



PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

EDUARDO SANTOS SILVA

Integrando Sensores para a Criação de Instrumentos Musicais Digitais



Universidade Federal de Pernambuco

posgraduacao@cin.ufpe.br

www.cin.ufpe.br/~posgraduacao

Recife
2015

EDUARDO SANTOS SILVA

**Integrando Sensores para a Criação de Instrumentos Musicais
Digitais**

Dissertação apresentada como requisito parcial para a obtenção do título de Mestre, pelo Programa de Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco.

Orientador: Geber Lisboa Ramalho, PhD.

Coorientador: Giordano Ribeiro Eulalio Cabral, PhD.

Recife
2015

Catálogo na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S586i Silva, Eduardo Santos da
Integrando sensores para a criação de instrumentos musicais digitais /
Eduardo Santos da Silva. – 2015.
134 f.:il., fig., tab.

Orientador: Geber Lisboa Ramalho.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CIn,
Ciência da Computação, Recife, 2015.
Inclui referências e apêndice.

1. Inteligência artificial. 2. Computação musical. I. Ramalho, Geber Lisboa
(orientador). II. Título.

006.3

CDD (23. ed.)

UFPE- MEI 2017-182

Dissertação de Mestrado apresentada por **Eduardo Santos Silva** à Pós-Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco, sob o título “**Integrando Sensores para a Criação de Instrumentos Musicais Digitais**” orientada pelo **Prof. Geber Lisboa Ramalho** e aprovada pela Banca Examinadora formada pelos professores:

Profa. Verônica Teichrieb
Centro de Informática / UFPE

Prof. João Paulo Cerquinho Cajueiro
Departamento de Engenharia Mecânica/ UFPE

Prof. Geber Lisboa Ramalho
Centro de Informática / UFPE

Visto e permitida a impressão.
Recife, 21 de agosto de 2015.

Profa. Edna Natividade da Silva Barros
Vice-Coordenador da Pós-Graduação em Ciência da Computação do
Centro de Informática da Universidade Federal de Pernambuco.

Aos meus pais.

Agradecimentos

Aos meus pais, por todo o apoio e por me incentivarem a correr atrás do que é importante para mim.

Às minhas irmãs Carolina e Júlia e ao meu irmão Renato pelas conversas, brigas, trelas e cumplicidades.

À minha família, por todo o apoio.

À Geber, por todas as discussões, pelas anotações incompreensíveis e por incentivar a realização do projeto.

À Giordano que sempre possui um olhar crítico e comentários pertinentes ao tema, e por buscar organizar as ideias malucas e discussões dentro de um método científico rigoroso.

À Filipe Calegário, João Tragtenberg e Jerônimo Barbosa por toda a ajuda durante o projeto, tirando dúvidas sobre programação, dúvidas sobre sensores, ou compartilhando informações pertinentes ao trabalho.

À Hugo Bessa que me ajudou bastante com a implementação do sistema nas etapas decisivas.

À Paulo Sérgio Nunes pelas conversas e pela ajuda na construção de um sistema físico para a versão final. Uma grande fonte de conhecimento sobre materiais.

À Catarina Apolônio pela ajuda na avaliação do artefato e pelos relevantes comentários sobre a qualidade do áudio do sistema.

À Veronica Teichrieb, pelo apoio no começo da pesquisa, pela disponibilidade quando foi preciso tirar dúvidas e por disponibilizar o uso do Leap Motion© no início do projeto, quando o dispositivo ainda não havia sido lançado.

À Lucas Figueiredo, do Voxar Labs, por toda a ajuda na tentativa de resolver os problemas encontrados e pela disponibilidade na discussão de possíveis soluções.

Ao meu primo Pedro Menezes pela ajuda na gravação de sons e por todo o incentivo.

À Jáder Abreu pela implementação do primeiro protótipo com o Leap Motion© e ter feito a gente tocar a música tema de “Missão Impossível” em um pedaço de vidro!

À Janiel Almeida pela parceria, pelas conversas sobre jogos e pelo apoio ao projeto.

Ao pessoal do grupo de pesquisa MustIC: Luca, João Paulo, Rodrigo, Sofia, Braz, Mariane, Scholz, Pedro Aléssio e Ricardo Brasileiro, por todas as contribuições.

À Anaís Simões, Pedro Reis, Danilo Coêlho, por todas as noites de filme e todas as saídas.

À Renato Gusmão, Cláudia Matos, Raíssa Simões e Pedro Sette, pelos eventos de Superbowl, Champions League e pelas noites mexicanas.

Ao pessoal do Nerd Soccer por sempre proporcionarem momentos épicos. Francamente...

Ao grupo de músicos que participaram do questionário online.

Resumo

Nos últimos anos, novas tecnologias de interface gestual emergiram com os lançamentos de eletrônicos como smartphones, tablets e controladores de videogames como o Nintendo Wii Remote® e o Microsoft Kinect®. O surgimento dessas novas interfaces proporcionou o crescimento em pesquisas na área de Interface Humano-Computador (HCI). Particularmente na área de Computação Musical, essas tecnologias oferecem a possibilidade de criação de novos Instrumentos Musicais Digitais (DMIs). Entretanto, apesar do grande número de demonstrações de DMIs (presentes em sites de compartilhamento de vídeos), os resultados da avaliação formal destas tecnologias muitas vezes não condizem com o que é apresentado ao público. Diversos trabalhos acadêmicos relatam grandes restrições de algumas dessas tecnologias de interface gestual, entre eles o reconhecimento de nuances gestuais, latência no sistema e ausência de *feedback* háptico. Estes são obstáculos na construção de um DMI que possa ser utilizado pelo público profissional. Neste contexto, este projeto busca avaliar novas tecnologias de interface gestual para analisar a capacidade de capturar as nuances gestuais visando aprimorar a expressividade musical destes instrumentos e solucionar problemas encontrados em pesquisas anteriores. Em particular, exploramos a ideia de integração de sensores como forma de proporcionar a captura de nuances expressivas a um baixo custo de desenvolvimento. Para testar e aprimorar esta abordagem, foi desenvolvido um DMI que busca simular um instrumento de percussão que integra 3 sensores: Leap Motion®, sensor piezoelétrico e um sensor de pressão FSR. Tal DMI foi desenvolvido seguindo a filosofia de design iterativo que envolve as etapas de inspiração, investigação, ideação, prototipação e avaliação do artefato. Também seguimos os princípios do design centrado no usuário que procura solução para os problemas de uma determinada área a partir do conhecimento das pessoas que estão inseridas naquele contexto. Os protótipos desenvolvidos foram avaliados de forma quantitativa, investigando a capacidade do sistema de solucionar os problemas encontrados em trabalhos anteriores. Além disso, o instrumento foi apresentado a um músico profissional, de forma a coletar informações complementares sobre a experiência do usuário. Os resultados mostram que a abordagem de integração de múltiplos sensores apresenta um grande potencial (com baixo custo de implementação) para reconhecimento de nuances de expressividade e apresentando baixa latência, que podem contribuir no desenvolvimento de DMIs expressivos. Todavia, novos avanços nessa área ainda podem contribuir para a resolução de todos os problemas levantados durante a análise da literatura.

Palavras-Chave: Interfaces Gestuais. Instrumentos Musicais Digitais. Avaliação de Tecnologias. Integração de Múltiplos Sensores. Expressividade Musical.

Abstract

In recent years, new gestural interface technologies emerged after the release of consumer electronics such as smartphones, tablets and videogame controllers like the Nintendo Wii Remote[®] and Microsoft Kinect[®]. The rise of these new interfaces provided the growth in the Human-Computer Interaction (HCI) research field. Particularly in the Computer Music area, those technologies offer the possibility of creating new Digital Music Instruments (DMIs). However, despite the wide number of DMI demonstrations (available at video sharing websites), the findings from the formal evaluation of these technologies are not consistent with what is presented to the public. Several academic research describe the limitations of some of these gestural interface Technologies, including the tracking of gestural nuances, system latency and the lack of haptic feedback which become obstacles in developing a DMI for professional musicians. In that context, this project aims to evaluate new gestural interface technologies to investigate their capabilities to capture gestural nuances in order to enhance their musical expressivity and solve those problems found in previous research. Particularly, we explore the idea of multisensor integration as an approach to achieve gestural nuance tracking at low implementation cost. To test and enhance this approach, a DMI that simulates an acoustic percussion instrument was implemented integrating 3 sensors: the Leap Motion Controller[®], a piezoelectric element and a FSR pressure sensor. This DMI was implemented following a philosophy of iterative design which involves the steps of inspiration, investigation, ideation, prototyping and evaluation of the artifact. It was also used a user-centered design approach which follows the principle that the solution of the problems of a certain area emerges from the knowledge of the people inserted in that context. The developed prototypes were evaluated following a quantitative approach, to investigate the system capabilities to solve problems found in previous research. Moreover, the instrument was presented to an professional musician to gather complementary information on the user's experience. The findings show that the multisensor integration approach presents great potential (with low implementation cost) to capture nuances of expression and providing low latency, that can contribute to the development of expressive DMIs. However, new advances in this field of research can contribute to solve the problems found during the literature review.

