

Da Assinatura Gestual à Expressividade Musical

José Fornari

Núcleo Interdisciplinar de Comunicação Sonora (NICS)

Universidade Estadual de Campinas (UNICAMP)
Cidade Universitária Zeferino Vaz – 13.083-872 – Campinas – SP – Brazil

tutifornari@gmail.com

Abstract. *Musicians express their art through gesture, understood here as "the movement with intention." The field of artistic technique is therefore seen as an intentional process of standardizing sign, consisting of the artist repetition-driven attention. Thus, a musical style is formed, and expressed by the recurrence patterns of unconscious gestures. These often convey a similarity that allow pattern identification, here referred as "gestural signature." Such patterns allow the recognition of a musical style found in a particular interpreter, with, in a broad range, contributes for the self-organization of a musical genre. Therefore, musical gestures are well correlated to its musical cognition. Patterns that constitute gestural signature can now be collected and identified using technological resources. This information can be gathered through motion sensors (i.e. cameras, accelerometers, etc.) and analyzed in real time by computer models, which determine its unique aspects. This paper presents an introductory research, currently in progress at NICS / UNICAMP, using data collected by gestural interfaces as a mechanism for the acquisition of musical patterns, as well as patches (computational models developed in PD programming environment) to control sound processing and the dynamic analysis of music expression. With that, it's here intend to lay the groundwork for the study of automatic identification of gestural signatures, and their correlation with musical expressiveness.*

Resumo. *O músico expressa a sua arte através do gesto, aqui entendido como "o movimento com intenção". O domínio da técnica artística é assim visto como um processo intencional de padronização gestual, constituído pela repetição guiada através da atenção do artista. Desse modo um estilo musical é formado e expresso pela recorrência inconsciente de padrões gestuais. Estes, muitas vezes, apresentam tamanho grau de auto-similaridade que constituem um padrão de identificação; aqui referido como "assinatura gestual". Tal padrão permite reconhecer o estilo musical de um dado interprete, ou, num macro senso, contribui para a formação de um gênero musical. O gesto é assim correlacionado à cognição musical. Os padrões que compõem a assinatura gestual podem atualmente ser coletados e identificados através de recursos tecnológicos. A informação gestual pode ser coletada através de sensores de movimento (ex: câmeras, acelerômetros, etc.) e analisados, em tempo-real, por modelos computacionais que colem seus aspectos únicos. Este trabalho apresenta uma pesquisa correlata, atualmente em andamento no NICS / UNICAMP, onde são utilizadas interfaces gestuais (equipamentos de aquisição de padrões gestuais) e patches (modelos computacionais do ambiente de programação PureData) para o controle dinâmico de processos sonoros e análise da expressividade musical. Com isso pretende-se lançar bases para um estudo posterior sobre a identificação automática da assinatura gestual, e a sua correlação com a expressividade musical.*

1. Introdução

Da mesma forma que gestos expressam manifestações artísticas, estas exprimem padrões gestuais. Nas culturas populares e folclóricas, freqüentemente gêneros musicais correspondem aos respectivos estilos de dança e são assim referenciados pela mesma denominação (ex: o Samba, a Polca, ou o teatro *Noh*). No Brasil, é rica a tradição da música popular, especialmente no que tange a existência de diversos ritmos e suas correspondentes danças. Estes estão intrinsecamente relacionados ao contexto do gesto corporal e à sua interveniente correspondência musical. Se o movimento com intenção é o gesto, a dança pode ser vista como uma autêntica “arte do gesto”. Do mesmo modo, a música pode ser definida como uma “arte do som”, ou conforme especificado pelo compositor Edgard Varèse, a arte do “som organizado”. A performance musical se assemelha à dança no fato de que ambas apresentam-se intrinsecamente relacionadas ao desenvolvimento de padrões gestuais desenvolvidos ao longo do tempo. Na música popular, o desenvolvimento de uma prosódia gestual de contextos artísticos pré-estabelecidos, através de estilos interpretativos embasados em gêneros musicais, permite o estabelecimento de um significado musical que permeia a comunicação emotiva entre compositor e audiência, mediado pela expressividade musical.

1.1. As Fronteiras do Tempo

Tradicionalmente, os elementos que compõe a música são categorizados em: Melódicos, Harmônicos e Rítmicos. Melodia é relacionada à percepção da altura musical (ou *pitch*), um fenômeno micro-temporal, que ocorre em diminutos intervalos de tempo (entre 50ms e 50 s, que equivalem às respectivas frequências de 20Hz e 20.000 Hz). Da mesma forma, a Harmonia corresponde à percepção da relação entre alturas musicais e, portanto, também ocorre primordialmente em diminutos intervalos de tempo. Já o Ritmo corresponde à percepção de maiores intervalos de tempo; acima dos 50ms, até aproximadamente 3s, que é considerado como o intervalo do agora, ou Presente Aparente (*Specious Present*), conforme definido pelo psicólogo William James. A figura 1 apresenta graficamente as fronteiras temporais da percepção musical que distinguem a percepção rítmica da tonal.



