

Metaflauta: design e performance de instrumento aumentado via suporte computacional

Gabriel Rimoldi¹, Jônatas Manzolli¹

¹Núcleo Interdisciplinar de Comunicação Sonora –
Universidade Estadual de Campinas (UNICAMP)
Rua da Reitoria, 165 – 13.083-872 – Campinas – SP – Brasil

{gabriel.rimoldi,jonatas}@nics.unicamp.br

Abstract.

This article describes the design, development and performance of Metaflauta, an augmented instrument based on a hybrid system of gestural capture. We explore both direct and indirect gestural acquisition of flutist as potential controls in multimodal environments. First we present the state of art on the use of sensors and gestural acquisition of flute and then discuss the criteria applied in the design and implementation of the computational system. Finally, we demonstrate the use of the system by means of sound synthesis and sound and visual processing modules developed.

Resumo.

Descrevemos neste artigo a concepção, desenvolvimento e performance através da Metaflauta, um instrumento aumentado baseado num sistema híbrido de captura gestual. Neste sistema, exploramos a captura direta e indireta dos gestos do flautista como potenciais controles gestuais aplicáveis a ambientes multimodais. Apresentamos primeiramente o estado da arte sobre o uso de sensores e aquisição gestual da flauta transversal e, em seguida, discutimos os critérios aplicados no design e na implementação do sistema computacional. Por fim, demonstramos a aplicação desse sistema no controle de síntese sonora digital, processamento sonoro e visual em tempo real.

1. Introdução

O desenvolvimento de novas Interfaces Musicais Digitais (Musical Digital Interfaces - MDI) tem propulsionado uma significativa multiplicação de meios de expressão musical nas últimas décadas. Nessa enorme e heterogênea quantidade de propostas de MDIs, os chamados “instrumentos aumentados” compreendem aqueles que utilizam aparatos externos (sensores) direta ou indiretamente acoplados a instrumentos acústicos, permitindo, assim, a desejada expansão de suas possibilidades sonoro-expressivas [1]. Há extensa literatura sobre a utilização da flauta transversal como instrumento aumentado [2] [3] [4] [5] [6]. Os propósitos do uso de sensores acoplados a esse instrumento variam desde a aplicação em sistemas musicais interativos [4], processamento e síntese sonora [3] até como ferramenta de pesquisa básica [5] [6]. Enquanto a maioria dos trabalhos têm explorado a aquisição gestual direta [2] [3] [4] [5] [6], isto é, através de um conjunto de sensores acoplados ao instrumento, outros

trabalhos tem utilizado o método de aquisição indireta [7] [8], através da análise e processamento do sinal de áudio capturado do instrumento [1] [9]. O trabalho que apresentamos neste artigo faz uso desses dois tipos de aquisição. Desta forma, denominamos a Metaflauta, como um “instrumento aumentado híbrido”. O sistema implementado associa sinais capturados por sensores, acoplados à flauta, ao processamento e extração de características do som utilizando descritores de áudio [10].

Apresentamos inicialmente um panorama sobre as diferentes aplicações da flauta transversal no contexto dos instrumentos aumentados e, em seguida, discutimos a concepção e implementação da Metaflauta. Por fim, descrevemos os resultados obtidos com a aplicação da Metaflauta no controle de sinais multimodais associados à improvisação do intérprete.

2. Flauta transversal como instrumento aumentado

Um dos trabalhos pioneiros no uso de sensores acrescidos a instrumentos acústicos foi a MIDI-flute desenvolvida no IRCAM em meados da década de 80 [2]. Seu protótipo inicial era composto de sensores óticos acoplados às chaves da flauta e um microfone que detectava a mudança de oitava a partir da análise do áudio. Sua versão final, composta de sensores de efeito Hall e imã acoplados às chaves, foi implementada por Michel Ducuor e Michel Starkier em parceria com o flautista Lawrence Beauregard. O objetivo de desenvolvimento da MIDI-flute foi a aplicação da mesma como guia para um sistema de score following [2]. Esse instrumento foi amplamente empregado em obras do repertório eletroacústico misto, tais como ...explosante fixe... (1991-93), de P. Boulez, Jupiter ou La Partition du ciel et de l'enfer (1989) de P. Manoury. Um dos aspectos mais interessantes dessa flauta é que ela não possuía cabos conectando-a à um processador externo, pois havia uma placa de circuito eletrônico acoplada no próprio instrumento.

A Digital Flute apresentada por Ystad e Voinier [3] é bastante similar à interface desenvolvida pelo IRCAM. As chaves dessa flauta foram acrescidas de 14 sensores de efeito hall e imã e um microfone foi posicionado junto à rolha no interior do tubo do instrumento. Diferentemente da MIDI-flute, os autores desenvolveram uma interface capaz de controlar síntese sonora digital até mesmo em situações em que o instrumentista não emitisse jato de ar através do tubo. O método de síntese utilizava tanto o sinal acústico como parâmetros de modelagem física. A extração de sinais produzidos pelo comportamento da coluna de ar no interior do tubo, permitia a manipulação de características do timbre do instrumento, simulando, por exemplo, o ressonador de um outro instrumento ou mesmo aumentar ou diminuir o tamanho do tubo virtualmente. Os autores utilizavam também quatro pedais para controlar parâmetros macroestruturais ou configurações dos sensores acoplados à flauta.

