity levels. Techniques of data classification allowed a semantic analysis of these timbral spaces, making possible the classification of groups of similar timbres in this space and ratifying PCA as an efficient tool for representing the dynamics of the timbre of musical instruments.

Key-words: timbre of musical instruments; timbral space; classification of timbre; musical acoustics.

1 - Introdução

1.1 - Timbre: Conceito Abstrato Multidimensional

Um instrumento musical é caracterizado por sua extensão de alturas e de níveis de intensidade e pela qualidade sonora ou timbre dos sons produzidos por ele. A representação sonológica de um instrumento musical envolve a estimação dos parâmetros físicos que contribuem para a percepção de cada um destes três atributos: *altura*, *intensidade* e *timbre*. Dentre eles, o timbre é o que apresenta maior complexidade na medição e na especificação dos parâmetros envolvidos na sua percepção. O conceito abstrato aparentemente simples de timbre refere-se comumente à cor ou à qualidade do som. É percebido a partir da interação de inúmeras propriedades estáticas e dinâmicas do som, agregando não apenas um conjunto extremamente complexo de atributos auditivos, mas também uma enorme gama de fatores que traduzem aspectos psicológicos e musicais. Sua definição oficial pela ASA (*American Standard Associa-*

tion) o dissocia dos conceitos de intensidade e altura: “atributo do sentido auditivo em termos do qual o ouvinte pode julgar que dois sons similarmente apresentados com a mesma intensidade e altura, são dissimilares” (RISSET e WESSEL, 1999). As variações de timbre são percebidas, por exemplo, como agrupamentos de sons tocados por um mesmo instrumento musical, ou falados por uma mesma pessoa, mesmo que estes sons possam ser bem distintos entre si, de acordo com sua 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 (RISSET, 1965; GREY, 1975; GORDON e GREY, 1978; GREY, 1978; MCADAMS e BREGMAN, 1979; WESSEL, 1979; RISSET, 1991). Estes trabalhos identificaram inúmeros fatores que formam o que podemos chamar de percepção do timbre, tais como: o volume do som (intensidade percebida); o envelope de amplitude (evolução da intensidade global), cuja parte inicial (ataque) pode assumir especial importância na discriminação do timbre de um instrumento musical dependendo da duração do som; flutuações de alturas e intensidades devido a vibratos ou trêmolos; estruturas dos formantes¹, que assumem maior importância na percepção de sons vocais; distribuição espectral (amplitudes das frequências dos componentes espectrais); evolução temporal da distribuição espectral. Devido à característica multidimensional deste atributo, a identificação da contribuição de cada um destes fatores concorrentes tem sido a principal questão levantada por pesquisas psicoacústicas sobre a percepção do timbre, que vêm se desenvolvendo desde o início do século.

Diferentemente de outros atributos do som musical, tais como altura, volume e duração, o timbre não pode ser associado a apenas uma dimensão física, não podendo ser especificado quantitativamente pelo sistema tradicional de notação musical como são o volume e a altura, descritos a partir de escalonamentos entre *fraco-forte* e de gamas de alturas. Isto torna mais complexa sua utilização numa composição musical, já que além de especificações de instrumentação, o timbre só pode ser especificado no sistema de notação musical tradicional a partir de instruções relacionadas a outros atributos (combinações de alturas; indicações de intensidade como *forte*, *crescendo*, que induzem o intérprete a uma variação do timbre), ou por instruções verbais que traduzem muitas vezes aspectos psicológicos.

Metodologias de testes subjetivos de similaridade, emprestadas da psicologia experimental permitiram uma redução da dimensionalidade deste atributo em representações de menor complexidade, com a introdução da noção de “taxa de similaridade” entre estímulos auditivos. Dados obtidos de respostas de julgamentos de similaridade/dissimilaridade em testes subjetivos, adequadamente manipulados através de técnicas de Estatística, tais como Escalonamento Multidimensional (MDS), possibilitaram as primeiras investigações quantitativas sobre a complexa estrutura deste atributo a partir de dados perceptivos (PLOMP, 1970; GREY, 1975; MILLER e CARTERETTE, 1975; GREY, 1977; WESSEL, 1979). Em um dos estudos mais clássicos em timbre musical, GREY (1975) mediu julgamentos subjetivos de similaridade entre pares de timbres de 16 instrumentos musicais diferentes, submeteu estes dados a um MDS e construiu

¹ Formantes são picos de amplitudes de frequências de componentes espectrais que mantêm uma certa constância, mesmo que a altura (frequência fundamental) varie.

um *espaço timbrístico* tridimensional no qual “valores” multidimensionais de timbre de instrumentos distintos foram posicionados de acordo com a similaridade/dissimilaridade entre eles. Além de mapear geometricamente o conceito de similaridade acústica, este estudo também mostrou a capacidade deste método de fornecer uma quantificação psicológica de uma estrutura relativamente complexa, a partir de dados bem simples - respostas de similaridade/dissimilaridade entre pares de timbres distintos, mostrando também perspectivas de relacionar estes dados de testes psicoacústicos a grandezas físicas e psicofísicas mensuráveis. Outros estudos propuseram tais *espaços timbrísticos*, nos quais a principal premissa foi a preservação da topologia natural de cada som representado: sons diferentes devem se colocar distantes entre si neste espaço, enquanto que sons similares se colocariam próximos uns dos outros; pontos intermediários entre estes valores poderiam representar interpolações entre timbres conhecidos (PLOMP, 1970; SHEPARD, 1972; GREY, 1975; WESSEL, 1979; PLOMP e RASCH, 1982; RISSET, 1991).

Estudos mais recentes conseguiram relacionar parâmetros físicos mensuráveis com as dimensões compartilhadas pelo timbre representado nestes espaços, combinando modelos quantitativos de relações perceptivas com explicações psicofísicas dos parâmetros identificados (HAJDA, KENDALL *et al.*, 1997; MISDARIIS, SMITH *et al.*, 1998). A possibilidade de estabelecer correlações entre fatores puramente perceptivos relacionados ao timbre e medidas acústicas extraídas diretamente do som, direcionou a pesquisa em timbre musical para abordagens mais quantitativas. Uma revisão histórica resumida da evolução da pesquisa em timbre musical pode ser encontrada em (MCADAMS, WINSBERG *et al.*, 1995).

1.2 - Correlatos Físicos do Timbre

Comprovadas as correlações entre dos dados perceptivos obtidos a partir de testes subjetivos e medições acústicas, surgem inúmeras outras abordagens na investigação de timbre baseadas em análise acústica, ou seja, a partir de dados extraídos diretamente do som. Uma metodologia adequada para este tipo de abordagem é a *Análise por Componentes Principais* (PCA), um método de *Análise Estatística de Multivariados*, que, como o MDS, constrói espaços multidimensionais a partir da variância de um conjunto de variáveis. No entanto, enquanto o MDS constrói uma representação multidimensional para relacionar variáveis embutidas em dados de julgamento de similaridade, a PCA utiliza métodos matemáticos para manipular variáveis medidas (acústicas) e construir para elas uma representação espacial similar. Trabalhos recentes aplicaram PCA a curvas de amplitude e frequência de componentes harmônicos, buscando relacioná-los com características perceptivas, produziram resultados semelhantes àqueles obtidos com testes subjetivos de similaridade aplicados aos mesmos sons e comprovaram a eficiência desta metodologia na investigação da variação dinâmica de timbre em performances musicais (COSI, DE POLI *et al.*, 1994; SANDELL e MARTENS, 1995; CHARBONNEAU, HOURDIN *et al.*, 1997a, 1997b; DE POLI e PRANDONI, 1997; ROCHEBOIS e CHARBONNEAU, 1997; BALIELLO, DE POLI *et al.*, 1998; DE POLI, RODÀ *et al.*, 1998; LOUREIRO, DE PAULA *et al.*, 2000, 2001, 2004a, 2004c).

Trabalhos de pesquisa já realizados na década de 60 mostraram ser bastante adequada a representação de sons de instrumentos musicais através de curvas de variações temporais das amplitudes e das frequências dos componentes harmônicos (LUCE, 1963; FREEDMAN, 1965; LUCE e CLARK, 1965; RISSET, 1965; LUCE e CLARK, 1967; STRONG e CLARK, 1967b,

1967a; FREEDMAN, 1968; GREY, 1975; GORDON e GREY, 1978; GREY, 1978; MCADAMS e BREGMAN, 1979). 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 freqüências, e em seguida vários envelopes que variavam de acordo com as freqüências dos harmônicos (STRONG e CLARK, 1967b). Risset analisou sons de trompete e obteve curvas individuais de evolução temporal das amplitudes e freqüências para cada harmônico. A partir destes dados ele conseguiu imitar os sons do trompete com o sistema de síntese MUSIC V, utilizando um envelope de controle para cada harmônico, construídos a partir de aproximações, por segmentos lineares, das curvas obtidas na análise. 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). A motivação de todos estes trabalhos em tomarem como ponto de partida sons de instrumentos musicais tradicionais procurando imitá-los com a maior fidelidade possível, certamente não está limitada ao propósito de duplicar eletronicamente estes sons, mas de explorar o mundo desconhecido desta propriedade do som que chamamos de timbre, procurando entender como ouvimos os sons destes instrumentos e principalmente como ouvimos a música produzida com eles.

1.3 - Variação Temporal do Timbre

Apesar dos instrumentos acústicos tradicionais oferecerem possibilidades de produzir e controlar com precisão uma vasta gama de timbres, dependendo da altura da nota, da maneira como é tocada e da habilidade do músico em seu controle dinâmico, não existe para este atributo um sistema de classificação de qualidade ou quantidade de seus estados de evolução ao longo do tempo. Investigações mais extensivas sobre a variação intencional do timbre e sobre a capacidade do ouvinte em perceber e compreender esta intenção, tornam-se cada vez mais necessárias para a compreensão da contribuição de parâmetros acústicos, entre eles a diferenciação de timbre para a condução e percepção da expressividade de uma performance musical.

Estudos sobre o timbre de instrumentos musicais, tais como os acima mencionados, têm se restringido à análise de notas musicais isoladas e quase-estáticas, em geral abordando comparações entre diferentes instrumentos musicais e fora de qualquer contexto musical, focalizando apenas o mecanismo perceptivo que discrimina um instrumento musical de outro. Poucos avanços têm sido alcançado em relação à discriminação perceptiva entre diferentes timbres produzidos em um único instrumento musical, seja em passagens musicais, ou até mesmo ao longo da extensão de uma mesma nota. Strawn e Piszczalski estudaram a transição entre as notas de um trecho musical, investigando amostras que incluíam mais de uma nota (PISZCZALSKI, 1979a, 1979b; PISZCZALSKI, GALLER *et al.*, 1981; STRAWN, 1985, 1986, 1987). Outros estudos investigaram a variação dos parâmetros do som em performances musicais procurando compreender o controle dinâmico e instantâneo que os instrumentistas detêm sobre o som de seus instrumentos (EGOZY, 1995; DUBNOV, TISHBY *et al.*, 1996; MASRI e BATEMAN, 1996; DUBNOV, 1997; DUBNOV e RODET, 1997; ROVAN, WANDERLEY *et al.*, 1997; ARCOS, DE MÁNTARAS *et al.*, 1998; BALIELLO, DE POLI *et al.*, 1998; DANNENBERG e DERENYI, 1998; DE POLI, RODÀ *et al.*, 1998; FRIBERG, BRESIN *et al.*, 1998)

Focalizado no timbre de um único instrumento, este estudo investigou modelos para representar a grande variedade de sonoridades produzidas por um único instrumento musical, a partir de

parâmetros espectrais extraídos de sons produzidos em execuções musicais, ao longo de toda a extensão de alturas e intensidades do instrumento, buscando identificar a variação do timbre de uma ou mais notas ao longo de execuções de frases musicais representativas do repertório do instrumento analisado e compreender o significado de parâmetros físicos determinantes da variação do timbre de um instrumento, na realização de um gesto “expressivo” numa interpretação musical. O conhecimento da estrutura complexa deste atributo, possibilitaria inferências, predições e classificações, facilitando inúmeras aplicações em análise musical, composição para instrumentos acústicos, composição eletroacústica (síntese, classificação e identificação).

2 - Bases Espectrais para a Caracterização de Timbre

2.1 - Delimitação do Conjunto de Timbres

Com o propósito de representar o timbre de um instrumento musical a partir de parâmetros acústicos, foi definido um conjunto adequado de sons que buscou incluir o máximo de timbres distintos produzidos ao longo de toda sua extensão de alturas. Duas simplificações foram consideradas na definição da base de dados para estes estudo: (1) foi limitada aos sons comumente tocados em instrumentos musicais em performance de música erudita ocidental tradicional, excluindo-se sonoridades produzidas no instrumento em contextos de outras tradições musicais, como também aqueles utilizados na música contemporânea conhecidos como técnicas estendidas; (2) foi apenas considerada a parte contínua de sons relativamente longos, afim de facilitar a estimação de parâmetros espectrais, excluindo ataque, decaimentos e transições entre notas sucessivas, limitando a investigação à percepção de variação lenta de timbre musical que geralmente acontece ao longo de notas mais longas durante uma performance musical.