Figura 1. Os limites da percepção temporal sonora.

O Ritmo é o elemento musical mais intimamente ligado ao gesto corporal, seja referindo-se à ação no movimento que gera o discurso rítmico, seja pela indução no ouvinte ao movimento corporal. A percepção auditiva está intimamente relacionada à sistemapercepção da orientação espacial, tanto em nível psicoacústico quanto cognitivo. A percepção sonora é regprincipalmente constituída pela sensibilidade do ouvido interno. Este é constituído por duas partes: 1) a Cóclea, que traduz os impulsos mecânicos do som em sinais elétricos; e 2) o Sistema Vestibular, responsável pelo equilíbrio corporal, orientado pela sensação de aceleração corporal. Alguns pesquisadores sugerem que seria esta a razão pela qual, ao escutarmos músicas que nos agradam, tendemos a marcar seu ritmo com movimentos corporais, bem como temos o impeto de aumentar o seu volume (intensidade sonora), uma vez que sons de maior intensidade estimulam mais integralmente o ouvido interno (TODD, 2009).

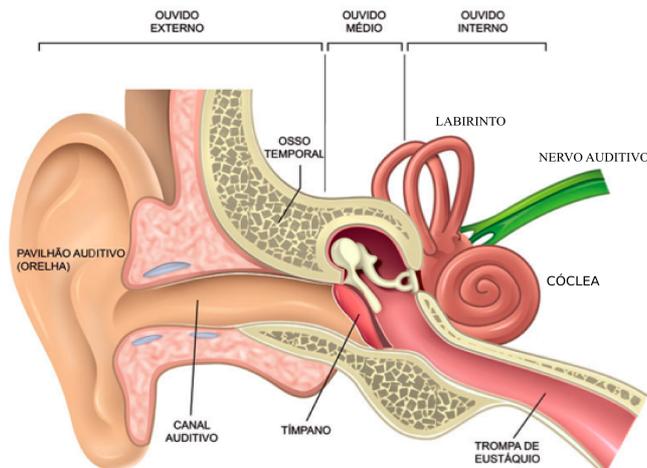


Figura 2. Diagrama do ouvido, onde é demonstrado o Ouvido Interno, constituído pela Cóclea e pelo Labirinto.

1.2. O Cérebro Musical

Em relação ao processamento cerebral da informação sonora da escuta musical, esta é feita simultaneamente por distintos circuitos neurais. Um deles é relacionado à razão, outro à emoção, e outro, por fim, ao gesto. O córtex auditivo é a área cerebral que processa o entendimento da informação sonora. Em termos evolutivos, o córtex é a área cerebral mais recente, e é relacionada ao raciocínio lógico e à cognição. Em paralelo, a informação sonora é também enviada às regiões mais antigas do desenvolvimento cerebral, relacionadas às sensações e às reações de sobrevivência, como “fugir ou lutar” (fight or flight). A informação sonora advinda da Cóclea pelo nervo auditivo, inicialmente chega, na forma de impulsos elétricos, na região do tronco encefálico, ou cerebral chamada de: núcleo coclear. Esta é a primeira região cerebral conectada ao nervo auditivo. O núcleo coclear é localizado na ponte do tronco cerebral que se conecta ao cerebelo onde é processada a informação de localização espacial (azimute, ou localização horizontal) juntamente com a sensação de equilíbrio, também captada no ouvido interno, pelo Sistema Vestibular. Em seguida, a informação é enviada ao Tálamo, uma estrutura cerebral constituinte do sistema Límbico, que se conecta ao Hipocampo (que regula a memória de longa-duração), e à Amígdala (que regula a agressividade e o medo no indivíduo). Do Tálamo, a informação processada segue para o Córtex Auditivo e para, a primeira região do córtex cerebral, localizado no lobo temporal, que recebe informação sonora. O córtex é composto pela região superficial, mais externa do cérebro, constituído pelas últimas camadas desenvolvidas durante a evolução biológica, e está presente apenas nos cérebros de mamíferos. O córtex regula sofisticadas funções cognitivas, como: atenção, memória, linguagem e consciência, o [KELLER, 2010]. A outra estrutura encefálica envolvida com o processamento de som é o Cerebelo [MATHIAK, 2002]. Esta é conhecida por regular funções motoras, como o equilíbrio e expressão corporal. No cerebelo é também processada informação referente à percepção temporal de eventos, como a noção de pulso e a percepção rítmica. A figura 3 demonstra as áreas cerebrais reacionadas aos processos de cognição, emoção e gesto musical.

