



## Identificação do intérprete a partir de informação de movimento corporal

MODALIDADE: COMUNICAÇÃO

*Davi Alves Mota*

*Escola de Música, Universidade Federal de Minas Gerais (UFMG) – daviavesmota@yahoo.com.br*

*Mauricio Alves Loureiro*

*Escola de Música, Universidade Federal de Minas Gerais (UFMG) – mauricio.alves.loureiro@gmail.com*

*Rafael Laboissière*

*Laboratoire de Psychologie et NeuroCognition - CNRS, Grenoble - rafael.laboissiere@upmf-grenoble.fr*

**Resumo:** Este estudo busca identificar padrões de informação de movimento corporal de músicos em performances de duos de clarineta em diferentes condições experimentais. Os resultados indicam que as regiões com valores mais altos de velocidade possuem informação suficiente para o reconhecimento dos músicos que realizaram as performances. Foi verificado também se houve tendência de modificação destas "assinaturas gestuais" a partir de diferentes situações de performance em conjunto, impostas por diferentes líderes.

**Palavras-chave:** Movimento corporal, Acoplamento expressivo, Performance em conjunto, Gesto expressivo, Gesto musical

### **Body Movement Information for Player Identification in Music Performance**

**Abstract:** This study seeks to identify information patterns of body movement of musicians in performances of clarinet duos in different experimental conditions. The results indicate that regions with higher velocity values have enough information for the recognition of the musicians who performed the excerpts. It was also verified whether there was a trend of change in these "gestural signatures" in different ensemble performance situations, imposed by different leaders.

**Keywords:** Body movement, Expressive coupling, Ensemble performance, Expressive gesture, Musical gesture

### **1. Introdução**

Diversos tipos de movimentos corporais não diretamente vinculados à produção sonora podem ser constantemente observados em qualquer execução instrumental. Gabrielsson (2003, p. 249) sugere que a movimentação realizada por um músico, além de comunicar informação relevante para a coordenação com outros, pode ainda assumir papéis diversos, como o de comunicar intenções expressivas, apresentar informações sobre a personalidade do artista ou simplesmente entreter a audiência. Nos últimos anos, grandes esforços foram dedicados ao estudo desses movimentos e, apesar de não haver consenso sobre sua origem ou função, sua existência é inegável (WANDERLEY; DEPALLE, 2001).

Vários estudos têm procurado quantificar gestos corporais envolvidos na performance musical e identificar como esta informação visual se relaciona com a música tocada e como é percebida por aquele que ouve (e vê) a música. Wanderley e colegas propuseram uma diferenciação entre gestos diretamente relacionados com a

produção do som (*instrumental gestures*), e gestos não diretamente envolvidos com a produção do som (*ancillary gestures*), sugerindo que estes últimos, para os quais propomos a designação *gestos acompanhantes*, podem estar mais relacionados com a intenção expressiva do intérprete (ROVAN, WANDERLEY *et al.*, 1997; WANDERLEY, 1999; CADOZ e WANDERLEY, 2000; WANDERLEY e VINES, 2006; JENSENIUS, WANDERLEY *et al.*, 2010). Tendo este tipo de movimento corporal como foco, este estudo buscou verificar se o movimento corporal do músico contém informação relacionada com sua intenção interpretativa em uma performance. Para testar esta hipótese, buscamos evidenciar: (1) se dados de movimento corporal capturados em performances musicais contêm informação suficiente para identificar seu intérprete a partir de recorrência de padrões; (2) se "assinaturas gestuais" de diferentes músicos poderiam ser influenciadas por diferentes situações interpretativas, como por exemplo em conjuntos instrumentais, onde é necessário seguir a concepção musical de um líder.