A Hyper-Flute, desenvolvida por Palacio-Quintin [4], baseou-se no sistema SensoLab, criado por J. Scherpenisse e A. J. Van den Broeck no Instituto de Sonologia de Haia. Essa interface foi utilizada em sistemas interativos musicais e dispunha de uma grande variedade de controles adicionados ao instrumento acústico. Essa diversidade implicava numa considerável complexidade de controle associado à performance instrumental, sendo necessário um aprendizado e treino para desenvolver os potenciais expressivos do instrumento. A autora apresenta ainda o desenvolvimento de uma Hyper-

Bass-Flute, que possui funcionamento similar à anterior, com a vantagem de possuir um maior espaço para o acoplamento de sensores.

Da Silva, Wanderley e Scavone [5] utilizaram sensores para mensurar aspectos relacionados à embocadura, mais especificamente à velocidade e direção do jato de ar. Os autores utilizaram dois sensores de pressão prolongados com duas pequenas sondas, cada qual posicionada numa extremidade lateral do porta-lábios da flauta. Esse sistema compreendia ainda um microfone, um sensor resistivo à pressão (FSR) e um atuador atado junto ao pescoço do instrumentista, o qual fornecia um feedback sobre a vazão do ar detectada pelos sensores de pressão.

Almeida e colaboradores [6], interessados em relacionar as transições de dedilhado da flauta transversal com aspectos acústicos decorrentes da dificuldade técnica de determinada passagem, analisaram a presença de transientes e a diminuição da amplitude entre duas notas ligadas. Para isto, utilizaram 16 sensores fotorresistores colados na borda inferior de cada chave. O sinal acústico do instrumento foi também gravado utilizando-se um microfone. Na Tabela 1, resumimos os principais tipos de sensores aplicados ao desenvolvimento de flautas transversais aumentadas encontradas na literatura.

Tabela 1: Utilização de diferentes sensores nos trabalhos que exploram a aquisição gestual direta da flauta transversal

Tipo de Sensores	Acelerômetro	FSR	LSR	Sensor de efeito hall	Ultrassônico	Chaves de mercúrio	Chave de contato	Microfone	Pedal	Sensor de Pressão	Atuador
MIDI-Flute [2]				●							
Ystad & Voinier [11]				●				●	●		
Palacio-Quintin [4]	●	●	●	●	●	●	●	●			
Silva et al. [5]		●								●	●
Almeida et al. [6]			●								

Outras pesquisas têm explorado ainda a extração de informações gestuais de maneira indireta, recuperando informação musical através da análise e processamento do sinal de áudio. Kereliuk e colaboradores [7] pesquisaram técnicas de aquisição indireta em tempo real para aferir posições de dedilhado da flauta na emissão de harmônicos. O objetivo da pesquisa foi detectar os diferentes padrões de dedilhado e pressão que resultam numa mesma frequência fundamental. Os autores extraíram informações temporais e espectrais dos sinais a partir da aplicação de análise de componentes principais (PCA) da energia dos sub-harmônicos de cada sinal. Verfaillie e colaboradores [8] realizaram experimento similar e utilizaram descritores de áudio para avaliar o efeito do aumento do fluxo de ar (overblow) na produção de notas harmônicas.

Encontramos ainda trabalhos que exploram a aquisição gestual direta através de sensores não acoplados ao corpo do instrumento ou instrumentista. É o caso, por exemplo, do Multimodal Music Stand (MMMS), um sistema de captura de gestos que utiliza sensores eletromagnéticos acoplados à estante, além de um conjunto de câmeras

e microfone que complementam a informação detectada pelos sensores [12]. A primeira experiência realizada com o sistema foi a obra interativa *timeandagain*, composição para flauta e computador de JoAn Kuchera-Morin. Para a referida obra, dois tipos de movimento distintos foram mapeados: a movimentação para cima e para baixo do corpo do instrumento e a rotação da cabeça no eixo transversal. O movimento do corpo da flauta no eixo longitudinal foi detectado pela câmera frontal e para o movimento da cabeça utilizaram a imagem capturada por uma câmera posicionada lateralmente. Visto que as informações extraídas com as câmeras eram pouco acuradas, devido ao ruído gerado pela variação de luz e outros fatores ambientais, os autores agregam informação originados através da interface de áudio e de sensores de campo elétrico acoplados ao MMMS.

3. Design e implementação da *Metaflauta*

A *Metaflauta* foi concebida como instrumento aumentado baseado num sistema híbrido de captura gestual, o qual associa informações diretas e indiretas produzidas a partir dos movimentos do flautista. A construção dessa interface teve como objetivo primordial a expansão das possibilidades de controle de sistemas interativos multimodais sem, entretanto, comprometer as possibilidades performáticas já propiciadas pelo instrumento acústico. A captura direta é realizada por meio de sensores acoplados ao instrumento enquanto que para a indireta utiliza-se processamento do sinal de áudio capturado via microfone.