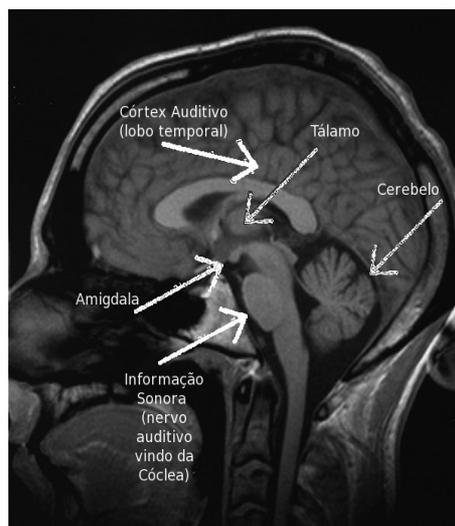


Figura 3. Áreas do cérebro relacionadas à audição.

Uma das características marcantes na música popular é a utilização ostensiva de um discurso rítmico. Este permeia a correspondência entre música e expressão corporal, como é o caso de uma coreografia ou de uma peça de teatro. Em diversas culturas, os gêneros musicais populares e folclóricos engendraram uma correspondência com a identidade cultural da comunidade social de onde surgiu. Estes não foram ditados ou proclamados por manifestos intelectuais, mas se auto-organizaram através da interação entre indivíduos pertencentes a mesma rede social. O pulso rítmico é, antes de tudo, um fenômeno acústico de percepção macro-estrutural do tempo, que comunica diretamente à emoção e, portanto, escapa ao escrutínio cognitivo e ao controle intelectual.

Como elemento mais primário da performance expressiva, o ritmo também trás outro importante aspecto do discurso musical. Em sua forma micro-estrutural, a variação da informação rítmica permite a caracterização do estilo musical de um interprete. Ao contrário do que pode-se pensar, tais variações temporais não ocorrem de modo aleatório. Estas ocorrem com determinada regularidade, o que acaba por compor uma identidade perceptual, aqui chamada de Assinatura Gestual. Este corresponde univocamente aos gestos do interprete, que compõem a prosódia a qual é executada muitas vezes com grande similaridade, normalmente pelo fato do interprete estar inconsciente de as estar regulando. Os gestos executados durante a performance expressiva do interprete carregam em si uma assinatura gestual que correspondem a sua expressividade musical. A coleta e análise de tal informação pode representar um rico material experimental para estudos etnomusicológicos, tais como: em pesquisas da variação contínua da expressividade musical pela prosódia emotiva do interprete, na correspondência do gesto com o estado emocional do interprete, na correlação da estrutura melódica e harmônica com o gesto executado, e nas distinções interpretativas relacionadas às diferenças socioculturais.

1.3. Interfaces Gestuais

Atualmente, as tecnologias computacionais permitem coletar dados do gesto artístico em tempo-real (dinâmica e continuamente, durante a performance). O gesto, cuja aquisição tem sido limitado pela informática, dada a limitada capacidade de expressão oferecida pelas tradicionais interfaces computacionais, tais como: o teclado alfanumérico do computador e o mouse, ressurgiu com a eclosão de interfaces gestuais baratas, acessíveis e precisas. Estas utilizam diferentes estratégias para coletar dados

gestuais através de sensores de movimento, tais como acelerômetros, giroscópios e câmeras de rastreamento, dados gestuais podem ser coletados e interpretados em tempo-real, por modelos computacionais (SCODITTI 2011). Interfaces gestuais comerciais, tais como o controle-remoto do videogame Wii (o wiimote), lançado em 2006, viabilizou o acesso geral para a utilização artística de tais interfaces gestuais. Utilizando seus sensores, é possível coletar dados dos gestos de interpretes; não apenas músicos, mas também de outras áreas artísticas, como: dança, teatro, desenho, artes plásticas, etc. e utilizá-los afim de viabilizar a interatividade de discursos artísticos multimodais, e sua análise, em termos de suas similaridades e distinções.

Este trabalho descreve uma pesquisa em andamento, atualmente realizada no NICS / UNICAMP, com a utilização da interfaces gestuais, tais como: wiimotes, web-cams (câmeras baratas, utilizadas em computadores) e microfones (que capturam o gesto através de seus respectivos sons) para a aquisição em tempo-real de dados do gesto musical, e o desenvolvimento de modelos computacionais de síntese sonora não-lineares, tais como modelos físicos de instrumentos de percussão, que são parametrizados em tempo-real por dados das interfaces gestuais. Tais modelos foram desenvolvidos no ambiente de programação *PureData*, ou PD (www.puredata.info). Com isso, pretendemos lançar as bases para um futuro projeto para a criação de um modelo computacional para a aquisição e análise da assinatura gestual da expressividade musical de instrumentos de percussão.