Variações intencionais de timbre, juntamente com flutuações de intensidade e duração são geralmente usadas pelo instrumentista para conduzir suas intenções expressivas. Embora o timbre possa variar independentemente da intensidade ou da duração, o alto grau de correlação entre timbre e intensidade facilita a amostragem de “valores” distintos de timbre para uma mesma nota a partir da especificação da intensidade. Assim, instruindo o músico a executar cada nota em quatro níveis de intensidade distintos (*pianíssimo*, *mezzo-piano*, *mezzo-forte* e *fortíssimo*), com o mínimo de variação possível, nos possibilitou amostrar quatro timbres diferentes para cada nota. Os níveis extremos deveriam ser estabelecidos como o mais *piano* ou mais *forte* possíveis, respectivamente, no âmbito das sonoridades comumente utilizadas na música ocidental tradicional e níveis intermediários definidos a partir de comparação com aqueles extremos. As amostras foram obtidas a partir de gravações de alta qualidade de todas as notas dos dois registros mais graves da Clarineta em Si bemol, variando de *Ré 3* (147 Hz) a *Lá 5* (880 Hz), executadas nos quatro níveis de intensidade definidos acima, com uma duração média de 3 segundos. As gravações foram feitas nos estúdios da Escola de Música da UFMG.

2.2 - Estimação de Parâmetros Espectrais

As curvas de amplitude dos componentes harmônicos foram usadas como parâmetros iniciais para o modelo de representação sonológica do timbre. O espectro harmônico variante no tempo foi extraído utilizando a Transformada de Fourier de Curta Duração, de acordo com o método de McAulay e Quatieri que determina os valores de picos de amplitude (*peak detection*) de uma Transformada de Fourier e estabelece uma correspondência entre os picos mais próximos de quadros de análise adjacentes (*peak continuation*), associando estes valores a valores instantâneos de frequência e amplitude de componentes harmônicos (MCAULAY e QUATIERI, 1986; SERRA, 1997). Com

a finalidade de reduzir a complexidade dos dados, foi considerado: (1) que os sons amostrados podem ser representados por uma soma de senóides cujos valores de amplitude e freqüência não variam abruptamente ao longo de sua duração; (2) que componentes com intensidade mais de 60 dB abaixo do nível de máximo podem ser descartados sem afetar a percepção do timbre como um todo; (3) e que as curvas de amplitude podem ser suavizadas por um filtro passa-baixa com freqüência de corte de 10 Hz sem afetar a qualidade sonora. Adotando um quadro de análise de tamanho relativamente pequeno de 1024 amostras (23,2 ms) e um salto de 512 amostras (11,6 ms), foi possível capturar flutuações espectrais na parte contínua do som, mesmo aquelas dos harmônicos de freqüência mais elevada, bem mais acentuadas. A Fig.1 exibe a distribuição espectral dos 18 primeiros harmônicos de quatro sons amostrados do *Si bemol 3* (233 Hz), na região central da clarineta. Além da diferença entre as intensidade das notas (o gráfico está em escala linear de amplitude), podemos notar também que a distribuição de intensidade dos harmônicos varia de maneira não uniforme entre as amostras. A ausência de harmônicos superiores no som *pianíssimo* é responsável por uma qualidade mais suave e aveludada, enquanto a presença acentuada do terceiro e quinto harmônicos no som *fortíssimo* produz um timbre mais brilhante.

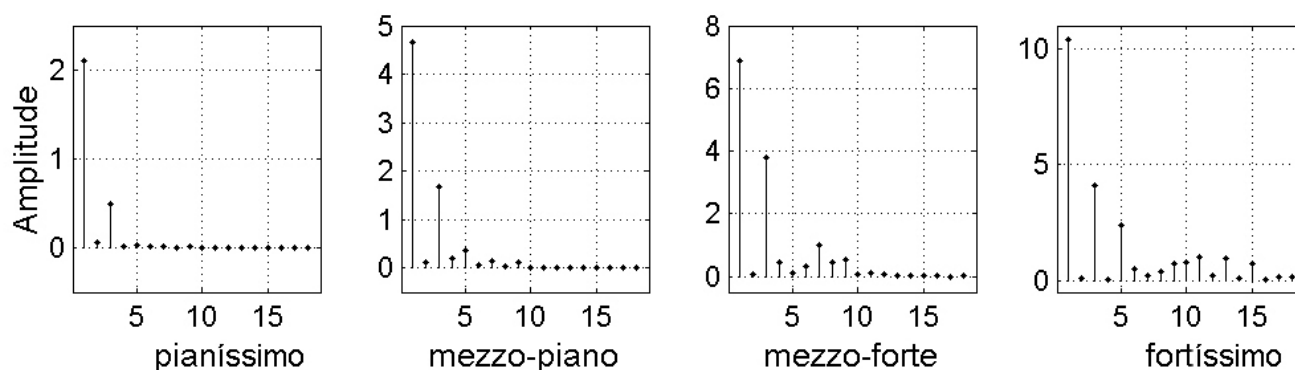


Fig.1: Espectros originais do Si bemol 3 nas intensidades *pp*, *mp*, *mf* e *ff* (região central da nota).

2.3 - Determinação das Bases Espectrais - Análise por Componentes Principais

O primeiro passo para a construção de um modelo capaz de representar a grande variedade de distribuições espectrais contidas nos sons que podem ser produzidos por um instrumento musical é reduzir o volume de dados obtidos da análise. O alto nível de correlação dos parâmetros espectrais apresentado, tanto no domínio da freqüência quanto no domínio do tempo, é uma característica comum da distribuição espectral de sons de instrumentos musicais e permitiu uma redução eficiente de dados a partir da *Análise por Componentes Principais* (PCA). Aplicada a um conjunto de variáveis multidimensionais, a PCA calcula uma base ortogonal determinada pelas direções de máxima variância dos dados analisados. As projeções dos dados originais nesta base, denominadas componentes principais (PCs), seguem trajetórias que acumulam a máxima variância em ordem decrescente, o que permite uma representação aproximada dos dados, utilizando apenas um número reduzido de dimensões (RENCHER, 1995; ROCHEBOIS e CHARBONNEAU, 1997; JOHNSON e WICHERN, 1998).

Dada a matriz de dados $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{N_Q}]$, que contém os N_Q valores de amplitudes dos N_H harmônicos ao longo do tempo, a projeção de \mathbf{X} em PCs é determinada por:

$$\mathbf{Y} = \mathbf{U}'\mathbf{X},$$

onde \mathbf{U} é a base ortogonal determinada por *Decomposição de Valor Singular* (SVD) da matriz de covariância de \mathbf{X} .

Os espectros originais podem ser recuperados pela soma das bases adequadamente ponderadas pelas amplitudes das trajetórias correspondentes, utilizando-se apenas um número reduzido de bases $N_B < N_H$, pela equação :

$$\mathbf{X} \approx \mathbf{U}\mathbf{Y},$$

que pode ser mais facilmente visualizada por sua representação matricial:

$$\begin{bmatrix} x_{11} & \cdots & x_{1N_Q} \\ x_{12} & \cdots & x_{2N_Q} \\ \vdots & \vdots & \vdots \\ x_{N_H1} & \cdots & x_{N_HN_Q} \end{bmatrix} \approx \begin{bmatrix} u_{11} & \cdots & u_{1N_B} \\ u_{12} & \cdots & u_{2N_B} \\ \vdots & \vdots & \vdots \\ u_{N_H1} & \cdots & u_{N_HN_B} \end{bmatrix} \mathbf{X} \begin{bmatrix} y_{11} & \cdots & y_{1N_Q} \\ y_{12} & \cdots & y_{2N_Q} \\ \vdots & \vdots & \vdots \\ y_{N_B1} & \cdots & y_{N_BN_Q} \end{bmatrix}$$

onde a matriz \mathbf{U} contém agora apenas as N_B bases utilizadas na reconstrução dos espectros contidos na matriz \mathbf{X} e \mathbf{Y} os N_B primeiros componentes principais (de maior variância). Pode-se pensar nestas bases como sendo um conjunto reduzido de configurações espectrais cuja soma ponderada por envelopes de amplitude definidos pelas trajetórias dos *componentes principais*, são capazes de reconstruir o sinal com uma boa aproximação (HORNER, BEAUCHAMP *et al.*, 1993; HORNER, 1995; HORNER e NGAI-MAN, 1996).

2.4 - Bases Espectrais de uma Nota Isolada

Primeiramente foi calculado, para cada som amostrado de cada nota, uma base ortogonal que, associada aos envelopes dos componentes principais correspondentes, foram capazes de reconstruir o som com grande precisão. Sub-espacos espectrais foram então construídos para representar as distribuições espectrais de todos os sons possíveis de uma única nota, calculando uma única base espectral tendo como dados de entrada a concatenação das quatro amostras desta nota nas quatro intensidades, *pp*, *mp*, *mf* e *ff*, como anteriormente definidas. Os quadros de tempo utilizados nesta concatenação foram extraídos da região de sustentação da nota, não contendo ataques nem decaimentos. As amostras foram normalizadas em amplitude e duração, contendo cada uma 75 quadros de tempo, equivalente a 870 ms. Trabalhos anteriores mostraram que os cinco primeiros componentes principais foram capazes de reconstruir todos os sons, sem qualquer perda perceptível das suas características timbrísticas (LOUREIRO, DE PAULA *et al.*, 2000).

A Tab.1 mostra a variância acumulada explicada pelos primeiros cinco Componentes Principais em cada execução individual da nota *Si bemol 3* (233 Hz), comparada com a variância obtida quando a PCA é aplicada a todas as execuções desta nota. O primeiro componente explica sozinho não menos que 74% da variância total para cada som isoladamente (primeiras quatro linhas da Tab.1), mas apenas 68,7% se a PCA é calculada a partir do conjunto de quatro sons (quinta linha da Tab.1). Uma reconstrução de 99% foi obtida com três PCs para cada som isolado, mas são necessários cinco PCs se a PCA é aplicada ao conjunto dos quatro sons.

| Número de PCS Utilizados | 1 | 2 | 3 | 4 | 5 |
|--------------------------|------|------|------|------|------|
| Si bemol 3 pp | 87,3 | 99,8 | 99,9 | 100 | 100 |
| Si bemol 3 mp | 96,3 | 99,7 | 99,9 | 99,9 | 100 |
| Si bemol 3 mf | 76,1 | 96,2 | 99,2 | 99,6 | 99,8 |
| Si bemol 3 ff | 74,4 | 97,2 | 98,6 | 99,5 | 99,7 |
| Si bemol 3 pp-mp-mf-ff | 68,7 | 89,5 | 94,4 | 97,2 | 99,0 |

Tab.1: Variância acumulada contida nos cinco primeiros PCs para diferentes bases de PCA: cada som isolado do Si bemol três (233 Hz) e todos os quatro sons do Si bemol 3.

Testes auditivos (*Mean Opinion Score*) de julgamento de similaridade foram realizados para verificar a perda perceptiva desta representação, nos quais notas originais foram comparadas em pares com as respectivas reconstruções obtidas com diferentes quantidades de PCs. Foi solicitado aos participantes que avaliassem a similaridade entre os sons numa escala de 1 a 5: (1) nenhuma relação com o som original foi identificada; (2) amostra sintetizada muito distorcida, mas com similaridade suficiente com o som original de modo a ser reconhecível; (3) identificação com o som original foi evidente; (4) uma discriminação entre os estímulos foi percebida, mas não foi possível identificar qual estímulo é o melhor ou o original; (5) não houve diferença percebida. Nenhum som comparado com sua reconstrução obtida com cinco PCs foi avaliado com pontuação menor que 4,2, com variância de 2%, ou seja, mesmo que diferenças entre o sinal original e o sinal sintético fossem percebidas, os ouvintes não souberam julgar qual deles era o som natural. Estudos anteriores testaram esta discriminação utilizando os mesmos sons, mas dentro de contextos musicais (DE PAULA, 2000), nos quais excertos de notas extraídas de gravações de performances musicais foram resintetizados usando diferentes quantidades de PCs. A partir de uma normalização adequada de amplitude e com uma sobreposição de 6 ms de *fade-in* e *fade-out*, os sons sintéticos foram reinseridos nas mesmas gravações, as quais foram submetidas a testes auditivos. Com cinco PCs, todos os pares foram avaliados com a pontuação 5. A partir destes resultados, o número de cinco PCs foi considerado suficiente para reconstruir os sons sem qualquer perda perceptível das características de timbre, o que permitiu uma taxa de redução de dados de 64:1.