Keywords: Gestural Interfaces. Digital Music Instruments. Technology Evaluation. Multisensor Integration. Musical Expressivity.

Lista de Figuras

Figura 2.1: Eletronic Numerical Integrator and Computer (ENIAC).	23
Figura 2.2: Xerox <i>Alto</i> (direita) e Xerox <i>Star</i> (esquerda).	24
Figura 2.3: Wii <i>Remote</i> (esquerda) e barra de captura (direita).	26
Figura 2.4: Steve Jobs apresentando o iPad.	27
Figura 2.5: Sensor Kinect da Microsoft.	28
Figura 2.6: Sensor Leap Motion.	29
Figura 2.7: Microsoft HoloLens.	30
Figura 2.8: Sintetizador Moog.	32
Figura 2.9: Controladores MIDI com diferentes interfaces.	33
Figura 2.10: Modelo de DMI.	34
Figura 2.11: ACPAD.	36
Figura 2.12: DMIs inspirados em instrumentos acústicos.	36
Figura 2.13: ReacTable.	37
Figura 4.1: o vBow (Extraído de Nichols, 2002).	56
Figura 4.2: MudPad (Extraído de Jansen, Karrer e Borchers, 2010).	57
Figura 4.3: MAX/MSP (em cima) e puredata (em baixo).	60
Figura 4.4: Matriz de coocorrência de sensores no NIME (Extraído de MEDEIROS; WANDERLEY, 2014a).	65
Figura 5.1: Microcontrolador arduino UNO.	71
Figura 5.3: Open Tone.	78
Figura 5.4: Muffled Tone.	78
Figura 5.5: Slap Tone.	79
Figura 5.6: Closed Slap.	79
Figura 5.7: Pressed Slap.	79
Figura 5.8: Bass Tone.	80
Figura 5.9: Heel Tone.	80
Figura 5.10: Tip Tone.	80
Figura 5.11: Gliss Tone.	81
Figura 6.1: Campo de Visão do Leap Motion SDK (v1.0).	83
Figura 6.2: Arquitetura do frame de reconhecimento.	85
Figura 6.3: Visualizador do Leap Motion (SDK v2.0).	86
Figura 6.4: Sensor piezoelétrico.	86
Figura 6.5: Esquema de conexão do piezo com o arduino.	87
Figura 6.6: Esquema de conexão do FSR com o arduino.	87
Figura 6.8: Diferentes microcontroladores arduino.	89
Figura 6.9: Ambientes de programação do Arduino (à esquerda) e Processing (à direita).	90
Figura 6.10: Interface do Ableton Live 9.	91
Figura 6.11: Drum Rack do Ableton Live (Extraído de http://en.wikiaudio.org/Ableton_Live:Drum_rack).	92
Figura 6.12: Configuração do mapeamento MIDI de parâmetros sonoros.	93

Figura 6.13: Folha de vidro com marcações.	95
Figura 6.14: Perspectiva do usuário.	95
Figura 6.15: Configuração do protótipo versão beta 1.	101
Figura 7.1: Formato de onda de percussão.....	106
Figura 7.2: Ritmo 1 – Four Stroke Ruff.....	113
Figura 7.3: Ritmo 2 – Flam Tap.....	113
Figura 7.4: Ritmo 3 – Heel-Tip Technique.....	113
Figura 7.5: Ritmo 4 – Tumbao.....	113
Figura 7.6: Cajón virtual (esquerda) e cajón acústico (esquerda).	115
Figura 7.7: Mapa mental da entrevista.	118

Lista de Tabelas

Tabela 1: Capacidade de rastreamento dos sensores Wii Remote e Kinect (Extraído de SILVA, 2012).....	40
Tabela 2: Sensores que aparecem em artigos da conferência NIME (Extraído de MEDEIROS; WANDERLEY, 2014).	64
Tabela 3: Movimentos capturados pelas interfaces.	84
Tabela 5: Reconhecimento dos gestos da conga com apenas uma mão.	110
Tabela 6: Reconhecimento dos gestos da conga com ambas as mãos.....	110

Sumário

1	INTRODUÇÃO	16
1.1	Motivação	16
1.2	Objetivos	18
1.3	Abordagem e método	19
1.4	Estrutura da Dissertação	20
2	EVOLUÇÃO DA INTERFACE HOMEM-MÁQUINA NAS TAREFAS COTIDIANAS E NA MÚSICA	22
2.1	Avanços Tecnológicos em HCI	22
2.1.1	<i>Primeiros computadores</i>	23
2.1.2	<i>O conceito do Desktop</i>	24
2.1.3	<i>Interfaces Gráficas e a implementação do modelo WIMP</i>	24
2.1.4	<i>Surgimento das Interfaces gestuais</i>	25
2.2	Novas formas de interação musical	30
2.2.1	<i>Descoberta da eletricidade e um grande salto tecnológico</i>	31
2.2.2	<i>Os avanços com o protocolo MIDI</i>	33
2.2.3	<i>O modelo de Instrumentos Musicais Digitais</i>	34
2.2.4	<i>Tipos de DMIs</i>	35
2.3	Conclusões	38
3	DESAFIOS NA CRIAÇÃO DE DMIS	39
3.1	Apresentando DMIs ao público profissional	39
3.2	Questionário sobre as tecnologias de percussão digital existentes	41
3.3	Problemas do mapeamento	45
3.4	Problema do entendimento e engajamento	47
3.5	Latência	48
3.6	Feedback háptico e relação física entre o músico e o instrumento	49
3.7	Expressividade do instrumento e virtuosidade do performer	50
3.8	Contextos de uso	53
3.9	Escopo da pesquisa	53
4	ESTADO DA ARTE	55

4.1	Latência	55
4.2	Feedback Háptico	55
4.3	Expressividade musical	57
4.3.1	<i>Protocolos de comunicação</i>	58
4.3.2	<i>Softwares de mapeamento</i>	59
4.3.3	<i>Síntese sonora</i>	61
4.3.4	<i>Sensores</i>	63
5	ABORDAGEM ADOTADA	68
5.1	Princípios utilizados	68
5.1.1	<i>Integração de múltiplos sensores</i>	68
5.1.2	<i>Baixo custo de implementação</i>	70
5.1.3	<i>Princípios de design</i>	71
5.1.4	<i>Conexão com tecnologias já existentes</i>	72
5.1.5	<i>Mobilidade</i>	73
5.2	Métodos de design adotados	73
5.2.1	<i>Método de inspiração</i>	73
5.2.2	<i>Método de ideação</i>	74
5.2.3	<i>Métodos de investigação</i>	74
5.2.4	<i>Método de prototipação</i>	75
5.3	Escolha do instrumento	75
5.4	Método de aquisição dos samples	81
6	CONSTRUÇÃO DO PROTÓTIPO	82
6.1	Sensores Utilizados	82
6.1.1	<i>Leap Motion®</i>	82
6.1.2	<i>Sensor piezoelétrico</i>	86
6.1.3	<i>Sensor de pressão</i>	87
6.2	Tecnologias utilizadas	88
6.2.1	<i>Integração dos sensores</i>	88
6.2.2	<i>Arduino</i>	88
6.2.3	<i>Ableton Live</i>	90
6.3	Protótipo preliminar com o Leap Motion®: Crystal Piano	93
6.3.1	<i>Tecnologias utilizadas</i>	94
6.3.2	<i>Configuração</i>	94