2. Coletando o Gesto Musical

Na literatura de área, existem vários exemplos de estudos, utilizando diferentes tipos de interações tecnológicas como meios para a aquisição de dados da expressividade musical. No século XVIII, por exemplo, houve o advento do primeiro instrumento musical elétrico, do qual existe registro histórico; o Denis d'Or, criado por P. Václav Divis [Tübingische 1754]. Este utilizava como sensores, arames eletrificados que sintetizavam e controlavam os parâmetros do som gerado em tempo-real, conforme visto na figura 4.

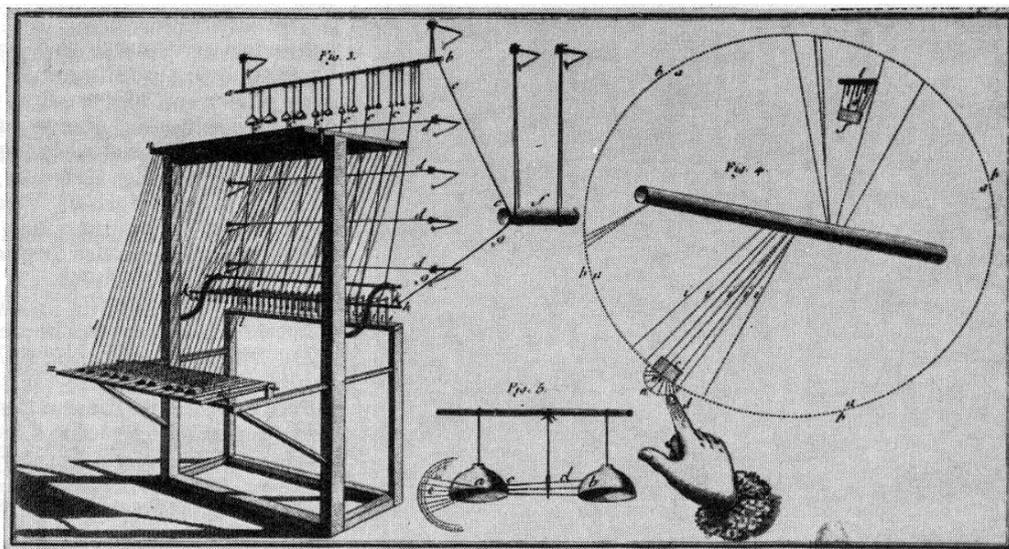


Figura 4. Diagrama do primeiro instrumento musical elétrico; o Denis d'Or.

Em 1920, houve a invenção do Theremin, de Lev S. Termen [Glinsky 2000]. Este instrumento é formado de uma antena que é usada como um sensor de proximidade, onde o músico (normalmente pela aproximação de suas mãos da antena, no entanto, sem tocá-la) controla dinamicamente a frequência e intensidade de uma onda senoidal contínua. Provavelmente inspirado neste instrumento eletrônico, Max V. Mathews

concebeu o Radio-Baton¹; também a proximidade de um dispositivo sensor capacitivo, que rastreia os movimentos das pontas de dois bastões, em um espaço tridimensional delimitado. A figura 5 apresenta o momento de uma performance antiga do Theremin



Figura 5. Momento de uma apresentação do Theremin.

Atualmente, a investigação nessa área é conhecida por diferentes nomenclaturas, tais como: Interação Tátil, Dispositivos de Entrada, ou Interfaces Gestuais. Muitas delas são apresentadas em eventos internacionais, tais como o NIME (*New Interfaces for Music Expression*) e diversos centros de pesquisa, tais como o IDMIL (*Input Devices and Music Interaction Laboratory*)² da McGill University.

Os sensores de aquisição gestual que funcionam de modo similar ao utilizado neste trabalho são: *BodyMusic*, um instrumento que permite aos músicos interagir de maneira intuitiva no controle da melodia musical, através de movimentos corporais [Horace 2005]; o sistema computacional desenvolvido por [Peng 2009] que usa um wiimote (o controle remoto do videogame Wii) como interface gestual para capturar os movimentos da regência de um maestro, em tempo real, onde a interface (wiimote) substitui a batuta.

Afim de desenvolver instrumentos de percussão virtual gestualmente controlados, buscou-se nesse trabalho uma forma de capturar os dados gestuais de modo dinâmico e transmiti-los sem fio (*wireless*) para um computador portátil (*laptop*), em tempo-real. Estes dados são utilizados para controlar dinamicamente os modelos computacionais de síntese sonora, que foram implementados em PD.