Comparando os espectros originais extraídos de um ponto situado na região central da nota (Fig.1) com as bases espectrais geradas pela PCA (Fig.2a), observamos que a Base 1, que acumula a maior variância do espectro, tende a preservar em média as características dos espectros originais, evidenciando uma predominância dos harmônicos ímpares de ordem inferior (1º, 3º e 5º), característica inerente dos sons da clarineta neste registro. Isso se deve ao fato de que aproximadamente 90% da energia e da variância do sinal estão concentradas nestes harmônicos. Nota-se que a trajetória dos envelopes associados a esta primeira base (Fig.2b) acompanha o envelope de amplitude do som. Harmônicos de ordem superior aparecem nos sons fortes e são captados pelas bases de ordem superior: o 9º harmônico, presente no *mezzo-forte* e no *fortíssimo*, já aparece na segunda base e o 13º harmônico, presente apenas no *fortíssimo*, é captado pela terceira base. Nota-se a presença de componentes harmônicos negativos nas bases de ordem superior, os quais são responsáveis pelo cancelamento de componentes à medida que as bases vão descrevendo a trajetória dos harmônicos de ordem superior.

Testes auditivos de discriminação mostraram a eficácia deste modelo de representação. Erros relativos quadráticos calculados para as reconstruções com bases globais mostraram que maiores distorções ocorrem nas regiões de ataque e decaimento das notas. Nestas regiões há maior ocorrência de inarmonicidade, variações bruscas de amplitude e frequência e baixos valores de amplitude (devido ao *fade-in* e *fade-out*), o que impede que estas variações sejam

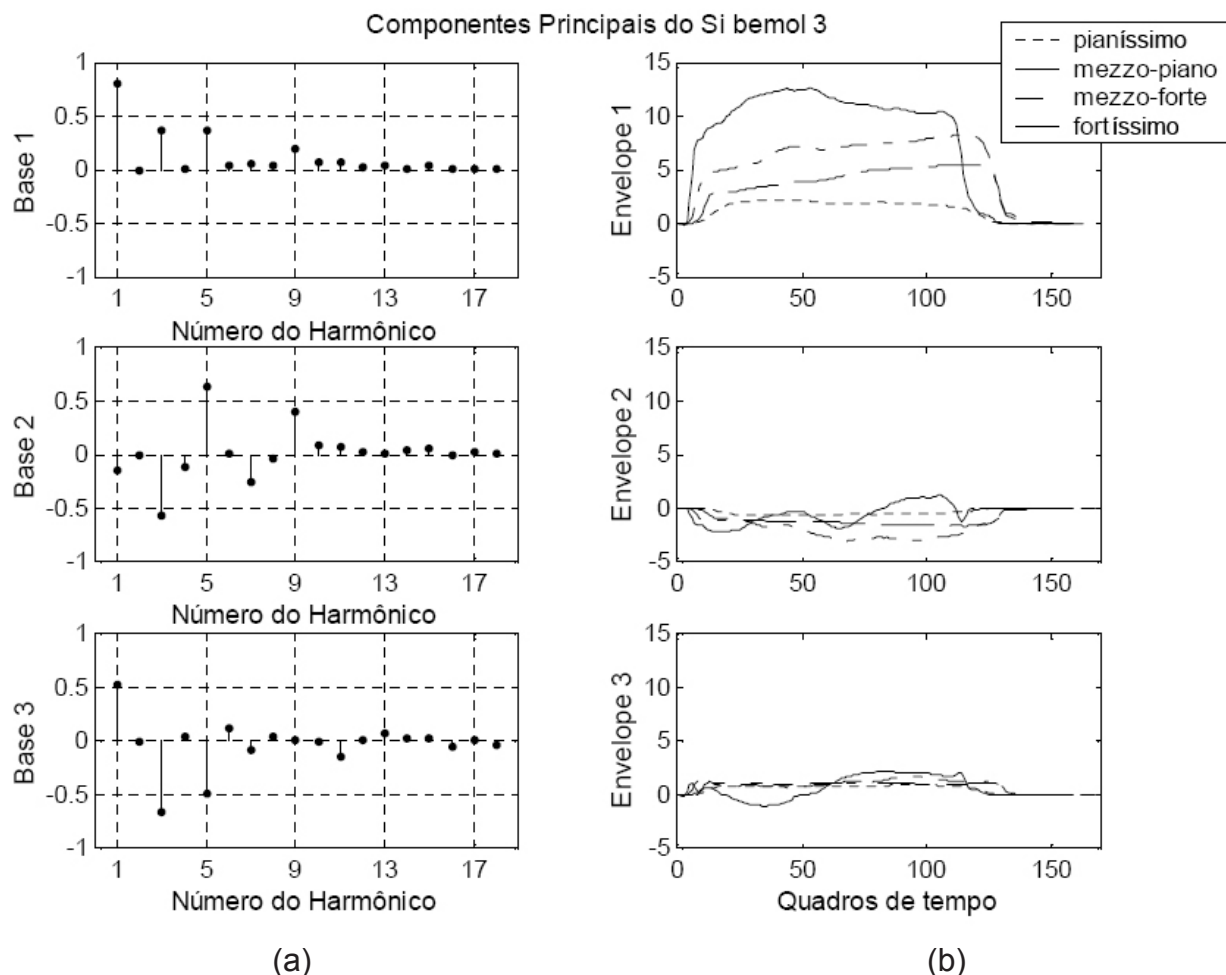


Fig.2: (a) três primeiras bases espectrais determinadas pela PCA da nota Si bemol 3 em *pianissimo*, *mezzo-piano*, *mezzo-forte* e *fortissimo*; (b) envelopes de amplitudes associados a cada base.

2.5 - Base Espectral de um Conjunto de Notas

Notas contíguas apresentaram bases individuais com características semelhantes, o que implica que a dinâmica destas notas estão sendo mapeadas para sub-espacos semelhantes, o que permitiu expandir o tamanho destes sub-espacos possibilitando a representação dos espectros de um conjunto de notas em todas suas configurações de intensidades. Assim como no caso de uma nota isolada, a PCA foi aplicada a um conjunto de espectros representativo destes sons, constituído pela concatenação de amostras normalizadas em intensidade e duração, primeiramente contendo quatro notas contíguas, de timbre semelhante e pertencentes a uma região bastante característica do instrumento. Foram tomadas as quatro notas, *Lá 3* (220 Hz), *Si bemol 3* (233 Hz), *Si 3* (247 Hz) e *Dó 4* (262 Hz), tocadas nos quatro níveis de intensidade definidos.

As bases assim geradas constituem o *Espaço Timbrístico* dessas quatro notas, no qual cada som ocupa uma posição única de acordo com sua configuração espectral. Uma comparação entre as curvas de amplitude dos harmônicos dos sons originais e de suas reconstruções geradas a partir desses sub-espacos espectrais, mostra que o modelo é eficiente para representação dos harmônicos de maiores amplitudes. A Fig.3 compara curvas de amplitude dos 1º, 3º, 5º e 7º harmônicos do *Si bemol 3* original com suas versões resintetizadas, calculadas com os primeiros cinco PCs do sub-espaco espectral definido por estas quatro notas (*Lá 3 - Dó 4*). A captura dos parâmetros relacionados a harmônicos de ordem superior e mesmo daqueles harmônicos de baixa ordem com baixos valores de amplitude apresentaram maiores margens de erro.

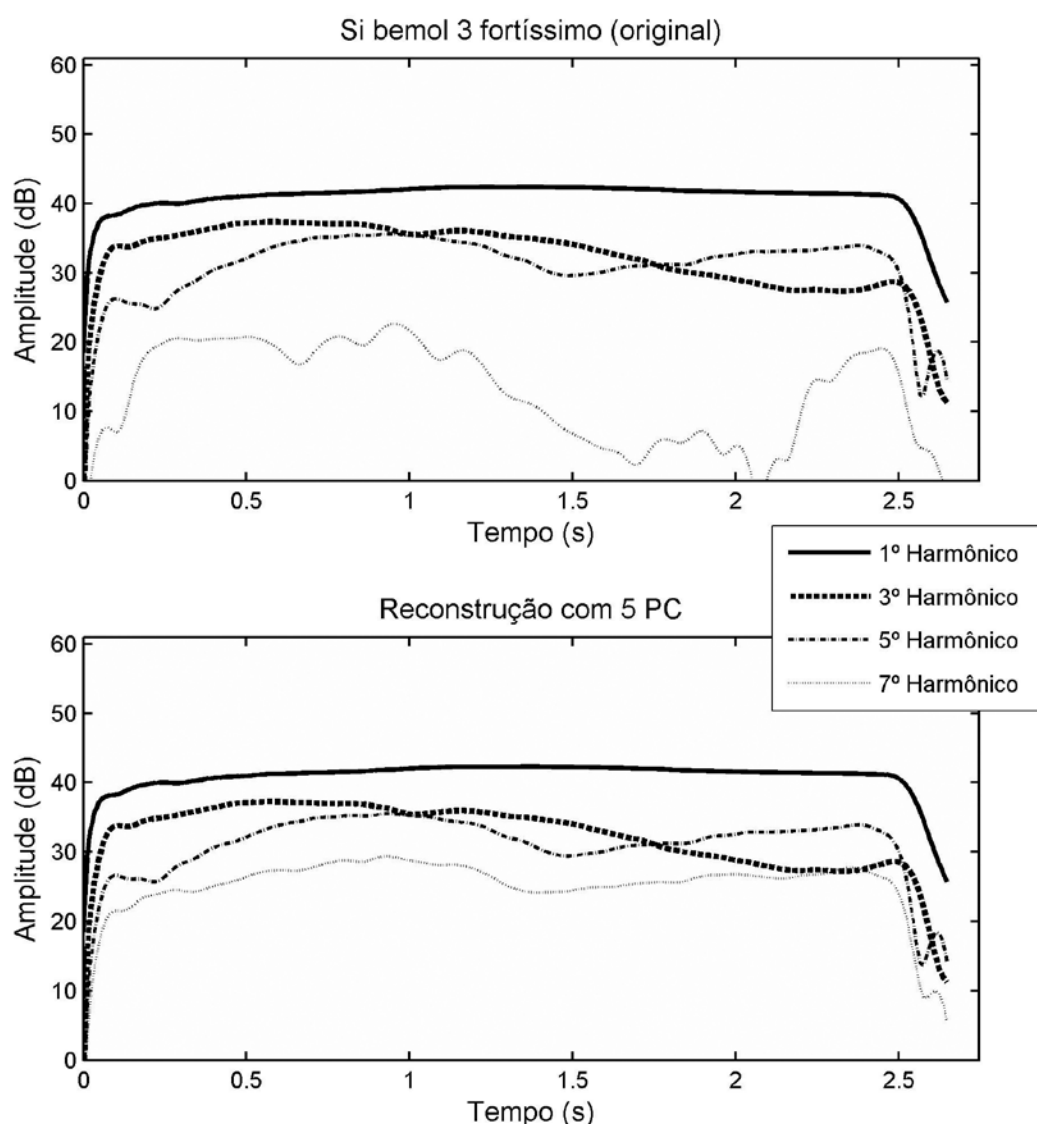


Fig.3: 1º, 3º, 5º e 7º harmônicos do Si bemol 3 *ff* original (acima) e sua versão de resintetizada (abaixo) usando o sub-espaco espectral das notas Lá 3 (220 Hz), Si bemol 3 (233 Hz), Si 3 (247 Hz) e Dó 4 (262 Hz) calculados com cinco componentes principais.

A Fig.4 mostra as três primeiras dimensões da base gerada para este conjunto de notas, ao lado das bases calculadas individualmente para três destas notas. Nota-se a semelhança entre as primeiras dimensões e a consistência da base global em capturar a predominância dos primeiros harmônicos

ímpares para todas as notas (1º, 3º, 5º e 7º harmônicos). A segunda dimensão deste *sub-espaço* parece concentrar-se na descrição dos 3º e 5º harmônicos seja por soma ou cancelamento. Sua semelhança com a segunda dimensão do *Si bemol 3* sugere que as características timbrísticas desta nota devem situar-se numa posição média em relação a este conjunto de notas.

Observando-se a consistência deste modelo de representação do timbre, uma única base espectral foi calculada para grupos maiores de notas em todos os quatro níveis de intensidade, com a finalidade de representar uma variedade cada vez maior de distribuições espectrais que o instrumento é capaz de produzir. Como se esperava, quanto maior o número de notas envolvidas, menos eficiente a representação se torna, do ponto de vista estatístico de representação de variância acumulada. Por outro lado, comparações entre notas de alturas distintas tornam-se mais complexas, tendo em vista que o timbre pode variar acentuadamente em função da nota tocada (altura), dependendo do instrumento. Sons de clarineta, tais como estes utilizados neste estudo, apresentam irregularidades na variação de timbre de nota para nota, que pode as vezes ser bastante acentuada, dependendo da região do instrumento, como por exemplo aquela mudança abrupta de timbre que ocorre entre os registros grave e médio, uma característica bem conhecida da clarineta.