## 2. Materiais e Métodos

A escolha de um instrumento de sopro foi motivada pelo fato de que movimentos corporais diretamente relacionados com a produção sonora, são limitados a movimentos de pequena amplitude, tais como dos dedos, mandíbula e lábios, o que facilita e induz o músico a se engajar em uma ampla variedade de *gestos acompanhantes* de maior amplitude, seja de membros, cabeça ou do próprio instrumento. Além disso, melhores perspectivas para aquisição de dados, devido a possibilidades de contato com um grande número de clarinetistas profissionais, foram decisivas na escolha do instrumento. Seis clarinetistas profissionais foram recrutados em Belo Horizonte, Minas Gerais, todos possuindo boa experiência anterior em prática de orquestra, e todos se conhecendo.

Foi utilizado no experimento um pequeno excerto da "*Dança do Camponês e o Urso*" do balé *Petrushka* de Igor Stravinsky, extraído do *Tableau IV*, N° 100 (três primeiros compassos), no qual primeira e segunda clarinetas tocam solo em uníssono (*solí a 2*). O excerto apresenta uma situação real de performance, que exige a sincronização de cada nota e permite tomadas isoladas com igualdade de condições.

As gravações foram realizadas em duas sessões, separadas por intervalos de dois e três dias. Na primeira sessão, os músicos foram instruídos a tocar 4 vezes o excerto como primeira clarineta. Foi solicitado ao músico que escolhesse, dentre as

quatro gravações, a que ele considerava a melhor. Esta gravação serviu mais tarde para a segunda sessão, na qual os músicos foram instruídos a tocar como segunda clarineta, acompanhando todas as gravações escolhidas da primeira sessão, incluindo as executadas por eles mesmos. A única instrução dada foi para acompanhar a primeira clarineta da melhor forma possível. Depois de ouvir a execução da primeira clarineta, cada clarinetista tocou quatro vezes o excerto, enquanto ouviam cada execução da primeira sessão através de um fone de ouvido em sua orelha direita, apresentadas de maneira aleatória. Três batidas de metrônomo foram incluídas no início de cada gravação a ser acompanhada, cujo andamento foi estimado a partir da duração do primeiro pulso completo, medido entre os *onsets* das notas 5 e 9, Fá# e Dó (Figura 1).



The image shows a musical score for Clarinet in Bb. The first staff is marked 'Sostenuto' and 'ff' and contains measures 1 through 14. The second staff contains measures 15 through 41. The music features complex rhythmic patterns with many beamed notes and dynamic markings.

Figura 1: Excerto da Dança do Camponês e o Urso do balé *Petrushka* de Igor Stravinsky, extraído do Quatrième tableau, No 100

Um *Optotrak Certus* (fabricado pela Norte Digital Inc.) foi usado para rastrear a posição tridimensional da clarineta e da cabeça dos instrumentistas. Três marcadores foram fixados na campana da clarineta para definir um corpo rígido (representado por suas posições espacial e angular), relacionado a ela e outros três igualmente para a cabeça do clarinetista, tal como utilizado por Wanderley et al. (2005), Keller, Knoblich e Repp (2007) e Keller e Appel (2010). Estudos anteriores demonstraram que o movimento da campana concentra grande parte do movimento global realizado durante a performance neste instrumento e por isso optamos por restringir a ela esta análise (WANDERLEY, 2002; PALMER et al., 2009).

O módulo da velocidade resultante de cada posição tridimensional, estimado pela distância euclidiana entre as posições do centroide de duas amostras subsequentes, foi utilizado para parametrizar o movimento da campana da clarineta, o qual denominamos *perfil de velocidade*:

$$v_i = f_a \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2}$$

onde  $f_a$  é a taxa de amostragem,  $x$ ,  $y$  e  $z$  as coordenadas espaciais e  $i$  o número da amostra.

Visando não descartar informação que possa ser relevante, tais como gestos de preparação e finalização, consideramos o movimento a ser analisado iniciando a um pulso métrico antes do início da primeira nota e finalizando a um pulso métrico após o final da última nota. Foi utilizado um filtro Butterworth passa-baixa, de ordem 3 e frequência de corte de 5Hz, com a finalidade de descartar movimentos de baixa amplitude, tais como aqueles causados pelo impacto dos dedos no instrumento, ou ajustes na embocadura.