Os sensores utilizados na *Metaflauta* focam a exploração de movimentos que não estão diretamente associados à produção sonora, o quais não podem ser recuperados por meio do processamento do áudio. Nosso principal propósito foi explorar possibilidades de controle através dos chamados *gestos auxiliares* [13] [9].

Em contrapartida, os gestos ou sonoridades advindas da produção sonora do instrumento são analisados em tempo real por um conjunto de descritores de áudio, os quais extraem as seguintes características: a) Temporais (energia média, zero crossing [14]), b) Espectrais (Centroide [15] e Fluxo [16]), c) Harmônicas (Chroma [17]) e d) Perceptivas (Loudness [18] e Escala Bark [19]). Tanto as informações gestuais capturadas diretamente via sensores como as capturadas indiretamente via análise e processamento do sinal de áudio são enviadas ao módulo de mapeamento, o qual combina e transforma as mesmas em múltiplos parâmetros de controle audiovisual.

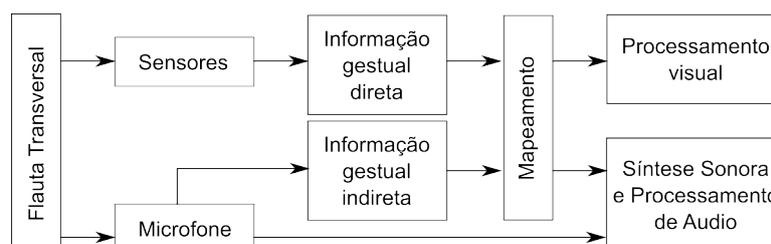


Figura 1: Diagrama de design da *Metaflauta*

3.1 Captura gestual direta

As informações capturadas pelos sensores são enviadas ao computador a partir de um protótipo Arduino¹, com processador Atmel ATmega2560. Na Tabela 2, descrevemos os tipos de sensores utilizados no protótipo desenvolvido. O sinal de cada sensor pode tanto ser utilizado de modo independente como também combinado ao sinal de outros sensores, afim de se extrair informações mais acuradas e de mais alto nível em relação aos movimentos do performer. É o caso, por exemplo, do giroscópio, acelerômetro e magnetômetro, que formam uma composição MARG (*Magnetic, Angular Rate and Gravity*), por meio da qual obtemos os valores de posicionamento angular do instrumento em três eixos, a dizer *pitch*, *roll* e *yaw*. Informações de ordem superior também podem ser utilizadas pelo sistema, como por exemplo, os momentos estatísticos de aceleração durante um certo período de tempo.

Tabela 2: Categorias e medida extraída de cada sensor na captura gestual direta da Metaflauta

Sensores	Função/Descrição
1 Acelerômetro triaxial	Mede a aceleração em cada eixo do instrumento.
1 Giroscópio triaxial	Mede a variação angular de cada eixo do instrumento.
1 Magnetômetro triaxial	Mede a intensidade e direção do campo magnético exercido sobre cada eixo do instrumento.
1 sensor ultrassônico	Detecta a distância da mão direita do instrumentista em relação ao corpo do instrumento.
1 FSR (<i>force sensitive resistor</i>)	Detecta a pressão do polegar direito.
6 chaves de contato	Controles discretos posicionados ao longo do corpo do instrumento sendo 3 controladas pelo polegar direito, 2 pelo polegar esquerdo e 1 pelo dedo mínimo esquerdo.