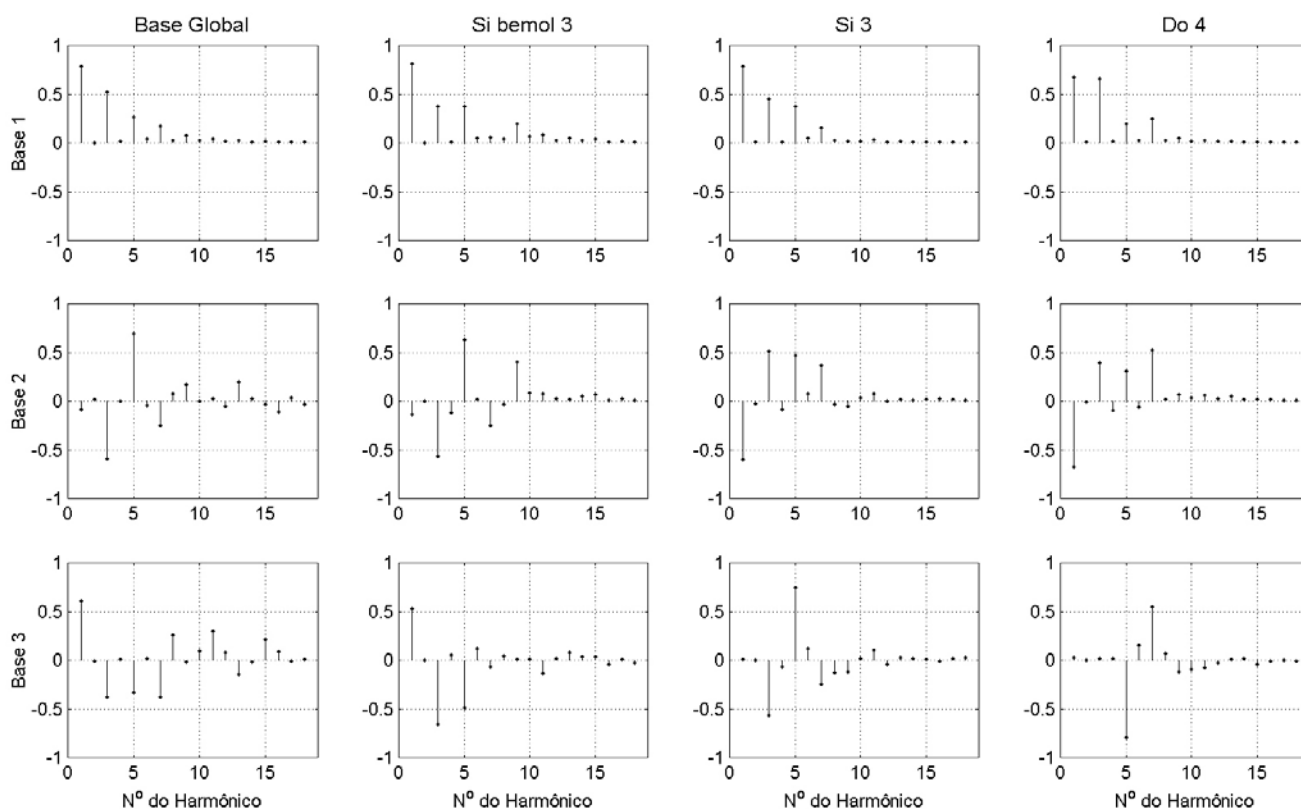


Fig.4: Bases espectrais do sub-espaço das notas Lá 3, Si bemol 3, Si 3 e Dó 4 (1ª coluna); bases espectrais individuais das notas Si bemol 3, Si 3 e Dó 4 (2ª coluna, 3ª e 4ª colunas).

Primeiramente calculou-se uma base para todas as 19 notas do registro grave do instrumento (76 amostras), de *Ré 3* (147 Hz) a *Lá bemol 4* (415 Hz) e em seguida para os dois registros mais graves, de *Ré 3* (147 Hz) a *Lá 5* (880 Hz), compreendendo 32 notas (128 amostras).

A Tab.2 compara a variância acumulada da base do *Si bemol 3* com a PCA calculadas para estes três grupos de notas: (1) as quatro notas contíguas acima mencionadas, de *Lá 3* a *Dó 4*; (2) as 19 notas do registro grave do instrumento; (3) e as 32 notas dos dois registros mais graves.

| Números de PCs | 1 | 2 | 3 | 4 | 5 |
|---|------|------|------|------|------|
| Si bemol 3 pp-mp-mf-ff | 68,7 | 89,5 | 94,4 | 97,2 | 99,0 |
| Lá 3 – Dó 4 | 56,4 | 80,4 | 90,3 | 94,2 | 96,9 |
| Ré 3 – Lá bemol 4 (registro mais grave) | 59,2 | 77,9 | 87,2 | 92,8 | 96,7 |
| Ré 3 – Lá 5 (registros grave e médio) | 53,4 | 71,8 | 84,1 | 89,7 | 94,2 |

Tab.2: Variância acumulada contida nos cinco primeiros PCs para diferentes bases de PCA: os quatro sons de Si bemol 3 (233 Hz); de Lá 3 (220 Hz) a Dó 4 (262 Hz); todo o registro mais grave, de Ré 3 (147 Hz) a Lá bemol 4 (415 Hz); ambos os registros grave e médio, de Ré 3 (147 Hz) a Lá 5 (880 Hz).

3 - Trajetórias do Espaço Timbrístico

A redução da dimensionalidade obtida com a PCA possibilitou representar a distribuição espectral em espaços de baixas dimensões. A Fig.5 mostra trajetórias tridimensionais dos quatro sons da nota *Si bemol 3* ao longo de seu próprio espaço². A correlação entre nível de intensidade e o primeiro PC é evidente: pontos espectrais pertencentes a cada som estão separados em grupos, posicionados em ordem crescente de intensidade, de *pp* a *ff*, ao longo da dimensão do primeiro PC. Quase confinados em suas posições ao longo do primeiro PC, os segundo e terceiro PCs variam diferentemente em direções distintas para cada som.

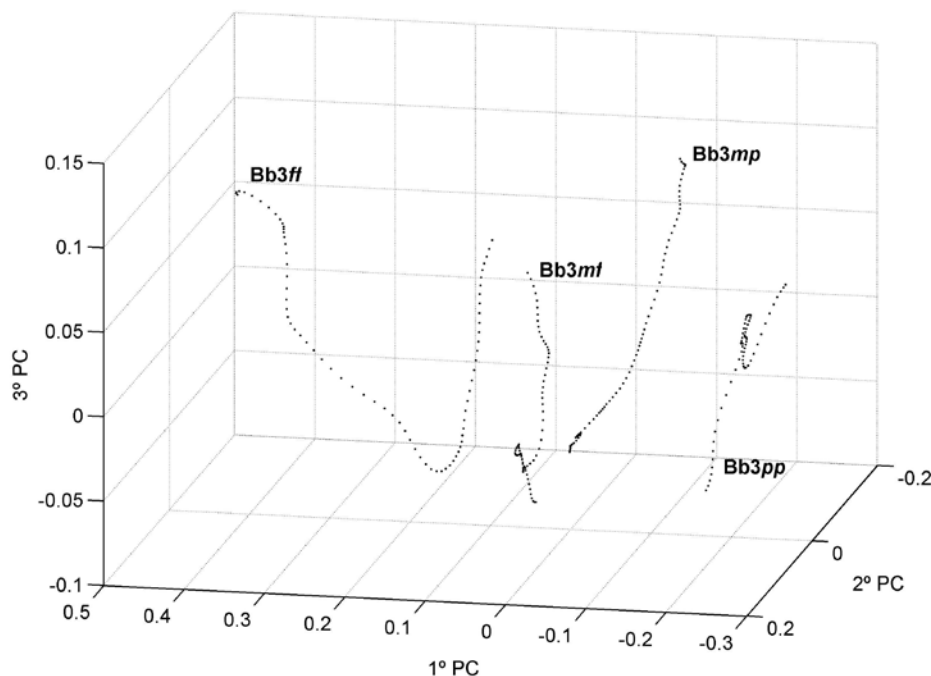


Fig.5: Trajetórias tridimensionais dos quatro sons da nota Si bemol 3 em seu próprio espaço (a nomenclatura inglesa foi adotada para nomear as notas musicais - A, B, C, D, E, F, G, “b” minúsculo para designar “bemol” e o símbolo “#” para sustenido).

² Na geração dos gráficos foi necessário abreviar os nomes das notas musicais para designar seus pontos correspondentes, adotando-se nas figuras a nomenclatura inglesa (A, B, C, D, E, F, G), “b” minúsculo para designar “bemol” e o símbolo “#” para sustenido.

A Fig.6 mostra as quatro notas contíguas acima mencionadas (*Lá 3*, *Si bemol 3*, *Si 3* e *Dó 4*), representadas no espaço espectral definido por elas. Observa-se aqui também que todos pontos pertencentes a um mesmo som se agrupam contiguamente e que a mesma correlação entre nível de intensidade e o primeiro PC, observada anteriormente é confirmada mesmo para sons de alturas distintas. Além disso, observa-se também que, enquanto notas *pp* e *mp* tendem a se compactar no lado esquerdo do espaço, notas de maior intensidade tais como *ff* de *Lá 3*, *mf* de *Lá 3* e *ff* de *Dó 4* têm suas trajetórias mais espalhadas no lado direito.

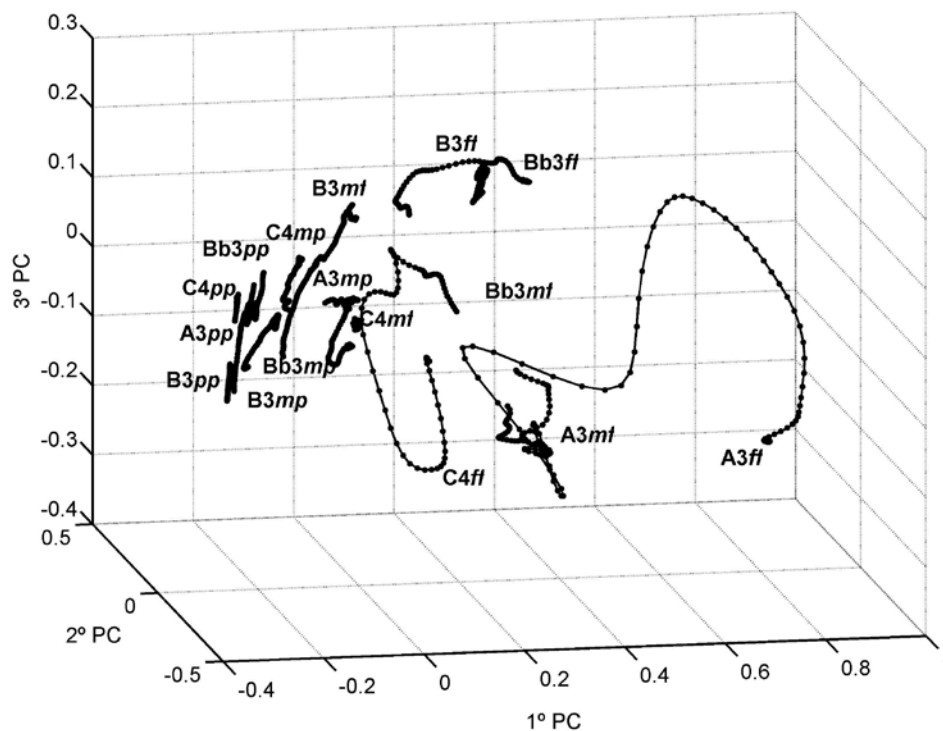


Fig.6: Trajetórias tridimensionais dos quatro sons das notas Lá 3, Si bemol 3, Si 3 e Dó 4 no espaço espectral definido por elas.

4 - Classificação de Timbres no Espaço Timbrístico da Clarineta

4.1 - Análise de agrupamentos (*clusters*) através de *K-means*

Uma análise de agrupamento utilizando usando o algoritmo *K-means* (KAUFMAN e ROUSSEEUW, 1989), foi realizada visando classificar a distribuição dos timbres ao longo de todo o instrumento. O objetivo deste algoritmo é organizar um determinado conjunto de N pontos \mathbf{X}_n de M dimensões em K agrupamentos (*clusters*), tendo cada um deles um ponto representativo \mathbf{z}_k , normalmente escolhido como o centróide destes pontos pertencentes a este agrupamento. Considerando que o algoritmo *K-means* pode convergir para um ponto de ótimo local, os valores iniciais destes centróides devem ser escolhidos adequadamente, ou o algoritmo iterativo de deve ser executado exaustivamente partindo-se aleatoriamente de pontos iniciais distintos. Este trabalho utilizou a Distância Euclidiana Quadrática para o cálculo da variância dos *clusters* (testes com outros tipos de distância mostraram resultados semelhantes). Toda vez que uma iteração produz um cluster vazio, o algoritmo cria um novo cluster constituído

pelo ponto mais afastado de seu centróide. Para evitar mínimos locais, o algoritmo das *K-means* foi executado 40 vezes com condições iniciais aleatórias e a melhor solução foi escolhida.

Inicialmente, os 16 sons das quatro notas da Fig.2 foram classificados a partir de suas trajetórias de 75 pontos obtidas com os cinco primeiros PC. A análise de cluster distribuiu todos os 16 sons em seis *clusters*, sendo que todas as notas tiveram todos seus 75 pontos espectrais alocados em um único *cluster*. Uma nova análise de *cluster* foi então executada com as 19 notas (76 sons) do registro mais grave da clarineta. Nove *clusters* forneceram a melhor correlação entre testes auditivos e a classificação obtida para este conjunto de sons (LOUREIRO, DE PAULA *et al.*, 2004b). Muito pouco destes sons tiveram seus pontos espectrais divididos em clusters distintos e quando isto aconteceu, em não mais que dois *clusters* e nesses casos, o *cluster* ao qual pertenceu a parte central do som sempre foi aquele que continha a maioria dos pontos.