6.3.3	<i>Discussões</i>	96
6.4	Versão alfa 1	96
6.4.1	<i>Tecnologias utilizadas</i>	96
6.4.2	<i>Configuração</i>	97
6.4.3	<i>Discussões</i>	97
6.5	Versão alfa 2	98
6.5.1	<i>Tecnologias utilizadas</i>	98
6.5.2	<i>Configuração</i>	98
6.5.3	<i>Discussões</i>	99
6.6	Versão beta 1	99
6.6.1	<i>Tecnologias Utilizadas</i>	99
6.6.2	<i>Configuração</i>	100
6.6.3	<i>Discussões</i>	102
6.7	Versão final	102
6.7.1	<i>Tecnologias Utilizadas</i>	103
6.7.2	<i>Configuração</i>	103
6.7.3	<i>Discussões</i>	103
7	RESULTADOS OBTIDOS	105
7.1	Métodos de avaliação	105
7.1.1	<i>Método de avaliação da latência</i>	105
7.1.2	<i>Método de avaliação do reconhecimento de gestos</i>	106
7.1.3	<i>Método de avaliação qualitativa</i>	107
7.2	Resultados da avaliação da latência	108
7.3	Reconhecimento de gestos	108
7.4	Resultados qualitativos	111
7.4.1	<i>Protocolo Experimental</i>	112
7.4.2	<i>Entrevista</i>	114
7.4.3	<i>Reconstrução do mundo do usuário e observações durante o experimento</i>	114
7.4.4	<i>Codificação e Mapa mental</i>	116
7.4.5	<i>Mapa mental</i>	118
7.4.6	<i>História Central</i>	119
8	CONCLUSÕES E TRABALHOS FUTUROS	121
8.1	Lições aprendidas com o Leap Motion©	122

8.1.1	<i>Vantagens</i>	122
8.1.2	<i>Desvantagens</i>	123
8.1.3	<i>Considerações Finais</i>	123
	Referências	125
	Apêndice A – Modelo do questionário online	133

1 INTRODUÇÃO

Nos últimos anos, diversos dispositivos de interface gestual como smartphones, tablets, smart TVs, controladores de videogame (como o *Nintendo Wii Remote*¹ e o *Microsoft Kinect*²), entre outros, despertam a atenção de pesquisadores da área de Interação Humano-Computador (HCI) e o público geral, com a possibilidade de se interagir com computadores de maneiras diferentes, rompendo com os paradigmas de interação consolidados ao longo do tempo nas interfaces compostas por monitores, teclados, mouses e o modelo WIMP (*Windows, Menus, Icons and Pointable Devices*) e possibilitando aplicações que não haviam sido imaginadas. O uso de interfaces gestuais vem ganhando espaço nas pesquisas acadêmicas, tendo em vista que interação pode beneficiar diversos campos de aplicação que abrange artes, auxílio na área de medicina, simulações, etc. (FREITAS et al., 2012; SINGER; LARKE; BIANCIARDI, 2003; WONG; YUEN; CHOY, 2008).

Neste contexto, a área de computação musical tem testemunhado um crescente número de novos instrumentos musicais digitais (DMIs, do inglês *Digital Musical Instruments*). Estes foram imensamente beneficiados por estas tecnologias de interfaces emergentes (JORDA et al., 2005; MILLER; HAMMOND, 2010; MONTAG et al., 2011).

1.1 Motivação

Diversos destes DMIs podem ser vistos por meio de documentação informal disponibilizada *online* em sites de compartilhamento de vídeos como o Youtube³ e Vimeo³. DMIs como o Kin Hackt⁴ e o Kinectar⁵, integram a capacidade de reconhecimento do sensor Kinect com o software Live⁶, da empresa Ableton, para composição musical e utilizam gestos como parâmetros de entrada. Diversos outros projetos, como o V motion project⁷ desenvolveram artefatos interativos musicais utilizando um ou mais sensores Kinect para aprimorar a precisão e reduzir o atraso sonoro do instrumento, ao mesmo tempo em que também fornecem *feedback* visual, em tempo real, para fins artísticos.

¹ Wii Remote é um controle de movimentos do console Nintendo® Wii, desenvolvido pela Nintendo® Company, Limited. Website do Nintendo® Wii: <<http://www.nintendo.com/wii>>.

² Microsoft® Kinect é um sensor de profundidade desenvolvido pela Microsoft® Corporation, inicialmente para o videogame Microsoft® Xbox 360 para servir como um controle de movimentos, porém adaptado para funcionar com os computadores Microsoft® Windows computers. Websites do Microsoft® Kinect: <http://www.xbox.com/en-US/kinect> and <<http://www.microsoft.com/en-us/kinectforwindows/>>.

³ Youtube e Vimeo são websites de compartilhamento de vídeos que permitem seus usuários divulgar, visualizar e compartilhar vídeos.

⁴ Vídeo disponível em: <<http://www.youtube.com/watch?v=YW86yyz0gj0>>.

⁵ Vídeo disponível em: <<http://www.youtube.com/watch?v=qXnLxi2nZrY>>.

⁶ Live é um software de composição musical criado pela empresa Ableton. Website do Ableton Live: <<https://www.ableton.com/en/live/>>.

⁷ O video do V motion project está disponível em: <<http://vimeo.com/45417241>>.

Novos instrumentos têm sido explorados utilizando diferentes controles de movimento de *videogame* como o Wii Remote (MILLER; HAMMOND, 2010; SILVA, 2012), projetos utilizando tablets multi-toque como o iPad da Apple (CALEGARIO, 2013), mesas multi-toque (JORDÀ et al., 2005) ou desenvolvendo seus próprios DMIs com interfaces inovadoras utilizando os mais diversos sensores, como o projeto “Hands”⁸ de Mike Waisvisz’s (na década de 1980).

Além disso, diversos pesquisadores abordam a avaliação de novas tecnologias e DMIs. Estudos avaliando a precisão da *performance* em propriedades musicais como *tempo*, latência em instrumentos construídos apenas com o controle *Wii Remote* (KIEFER; COLLINS; FITZPATRICK, 2008), com a barra de sensor do console Wii (PENG; GERHARD, 2009) e com diversas plataformas multi-toque disponíveis naquele momento (MONTAG et al., 2011). Ademais, outros analisaram os movimentos realizados em um piano utilizando sensores inerciais conectados ao usuário (HADJAKOS; MÜHLHÄUSER, 2010).

É importante destacar que pelo fato de instrumentos musicais estarem inseridos num contexto, o qual demanda altamente de precisão temporal, resposta ao usuário e que, para se tocar os diversos instrumentos existentes é necessária uma alta precisão e expertise nas técnicas gestuais utilizadas, diversas propriedades essenciais das tecnologias gestuais podem ser testadas ao se avaliar DMIs.

Apesar da crescente exploração nesta área e o grande número de demonstrações *online*, a maioria destes novos DMIs e destas novas interfaces não são utilizadas por músicos profissionais. Muitas das tecnologias utilizadas para construção destes DMIs, quando avaliadas rigorosamente ou quando apresentadas ao público profissional, não apresentam a mesma qualidade que prometem e, pelo contrário, possuem diversos problemas essenciais para a utilização destes instrumentos (SILVA, 2012), levantando dúvidas sobre as qualidades e reais capacidades destas tecnologias emergentes. Dentre as limitações, o trabalho citado revelou latência além do desejado, falta de feedback tátil e/ou visual, ausência de controle de dinâmica do sistema desenvolvido e a carência de timbres do instrumento.

Estes problemas levantados pelos músicos profissionais estão diretamente relacionados com a capacidade de expressividade do instrumento desenvolvido (FELS; GADD; MULDER, 2002; JORDA, 2005; O’MODHRAN; CHAFE, 2000).

Como diversas vezes, nos trabalhos acadêmicos, o usuário final do instrumento é o próprio desenvolvedor, pode existir uma certa adaptação do usuário ao utilizar o sistema, o que pode não revelar suas limitações frente a músicos profissionais.

⁸ “Hands” é um Instrumento Musical Digital desenvolvido no STEIM (steim.org). Descrição disponível em: <<http://www.crackle.org/TheHands.htm>>.

Além disso, diversos trabalhos acadêmicos levantam problemas na construção de DMIs (CALEGARIO, 2013; JORDA, 2005; MEDEIROS et al., 2014; WANDERLEY; ORIO, 2002). Várias dessas pesquisas levantam limitações de sensores para o *design* de instrumentos que contradizem a qualidade das demonstrações disponíveis *online*, principalmente com o sensor Kinect© (HSU et al., 2013; ODOWICHUK et al., 2011; OKADA et al., 2014; ROSA-PUJAZÓN et al., 2013; TODOROFF; LEROY; PICARD-LIMPENS, 2011).