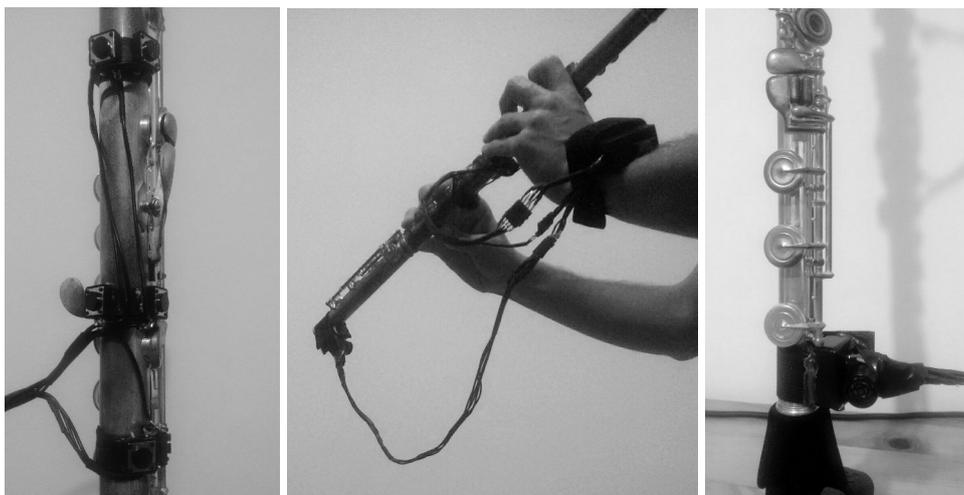


Figura 2: Detalhes dos sensores acoplados à Metaflauta

¹ <https://www.arduino.cc/>

3.2 Captura gestual indireta

A recuperação de informações gestuais relacionadas à produção sonora do instrumento foi realizada a partir da análise e processamento do sinal de áudio. As características extraídas do sinal de áudio foram agrupadas em quatro categorias, a dizer: temporais, espectrais, harmônicas e perceptivas. Como demonstrado pelo diagrama da Figura 3, o vetor de características temporais é obtido através do janelamento do sinal capturado pelo microfone. Cada janela temporal é processada via Transformada Rápida de Fourier (*Fast Fourier Transformation - FFT*), através da qual são extraídas as características espectrais do sinal. As características harmônicas e perceptivas são extraídas através do sinal obtido pela FFT por meio de modelos sinusoidais e relacionados à percepção auditiva humana, respectivamente.

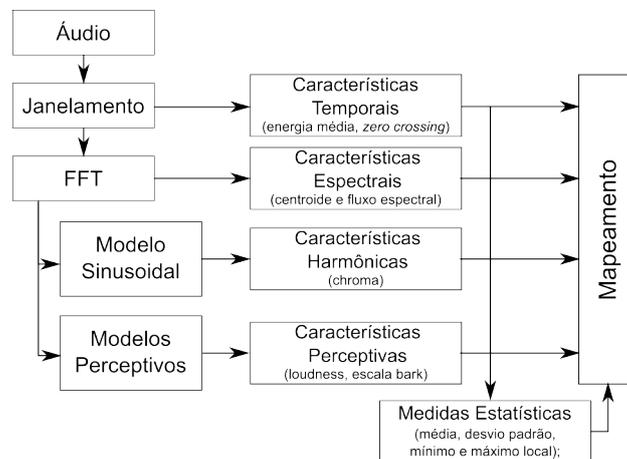


Figura 3: Diagrama de aquisição gestual indireta através da análise e processamento do sinal de áudio capturado.

Os sinais obtidos pela análise do áudio são então enviados ao módulo de mapeamento, responsável pela associação dos mesmos aos parâmetros de controle de processamento audiovisual. Medidas estatísticas que descrevem o comportamento de cada vetor de características ao longo do tempo são enviadas ao módulo de mapeamento, configurando-se as mesmas também como variáveis de controle do sistema. Para a análise e extração das características supracitadas, utilizamos a biblioteca *PDescriptors* implementada em Pure Data² por Monteiro e Manzolli [10].

4. Módulos de síntese sonora e processamento de áudio

4.1 Phase Vocoder

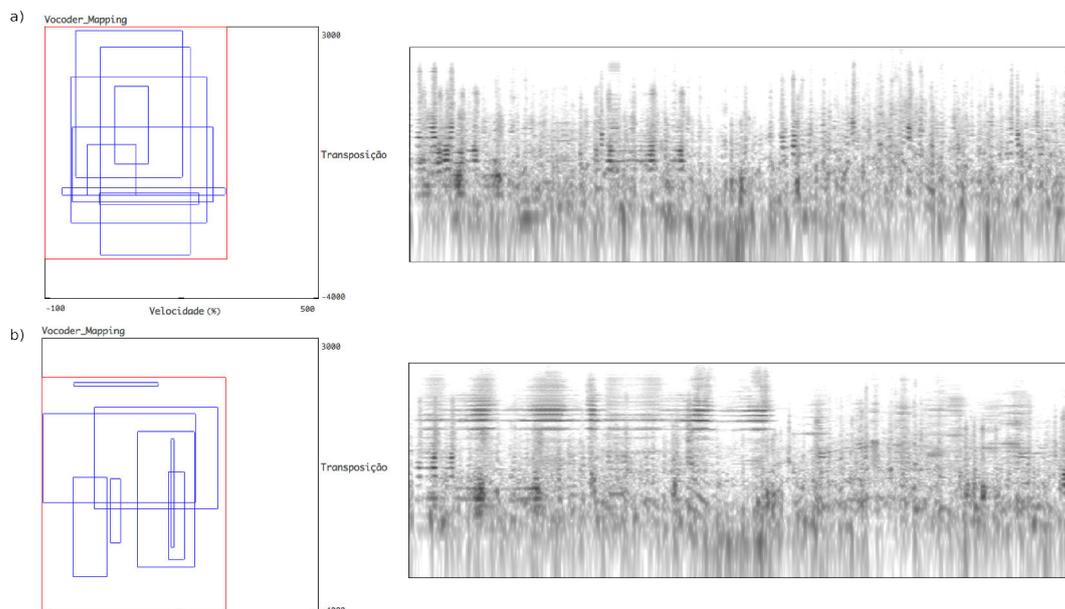
Neste módulo, o som é gravado e armazenado pelo performer em um *buffer*. O sistema compreende oito módulos independentes de síntese via *phase vocoder* que processam o sinal armazenado no *buffer*. O performer pode armazenar o som continuamente através