A Fig.7 mostra as 11 notas mais graves da clarineta (de *Ré 3* a *Dó 4*), representadas pela localização de seu ponto central no *espaço timbrístico* calculado para o registro mais grave do instrumento. A figura mostra um grande agrupamento de sons de sons próximo à origem do espaço (à esquerda) que inclui as notas *pp* e *mp* (ou ambos) de todas as notas de *Ré 3* a *Dó 4*, com exceção do *Mi bemol 3*. Sons *mf* e *ff* ficaram mais espalhados ao longo de todas as três dimensões, confirmando a tendência já observada de que a diferenciação de nível de intensidade resulta num espalhamento mais acentuado do que a diferenciação de altura. Essa tendência pode ser observada com maior clareza na Fig.8 que ordena por altura todos os 76 sons do registro mais grave da clarineta mostrando os *clusters* os quais pertencem. Esta figura é uma projeção planar da Fig.7, que realça a correlação do *cluster* com o nível de intensidade e mostra também que as notas são espalhadas mais acentuadamente pelo espaço em decorrência da variação de intensidade do que da variação de altura.

Testes auditivos mostraram uma forte relação entre o brilho percebido e o *cluster* ao qual pertence a nota. Devido à relação conhecida entre centróide espectral e a percepção de brilho, os *clusters* foram ordenados de acordo com a média do centróide espectral dos sons contidos neles. Pode-se observar que os três primeiros *clusters* agrupam quase todas as notas *pp* e *mp*. Além disso, notas mais agudas em *mf* e *ff* foram também alocadas nestes clusters. Enquanto que as notas mais agudas foram compactamente posicionadas nestes *clusters*, os quatro últimos clusters contêm apenas notas *mf* e *ff* da oitava mais grave do instrumento, com exceção das notas *Mi 4 mf* e *Fá 4 ff*. Isto indica que as notas mais agudas tendem a variar menos o timbre em função da intensidade, ao passo que notas mais graves exibem uma maior variação de timbre.

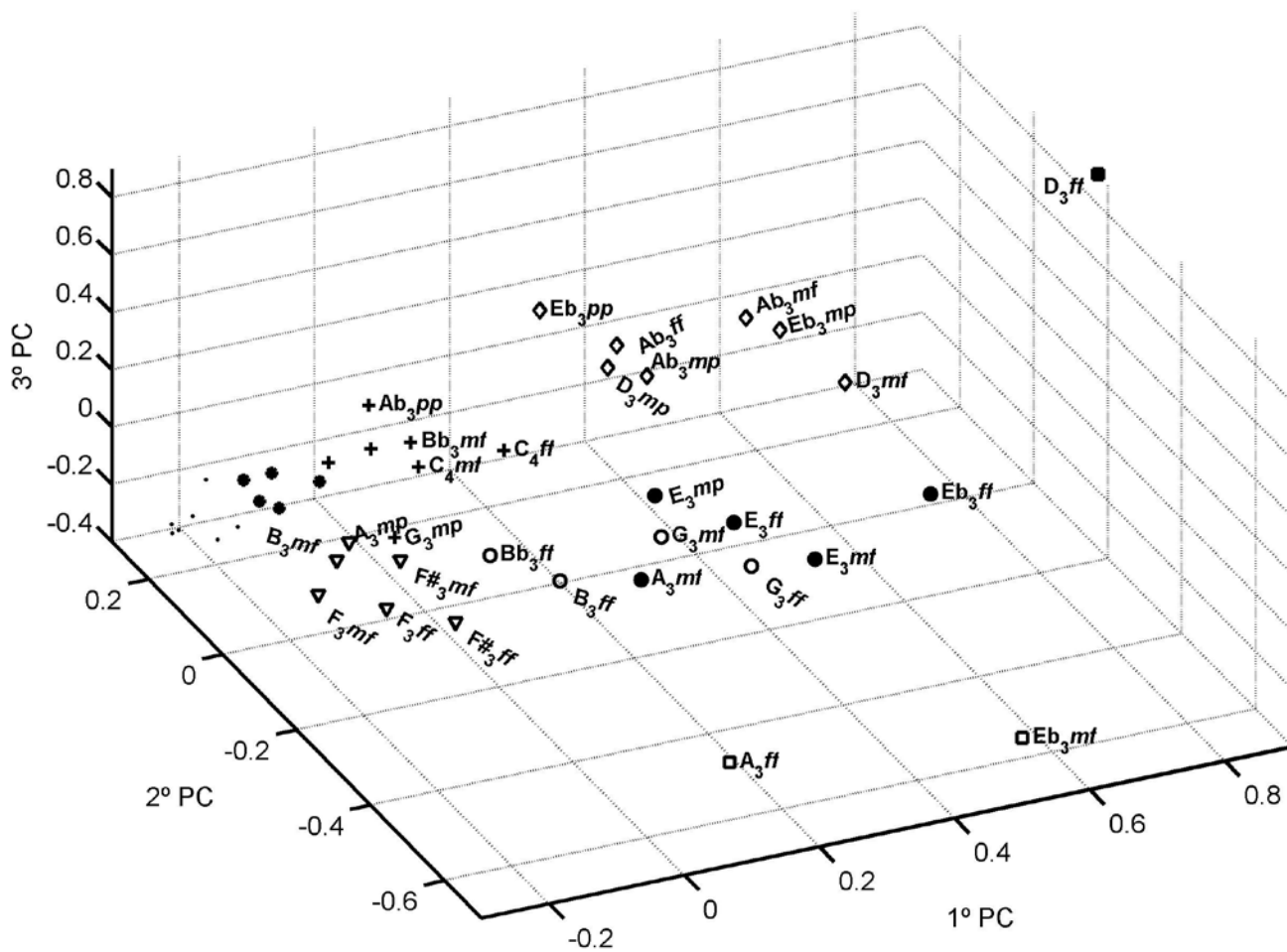


Fig.7: Trajetórias tridimensionais dos quatro sons de cada uma das 11 notas mais graves da clarineta, de Ré 3 a Dó 4, representadas no espaço espectral definido pelo registro grave inteiro (pontos não rotulados na parte esquerda da Figura correspondem às notas: Ré 3 *pp*, Mi 3 *pp*, Fá 3 *pp*, Fá 3 *mp*, Fá sustenido 3 *pp* e *mp*, Sol 3 *pp*, Lá bemol 3 *pp*, Lá 3 *pp*, Si bemol 3 *pp*, *mp* e *mf*, Si 3 *pp* e *mp*, Dó 4 *pp* e *mp*).

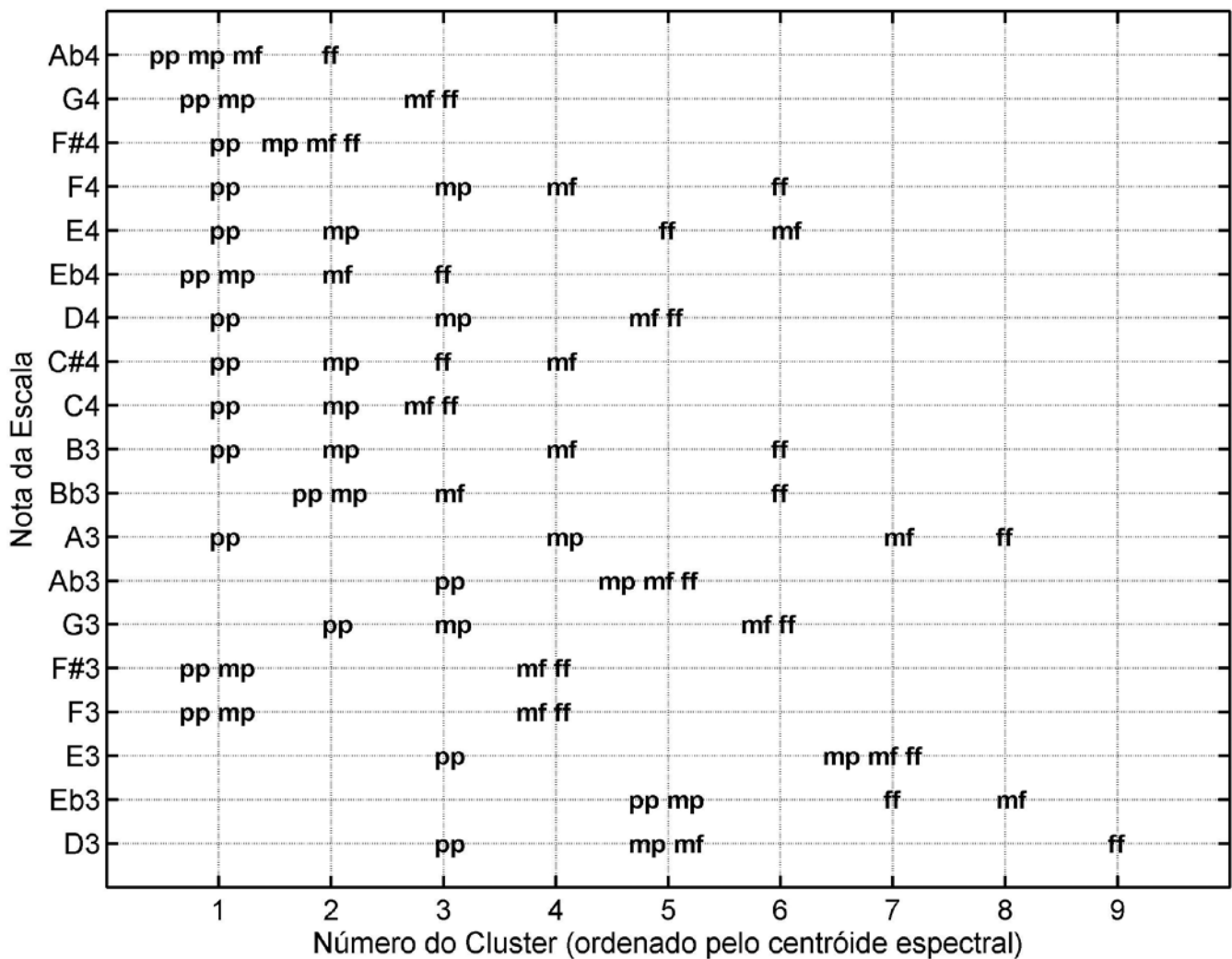


Fig.8: Cluster das 19 notas do registro mais grave da clarineta. Notas estão ordenadas por altura e cluster pela média dos centróides espectrais.

A Fig.9 mostra curvas de *clusters* dos quatro sons de cada uma das três notas mais graves e das três notas mais agudas do registro grave. As notas estão ordenadas por altura e os *clusters* representados pela coordenada do eixo vertical, numa representação bi-dimensional de linhas individuais da Fig.8, que realça a tendência das notas mais graves apresentarem maiores variações e as notas mais agudas apresentarem timbres mais estáveis e concentrados nos clusters de ordem inferior (valores inferiores de centróides espectrais).