O sensor de movimento utilizado neste trabalho, é o acelerômetro tridimensional do wiimote. A razão de se optar por usar este controle remoto é o fato de que este equipamento é preciso, barato, tem comunicação dinâmica, rápida e sem-fio (*wireless*, em protocolo *bluetooth*). O wiimote é uma interface leve e portátil que, além do já mencionado acelerômetro tridimensional, vem também equipada com uma câmera infravermelha (para rastrear simultaneamente até quatro pontos de luz), diversos botões extra (ex: A, B, -, +, 1, 2, *Up*, *Down*, *Left*, *Right* e *Home*) e aceita, como acessório opcional, uma complementação (*WiiMotionPlus*) onde se encontra um giroscópio integrado (*tuning fork gyroscope*) que aumenta a capacidade de captação de movimentos de rotação, nas três coordenadas do espaço.

O acelerômetro embutido neste dispositivo é o Analog Devices ADXL330; um acelerômetro de 3 eixos lineares e não-colineares, com *motion-sensing*. Este

1 <http://emfinstitute.emf.org/exhibits/radiobaton.html>

2 <http://www.idmil.org>

equipamento tem cerca de 3g de sensibilidade, 8 bits de resolução por eixo, a uma taxa de amostragem de 100 Hz. Os dados são transmitidos através do protocolo bluetooth que usa um IC Broadcom 2042, projetado para dispositivos que estejam em conformidade com o padrão *Bluetooth Human Interface Device*, tais como: teclados e mouses sem-fio, para computadores pessoais [LEE, 2008]. Os dados do acelerômetro são transmitidos em sete canais: 3 canais para a aceleração : x, y e z; 3 canais para seus respectivos ângulos de rotação: *yaw*, *pitch* e *roll*; e um canal para a aceleração geral.

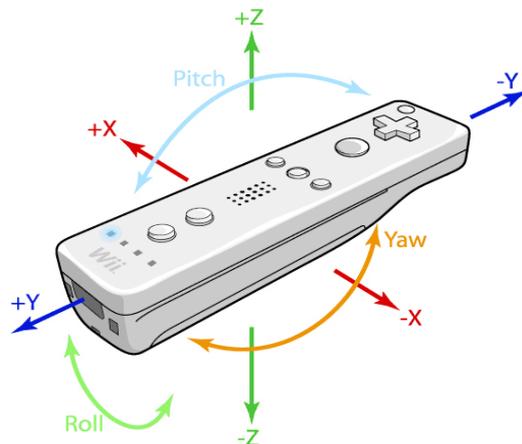


Figura 6. O controle remoto do videogame Wii; o wiimote.

Neste trabalho, os dados do wiimote transmitidos por bluetooth são coletados pelo software OSCulator³. Um programa desenvolvido para o MAC OS X, que vincula dispositivos bluetooth, como o wiimote, com diversos softwares musicais, através de redes no protocolo OSC (*Open Sound Control*)⁴. Segundo seus desenvolvedores, o OSCulator, este é capaz de conectar até 8 wiimotes simultaneamente. Neste experimento, testamos com sucesso a conexão com até 4 wiimotes, cada um controlando independentemente um modelo de percussão virtual distinto.

3. Modelos de Percussão Virtual

Foram desenvolvidos três modelos computacionais de síntese sonora em tempo real. Estes foram inspirados em instrumentos de percussão bastante conhecidos e utilizados na música popular brasileira: o Agogô, o Chocalho e o Tambor. O agogô e o chocalho são idiofones, sendo o chocalho um idiofone de agitação; enquanto que o tambor é um membranofone. Diferente do tambor e do chocalho, o agogô apresenta uma altura definida (*pitch*), o que é dado pela presença de um reduzido e uniformemente distribuído número de parciais em seu espectro sonoro, que compõem a parte determinística do som e se destacam em termos de amplitude, quando comparados ao restante de parciais que compõem a porção estocástica do sinal sonoro deste instrumento. Assim, é possível reconhecer auditivamente uma nota da escala musical do som gerado pelo agogô, enquanto que não é tão possível fazer o mesmo para o som gerado pelo tambor, e ainda menos para o chocalho. Estes últimos geram sons primordialmente estocásticos, enquanto que o o agogô gera sons primariamente determinísticos [SERRA, 1989].

O som gerado pelo modelo computacional e controlado pelos dados gestuais deve ser sintetizado de forma contínua e em tempo-real, sem qualquer atraso perceptual, de modo que o músico se sinta confortável tocando tal instrumento virtual, onde o som é percebido como sendo imediatamente sintetizado e controlado pelos gestos do músico.