Ainda assim, existem empresas que desenvolvem instrumentos musicais digitais para o público profissional. Esse é o caso de controladores MIDI como teclados musicais, baterias eletrônicas, violões MIDI, interfaces para DJs e pads de percussão digital. Porém, um questionário realizado com percussionistas profissionais sobre instrumentos digitais de percussão mostra que mesmo que o uso desses instrumentos possua certas vantagens em alguns contextos de uso, os mesmos ainda possuem características que não agradam o público profissional.

Em 2012, o dispositivo Leap Motion Controller© foi revelado (Leap Motion inc., 2015). Este dispositivo de reconhecimento de gestos fornece um campo de visão de aproximadamente 150° e utiliza um sensor de profundidade para rastrear características das mãos com precisão de até um centésimo de milímetro. Um controle de alta-granularidade pode representar uma nova oportunidade para criação de uma nova geração de DMIs. Entretanto, uma avaliação relacionada com latência e precisão (entre outros problemas comuns) do sensor deve ser realizada para analisar as possibilidades de uso.

Outro fator importante na popularização destes DMIs é a construção de um sistema a um baixo custo de implementação. Isto parte da filosofia de movimentos como o *Makers* e o *Do-it-Yourself* de compartilhamento de informações, diminuindo as barreiras de desenvolvimento para iniciantes.

1.2 Objetivos

Este projeto apresenta uma avaliação de tecnologias para construção de DMIs.

Partindo dos resultados da pesquisa realizada por (SILVA, 2012) e buscando os princípios do processo de design utilizados por (CALEGARIO, 2013), este projeto busca uma nova iteração no processo de avaliação das tecnologias existentes para a construção de DMIs, visando o baixo custo de implementação computacional e a integração de múltiplos sensores, de forma a responder a seguinte pergunta de pesquisa:

“Como seria possível superar alguns desses obstáculos de construção de Instrumentos Musicais Digitais, particularmente a captura de nuances de expressão musical e a precisão temporal, utilizando tecnologias de baixo custo de desenvolvimento?”

Para isto, deve-se desenvolver uma “plataforma” (abordagem) de criação de DMIs expressivos que possa solucionar os problemas encontrados em pesquisas anteriores (SILVA, 2012) e considerar as dificuldades de construção de DMIs no design do artefato deste projeto. Este projeto foca, principalmente, os problemas de captura de nuances gestuais e precisão temporal (latência), bem como fornecer algum tipo de *feedback* visual/háptico para o músico. Assim, um DMI deve ser desenvolvido e avaliado, e tecnologias e abordagens diferentes devem ser investigadas com intuito de solucionar este problema.

Assim, buscou-se focar um instrumento de percussão. Visamos os instrumentos de percussão pela sua grande demanda de precisão espacial e temporal e buscamos saber o quão perto a interface apresentada está de reproduzir a vasta expressividade que um instrumento acústico possui. Entretanto, mesmo que este DMI apresente algumas vantagens em relação ao seu análogo digital, e.g. ser mais portátil que um kit de percussão ou ser possível tocar diversos instrumentos diferentes em um único sistema, o objetivo do projeto é construir um sistema que possua a capacidade de captura de nuances de expressividade musical, mesmo que não seja um instrumento de percussão. Ou seja, uma plataforma expressiva para construção de novos DMIs.

1.3 Abordagem e método

Para isso, buscou-se construir um novo DMI destinado ao público profissional, denominado V-Conga, utilizando essa abordagem. Este DMI utiliza uma combinação de sensores e novas tecnologias de forma a simular instrumentos de percussão já existentes e consolidados no mercado. Instrumentos de percussão foram escolhidos por serem altamente demandantes de precisão espacial e temporal (para acompanhamento musical).

A abordagem utilizada visou a integração de múltiplos sensores. Assim, procura-se utilizar as informações provindas de diferentes sensores de forma complementar, visando tirar o melhor proveito de cada um deles. Diferente da filosofia de utilização de um único sensor, e.g. uma câmera RGB convencional, que demandaria de um alto nível técnico para implementação de algoritmos complexos de processamento de imagem, a integração de múltiplos sensores pode se munir de tecnologias que seguem a filosofia DIY, e.g. o Arduino, que diminuem o custo de implementação desses sistemas.

Entretanto, apesar de possíveis vantagens que o sistema possa trazer quando comparado com os instrumentos acústicos análogos, o projeto não busca a criação de um novo produto que funcione como instrumentos já existentes. A prioridade é saber se as tecnologias emergentes podem resolver problemas encontrados na Computação Musical para aprimorar a expressividade de DMIs e como estas tecnologias podem ser usadas para isso. Além disso, esta

contribuição pode também beneficiar o campo de HCI com a avaliação de tecnologias para construção de novas interfaces.

Utilizando uma abordagem de *design* centrado no usuário, na qual partimos do princípio de que o usuário que possui a solução para o problema, o processo partiu de pesquisas já nesta área utilizando a mesma linha de avaliação (SILVA, 2012) e seguiu um ciclo envolvendo investigação, concepção, prototipação e avaliação do produto.

Este processo de *design* fez com que o método utilizado para criação do protótipo final fosse aprimorado. O mesmo incorpora múltiplos sensores, o mínimo de programação necessária e um custo relativamente baixo, para facilitar a adesão de novos pesquisadores e principalmente artistas que pretendem expandir a expressividade de suas performances.

Os resultados mostram que, apesar de certos problemas na utilização de sensores com câmera, a integração de múltiplos sensores pode contribuir para solucionar os obstáculos encontrados e são capazes de reproduzir a maioria das nuances encontradas em instrumentos acústicos, provando que esta é uma direção interessante a ser tomada na criação de novos instrumentos digitais musicais.

Entretanto, apesar de o sistema conseguir capturar a maioria das nuances gestuais necessárias na utilização do instrumento e possuir uma baixa latência, a instabilidade na robustez apresentada pelo sensor de rastreamento (Leap Motion©), neste estágio de desenvolvimento do kit de desenvolvimento de software (SDK, do inglês *Software Development Kit*), permite identificar que novos avanços na área de rastreamento devem acontecer para atender as demandas do contexto dos percussionistas profissionais.

1.4 Estrutura da Dissertação

A estrutura da dissertação é apresentada abaixo:

O capítulo 2 apresenta um breve histórico sobre como os avanços tecnológicos proporcionaram o desenvolvimento de novas interfaces de interação homem-máquina e como elas modificaram os paradigmas de interação computacional – da manipulação direta de componentes eletrônicos até as interfaces gestuais. Além disso, mostra como os avanços tecnológicos possibilitaram novas formas de interação com a música – o desenvolvimento de instrumentos eletrônicos, a mídia digital e as criações de protocolos de comunicação entre instrumento e computador e instrumentos digitais musicais (DMIs).

Os capítulos 3 e 4 apresentam informações obtidas a partir da análise da literatura. O capítulo 3 descreve as dificuldades e problemas encontrados em pesquisas anteriores (relativas à construção de instrumentos musicais digitais),

explicitando o escopo da pesquisa. Já o capítulo 4 apresenta o estado da arte do artefato desenvolvido neste projeto.

No capítulo 5, é apresentada a abordagem adotada durante o projeto. São descritos, os princípios seguidos durante esse processo, os métodos de design adotados e a justificativa da escolha do instrumento.

O capítulo 6 descreve o processo de construção do sistema final (V-Conga), explicitando as escolhas dos sensores e tecnologias e apresentando a evolução cronológica do sistema, detalhando as abordagens utilizadas e discutindo cada etapa de iteração do processo.

O capítulo 7 apresenta os resultados obtidos na avaliação do sistema. São explicitadas as avaliações quantitativas de latência e reconhecimento de nuances gestuais e a avaliação qualitativa feita a partir do experimento realizado com o músico profissional.

Por fim, no capítulo 8 são apresentadas as conclusões obtidas durante o projeto de pesquisa e são discutidos os trabalhos futuros.