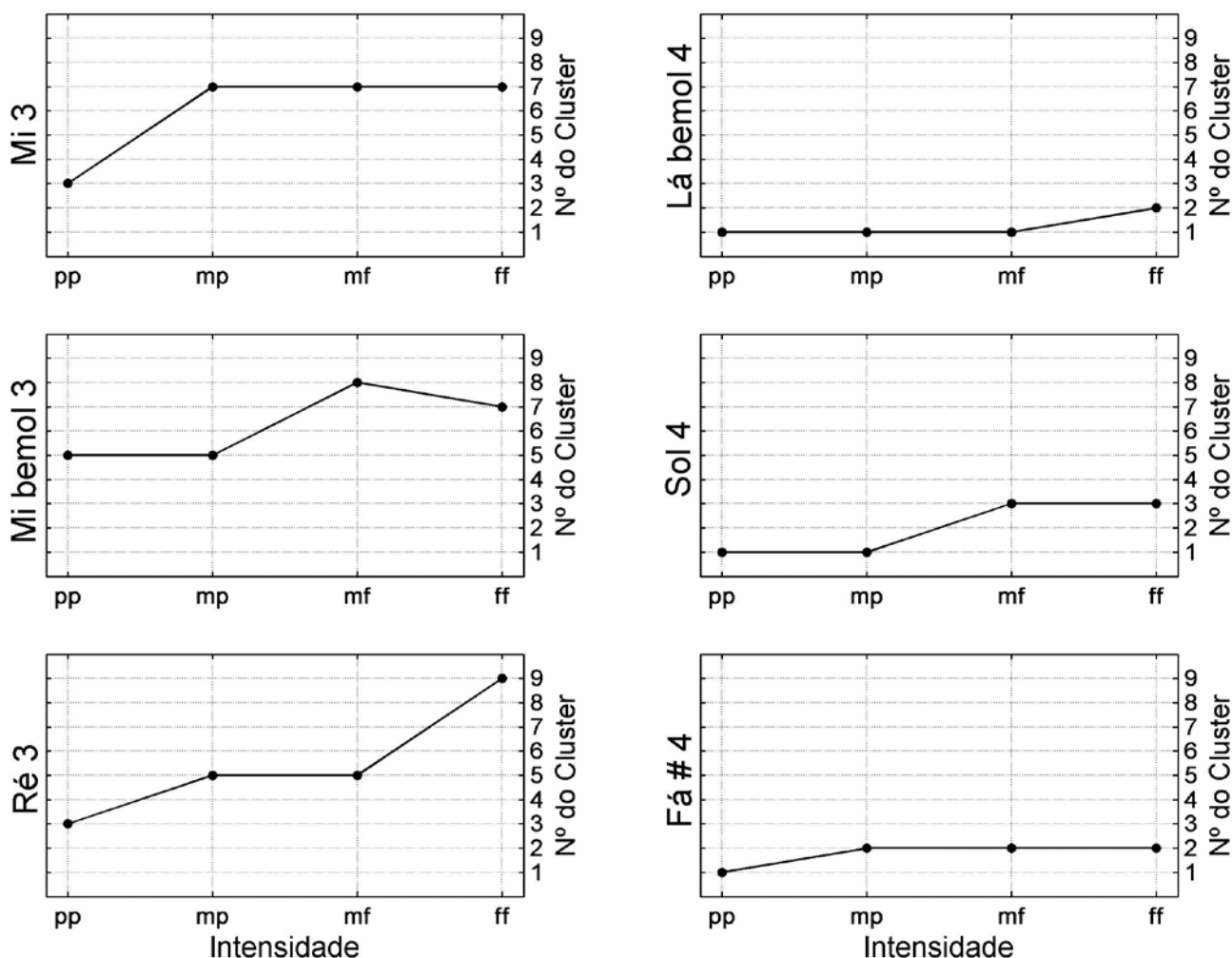


Fig.9: Curvas de *clusters* das três notas mais graves e das três notas mais agudas do registro grave da clarineta, Ré 3, Ré bemol 3, Mi 3, Fá sustenido 4, Sol 4 e Lá bemol 4. As notas estão ordenadas por altura e *clusters* pela média dos centróides espectrais.

Realizou-se, então, uma nova análise de agrupamento das 33 notas (132 sons) dos dois registros mais graves do instrumento, de Ré 3 até Si bemol 5, que resultou numa representação mais adequada através de doze clusters. A Fig.10 mostra um gráfico semelhante ao da Fig.8, incluindo agora as 14 notas do segundo registro, Lá 4 a Si bemol 5. Com a exceção de uma *mf* e quatro *ff* (Si bemol 4 *mf* e *ff*, Dó sustenido 4 *ff*, Si 4 *ff* e Si bemol 5 *ff*), todos os sons do segundo registro estão contidos nos cinco primeiros *clusters*, juntamente as notas *pp* e *mp* do registro mais grave, confirmando a tendência já observada de notas mais agudas se agruparem mais compactamente e reforçando a correlação da classificação com a variação do nível de intensidade, assim como a diminuição da variação de timbre com o aumento da altura.

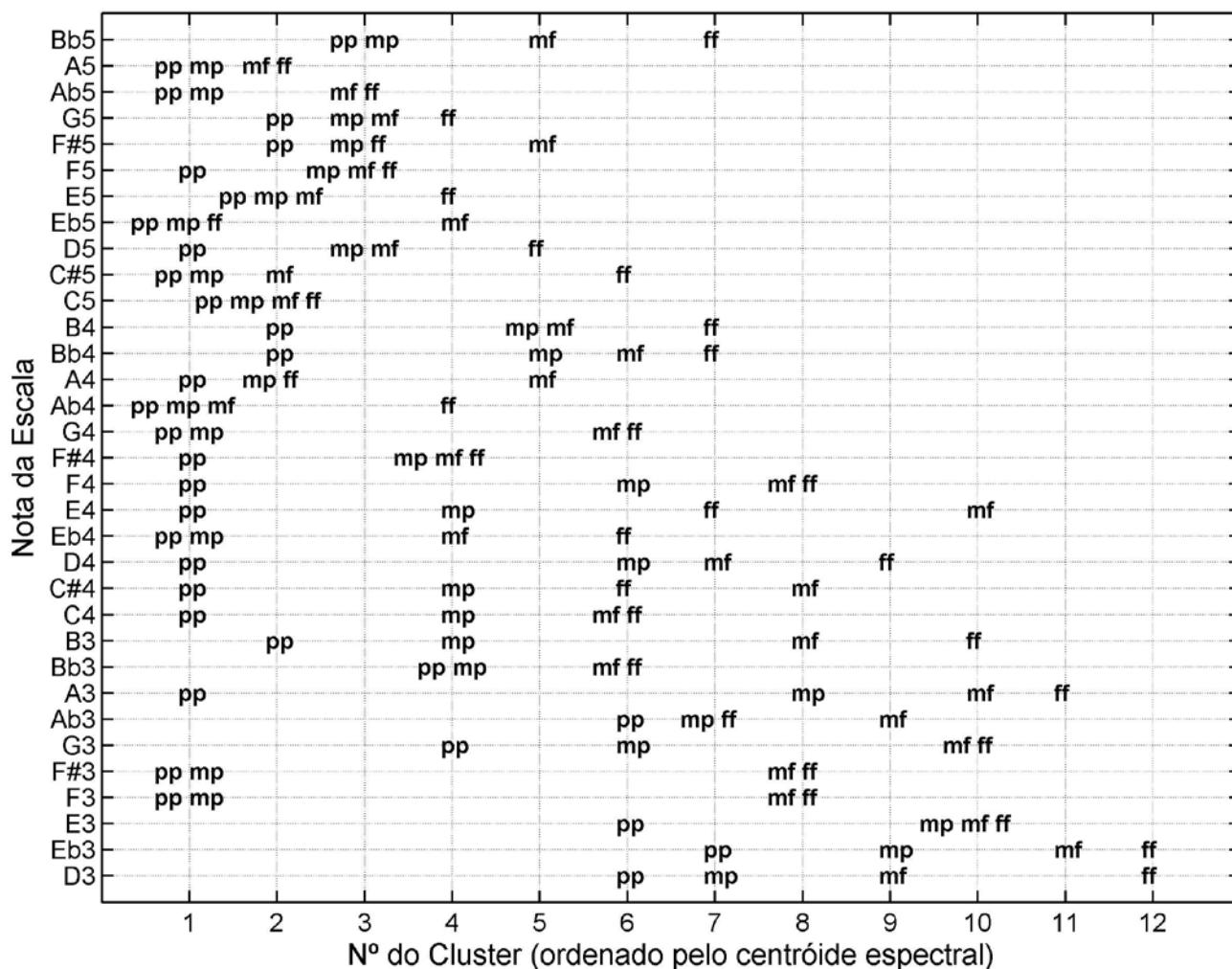


Fig.10: Clusters dos quatro sons de cada uma das 33 notas dos primeiros dois registros da clarineta. As notas estão ordenadas por altura e clusters pela média dos centróides espectrais.

A utilização de técnicas de agrupamento em *clusters* aplicadas ao espaço timbrístico de baixa dimensão produzido pela análise por componentes principais nos permite visualizar semelhanças entre timbres de notas, mas não permite uma boa análise da dinâmica de variação do timbre destas notas. Isto ocorre porque a variação do timbre não é linear como a representação em *clusters* nos leva a acreditar e porque o número de *clusters* é, em geral, reduzido, o que não permite uma representação adequada da topologia da variação do som.

4.2 - Mapas Auto-Organizativos de Kohonen

Mapas Auto-Organizativos de Kohonen, conhecidos como SOM (*Self-Organizing Maps*), são algoritmos de redes neurais não supervisionados, capazes de mapear dados de entrada de grandes dimensões em espaços de baixa dimensão, preservando as relações topológicas essenciais dos dados originais, promovendo assim uma projeção não-linear destes dados. Aplicações de SOM incluem uma larga gama de problemas de

engenharia (robótica, análise de imagem, som e processos), física, lingüística, medicina e psicologia. Por não se fundamentar em suposições *a priori* sobre as características dos dados analisados, SOM se mostra como uma ferramenta poderosa para experimentos que envolvem análise de dados de grande complexidade e com altos índices de não linearidade, tais como sons musicais.

Leman propôs uma comparação entre os resultados obtidos por mapeamentos timbrísticos resultantes de redes SOM com aqueles espaços construídos por MDS a partir de medições psicológicas, como referência básica para a pesquisa cognitiva em música em geral (LEMAN, 1994). Toiviainen comparou a eficiência das representações do timbre musical em espaços construídos por distâncias topológicas computacionalmente calculadas por SOM e por medições de estimativa subjetiva de similaridade (TOIVIAINEN, KAIPAINEN *et al.*, 1995). Os resultados deste trabalho comprovaram um alto grau de correlação entre os dois domínios, sugerindo uma adequação do modelo de Kohonen para projetar complexos perceptivos multidimensionais neste tipo de representação. A equipe de pesquisa do CSC da Universidade de Pádua desenvolveu estudos de classificação de timbres musicais utilizando Mapas Auto-Organizativos de Kohonen (DE POLI, PRANDONI *et al.*, 1993; DE POLI e TONELLA, 1993). Faiten pré-processou parâmetros espectrais por modelos auditivos simplificados e obteve espacializações timbrísticas utilizando Mapas Auto-Organizativos de Kohonen (FAITEN e GUNZEL, 1994). Leman utilizou Modelos Auditivos e Mapas de Kohonen com o objetivo de estabelecer uma plataforma que incorporasse fenômenos auditivos e neuronais para a pesquisa musicológica (LEMAN, 1991, 1992).

Kohonen formalizou um algoritmo para redes neurais auto-organizativas a partir de um processo numérico simples, que resulta numa projeção modificada da estrutura interna de uma rede neural em um mapa de dimensão menor, capaz de isolar as características mais relevantes do sinal de entrada, fornecendo um mapeamento daquilo que contem sua maior variância numérica (KOHONEN, 1995; BRAGA, LUDERMIR *et al.*, 2000). Os arranjos destes mapas neurais artificiais podem tomar várias formas, o que vai definir as relações de vizinhança entre os neurônios. Um mapa SOM hexagonal de tamanho 16x9 foi usado para mapear os 76 sons (19 notas) do registro mais grave da clarineta. A Fig.11 mostra a relação entre este mapeamento e os níveis de intensidade. Como na classificação feita com *k-means*, notas *ff* e *mf* tenderam a se agrupar. Além disso, notas *pp* e *mp* também foram agrupadas mais compactamente que notas *mf* e *ff*, como já observado no espaço timbrístico da Fig.3. Isto pode ser verificado pela distribuição de distância métrica do mapa, mostrada no gráfico do lado direito da Fig.11, no qual distâncias entre hexágonos representam distâncias entre células do mapa.

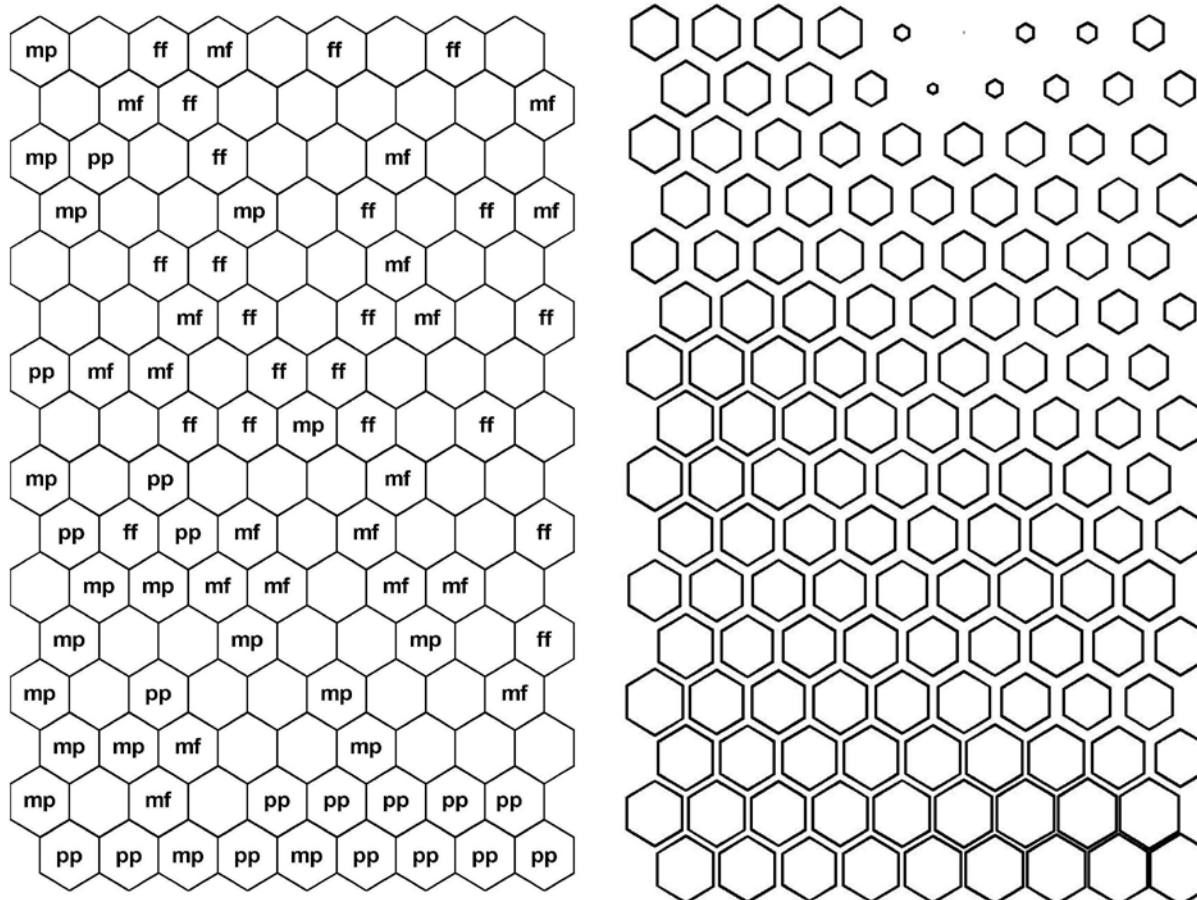


Fig.11: Mapeamento SOM dos níveis de intensidade (esquerda) dos 76 sons (19 notas) do registro grave da clarineta e a distribuição da distância entre as células do mapa (direito) - tamanho das células correspondem a uma maior proximidade entre elas.