A diferença entre massas corporais, alturas e idades dos participantes, parecem ter influenciado na amplitude e velocidade dos movimentos, resultando em maior variabilidade de valores absolutos da velocidade em comparações entre sujeitos. Com a finalidade de minimizar a variabilidade relacionada a características corporais, e otimizar a detecção das “assinaturas gestuais” por detalhes temporais que emergem de intenções interpretativas dos músicos, a amplitude do perfil de velocidade de cada execução foi normalizada pelo seu valor quadrático médio, definido como:

$$V_Q = \sqrt{\frac{1}{N} \sum_{i=1}^N v_i^2}$$

onde  $v_i$  é a amplitude de cada uma das  $N$  amostras.

Devido à variabilidade dos andamentos de cada execução, as curvas de velocidade foram ajustadas para um mesmo tamanho, utilizando a técnica de *time warping*, como sugerido por Wanderley (2002) e Wanderley et al. (2005), resultando em representações de mesmo andamento médio para todas as execuções. A média de todas as gravações dos valores dos *onsets* de cada nota foi utilizada para modelar o *time warping*, visando também minimizar os desalinhamentos das curvas de velocidade com a estrutura musical.

### 3. Análise dos dados

Uma Análise Discriminante Linear (*Linear Discriminant Analysis - LDA*) (FISHER, 1938) foi utilizada para identificar tanto o clarinetista executante de cada performance quanto o clarinetista que liderava cada performance. A técnica propõe a transformação linear de observações multivariadas  $\mathbf{x}$  em observações univariadas  $\mathbf{y}$ ,

pertencentes a uma população  $\pi_i$ ,  $i = 1, 2, 3, \dots, k$ , de tal forma que sejam o mais separadas possível dos valores de  $\mathbf{y}$  pertencentes à população  $\pi_j$ , para  $j \neq i$ . A LDA encontra, portanto, a combinação linear  $\mathbf{y}$  de  $\mathbf{x}$ , de média  $\mu$ , expressa pela matriz  $\mathbf{U}$ , de dimensão  $k-1$ , que proporciona a maior separação entre os agrupamentos sugeridos (FISHER, 1938; DUDA & HART, 1996):

$$\mathbf{y} = \mathbf{U}(\mathbf{x} - \mu)$$

Para a identificação dos clarinetistas, dois agrupamentos dos dados foram propostos, cada um dividido em 6 classes, correspondentes ao número de clarinetistas. Um agrupamento, denominado *executante*, compreendeu todas as 24 execuções realizadas por cada músico, correspondente às 4 gravações acompanhando cada um dos outros 5 músicos, e a si próprio. Um segundo agrupamento, denominado *líder*, compreendeu todas as 24 execuções lideradas por cada músico como primeiro clarinetista acompanhado pelos outros 5 músicos, inclusive por si mesmo. O objetivo do agrupamento *líder*, foi verificar evidências de que características dos gestos do primeiro clarinetista possam ser transmitidas ao clarinetista que o acompanha.

A técnica LDA é aplicada primeiramente para treinar um classificador com as observações dos diferentes agrupamentos, para em seguida classificar novas observações (JOHNSON; WICHERN, 1998, p. 683). Considerando a dimensão do espaço dos dados igual ao número de amostras (cerca de 1500), uma redução da dimensionalidade, foi obtida por Análise por Componentes Principais (PCA), técnica que estima combinações lineares de menor dimensão das variáveis, que explica ao máximo da variância (JOHNSON; WICHERN, 1998, p. 458). Visando determinar o número ideal de componentes principais para a discriminação, a PCA foi processada iterativamente com a LDA, estimando-se o número de erros de classificação para cada iteração, tal como aplicado em sistemas de reconhecimento de face por Marcialis e Roli (2002) e Lu, Wang e Jain (2003). O uso combinado das técnicas PCA e LDA permite portanto reduzir o número de dimensões utilizadas na análise e ao mesmo tempo revelar as variáveis mais relevantes para a discriminação dos grupos.

A LDA foi realizada utilizando o pacote MASS (VENABLES; RIPLEY, 2002) do ambiente de computação estatística R (R Development Core Team, 2008). Uma representação gráfica do procedimento de análise é mostrada na Figura 2.