2 EVOLUÇÃO DA INTERFACE HOMEM-MÁQUINA NAS TAREFAS COTIDIANAS E NA MÚSICA

A palavra “tecnologia” é comumente associada com novos dispositivos e sistemas criados. Mesmo exercendo impacto em todas as áreas de conhecimento, muitas vezes é apenas relacionada com o campo da Computação ou das engenharias. Entretanto, a utilização de técnicas aprimoradas desenvolvidas com os avanços de conhecimento, também fazem parte do conjunto que engloba a etimologia da palavra. Este é um campo que funciona por retroalimentação, ou seja, o avanço do conhecimento proporciona evolução nas técnicas de pesquisa e de criação de novos dispositivos que, por sua vez, podem aumentar a capacidade de entendimento sobre uma certa área que irá aumentar o conhecimento dos pesquisadores e desenvolvedores e assim por diante.

Neste capítulo, veremos o qual foi o impacto dos avanços tecnológicos na evolução das interfaces de interação homem-máquina e no meio musical, e como estas interfaces possibilitaram a criação de novos instrumentos musicais.

2.1 Avanços Tecnológicos em HCI

Nesta seção, vemos como a interação homem-máquina se beneficiou com os avanços científicos ao longo da história, apresentando a evolução nos dispositivos de interface e como o surgimento dos mesmos mudaram os modelos de interação em HCI.

A tecnologia sempre exerceu um importante papel na interação homem-máquina. O homem utiliza os avanços tecnológicos para extrair máxima capacidade de artefatos já conhecidos e aprimorar modelos já consolidados. Em 1911, Frederick Winslow Taylor (1856-1915) publicava o *Princípio da Administração Científica*, obra que revolucionou a economia mundial. Para Chiavenato (CHIAVENATO, 2003, p. 55):

“O ponto de partida de Taylor foi a aplicação dos princípios da tecnologia de sua época ao trabalho manual. Procurou aplicar às operações manuais os mesmos princípios que os projetistas aplicavam às operações das máquinas no século XIX. Para tanto, ele identificava o trabalho a ser feito, decompunha-o em suas operações individuais, designava a maneira certa de realizar cada operação e, finalmente, reunia as operações na sequência que permitia realizá-lo mais rapidamente e com maior economia de tempo e movimentos.”
(CHIAVENATO, 2003, p. 55).

Chiavenato descreve a aplicação do método cartesiano criado por René Descartes e utilizada por Taylor, na administração, de forma a aprimorar o modelo econômico industrial de sua época. Apesar de diversos problemas relacionados com o modelo de Taylor, o mesmo serviu de base para a administração atual e também serviu como inspiração para os novos princípios que o sucederam, como os aplicados por Henry Ford (1863-1947) na *Ford Motor Co.* fundada em 1903 e publicadas em seu livro *Minha filosofia de indústria*, os quais aplicou novos avanços tecnológicos de forma a padronizar não só o maquinário utilizado na indústria, mas também o produto e seu desenho, o material e a mão de obra.

2.1.1 Primeiros computadores

O sucesso dos computadores *mainframe*, como o Mark I⁹ desenvolvido pela universidade de Harvard, proporcionou que, em 1946, novos avanços tecnológicos levassem à criação do primeiro computador programável para uso geral: o *Eletronic Numerical Integrator and Computer* (ou ENIAC). Desenvolvido por John Eckert e John Mauchly, na escola de engenharia elétrica *Moore School of Electrical Engineering*, com apoio do departamento militar, o ENIAC teve o propósito de ser utilizado prioritariamente como instrumento de pesquisa balística para as forças armadas (GOLDSTINE; GOLDSTINE, 1946).

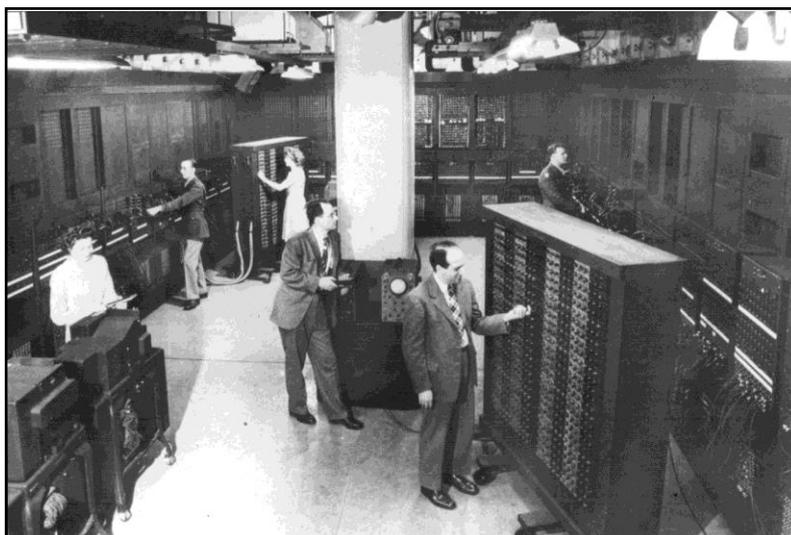


Figura 2.1: Eletronic Numerical Integrator and Computer (ENIAC).

A interface desenvolvida para ele ser operado era baseada em manipulações de componentes eletrônicos como cabos, switches, válvulas e relês, mas também a partir da leitura de cartões numéricos desenvolvidos pela empresa IBM. O problema computacional e as estimativas iniciais para o mesmo eram introduzidos na unidade de processamento central chamada *master programmer* que eram realizadas pelas unidades aritméticas da máquina (HARTREE, 1946).

⁹ Mais informações no site: <<http://chsi.harvard.edu/markone/function.html>>.

2.1.2 O conceito do Desktop

Contrária a ideia de utilizar estes sistemas para aplicações numéricas estava a visão de utilizar estes sistemas para tarefas mais cotidianas. Na mesma década do ENIAC, Vannevar Bush já imaginava o conceito de um computador compacto que se assemelhava a uma mesa de trabalho, ou *desktop*: o MEMEX (BUSH, 1945).

Ligada a essa possibilidade de utilizar estes sistemas de forma a auxiliar as tarefas que precisavam ser realizadas, surgiu a necessidade de aprimorar a forma com a qual os homens interagiam com estas máquinas. Esta necessidade é um dos grandes pilares das pesquisas desenvolvidas por Douglas Carl Engelbart (1925-2013) em seus estudos para aprimorar a capacidade do usuário em realizar uma determinada tarefa utilizando os computadores (ENGELBART, 1962; ENGELBART; ENGLISH, 1968). Estes estudos levaram a diversas descobertas tecnológicas apresentadas alguns anos depois na demonstração conhecida por “Mother of All Demos”¹⁰, em 9 de dezembro de 1968, como o conceito de hipertexto, teleconferência e o primeiro protótipo do mouse que usamos atualmente.

2.1.3 Interfaces Gráficas e a implementação do modelo WIMP

Outro grande marco na área de HCI foram os lançamentos do *Xerox Alto* e do *Xerox Star*. Sistemas compactos composto por uma unidade de processamento que incluía um display, um teclado - interface de escrita que se assemelhava a uma máquina de escrever (que já era amplamente utilizada e já consolidada no mercado), um mouse e ainda apresentava as primeiras interfaces gráficas do usuário (GUIs, do inglês Graphical User Interface).

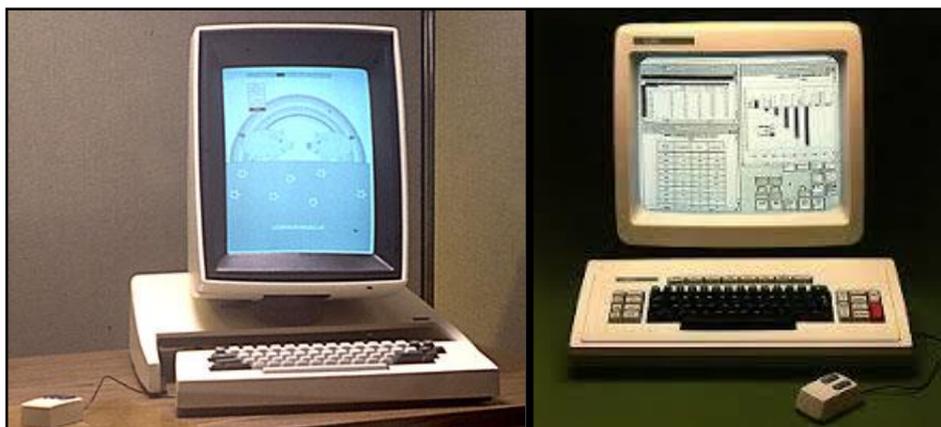


Figura 2.2: Xerox Alto (direita) e Xerox Star (esquerda).