² www.puredata.info.

de um *buffer* circular ou selecionar trechos específicos ao longo da performance. Cada módulo possui dois parâmetros de que são controlados interface gestual, a dizer: a velocidade de leitura do *buffer* e a taxa de transposição de altura, associados respectivamente ao *roll* e *yaw* da composição MARG. Os dados originados da interface gestual são mapeados de modo independente para cada módulo de síntese através de uma transformação afim. Para cada módulo k , o mapeamento entre os parâmetros controle transposição e velocidade e as variáveis *roll* e *yaw* são respectivamente mapeados como:

$$f(\text{roll})_k = A_k \text{roll} + B_k \quad g(\text{yaw})_k = C_k \text{yaw} + D_k$$

A manipulação dos valores de A, B, C e D associados a cada módulo k na transformação afim configura-se um controle macroestrutural que pode ser manipulado pelo instrumentista ao longo da performance. A partir de um espaço bidimensional de controle paramétrico, o performer controla o domínio sob o qual serão gerados randomicamente valores que modularão as variáveis de controle para cada módulo k . A alternância entre domínios previamente configurados pode ser realizada através de chaves de contato acopladas à flauta. Na Figura 4 demonstramos diferentes domínios estabelecidos entre as variáveis de controle e os parâmetros para cada módulo de síntese. Em cada um dos mapas (denotados pelas letras de *a* até *d*), o quadrado vermelho representa o domínio estabelecido no espaço paramétrico e os quadrados azuis o mapeamento atribuído a cada um dos oito módulos de síntese. Os espectrogramas³ apresentados junto a cada mapa paramétrico apresentado pela figura demonstram as diferentes respostas do módulo através simulação de um mesmo vetor gestual armazenado da interface e de uma mesma amostra sonora utilizada no processamento⁴.



³ Os espectrogramas foram gerados com o auxílio do software *Sonic Visualizer* (<http://www.sonicvisualiser.org/>), com taxa de amostragem de 44.1KHz, janelamento 4096 amostras, taxa de sobreposição de 512 amostras e escala logarítmica na representação das frequências.

⁴ Os arquivos de áudio gerados na simulação apresentada pela Figura 4 encontram-se em anexo.

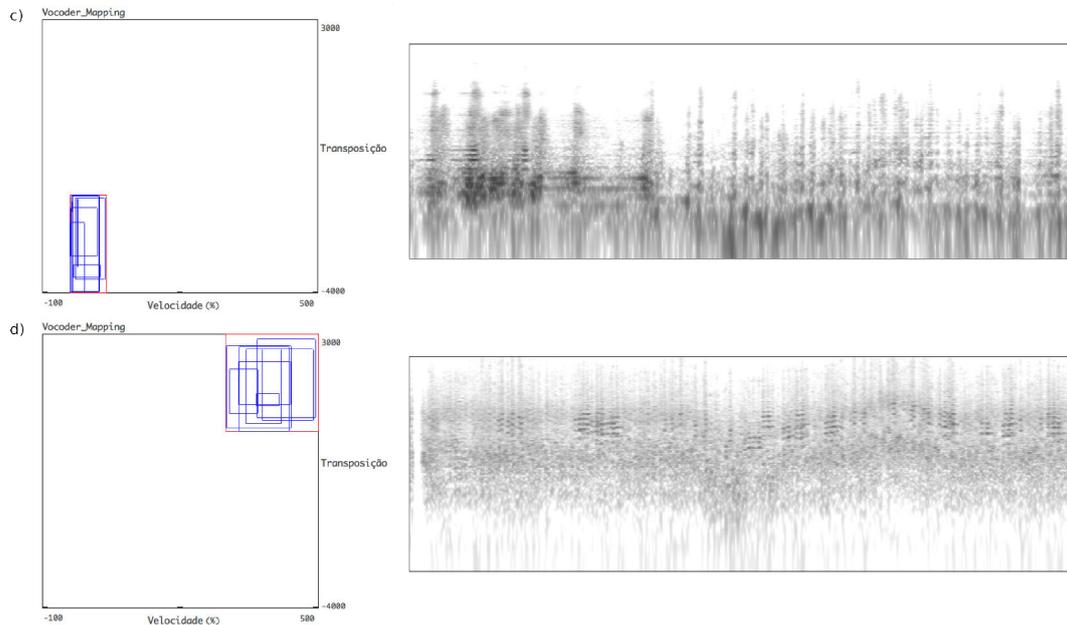


Figura 4: Mapa de controle macroestrutural das transformações afins associadas a cada módulo de síntese e as respectivas respostas de espectro obtidas pela simulação de um mesmo vetor gestual e mesma amostra sonora processada.