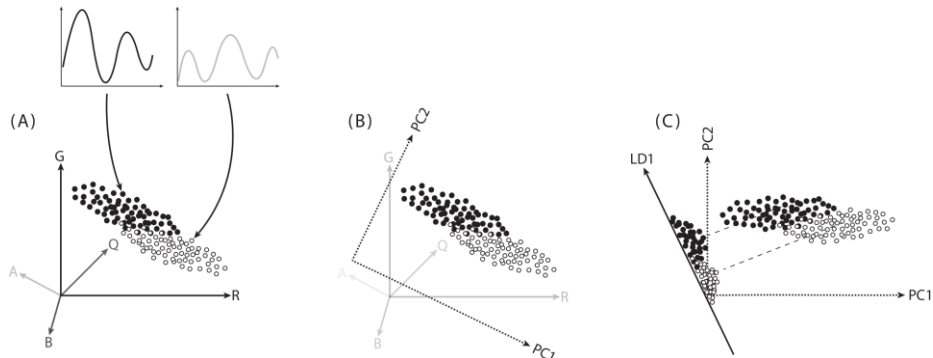


Figura 2: Uso combinado de PCA e LDA: (A) cada ponto amostrado nos *perfis de velocidade* é tratado como uma coordenada espacial; (B) a PCA projeta os dados nas direções que melhor explicam a variância; (C) a LDA projeta os dados nas direções que maximizam a separação entre as classes.

#### 4. Resultados

A Figura 8 mostra o número de erros de classificação e a variância acumulada em função do número de PCs utilizados. Para o agrupamento *executante*, apenas 10 PCs foram necessários para uma separação completa das classes (gráfico superior), enquanto que o agrupamento *líder* exigiu 86 PCs (gráfico superior).

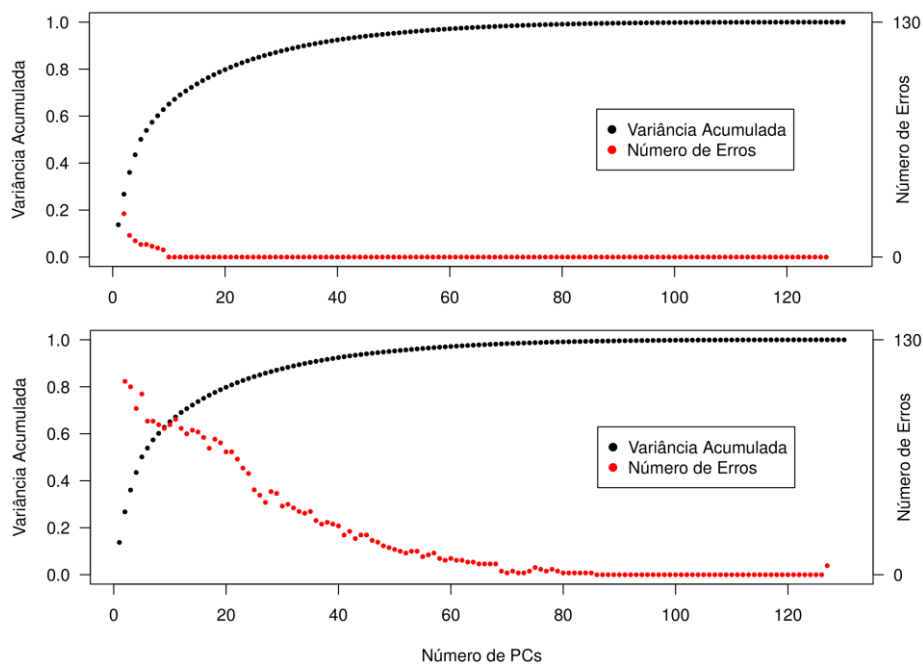


Figura 3: Número de erros de classificação (vermelho) e a variância acumulada (preto) em função do número de PCs utilizados da análise combinada PCA-LDA para os agrupamentos *executante* (gráfico superior) e *líder* (gráfico inferior).