Entretanto, introduzir as interfaces gráficas foi apenas um dos avanços tecnológicos dos sistemas. O software do *Alto* foi pioneiro ao construir a ideia

¹⁰ Vídeos disponíveis em: <<http://dougengelbart.org/events/1968-demo-highlights.html>>.

arquivos e janelas, nas quais os variados programas seriam executados, além de já incorporar a ideia de compartilhamento de informações e rede de conexão entre sistemas (LAMPSON; TAFT, 1976; WADLOW, 1981). Este sistema proporcionou a evolução do paradigma de interface que utilizamos atualmente.

Já o *Star*, lançado em 1981, pode ser considerado mais um salto tecnológico na direção das interfaces utilizadas atualmente. O mesmo eliminava diversos problemas encontrados em sistemas anteriores e aumentava a capacidade de o usuário final focar mais nas tarefas a serem realizadas. Notoriamente, a introdução do conceito de ícones junto aos outros anteriormente consolidados pelo *Alto* foi revolucionária e rapidamente difundida para os outros sistemas (JOHNSON et al., 1989). O sucesso deste sistema de interação se deve indubitavelmente aos dispositivos que eram utilizados para operar o *Xerox Star* – uma interface visual construída suportada pelo teclado e mouse - serviu como base para a elaboração do paradigma de interação utilizado na maioria das interfaces utilizadas atualmente: o paradigma conhecido como WIMP (“Windows, Icons, Menus and Pointable devices” – que pode ser traduzido em “Janelas, Ícones, Menus e Dispositivos apontadores”).

Antes do lançamento do *Star*, diversos pesquisadores já documentavam novas formas de se interagir com os computadores. Estas inovações buscavam uma alternativa ao modelo de interfaces com botões (que foi consolidada com o modelo WIMP). Particularmente, a primeira interface gestual digital formalmente documentada, criadas para servir como dispositivo de entrada para computadores, foi criada em 1982 na universidade de Toronto, pelo cientista Nimish Mehta, e utilizava uma câmera e princípios básicos de iluminação para detectar os gestos realizados pelos usuários, e alavancou as pesquisas na mesma área. Pouco tempo depois, em 1983, Myron Krueger lançou a demo da *Video Place/Video Desk*¹¹, um sistema rico em interações gestuais com tecnologia baseada câmeras e visão computacional desenvolvido alguns anos antes (BUXTON, 2007). Diversas outras interfaces ainda surgiram na mesma década, incluindo pesquisas que mostram superfícies de contato multi-toque (BUXTON; HILL; ROWLEY, 1985; LEE; BUXTON; SMITH, 1985; WESTERMAN; ELIAS; HEDGE, 2001) com um ou mais usuários realizando gestos independentemente (DIETZ; LEIGH, 2001). Entretanto, estas interfaces ainda eram produtos altamente custosos de forma que seu acesso era limitado.

2.1.4 Surgimento das Interfaces gestuais

Em 2006, Jefferson Y. Han apresentou uma palestra no TED Talks na qual apresentava uma demonstração intitulada “*The radical promise of the multi-touch interface*”¹² da sua mais nova interface de interação: um protótipo de baixos custo

¹¹ Vídeo disponível em: <<https://www.youtube.com/watch?v=dqZyZrN3PI0>>.

¹² Vídeo disponível em: <<https://goo.gl/emHhNV>>.

e tempo de resposta de uma superfície multi-toque facilmente replicável (HAN, 2005).

Saffer (SAFFER, 2008, p.2) acredita que este foi um dos grandes momentos de transição na forma em que as grandes indústrias e pesquisadores pensavam sobre as novas interfaces:

“Since then, consumer electronics manufacturers such as Nintendo, Apple, Nokia, Sony Ericsson, LG, and Microsoft have all released products that are controlled using interactive gestures. Within the next several years, it’s not an exaggeration to say that hundreds of millions of devices will have gestural interfaces.” (SAFFER, 2008, p.2).

Certamente, em 2005, a empresa japonesa de videogames Nintendo anunciou o, até então, mais novo console de última geração: o Nintendo Wii, que viria a ser lançado no ano seguinte. O lançamento do Nintendo Wii pode servir como marco para o lançamento de novas interfaces gestuais produzidas em larga escala e disponibilizada para o público.

O Wii trazia consigo uma tecnologia inovadora de interação gestual, vista poucas vezes anteriormente no contexto de jogos digitais, se afastando, mas também incorporando, o modelo de interação por botões utilizado até então. Diferentemente de interfaces anteriores, o *Wii Remote* (como é conhecido o controle do console) é um controle de alta precisão que contém um acelerômetro e um sensor infravermelho (IR, do inglês Infrared) que é detectado por uma pequena barra de captura que contém sensores IR.



Figura 2.3: Wii Remote (esquerda) e barra de captura (direita).

Fazendo uma triangulação a partir dos sensores IR contidos no controle e na barra de captura, pode-se localizar a posição espacial (x, y, z) do sensor contido no Wii Remote e, conseqüentemente, a posição da mão do usuário. Além disso,

o acelerômetro contido no controle consegue identificar a rotação realizada pelo usuário no controle.

Aproximadamente 3 meses após o lançamento do Wii (janeiro de 2007), Steve Jobs (1955-2011), junto à Apple Inc., lançava a primeira geração do iPhone (KERRIS; DOWLING, 2007). O iPhone possuía diversos sensores dos quais pode-se destacar a tela multi-toque, um acelerômetro e um dispositivo Wi-fi. Versões mais recentes já possuem um giroscópio e GPS, entre outros sensores. No ano seguinte, a Apple lançou a loja virtual de aplicativos (*Apple App Store*) e o kit de desenvolvimento de software (SDK, do inglês *Software Development Kit*), o suporte para os desenvolvedores da plataforma iOS (sistema operacional utilizado pela empresa) baseado na linguagem de programação Objective-C. Mais tarde, em 2010, a Apple lançaria o iPad: seu primeiro tablet que possuía diversos dos sensores embutidos no iPhone, porém com um display maior (SMITH; EVANS, 2010).



Figura 2.4: Steve Jobs apresentando o iPad.

Até então, a área de reconhecimento de gestos em três dimensões, era feito majoritariamente utilizando câmeras e técnicas de visão computacional, padrões de textura e cor da pele e modelos estatísticos como o Modelo Oculto de Markov (ou *Hidden Markovian Model* – HMM) (MITRA; ACHARYA, 2007). Entretanto, todo o processamento de dados era realizado pelo próprio computador, o que pode demandar tempo e prejudicar as aplicações (dependendo da capacidade do processador do mesmo).

Apenas em 2010 a Microsoft lançaria o Kinect®, um sensor de reconhecimento de gestos, reconhecimento facial e também podia realizar tarefas por comando de voz que possuía a capacidade de processar essas informações internamente. Este dispositivo foi lançado junto ao console da própria empresa, o Xbox 360, com o intuito de ampliar a experiência do usuário introduzindo novas formas de interação com os jogos digitais, como forma de concorrência ao Nintendo Wii e

o Playstation *Move* (Interface gestual lançada em 2009 pela Sony como forma de concorrência do Nintendo Wii Remote).



Figura 2.5: Sensor Kinect da Microsoft.

O Kinect possui uma câmera RGB, um sensor de profundidade IR, além de microfones para captura de voz. Estes sensores são utilizados de modo a capturar uma cena 3D numa forma de nuvem de pontos, na qual cada um desses pontos possui uma informação IR com a qual o *software* do dispositivo é capaz de calcular sua posição espacial (x, y, z) e já é pré-programado para reconhecer 48 pontos da articulação do corpo humano (Silva, 2012), o que facilita o desenvolvimento de novas interfaces de interação.

Interfaces de interação com o Kinect se tornaram bastante populares não só pelo fato de o sensor ter sido popularizado com as vendas do videogame, com o qual ele vinha atrelado, mas também pelo fato do dispositivo suportar aplicações desenvolvidas em C++, C#, Visual Basic ou qualquer outra linguagem de programação .NET.