4.2 Filtragem espectral

Neste outro módulo, o som do instrumento é processado em tempo real por um conjunto de 24 filtros que recombina informações espectrais de amostras captadas e armazenadas do próprio instrumento. Ao longo da improvisação, o performer amostra o áudio executado utilizando uma das chaves acopladas ao instrumento. Destas amostras, são armazenadas tanto as informações de magnitude do espectro sonoro, bem como as informações gestuais diretas e indiretas associadas ao momento de execução das mesmas. O módulo possui a capacidade de armazenamento de até 500 quadros de análise de amostras. As informações gestuais diretas utilizadas neste módulo foram o deslocamento azimutal (*yaw*) e de elevação (*roll*) do corpo do instrumento, que constituem um espaço bidimensional de representação do movimento. As informações indiretas utilizadas foram o *loudness* e *chroma*, ambos obtidos via Transformada de Fourier (FFT) do sinal amostrado.

Cada filtro processa o sinal em tempo real a partir de uma combinação da magnitude espectral de diferentes amostras armazenadas. Tendo em vista o estabelecimento de um controle macroestrutural do performer em relação à seleção destas amostras, utilizamos o algoritmo *boids* como modelo de controle da interface. O algoritmo, desenvolvido por Craig Reynolds [20] e implementado em Pure Data por Eric Singer, Jasch e Andr Sier, simula o comportamento emergente do deslocamento espacial de pássaros e baseia-se num conjunto de regras simples e locais, tais como: separação, alinhamento e coesão entre vizinhos.

No presente módulo, cada filtro é associado a um *boïd*, que navega sobre o espaço bidimensional de representação do movimento, constituído por sua vez das informações gestuais diretas (*yaw* e *roll*) amostradas em *buffer*. A posição do *boïd* associado a cada filtro determinará quais amostras este utilizará no processamento do sinal do instrumento, constituindo deste modo o filtro uma soma ponderada da magnitude espectral das três amostras mais próximas ao referido *boïd* no espaço de representação do movimento. Matematicamente, denotamos por $B_t = \{x_{bt}, y_{bt}\}_{b=1,2,\dots,24}$ o conjunto de coordenadas de cada *boïd* b no tempo t e $A = \{x_a, y_a\}_{a=1,2,\dots,500}$ o conjunto de coordenadas espaciais associadas a cada amostra armazenada. A magnitude espectral associada ao filtro b resultará da soma ponderada das três menores distâncias entre A e B_t , de forma que:

$$f(b)_t = \frac{1}{\sum_{j=1}^3 \bar{d}(\min_j)} \sum_{j=1}^3 \left[\sum_{k=0}^N \text{mag}(\min_j D(A, B_t)) \bar{d}(\min_j D(A, B_t)) \right]$$

A utilização deste algoritmo no controle de parâmetros musicais propicia o surgimento de comportamentos emergentes, que possuem um certo grau de complexidade e diversidade, sem, no entanto, sobrecarregar o performer com um demasiado número de parâmetros de controle. As variáveis de controle em tempo real do sistema estão associadas a parâmetros globais, tais como: a posição e força de um ponto atrator aos *boïds*, a velocidade global, mínima e máxima desempenhada pelos mesmos e a inércia ou prontidão para mudança de velocidade e aceleração dos *boïds*. Na Tabela 3 apresentamos as variáveis de controle associado ao controle de parâmetros globais do algoritmo.

Tabela 3: Parâmetros globais do algoritmo *boïds* associado às variáveis de controle da *Metaflauta*

Parâmetros de controle do algoritmo	Mapeamento
Posição (x, y) do ponto atrator dos <i>boïds</i>	<i>Roll</i> e <i>Yaw</i> da composição MARG
Força de atração do ponto atrator	Magnitude da aceleração obtida do acelerômetro + <i>Loudness</i>
Velocidades global, máxima e mínima desempenhada pelos <i>boïds</i>	Valores de média, mínimo e máximo de velocidade integrada da composição MARG
Inércia (prontidão para mudança de velocidade e aceleração dos <i>boïds</i>)	<i>Fluxo Espectral</i> (desvio padrão local)

O algoritmo *boïds* é também utilizado como controle da espacialização sonora. A posição no eixo azimutal do sinal sonoro resultante de cada filtro é associado ao movimento do respectivo *boïd* no espaço de representação do movimento.

5. Processamento visual em tempo real

Nossa proposta de processamento visual controlado pela *Metaflauta* foi o estabelecimento de uma associação direta com os parâmetros de controle aplicados aos módulos de síntese e processamento do som. A composição gráfica foi realizada com o auxílio da biblioteca GEM implementa em Pure Data. As amostras sonoras armazenadas em *buffer* e utilizada no processamento de ambos os módulos de síntese sonora

referidos anteriormente são representadas pela renderização de múltiplas texturas. A posição de cada textura na imagem é definida pelas coordenadas da composição MARG registradas no momento de captura da amostra. A cor de cada da textura é definida pelo centro de massa do *chroma* e o tamanho e taxa de transparência da textura definida pelo *loudness*, ambas medidas extraídas da análise da referida amostra.

Associado ao controle de parâmetros do módulo de síntese via filtragem espectral, o deslocamento dos *boids* foi representado por meio de círculos e os filtros selecionados ao longo do tempo por cada *boid* representado por linhas curvas que conectam os referidos pontos no espaço de representação do movimento. Na Figura 5 apresentamos alguns resultados visuais obtidos pelo processamento em tempo real dos dados originados da *Metaflauta*.