O desempenho de classificação deste modelo foi avaliada, visando confirmar sua capacidade de classificar corretamente observações não usadas no treinamento (DUDA; HART, 1996, cap. 9, p. 32). De acordo com Stone (1974), uma das técnicas de validação mais utilizadas é a *validação-cruzada*, que consiste em dividir os dados disponíveis aleatoriamente em conjuntos de treinamento e teste. O

primeiro deles é utilizado para ajustar os parâmetros do modelo classificador, e o segundo para estimar o erro de classificação. A performance do classificador é estimada pela média dos  $k$  erros de cada iteração.

A validação-cruzada mostrou uma taxa média de discriminação de 95,40 % (d.p. = 5,96) para o agrupamento *executantes* (Tabela 1). Este resultado evidencia a existência de características gestuais próprias à cada indivíduo testado, o que poderíamos chamar de uma "assinatura gestual" dos intérpretes. Estas características são fortemente recorrentes entre as execuções dos intérpretes e não parecem sofrer efeitos de diferentes condições de execução, isto é, quando estes intérpretes acompanham outros músicos seu padrão de movimentação individual se mantém na maior parte inalterado.

	<i>Clarinetista</i>						<i>classificações corretas (%)</i>	
	C1	C2	C3	C4	C5	C6	<i>média</i>	<i>desvio padrão</i>
<i>N</i>	24	22	24	24	21	15	-	-
<i>probabilidade a priori (%)</i>	18,46	16,92	18,46	18,46	16,15	11,53	-	-
<i>classificações corretas (%)</i>	100	90,90	95,83	100	85,71	100	95,40	5,96

Tabela 1: Resultado da validação cruzada para a classificação do agrupamento *executante*.

A taxa média de discriminação obtida para o agrupamento *líder* apresentou baixos valores de acerto, variando de 12,50 % a 35,00 %. Este resultado sugere que os clarinetistas acompanhantes não se aproximaram do padrão gestual dos líderes. Este resultado era esperado, dado o número elevado de dimensões necessárias para a separação completa das classes (86 PCs).

## 5. Discussão

Os resultados sugerem que mesmo excertos de curta duração (13 a 15 segundos) possuem informação gestual suficientemente relevante para discriminar os músicos. A tabela 1 mostra ainda para o agrupamento *executante*, que os valores de classificações corretas variam entre os diferentes clarinetistas, três deles apresentando 100 % de acerto. Este resultado sugere que a hipótese de “assinatura gestual” do intérprete pode se dar em duas dimensões: (1) músicos com menor índice de classificações corretas variam mais os seus gestos que os outros, ou seja, se afastam de seus padrões com mais facilidade; (2) músicos distintos podem apresentar similaridade entre seus padrões gestuais, o que dificultaria a discriminação entre eles.

O agrupamento *líder* teve como objetivo principal testar o efeito das diferentes condições interpretativas impostas por cada primeiro clarinetista, ao clarinetista que o acompanha. Valores altos de classificações corretas para este agrupamento indicariam que os clarinetistas acompanhantes adaptariam seu padrão gestual ao padrão gestual dos primeiros clarinetistas, o que poderia ser uma evidência de acoplamento entre gesto e expressividade acústica, já que não houve contato visual entre os músicos durante as gravações. No entanto, foram observadas baixas taxas de classificações corretas para este agrupamento (22.25%). Apesar dos resultados mostrarem a existência de pequenas variações em regiões críticas do excerto musical, estas não foram suficientes para permitir que os primeiros clarinetistas fossem identificados a partir dos gestos dos clarinetistas que o acompanhavam.