Em outubro de 2012 a companhia *startup* Leap Motion Inc. lançou o programa de desenvolvedores de um projeto que vinha sendo desenvolvido, por seus fundadores, desde 2008. O Leap Motion Controller© (“Leap Motion”, 2015), que seria lançado para o público consumidor no ano seguinte (2013), é um sensor de reconhecimento gestual que visa aprimorar tanto a precisão de rastreamento quanto a latência relativa ao mesmo.



Figura 2.6: Sensor Leap Motion®.

Tentando combinar a capacidade de reconhecimento espacial de dispositivos como o Kinect da Microsoft ou o Xtion® da Asus, com a baixa latência e robustez de interfaces multi-toque de smartphones e tablets top de linha como o iPad® da Apple e o Galaxy® da Samsung, o Leap Motion Controller® utiliza novas técnicas de rastreamento em três dimensões para reduzir o tempo de processamento e, conseqüentemente, a latência nas aplicações.

Estes avanços proporcionaram um crescente número de interfaces incorporadas (*embodied interfaces*), ou seja, interfaces em que a comunicação homem-computador utiliza o próprio corpo do usuário como parâmetros de entrada. Atualmente, eletrônicos para consumo geral, como *smartphones*, *notebooks*, *tablets* e televisores (como as *smart TVs*), já utilizam gestos para controlar suas interfaces. Principalmente interfaces multi-toque, às vezes combinadas a gestos no ar, ganharam bastante espaço com o sucesso dos *smartphones*.

Além disso, com os avanços nas áreas de realidade virtual (VR) e realidade aumentada (AR), diversas pesquisas buscam combinar interações gestuais com um ambiente de VR ou AR. Variando entre uso de luvas e marcadores para reconhecer movimentos como o FingARtips (BUCHMANN et al., 2004), utilizando objetos tangíveis como o *Geomagic Touch Haptic Device* – uma caneta que possui *feedback* de força (*haptic feedback*) – para ser utilizado em ambientes virtuais, ou utilizando movimentos do corpo como o Microsoft Hololens® que utiliza gestos para interagir com ambientes de realidade aumentada.



Figura 2.7: Microsoft HoloLens.

Paralelamente à chegada desses novos sensores, diversos algoritmos de visão computacional continuam a ser aprimorados visando reconhecimento facial e de gestos, utilizando uma ou mais câmeras, para aplicações em HCI (SATO; SAITO; KOIKE, 2001; WACHS et al., 2011).

Neste contexto, as interfaces gestuais incorporam uma gama de possibilidades que não eram possíveis ou não eram tão eficazes com as interfaces desenvolvidas apenas com botões. Algumas dessas possibilidades são a flexibilidade nas formas de interação, a ausência da dependência de diversos tipos de *hardware*, interações mais naturais e principalmente a capacidade de expressar as nuances dos movimentos (SAFFER, 2008).

Apesar de apresentarmos as diversas tecnologias de interfaces e como foi a evolução nessa área de pesquisa, o campo de HCI não se beneficia puramente das interfaces, e sim em como será a interação entre o homem e os computadores. Para isso, as tecnologias e interfaces emergentes devem ser avaliadas e testadas com o público alvo. Todavia, avaliar estas tecnologias proverá expertise aos designers ao lidar com as mesmas para projetar novas interfaces, alertar possíveis problemas ao trabalhar com elas e como estes problemas podem ser contornados.

2.2 Novas formas de interação musical

Esta seção mostra os avanços tecnológicos que influenciaram a criação de novos instrumentos musicais e como os mesmos permitiram que o homem interagisse com o som de maneiras diferentes. Além disso são apresentados os Instrumentos Digitais Musicais (DMIs), os quais são o objeto de estudo deste projeto.

Blades (BLADES, 1992, p. 34–35) acredita que os primeiros passos na descoberta de instrumentos de percussão data da era paleolítica. Mesmo que a maioria das descobertas arqueológicas recuperaram diversos instrumentos de ossos, pedra, argila e metal, acredita-se que os primeiros passos da descoberta

da percussão bem como os primeiros instrumentos a serem aprimorados foram a partir de colisão de membros corporais com outros objetos ou com partes do próprio corpo:

“The seeds of the first instrument were sown unconsciously by an early man as he stamped upon the ground, beat upon his throat, clapped his hands, or slapped his body. He produces contrasting sounds with hollowed hands, flat palms, heels or toes, or by striking either bony or fleshy parts of the body.

[...] Percussion – the act of striking – was an art in which primitive man was well skilled. He survived by the dexterity of his blow; from which it is fair to assume that the first instruments to augment the hand clap and stamp of the foot may have been the implements or weapons upon which he relied for food or survival. ”
(Blades, 1992 pp. 34-35).

Novas formas de tratamento do couro de animais para criação de novos instrumentos servem como evidências do aperfeiçoamento da luteria (termo atribuído ao processo de criação e manutenção de instrumentos musicais). Além disso, o domínio de técnicas para se produzir metais, bronze e diversos outros instrumentos musicais precursores dos instrumentos musicais acústicos atuais.

2.2.1 *Descoberta da eletricidade e um grande salto tecnológico*

A descoberta da eletricidade fez com que novas formas de geração sonora fossem possíveis. Particularmente, Gohn (GOHN, 2001) considera a invenção do telefone, por Alexander Graham Bell (1847-1922) em 1876, como peça fundamental na evolução dos instrumentos e criação de instrumentos eletrônicos. Pela primeira vez, existia a possibilidade de se transformar ondas sonoras em eletricidade e vice-versa.

Os impactos causados pela invenção de Bell repercutiram na criação de diversos instrumentos musicais que buscaram utilizar a eletricidade para aprimorar instrumentos acústicos, criar novas formas de produzir som e explorar novas texturas sonoras. O primeiro instrumento que utilizava a eletricidade para geração sonora data do começo século XX, quanto Thaddeus Cahill (1867-1934) criou o *Telharmonium* ou *Dynamophone*. O *Dynamophone* utilizava componentes elétricos para produzir diferentes frequências que eram passadas, por uma interface de piano, para receptores de telefone (SHEPARD, 2013).

Vale ressaltar que diversos instrumentos elétricos que utilizamos atualmente foram concebidos nesta época, já que muitos inventores tentaram combinar a eletricidade para amplificar o som de instrumentos aústicos. Além disso, um dos instrumentos que, discutivelmente, estava mais à frente do seu tempo também foi inventado pouco tempo após o Dynamophone: o Thérémin. Leon Thérémin (1896-1993), em 1920, construiu um instrumento elétrico que utilizava a posição das mãos do músico em relação a duas antenas metálicas para determinar a frequência e o volume do som produzido (MANNING, 2013). O conceito de utilizar a eletricidade para criar novos sons, por trás do Dynamophone e do Thérémin, já remete aos primeiros sintetizadores.

A década de 1960 foi marcada pela criação dos primeiros sintetizadores controlados por voltagem. Em 1964, Robert Moog (1934-2005) construiu um oscilador e amplificador controlado por voltagem que só depois atribuiria o nome de sintetizador. Talvez pela inovação e pelo sucesso do Mixtur-Trautonium na concepção de trilhas sonoras para filmes em Hollywood, e também a facilidade de interação do sistema somada a familiaridade da interface do piano, os sintetizadores controlados por voltagem fizeram enorme sucesso e no final da década de 1960, quatro grandes empresas já concorriam pela liderança deste mercado (MANNING, 2013).

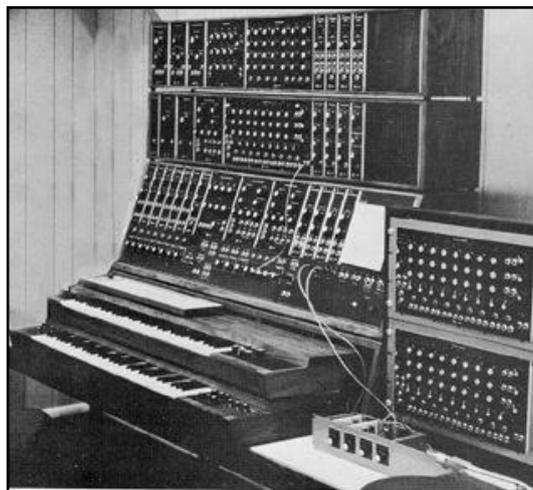


Figura 2.8: Sintetizador Moog.