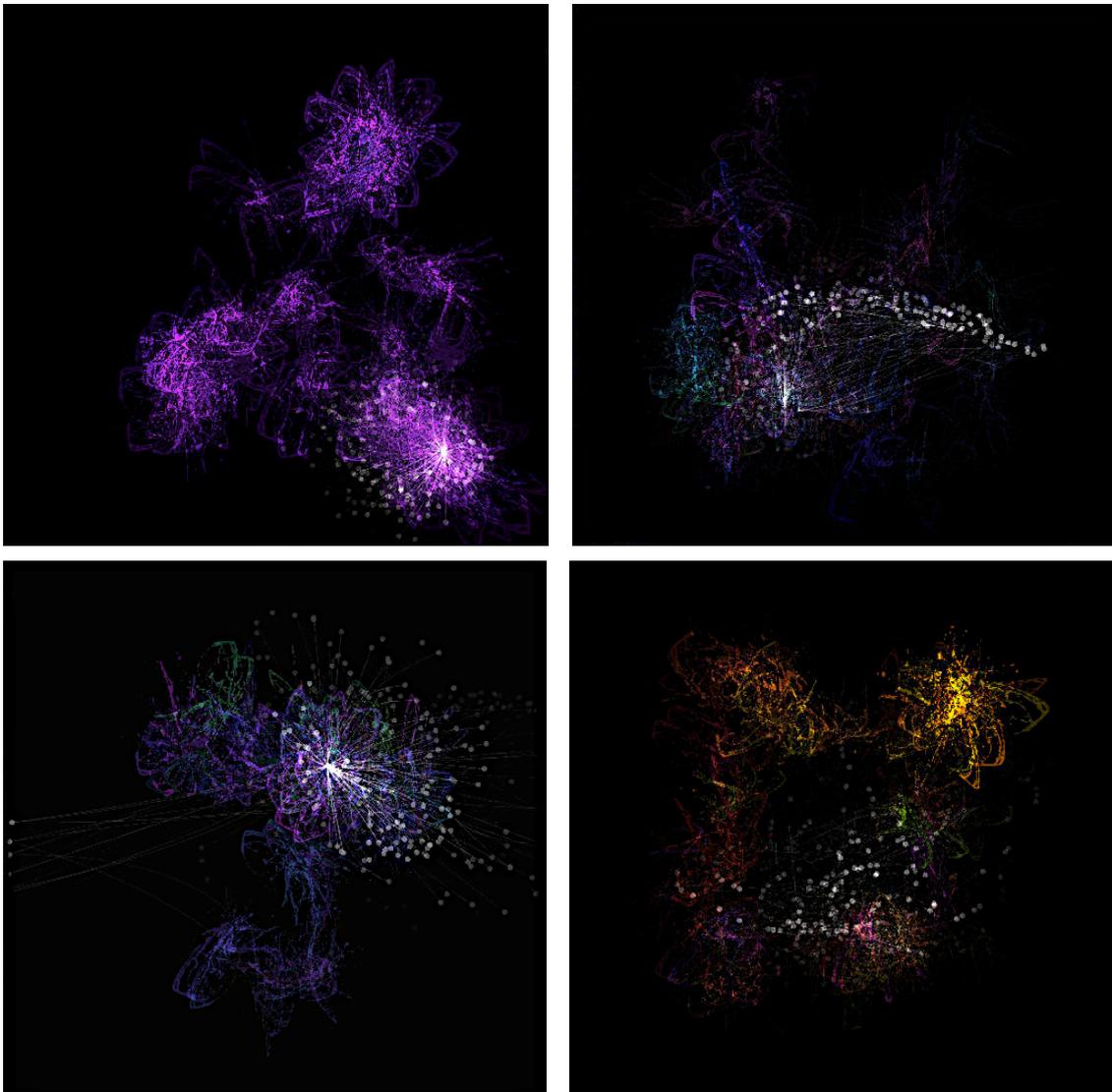


Figura 5: Amostras do processamento visual em tempo real gerado a partir dos dados extraídos da interface gestual da *Metaflauta*

6. Considerações Finais

Apresentamos neste trabalho a concepção e desenvolvimento da *Metaflauta*, um instrumento aumentado baseado num sistema híbrido de captura gestual. Em nosso modelo, a aquisição direta foi aplicada como estratégia para a captura de gestos auxiliares, ao passo que alguns gestos diretamente correlacionados à produção sonora foram capturados indiretamente, isto é, por meio da análise e processamento do sinal de áudio do instrumento. Discutimos ainda as estratégias de mapeamento das variáveis obtidas da interface no controle de parâmetros de processamento de som e imagem. Consideramos que a associação de informações gestuais capturadas direta e indiretamente demonstrou-se eficaz como estratégia de mapeamento de múltiplos parâmetros de controle dos módulos de processamento audiovisual.