## 6. Conclusão

Foi utilizada uma combinação de PCA e LDA, que permitiu reduzir a dimensionalidade dos dados ao mínimo necessário para uma separação entre as classes. O objetivo foi mostrar a recorrência dos padrões gestuais dos participantes e a invariância de padrões individuais para cada participante. Para isso, a discriminação foi aplicada para dois agrupamentos: (1) *executante*, compreendendo todas as execuções realizadas por cada músico, independente de quem estava acompanhando; e (2) *líder*, compreendendo todas as execuções lideradas por cada músico como primeiro clarinetista. Os resultados para o agrupamento *executante*, mostraram que os padrões gestuais dos músicos possuem uma recorrência entre as diferentes performances, que permitiu a separação das execuções de cada clarinetista utilizando apenas os dados gestuais, com uma taxa de 95,4% de acertos, sugerindo a existência de uma “assinatura gestual” de cada intérprete.

A comunicação humana não se limita ao uso de apenas uma das vias de informação sensorial. Assim como na comunicação verbal, a música faz uso de outras modalidades, tais como movimentos corporais e expressões faciais e gestuais não diretamente associados à produção sonora, mas que podem estar diretamente associados à comunicação, por exemplo, de andamento ou intenção expressiva. Acreditamos que a recorrência dos padrões de movimentação observada entre as execuções de um mesmo músico seja um forte indicativo da existência de informação musical codificada no gesto corporal do intérprete. A possibilidade de identificar artistas a partir de informação extraída do sinal de áudio e vídeo de suas performances,





pode abrir perspectivas promissoras para o reconhecimento de conceitos estilísticos relacionados, por exemplo, a diferentes escolas ou mestres e oferecer grande potencial para a manipulação e gerenciamento de base de dados musicais.

## 7. Referências

- CADOZ, C.; WANDERLEY, M. et al. Gesture-music. *Trends in Gestural Control of Music*, Paris: Ircam, p. 71–94, 2000.
- DELALANDE, F. La gesticque de gould. *Glenn Gould Pluriel*, p. 85–111, 1988.
- DUDA, R.; HART, P. *Pattern classification and scene analysis*. second. [S.l.]: Wiley, 1996.
- FISHER, R. A. The statistical utilization of multiple measurements. *Annals of Eugenics*, Blackwell Publishing Ltd, v. 8, n. 4, p. 376–386, 1938.
- GABRIELSSON, A. Music performance research at the millennium. *Psychology of music*, Sage Publications, v. 31, n. 3, p. 221–272, 2003.
- JOHNSON, R.; WICHERN, D. *Applied multivariate statistical analysis*. 4. ed. [S.l.]: Prentice hall Upper Saddle River, NJ, 1998.
- KELLER, P.; APPEL, M. Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, v. 28, n. 1, p. 27–46, 2010.
- KELLER, P.; KNOBLICH, G.; REPP, B. Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and cognition*, v. 16, n. 1, p. 102–111, 2007.
- LU, X.; WANG, Y.; JAIN, A. Combining classifiers for face recognition. In: IEEE. *Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on*. [S.l.], 2003. v. 3, p. III–13.
- MARCIALIS, G.; ROLI, F. Fusion of lda and pca for face verification. In: TISTARELLI, M.; BIGUN, J.; JAIN, A. (Ed.). *Biometric Authentication*. Berlin: Springer, 2002, p. 30–37.
- PALMER, C. et al. Synchronization of motion and timing in clarinet performance. In: *International Symposium on Performance Science*. [S.l.: s.n.], 2009. p. 1–6.
- R DEVELOPMENT CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2008. Disponível em: <<http://www.R-project.org>>.
- STONE, M. Cross-validators choice and assessment of statistical predictions. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 111–147, 1974.
- VENABLES, W. N.; RIPLEY, B. D. *Modern Applied Statistics with S*. Fourth. New York: Springer, 2002.
- WANDERLEY, M. Quantitative analysis of non-obvious performer gestures. In: WACHSMUTH, I.; SOWA, T. (Ed.). *Gesture and Sign Language in Human-Computer Interaction*. Springer Berlin Heidelberg, 2002, p. 241–253.
- WANDERLEY, M. et al. The musical significance of clarinetists' ancillary gestures: An exploration of the field. *Journal of New Music Research*, v. 34, n. 1, p. 97–113, 2005.
- WANDERLEY, M.; DEPALLE, P. Gesturally-controlled digital audio effects. In: CITESEER. *Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01)*. [S.l.], 2001.