Embora mapas SOM exibidos nas figuras sejam uma projeção dos dados em duas dimensões, foram identificadas características consistentes entre os mapeamentos obtidos com o SOM e o algoritmo *K-means*. Como no *K-means*, o SOM foi capaz de mapear todos os pontos de uma única amostra de som em uma ou no máximo duas células. A Fig.12 mostra o mapeamento dos sons de duas notas contíguas da oitava mais grave da clarineta, *Fá 3* e *Fá susenido 3*. A figura indica em cada célula a porcentagem da extensão total de cada som nela mapeada. Observa-se que todos os sons destas notas se agruparam bem juntos no canto inferior do mapa, similarmente ao agrupamento observado na representação do espaço espectral PCA. Este fato sugere que os sons destas notas não variam consideravelmente com nível de intensidade. A Fig.12(a) mostra o mapeamento de duas outras notas contíguas, *Lá bemol 3* e *Lá 3*. Apesar de serem de alturas consecutivas, elas ficaram amplamente espalhadas no mapa, um fato que também pode ser confirmado pelas trajetórias destas mesmas no espaço espectral representado na Fig.3. Comparando as Fig.6 e Fig.12(a), é possível notar a alta correlação entre a análise *K-means* e o mapa de Kohonen. Em ambas as classificações, o *Lá bemol 3* foi mapeado em *clusters* de ordem superior (células do topo do mapa) e seus sons ocuparam uma mesma região do mapa, enquanto que o *Lá 3* foi espalhado em um maior número de *clusters* e células. Essa correlação é

esperada, uma vez que a fonte de dados para ambos os algoritmos de classificação foi as trajetórias da PCA, que acumulavam a maior variância do sinal nas primeiras componentes. Ou seja, a topologia representada pelo SOM foi derivado de uma fonte de dados com energia concentrada em poucas dimensões.

A Fig.12(c) mostra as trajetórias de seis notas da oitava mais grave do instrumento, que inclui aquelas notas mostradas nas Fig.12(a) e 12(b): Ré 3, Mi bemol 3, Fá 3, Fá sustenido 3 e Lá bemol 3. Apesar de serem notas de alturas consecutivas, elas foram mapeadas em dois grupos distintos em lados opostos do mapa: Ré 3, Mi bemol 3 e Lá bemol 3 no canto superior esquerdo e Fá 3, Fá sustenido 3 e Lá 3 no canto inferior direito. Comparando as Fig.4 e Fig.12(c), podemos observar que as notas destes dois grupos foram também espalhadas em lados opostos do espaço espectral da Fig.3 e classificadas pela análise *K-means* como pertencentes ao mesmo *cluster*. Além disso, em ambas as classificações os sons do Lá bemol 3 foram mapeados compactamente, enquanto que o Lá 3 foi amplamente espalhado ao longo de clusters e células, confirmando a alta correlação entre a análise *K-means* e o mapa de Kohonen.

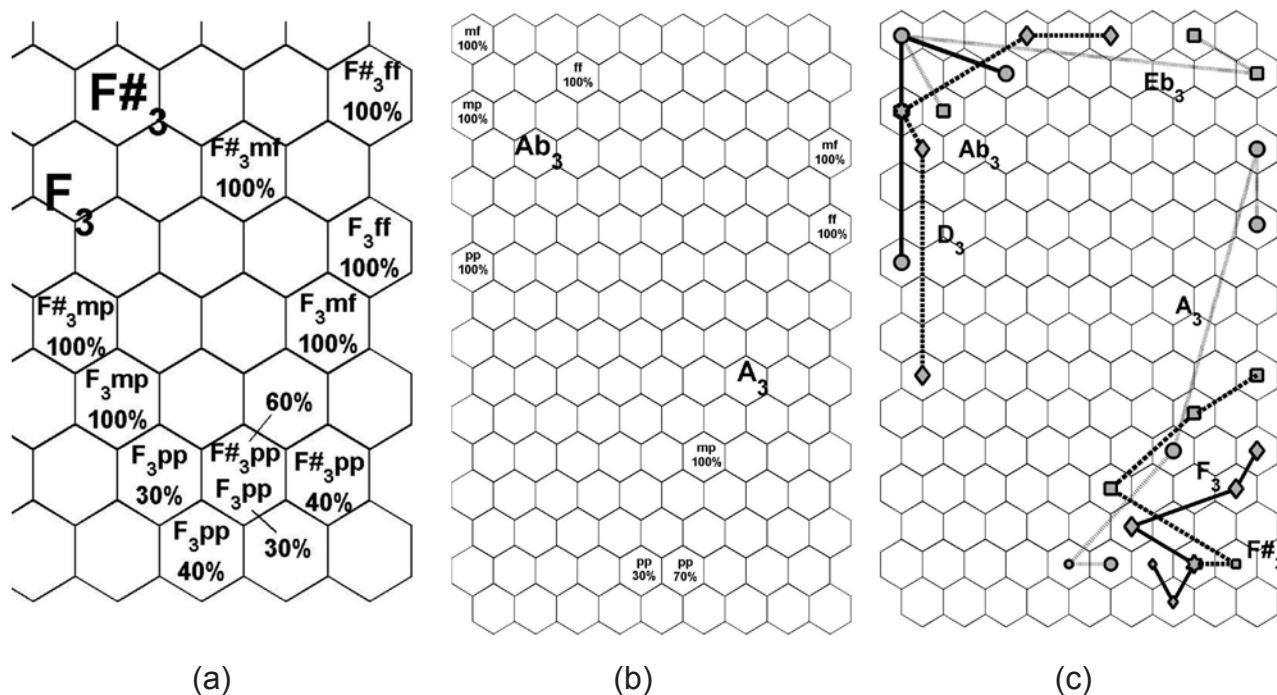


Fig.12: Mapeamento SOM: (a) notas Fá 3 e Fá sustenido 3; (b) notas Lá 3 e Lá bemol 3; (c) trajetórias das notas Ré 3, Mi bemol 3, Fá 3, Fá sustenido 3 e Lá bemol 3.

5 - Conclusão

O timbre é um atributo complexo e de difícil representação, especialmente quando se deseja representar sua dinâmica em um único instrumento musical. Este trabalho mostrou uma metodologia para o estudo do timbre em um instrumento musical. Os resultados obtidos até o momento evidenciaram o potencial dos métodos utilizados para o estudo das características estáticas e dinâmicas do timbre de instrumentos musicais. Inicialmente, o timbre foi caracterizado num nível físico pelo seu conteúdo espectral. Em seguida, a análise por componentes principais permitiu uma representação em um nível mais elevado, a partir

das “*tendências de variação*” de seu espectro. Finalmente, técnicas de classificação como o *K-means* e o SOM permitiram uma descrição semântica de grupos timbrísticos e ratificou a possibilidade de se utilizar a PCA como uma forma de criação de espaços timbrísticos de conjuntos notas e do instrumento como um todo. Com vistas a uma classificação das características estáticas e dinâmicas ao longo de toda a extensão do instrumento, notas pertencentes a regiões que apresentam características semelhantes de timbre foram agrupadas para gerar espaços timbrísticos específicos destas regiões. O conjunto destes espaços constituiu o que podemos chamar de *Espaço Timbrístico da Clarineta*, onde todas as notas do instrumento em suas configurações mais variadas puderam ser representadas de maneira a situar timbres semelhantes em regiões próximas deste espaço.

Foi observado em todos os resultados apresentados que: (1) classes de timbre tenderam a se dividir como uma função do brilho espectral que é conhecidamente correlacionado com o nível de intensidade em instrumentos de sopro; (2) a oitava mais baixa da clarineta exibiu, em geral, maior riqueza de diferenciação de timbre que notas mais altas; (3) a oitava mais alta do registro médio (de *Dó 5* para cima), exibiu menor brilho espectral e menor variação de timbre.

Esta representação gráfica do timbre e de sua evolução temporal oferece caminhos para a compreensão do controle dinâmico e instantâneo que o instrumentista detém sobre o som dos instrumentos acústicos e de como este controle é percebido pelo ouvinte, o que pode nos conduzir a formulações de plataformas que forneçam outros tipos de controle de outros tipos de estruturas musicais.

6 - Agradecimentos

Este trabalho foi apoiado em parte pela CAPES (Ensino superior brasileiro que funda agência), e CNPq (Conselho Nacional para Desenvolvimento Científico e Tecnológico), Brasil.

7 - Referências Bibliográficas

- ARCOS, J. L., DE MANTARAS, R. L. e SERRA, Xavier. SaxEx: A Case-Based Reasoning System for generating Expressive Musical Performances. *Journal of New Music Research*, v.27, n.3, 1998.
- BALIELLO, S., DE POLI, Giovanni e NOBILI, R. The Color of Music: Spectral Characterization of Musical sounds Filtered by a Cochlear Model. *Journal of New Music Research*, v.27, n.4, 1998.
- BRAGA, A. P., LUDERMIR, T. B. e CARVALHO, A. C. P. L. F. *Redes Neurais Artificiais: Teoria e Aplicações*. Rio de Janeiro, RJ: Editora LTC, 2000.
- CHARBONNEAU, G., HOURDIN, C. e MOUSSA, T. A Multidimensional Scaling Analysis of Musical Instrument's Time-Varying Spectra. *Computer Music Journal*, v.21, n.2, 1997a. p.40-55.
- _____. A Sound Synthesis Technique Based on Multidimensional Scaling of Spectra. *Computer Music Journal*, v.21, n.2, 1997b. p.56-68.
- COSI, Piero, DE POLI, Giovanni e LAUZZANA, Gianpaolo. Auditory Modeling and Self-Organizing Neural works for Timbre Classification. *Journal of New Music Research*, v.23, 1994. p.71-98.
- DANNENBERG, Roger B. e DERENYI, I. Combining Instrument and Performance Models for High-Quality Music Synthesis. *Journal of New Music Research*, v.27, n.3, 1998.
- DE PAULA, Hugo Bastos. *Análise e Re-síntese de Som Natural de Clarineta Utilizando Análise por componentes Principais*. Belo Horizonte: Departamento de Engenharia Elétrica, Universidade Federal de Minas Gerais, 2000 (Dissertação, Mestrado em Engenharia Elétrica).
- DE POLI, Giovanni e PRANDONI, P. Sonological Models for Timbre Characterization. *Journal of New Music Research*, v.26, 1997. p.170-197.