Utilizou-se o ambiente de programação de *PureData*, também conhecido por suas

3 <http://www.osculator.net/>

4 <http://opensoundcontrol.org/>

inicias, PD, para desenvolver tais modelos computacionais. PD é um ambiente de programação visual, multi-plataforma e de código-livre, criada especialmente para o processamento e controle em tempo real de dados midiáticos, como o som, a imagem e seus respectivos controles [Puckette, 1997]. Semelhante ao MAX/MSP, os algoritmos em PD também são programados pela conexão de blocos, dentre as categorias seguintes: *object*, *message*, *number*, *symbol*, *bang*, *toggle*, *array*, dentre outras. Tais estruturas interligadas são chamadas de "patches" e assemelham-se a um diagrama de fluxo de dados, no qual ocorre o processamento em tempo-real dos "dados puros", ou seja, dados que podem independentemente ser de áudio, imagem, vídeo, MIDI, texto de controle, etc. Existem duas versões do PD: 1) a versão oficial, chamada de *vanilla*, que é mantida por seu criador – Miller Puckette, e a versão estendida, a *PD-extended*, que é a versão oficial agregada com diversas bibliotecas de objetos que não são encontrados na versão oficial. Esta versão é desenvolvida por uma comunidade de programadores voluntários, onde várias expansões são constantemente testadas e eventualmente incluídas na versão mais recente do *PD-extended*. Usou-se aqui esta versão, para a criação dos instrumentos de percussão virtual aqui mencionados.

Entre outras inclusões, o *PD-extended* possui objetos para receber e transmitir dados OSC (*Open Sound Control*). O objeto "*dumpOSC*" recebe este protocolo a partir de uma porta UDP previamente especificada, e expõe todos os dados recebidos nesta porta em uma única saída. Nós usamos este objeto para receber os dados vindos do *OSCulator*; contendo as informações gestuais recuperadas pelo acelerômetro do wiimote, que controla os modelos computacionais das percussões virtuais.

Cada modelo é escrito na forma de um patch do PD usando um método de síntese sonora para gerar diferentes sonoridades de percussão musical. Na implementação dos instrumentos percussivos virtuais não existem amostras de áudio previamente gravadas. Cada som é gerado em tempo real, a partir de equações que simulam comportamento físico dos instrumentos musicais acústicos reais, em que foram inspirados. Desta forma, estes modelos também podem facilmente permitir que o usuário explore técnicas sonoras estendidas que seriam impossíveis de serem alcançadas por instrumentos acústicos reais, dado o limite natural de seus atributos físicos.

Estas percussões virtuais têm algumas características em comum, em relação à ordenação de seu controle pelo gesto do músico. O ângulo rotacional (o *pitch* rotacional), por exemplo, foi utilizado para controlar a altura musical (o *pitch* musical). O ângulo de rotação "*roll*" foi utilizado para o controle da localização espacial do som, através da parametrização das funções de localização ITD ILD [Blauert 1983]. O botão "A" foi usado para permitir a mudança da frequência sonora, através dos dados do ângulo de rotação "pitch". O botão "B", foi usado para interromper o decaimento do envelope de amplitude de som sintetizado.

4. Análise e Identificação do Gesto

Os gestos dos músicos são capturados pela interface gestual, no caso, o wiimote, e [arametrizam a síntese do som correspondente à percussão, a qual foram inspirados. Tudo isso ocorre dentro de um intervalo de tempo menor que aquele percebido pela audição, o que corresponde a dizer que durante a performance os instrumentos virtuais aparentam gerar seus respectivos sons imediatamente. Isto faz com que tais instrumentos sejam "sentidos" pelos músicos executantes como se fossem instrumentos musicais acústicos. Os três modelos desenvolvidos permitem que o gesto seja capturado durante uma performance, de uma maneira lúdica e intuitiva, onde os dados destes gestos podem ser gravados pelo modelo computacional em tempo-real. O primeiro modelo testado foi aquele inspirado no "agogô". Este se baseia num método de síntese aditiva, e utiliza 11 osciladores senoidais para criar um número equivalente de parciais sonoros. Este modelo foi inspirado num método e síntese aditiva para emulação de sons

de sinos [RISSET, 1969]. Adaptando este modelo de sinos para o agogô, pudemos criar um modelo de síntese sonora que gera sons similares a este instrumento, uma vez que o agogô e o sino possuem similaridades físicas, na geração de seus respectivos sons.