Em trabalhos futuros, pretendemos ampliar a utilização das informações gestuais capturadas como variáveis de controle para os módulos de processamento audiovisuais. Tendo em vista o bom desempenho no controle de parâmetros através de informações de mais alto nível obtidas pela fusão de diferentes sensores, consideramos a utilização de procedimento similar aos outros dados capturados. A interpretação de informações de mais alto nível, sobretudo em relação às informações gestuais capturadas indiretamente, poderá assegurar uma maior robustez e precisão no controle de múltiplos parâmetros sem, no entanto, dispender um considerável aumento na complexidade de performance do instrumento. Numa perspectiva próxima, pretendemos ainda explorar a utilização de métodos de aprendizado não supervisionado de máquina no mapeamento da interface que possibilite um ajuste dinâmico das variáveis de controle mapeadas aos diferentes parâmetros de controle sonoro e visual.

Referências

- [1] MIRANDA, E. R.; WANDERLEY, M. M. *New digital musical instruments: control and interaction beyond the keyboard*. Middleton, USA: A-R Editons, 2006.
- [2] POUSSET, D. "La flute-midi, l'historique & quelques applications" *Master's thesis, Université de Paris - Sorbonne*, 1992.
- [3] YSTAD S.; VOINIER, T. "A virtually real flute". *Computer Music Journal*, vol. 25, no. 2, pp. 13-24, 2001.
- [4] PALACIO-QUINTIN, C. "The hyper-flute" in *Proceedings of the International Conference New Interfaces for Musical Expression, NIME-03*, Montreal, 2003, pp. 206–207.
- [5] SILVA, T.; WANDERLEY, M. M.; SCAVONE, G. "On the Use of Flute Air Jet as A Musical Control Variable ," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME05)*, Vancouver (Canada), 2005, pp. 105-108.
- [6] ALMEIDA, A.; CHOW, R.; SMITH, J.; WOLFE, J. "The kinetics and acoustics of fingering and note transitions on the flute" *Journal of Acoutical Society of America*, vol. 126, no. 3, pp. 1521-1529, 2009.
- [7] KERELIUK, C.; SCHERRER, B.; VERFAILLER, V.; DEPALLE, P.;

- WANDERLEY, M. M. "Indirect Acquisition of fingerings of harmonic notes on the flute" in *Proceedings of the International Computer Music Conference (ICMC'07)*, Copenhagen, Denmark, 2007, pp. 263-266.
- [8] VERFAILLE, V.; DEPALLE, P.; WANDERLEY, M. M. "Detecting overblown flute fingerings from the residual noise spectrum" *Journal of Acoustical Society of America*, vol. 127, no. 1, 2010.
- [9] TRAUBE, C.; DEPALLE, P.; WANDERLEY, M. M. "Indirect Acquisition of Instrumental Gesture Based on Signal, Physical and Perceptual Information" in *Proceedings of the 2003 Conference on New Interfaces for Musical Expression (NIME-03)*, Montreal (Canada), 2003.
- [10] MONTEIRO, A. C.; MANZOLLI, J. "A Framework for Real-time Instrumental Sound Segmentation and Labeling," in *Proceedings of IV International Conference of Pure data*, Weinmar, 2011.
- [11] YSTAD, S.; VOINIER, T. "Design of a Flute Interface to Control Synthesis Models," in *Proceedings of the International Computer Music conference ICMC99*, Beijing (China), 1999.
- [12] BELL, B. et al., "The Multimodal Music Stand," in *Proceeding of New Interfaces for Music Expression Conference, NIME07*, New York, 2007.
- [13] WANDERLEY, M. M.; VINES, B. "Origins and functions of clarinetist's Ancillary Gestures," in *Music and Gesture*, Ashgate, Hampshire, England, 2006.
- [14] GOUYON, F.; PACHET, F.; DELERUE, O. "On the use of zero-crossing rate for an application of classification of percussive sounds.," in *Proceedings of the 6th conference on digital audio effects*, Verone, Italy, 2000, pp. 147-152.
- [15] WESSEL, D. L. "Timbre space as a musical control structure," *Computer Music Journal*, vol. 3, no. 2, pp. 45-52, 1979.
- [16] ONG, B. S. "Towards automatic music structural analysis identifying characteristic within-song excerpts in popular music" *Master's thesis, Music Technology Group, UPF*, 2005.
- [17] FUJISHIMA, T. "Realtime Chord Recognition of Musical Sound: a System Using Common Lisp Music," in *CCRMA*, 1999.
- [18] ZWICKER, E.; FASTL, H. *Psychoacoustics: Facts and models*. Secaucus, NJ, USA: Springer-Verlag New York, 1990.
- [19] ZWICKER, E. "Subdivision of the audible frequency range into critical bands" *Journal of the Acoustical Society of America*, vol. 33, pp. 248-249, 1961.
- [20] REYNOLDS, C. "Flocks, herds and schools: A distributed behavioral model," in *SIGGRAPH '87: Proceedings of the 14th annual conference on Computer graphics and interactive techniques (Association for Computing Machinery)*, New York, 1987, pp. 25-34.