- DE POLI, Giovanni, PRANDONI, P. e TONELLA, P. Timbre Clustering by Self-Organizing Neural Networks. *Anais do X Colloquium on Musical Informatics*. Milan: University of Milan, 1993.
- DE POLI, Giovanni, RODÀ, Antonio e VIDOLIN, Alvise. Note-by-Note Analysis of the Influence of Expressive Intentions and Musical Structure in Violin Performance. *Journal of New Music Research*, v.27, n.3, 1998. p.293-321.
- DE POLI, Giovanni e TONELLA, P. Self-Organizing Neural Networks and Grey's Timbre Space. *Anais do International Computer Music Conference*. Tokyo: International Computer Music Association, 1993. p. 441-444.
- DUBNOV, Shlomo. Emotion - Is it measurable? *Anais do KANSEI - The Technology of Emotion, AIMI International Workshop*. Genova, 1997.
- DUBNOV, Shlomo e RODET, Xavier. Statistical Modeling of Sound Aperiodicities. *Anais do International Computer Music Conference*. Tesseloniki, Grece: International Computer Music Association, 1997.
- DUBNOV, Shlomo, TISHBY, N. e COHEN, D. Hearing Beyond The Spectrum. *Journal of New Music Research*, v.25, 1996.
- EGOZY, E. B. *Deriving Musical Control Features from a Real-Time Analysis of the Clarinet*. Cambridge, Massachusetts, Massachusetts Institute Technology, 1995 (Dissertação, Mestrado em Engenharia Elétrica).
- FAITEN, B. e GUNZEL, S. Automatic Indexing of a Sound Database Using Self-Organizing Neural Nets. *Computer Music Journal*, v.18, n.3, 1994. p.53-65.
- FREEDMAN, M.D. *A technique for Analysis of Musical Instrument Tone*. Urbana, Illinois, University of Illinois, 1965 (Tese, Doutorado em Engenharia Elétrica).
- _____. A Method for Analysing Musical Tones. *Journal of The Audio Engineering Society*, v.16, 1968. p.419-425.
- FRIBERG, Anders, BRESIN, Roberto, FRYDÉN, Lars e SUNDBERG, Johan. Musical Punctuation on the Micro-level: Automatic Identification and Performance of Small Melodic Units. *Journal of New Music Research*, v.27, n.3, 1998.
- GORDON, J. W. e GREY, J. M. Perceptual Effects of Spectral Modifications on Music Timbres. *Journal of the Acoustics Society of America*, v.63, 1978. p.1493-1500.
- GREY, John M. *An exploration of musical timbre*. CCRMA, Dept. of Music Stanford University. Stanford, Calif., (STAN-M-2), 1975.
- _____. Multidimensional Perceptual Scaling of Musical Timbres. *Journal of the Acoustics Society of America*, v.5, 1977. p.1270-1277.
- _____. Timbre Discrimination in Musical Patterns. *Journal of the Acoustics Society of America*, v.64, 1978. p.467-472.
- HAJDA, J. M., KENDALL, Roger A., CARTERETTE, Edward C. e HARSHBERGER, M. L. Methodological Issues in Timbre Research. In: I. Deliège e J. A. Sloboda (Ed.). *Perception and Cognition of Music*. Hove: Psychology Press, 1997, p.253-306.
- HORNER, A. Wavetable Interpolation Synthesis Based on Time-Variant Spectral Analysis of Musical Sounds. *Journal of The Audio Engineering Society*, 1995.
- HORNER, A., BEAUCHAMP, J. W. e HAKEN, L. Methods for Multiple Wavetable Synthesis of Musical Instrument Tones. *Journal of The Audio Engineering Society*, v.41, n.5, 1993. p.336-356.
- HORNER, A. e NGAI-MAN. Group Synthesis with Genetic Algorithms. *Journal of The Audio Engineering Society*, v.44, n.3, 1996. p.130-147.
- JOHNSON, R. e WICHERN, D. W. *Applied Multivariate Statistical Analysis*. Upper Sadlle, New Jersey, 1998.
- KAUFMAN, L. e ROUSSEUW, P. J. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons, 1989.
- KOHONEN, Teuvo. *Self-Organizing Maps - Springer Series in Information Sciences*. Berlim: Springer Verlag. v.30, 1995.
- LEMAN, Marc. Emergent Properties of Tonality Functions by Self-Organization. *Interface - Journal of New Music Research*, v.19, n.2-3, 1991. p.85-106.
- _____. Tone Context by Pattern Integration over Time. In: D. Baggi (Ed.). *Computer Generated Music*. Los Alamitos, California: IEEE Computer Science Press, 1992.

- _____. Schema-Based Tone Center Recognition of Musical Signals. *Journal of New Music Research*, v.23, n.2, 1994. p.169-204.
- LOUREIRO, Mauricio Alves, DE PAULA, Hugo Bastos e YEHIA, Hani Camille. Study of the Timbre Dynamics of an Expressive Performance Using Principal Component Analysis of Spectral Parameters. *Anais do VII Simpósio de Computação e Música*, Ed. C. V. D. L. Palombini. Curitiba, PR: PUC - PR, 2000.
- _____. Sonological Representation of a Musical Instrument by Sub-spaces of Spectral Component. *Mikropolyphonie - The Online Contemporary Music Journal*, v.7, 2001.
- _____. Representing Timbre Dynamics of a Musical Instrument: Comparison Between GA and PCA. *Anais do IEEE International Workshop on Machine Learning for Signal Processing*. São Luiz, Maranhão: Universidade Federal do Maranhão, 2004a. p. 381-390.
- _____. Timbre Classification of a Single Music Instrument. *Anais do 5th International Conference on Music Information Retrieval - ISMIR*. Barcelona: University Pompeu Fabra, 2004b. p. 546-549.
- _____. Timbre Representation of a Single Music Instrument. *Anais do International Computer Music Conference*. University of Miami, USA: International Computer Music Association, 2004c. p. 394-397.
- LUCE, D. A. *Physical Correlates of Nonpercussive Musical Instrument Tones*. Cambridge, Massachusetts: Department of Physics, MIT, 1963 (Tese, Doutorado em Física).
- LUCE, D. A. e CLARK, M. Duration of Attack Transients of Nonpercussive Orchestral Instruments. *Journal of The Audio Engineering Society*, v.13, 1965. p.194-199.
- _____. Physical Correlates of Brass-Instrument Tones. *Journal of the Acoustics Society of America*, v.42, 1967. p.1232-1243.
- MASRI, Paul e BATEMAN, A. Improved Modelling of Attack Transients in Music Analysis-Resynthesis. *Anais do International Computer Music Conference*. Hong-Kong: The International Computer Music Association (ICMA), 1996. p. 100-103.
- MCADAMS, Stephen e BREGMAN, Albert S. Hearing Musical Streams. In: C. Roads e J. Strawn (Ed.). *Foundations of Computer Music*. Cambridge, Massachusetts: MIT Press, 1979, p.658-698.
- MCADAMS, Stephen, WINSBERG, S., DONNADIEU, S., DE SOETE, G. e KRIMPHOFF, J. Perceptual Scaling of Synthesized Musical Timbres: Common Dimensions, Specificities and Latent Subject Classes. *Psychological Research*, v.58, 1995. p.177-192.
- MCAULAY, R. J. e QUATIERI, T. F. Speech Analysis/Synthesis Based on a Sinusoidal Representation. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, v.34, n.4, 1986. p.744-754.
- MILLER, J. R. e CARTERETTE, Edward C. Perceptual Spaces for Musical structures. *Journal of the Acoustics Society of America*, v.58, 1975. p.711-720.
- MISDARIIS, N. R., SMITH, B. K., PRESSNITZER, D., SUSINI, P. e MCADAMS, Stephen. Validation of a Multidimensional Distance Model for Perceptual Dissimilarities Among Musical Timbres. *Anais do 16th International Congress on Acoustics*. Woodbury, New York: ASA - The Acoustical Society of America, 1998.
- PISZCZALSKI, M. Spectral Surfaces from Performed Music, Part 1. *Computer Music Journal*, v.3, n.1, 1979a. p.18-24.
- _____. Spectral Surfaces from Performed Music, Part 2. *Computer Music Journal*, v.3, n.3, 1979b. p.25-27.
- PISZCZALSKI, M., GALLER, B., BOSSEMEYER, R., HATAMIAN, M. e LOOFT, F. Performed Music: Analysis Synthesis and Display by Computer. *Journal of The Audio Engineering Society*, v.29, 1981. p.38-46.
- PLOMP, Reinier. Timbre as a Multidimensional Attribute of Complex Tones. In: R. Plomp e G. F. Smoorenburg (Ed.). *Frequency Analysis and Periodicity Detection in Hearing*. Leiden: A. W. Sijthoff, 1970.
- PLOMP, Reinier e RASCH, Rudolf A. The Perception of Musical Tones. In: D. Deutsch (Ed.). *Psychology of Music*. San Diego, California: Academic Press, 1982, p.1-24.
- RENCHER, A. C. *Methods of Multivariate Analysis*. New York: John Wiley & Sons, 1995.
- RISSET, Jean-Claude. Computer Study of Trumpet Tones. *Journal of the Acoustical Society of America*, v.38, 1965. p.912.
- _____. Timbre Analysis by Synthesis: Representations, Imitations and Variants for Musical Composition. In: G. De Poli, A. Piccialli, et al (Ed.). *Representation of Musical Signals*. Cambridge, Massachusetts: MIT Press, 1991, p.7-43.

- RISSET, Jean-Claude e Mathews, Max V. Analysis of Musical Instrument Tones. *Physics Today*, v.22, n.2, 1969. p.23-40.
- RISSET, Jean-Claude e WESSEL, David L. Exploration of Timbre by Analysis and Synthesis. In: D. Deutsch (Ed.). *Psychology of Music*. San Diego, California: Academic Press, 1999, p.25-58.
- ROCHEBOIS, T. e CHARBONNEAU, G. Cross-Synthesis Using Interverted Principal Harmonic Sub-Spaces. In: M. Leman (Ed.). *Music, Gestalt and Computing: Studies in Cognitive and Systematic Musicology*. Berlin-Heidelberg: Springer Verlag, 1997, p.375-385.
- ROVAN, J. B., WANDERLEY, Marcelo, DUBNOV, Shlomo e DEPALLE, P. Instrumental Gestural Mapping Strategies as Expressivity Determinants in Computer Music Performance. *Anais do KANSEI - The Technology of Emotion, AIMI International Workshop*. Genova, 1997.
- SANDELL, G. J. e MARTENS, W. Perceptual Evaluation of Principal-Component-Based Synthesis of Musical Timbres. *Journal of The Audio Engineering Society*, v.43, n.12, 1995. p.1013-1028.
- SERRA, Xavier. Musical Sound Modeling with Sinusoids plus Noise. In: A. Piccialli, C. Roads, et al (Ed.). *Musical Signal Processing*: Swets & Zeitlinger Publishers, 1997.
- SHEPARD, R. N. *Psychological Representations of Speech Sounds in Human Communication*. New York: McGraw-Hill, 1972.
- STRAWN, John. *Modelling Musical Transitions*. Palo Alto, California: CCRMA, Department of Music, Stanford University, 1985 (Tese, Doutorado em Música).
- _____. Orchestral Instruments: Analysis of Performed Transitions. *Journal of The Audio Engineering Society*, v.34, 1986. p.867-880.
- _____. Analysis and Synthesis of Musical Transitions Using the Discrete Short-Time Fourier Transform. *Journal of The Audio Engineering Society*, v.35, 1987. p.3-14.
- STRONG, W. e CLARK, M. Perturbations of Synthetic Orchestral Wind-Instrument Tones. *Journal of the Acoustics Society of America*, v.41, n.2, 1967a. p.277-285.
- _____. Synthesis of Wind-Instrument Tones. *Journal of the Acoustics Society of America*, v.41, n.1, 1967b. p.39-52.
- TOIVIAINEN, Petri, KAIPAINEN, Mauri e LOUHIVUORI, Jukka. Musical Timbre: Similarity Ratings Correlate with Computational Feature Space Distances. *Journal of New Music Research*, v.24, 1995. p.282-298.
- WESSEL, David L. Timbre Space as a Musical Control Structure. In: C. Roads e J. Strawn (Ed.). *Foundations of Computer Music*. Cambridge, Massachusetts: MIT Press, 1979, p.640-657.

Maurício Alves Loureiro é Bacharel em Engenharia Aeronáutica pelo ITA (1976), Bacharel em Música (clarineta) pela Staatliche Hochschule für Musik Freiburg, Alemanha (1983), onde estudou com Dieter Klöcker e Doutor em Música pela University of Iowa, EUA (1991). Foi primeiro clarinetista da orquestra de Campinas e da OSESP. Atuou como solista frente a inúmeras orquestras sinfônicas e renomados conjuntos de câmara, participou como clarinetista e professor de inúmeros festivais de música no Brasil e exterior e tem desenvolvido intensa atividade como solista e camerista, especialmente dedicado à interpretação da música contemporânea e da música eletroacústica. Foi professor adjunto do Instituto de Artes da UNESP (1984-1992) e é atualmente professor titular da Escola de Música da UFMG, onde é coordenador da Pós-Graduação e pesquisador no âmbito dos Grupos de Pesquisa *Sonolgia* e *CEFALA - Centro de Estudos da Fala Acústica Linguagem e Música*.

Hugo Bastos de Paula é atualmente professor do Instituto de Informática da PUC Minas, membro da coordenação didática e coordenador de pesquisa do curso de Sistemas de Informação. É graduado em Ciência da Computação, com mestrado e Doutorado em Engenharia Elétrica pela UFMG. Estudou clarineta com Walter Alves de Souza, primeiro clarinetista da Orquestra Sinfônica de Minas Gerais. Teve seu primeiro contato com a área de Computação Musical em 1996, quando participou de um projeto voltado para a implementação de uma linguagem de composição no Csound e desde 1998 trabalha com modelagem do timbre de instrumentos musicais, assunto no qual focalizou sua dissertação de mestrado. É membro dos grupos de pesquisa *Sonolgia* e *CEFALA - Centro de Estudos da Fala Acústica Linguagem e Música*, na UFMG e do VIPLAB -Visual Information Processing Lab, na PUC Minas.