O segundo modelo de percussão foi inspirado no membranofone “tambor”. Para isso, utilizamos um método computacional de síntese subtrativa, contendo um gerador de ruído-branco dinamicamente processado por um filtro de ressonância. Em termos de sinais de áudio, o ruído-branco equivale ao sinal sonoro que apresenta um espectro de frequência contínuo. Isto corresponde ao som que possui todos os parciais audíveis. O som do ruído-branco é similar ao som gerado por uma quantidade imensa de parciais independentes porém similares, como o som gerado pelas ondas do mar (formada pela imensa quantidade do som de bolhas de ar estourando), pela chuva (formado pela imensa quantidade de sons de pingos de chuva colidindo com superfícies diversas), ou mesmo o chiado emitido por um rádio ou televisor fora de sintonia (formado pelo som da imensa quantidade de perturbações eletromagnéticas). No modelo do tambor, o ruído-branco é filtrado por um filtro de ressonância. Este é aqui utilizado por apresentar um comportamento variável ao longo do tempo, como é o caso dos instrumentos membranofones reais (acústicos), onde existe uma rápida variação de estados entre o momento de impacto - do golpe na pele - até a dissipação da energia sonora, que ocorre normalmente em poucos centésimos de segundo.

O terceiro modelo é o do “chocalho”. Este é um instrumento de percussão da família dos idiofones, como o agogô, porém, seu som não é primordialmente determinístico (tonal), mas estocástico (sem clareza tonal, ou seja, ruidoso). Por isso, o modelo computacional utilizado para emular o chocalho é também o método de síntese subtrativa sobre o ruído-branco, porém, sem a necessidade de utilização de um filtro de ressonância, mas apenas um filtro passa-faixa. Através da parametrização da largura e região do espectro da faixa de passagem deste filtro, pode-se ajustar o timbre deste instrumento, de modo a corresponder mais adequadamente ao propósito de emular o som gerado por um chocalho acústico. A figura 7 mostra o patch desenvolvido em PD para a aquisição dinâmica dos dados coletados pelo wiimote (apresentados na tabela abaixo) ao mesmo tempo em que estes parâmetros controlam a síntese sonora subtrativa, que gera em tempo-real o som deste chocalho virtual.

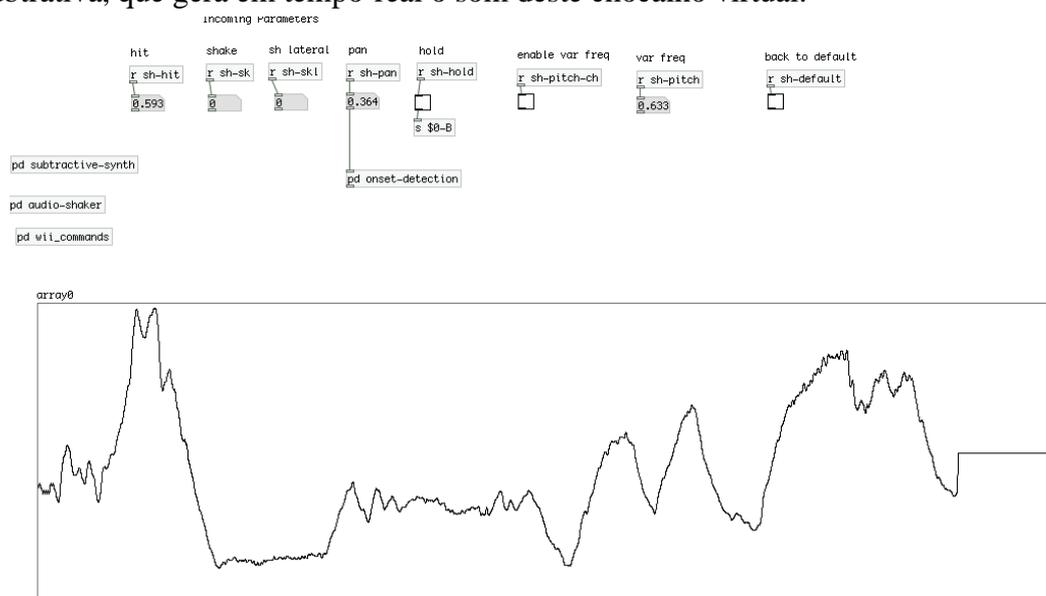


Figura 6. Modelo de aquisição de dados gestuais e síntese sonora em PD.

5. Conclusões

Neste artigo descrevemos um método de análise musical de performances expressivas através da coleta de informação musical relacionada à assinatura gestual do intérprete. Ao invés de utilizarmos descritores acústicos e simbólicos, como aqueles desenvolvidos pelos diversos grupos de MIR (*Music Information Retrieval*), que desenvolvem modelos computacionais para analisar a informação musical diretamente de arquivos de áudio (dados acústicos), ou da notação musical (dados simbólicos); o método aqui apresentado analisa a informação musical a partir da assinatura gestual capturada através do gesto musical do intérprete durante uma performance expressiva. Isto permite que a informação de tempo, da ação gestual relacionada à emissão de um evento musical - o *onset* - seja coletado de maneira bastante eficiente e precisa, mesmo que o instrumento seja tocado de modo lúdico e intuitivo. O instrumento que apresentamos aqui para realizar a coleta de dados gestuais foi o chocalho. Como pode-se observar na tabela da figura 7, os dados gestuais são coletados de modo minucioso. Os menores gestos são percebidos pelo sensor utilizado (o acelerômetro integrante do wiimote) e possíveis de serem facilmente dispostos de modo gráfico. Com isto é possível posteriormente analisar estas séries temporais afim de detectar padrões de similaridade, onde é expressada a assinatura gestual do intérprete.

Apresentamos aqui uma pesquisa ainda se iniciando, realizada com instrumentos virtuais de percussão, e que, até o momento, trata apenas da coleta dinâmica de *onsets* de performances expressivas. Em pesquisas futuras, pretendemos explorar a coleta de informação de instrumentos virtuais melódicos e harmônicos, onde a assinatura gestual pode ser coletada não somente destes *onsets*, como também de outros movimentos involuntários, como o deslocamento do corpo do instrumento musical, o rastreamento do movimento dos dedos do músico, de seus deslocamentos corporais extra-musicais, como sua cabeça, ombros, bem como de expressões faciais.

Com isto, a pesquisa nessa área permitirá a criação de modelos computacionais de emulação gestual que possam conter a assinatura gestual de intérpretes proeminentes; bem como o futuro desenvolvimento de modelos computacionais mais avançados, para a descrição micro-gestual da interpretação automática e autônoma de algoritmos sequenciadores adaptativos, que interpretem a notação musical de forma mais humana, inclusive com a possibilidade de interpretação melódica de instrumentos solistas e cantores. Com isto, tais padrões de assinaturas gestuais poderão ser armazenados em bancos de dados que permitam a re-interpretação de notação musical de acordo com o estilo musical de um intérprete humano.

Alguns vídeos da execução dos modelos computacionais de percussão virtual controlados por dados gestuais, conforme aqui apresentados, podem ser vistos no link: www.youtube.com/nicsunicamp.

Referencias Bibliográficas

- BLAUERT, J.: Spatial hearing - the psychophysics of human sound localization; MIT Press; Cambridge, Massachusetts. 1983.
- GLINSKY, A., 2000. Theremin: Ether Music and Espionage. Urbana, Illinois: University of Illinois Press.
- HORACE, H. S., Kwong, B., Law, K. C. K., 2005. Bodymusic: a novel framework design for body-driven music composition. In ACE'05: Proceedings of the 2005 ACM SIGCHI. International Conference on Advances in Computer Entertainment Technology. ACM Press. pages 342-345, New York, NY, USA, 2005.
- KELLER, Damián, Burdasz R. (organizadores). Criação musical e tecnologias: teoria e prática interdisciplinar. Livro da Anppom. Série Pesquisa em Música no Brasil. ISBN: 978-85-63046-01-7. Goiânia: Anppom, 2010.
- LEE, Jonh. Hacking the nintendo wii remote. Pervasive Computing, IEEE, p. 39-45. 2008.
- MATHIAK, Klaus, Ingo Hertrich, Wolfgang Grodd, and Hermann Ackermann. Cerebellum and Speech Perception: A Functional Magnetic Resonance Imaging Study. Journal of Cognitive Neuroscience, August 15, 2002, Vol. 14, No. 6 , Pages 902-912.
- PENG, L., Gerhard, D., 2009. A Wii-based gestural interface for computer-based conducting systems. International Conference on New Interfaces for Musical Expression.
- PUCKETTE, M. S., 1997. Pure data. In: Proceedings of the International Computer Music Conference, pp. 224–227. International Computer Music Association.
- RISSET, J. C., An Introductory Catalog of Computer- synthesized Sounds, Murray Hill, New Jersey: Bell Laboratories, 1969.
- SCODITTI, A., et al., A Novel Taxonomy for Gestural Interaction Techniques Based on Accelerometers. Proceedings of the 2011 conference on Intelligent User Interfaces. 2011.
- SERRA, X. A System for Sound Analysis/Transformation/Synthesis based on a Deterministic plus Stochastic Decomposition. Ph.D. Thesis. Stanford University. 1989.
- TODD, N., and F. Cody. Vestibular Responses to Loud Dance Music: A Physiological Basis for the 'Rock and Roll threshold'? Journal of the Acoustical Society of America 107:1, pp. 496–500. 2000.
- TUBINGISCHE, 1754. Berichte von gelehrten Sachen, XXX, p. 395.