Paralelamente a este avanço instrumental, as descobertas feitas no campo de armazenamento de som, provenientes das invenções do fonógrafo, grafafone e gramofone, trouxeram novas formas de interação com áudio (GOHN, 2001) – como a fita magnética, compact disks e LP – que abriram as portas para que, após a criação dos computadores, o conceito de gravação de áudio culminasse na criação da mídia digital.

2.2.2 Os avanços com o protocolo MIDI

A chegada de microprocessadores e a incorporação do som digital trouxeram novas possibilidades de se interagir com som. (MANNING, 2013) destaca que muitos dos processamentos de som e as descobertas feitas anteriormente à década de 1980 estavam sendo transferidas para o meio digital. Nesta mesma época, houve uma revolução na forma em que os instrumentos eram construídos e um protocolo de comunicação foi desenvolvido para facilitar a integração dos instrumentos musicais na era digital: o protocolo MIDI ou (“Musical Instruments Digital Interface” ou “Interface Digital para Instrumentos Musicais”).

O protocolo serial MIDI simplifica a forma com que instrumentos digitais possam se comunicar com computadores. Mandando informações numéricas e simbólicas, o instrumento pode comunicar quais notas estão sendo tocadas, mudanças em parâmetros de controle, como alterações em efeitos e volume, velocidade na qual as notas são tocadas, pressão feita ao tocar uma ou mais notas, a partir de valores entre 0 e 127 para cada um dos 16 canais disponíveis. O computador recebe estes valores por uma interface física, geralmente uma interface MIDI-USB, ou sem fio.

A tecnologia MIDI foi bem aceita pela comunidade e levou à criação de diversos controladores que, diferentemente dos instrumentos elétricos convencionais, o *output* de saída é um sinal MIDI. Os controladores MIDI são construídos das mais variadas formas. Como o sinal MIDI pode ser processado de qualquer forma, o som que sai do instrumento não é necessariamente o som do instrumento físico que está sendo tocado, ou seja, um violão MIDI pode gerar o som de uma flauta, saxofone, vozes, ou até ser utilizado para controlar outros parâmetros como iluminação, ser utilizado como teclado de um computador, etc. O que vai diferenciar a utilidade do controlador é como o sinal será processado pelo computador. Atualmente, diversos instrumentos elétricos (mais comumente nos teclados e instrumentos de percussão) possuem também uma saída MIDI.



Figura 2.9: Controladores MIDI com diferentes interfaces.

O aumento na capacidade de processamento dos computadores junto à possibilidade de utilização destes para realiza múltiplas tarefas e a difusão dessa tecnologia proporcionaram o desenvolvimento de novos algoritmos de

processamento de áudio e tecnologias, possibilitaram a invenção de interfaces de interação inovadoras, novas formas de se interagir com áudio e novas possibilidades de criação de instrumentos (MIRANDA; WANDERLEY, 2006). As chamadas interfaces gestuais podem utilizar gestos, movimentos e/ou posições estáticas em uma cena em duas ou três dimensões, reconhecidos por meio de diversos tipos de sensores emergentes, para interagir com estes instrumentos.

2.2.3 O modelo de Instrumentos Musicais Digitais

Existem diferentes visões entre como definir Instrumentos Musicais Digitais (DMIs, do inglês *Digital Musical Instruments*), que valem a pena ser consideradas (CALEGARIO, 2013), porém, para o escopo deste projeto, os DMIs são artefatos construídos com a finalidade de produzir um resultado musical, nos quais a interface de controle que vai receber os parâmetros de entrada (*inputs*), sejam eles puramente gestos, ou interagindo com algo tangível, é independente da saída (*output*) sonora, porém ambas estão relacionadas por um mapeamento (*mapping*) escolhido pelo designer. Sendo assim, Rován et. al. (ROVAN et al., 1997) propõem o seguinte modelo para um DMI:

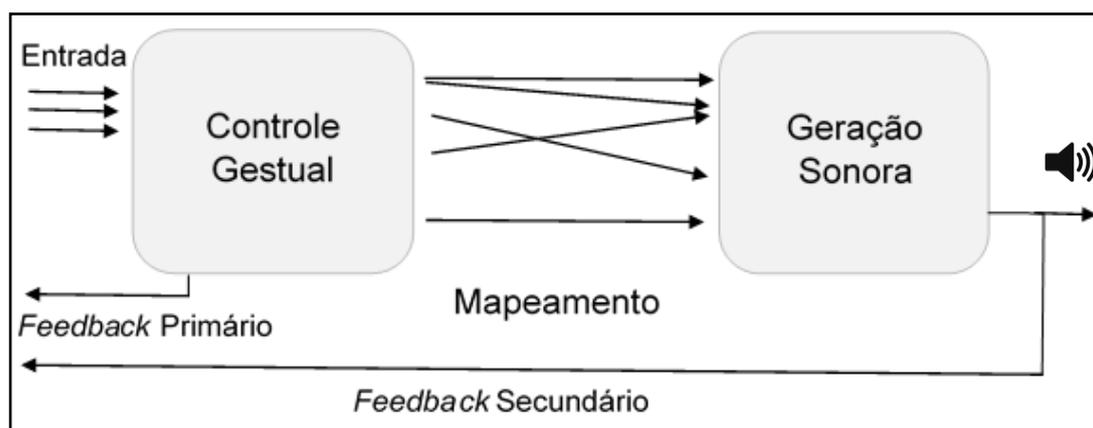


Figura 2.10: Modelo de DMI.

Miranda e Wanderley (2006, p. 4) nos falam que:

“The idea behind our approach to DMIs is analogous to breaking apart an acoustic instrument so that one could separate the two functionalities of gestural interface and sound generator and combine them in any desired way. Therefore, the notion of a fixed causality between an action and a sound in a DMI does not necessarily apply.” (MIRANDA; WANDERLEY, 2006, p. 4).

Isto quer dizer que, nos instrumentos acústicos, os sons reproduzidos por movimentos específicos estão diretamente relacionados com a mecânica

vibracional do ar, levando em conta o material do instrumento e o seu formato, os quais irão modelar a dinâmica do ar, fazendo com que estes resultados sonoros possuam características únicas. Por outro lado, no caso dos DMIs, o controle gestual é separado da produção sonora, e cada designer fará suas escolhas de mapeamento, então, um mesmo gesto pode ser mapeado em diferentes sons ou, num caso mais extremo, diferentes parâmetros de controle, i.e., uma certa pose (ou movimento) num espaço de coordenadas (X, Y, Z) pode disparar um determinado som ou mudar certa camada de *loop* musical, ou ainda pausar um fundo musical.

2.2.4 Tipos de DMIs

Com esta independência, as possibilidades de criação de novos instrumentos são imensas. Ao combinar os diversos sensores de captura gestual com os mais variados parâmetros de controle sonoro, nota-se uma grande variedade de DMIs.

Algumas características específicas podem distinguir certos DMIs de outros. É possível, assim, classificá-los (MIRANDA; WANDERLEY, 2006; WANDERLEY, 2001) da seguinte maneira:

- **Instrumentos musicais aumentados:** Instrumentos acústicos que utilizam diversos sensores para aprimorar ou complementar sua capacidade sonora.
- **Simuladores de instrumentos acústicos:** São DMIs contruídos para reproduzir tanto o controle gestual quanto o resultado sonoro de instrumentos acústicos já existentes.
- **DMIs inspirados em instrumentos acústicos:** DMIs que se munem do controle gestual de instrumentos acústicos já existentes, porém com finalidade sonora diferente. Estes podem ser criados também para solucionar limitações existentes nos instrumentos originais, mesmo que não pretendam reproduzi-los.
- **Controles Gestuais Alternativos:** Instrumentos que não possuem semelhança com instrumentos acústicos. Mesmo assim podem possuir o resultado sonoro de instrumentos acústicos.

É importante ressaltar que esta não é uma forma de classificação rígida, ou seja, um determinado DMI não está inserido em uma classe e apenas nela. Ele está inserido em um espectro contínuo que é relativo à sua semelhança com instrumentos acústicos já existentes.

Um exemplo de instrumento aumentado (*augmented instruments*) é o ACPAD¹³, criado pelo músico Robin Sukroso em parceria com Indian Institute of Technology Bombay. Desenvolvido para funcionar em instrumentos acústicos

¹³ Mais detalhes em: <http://acpad.com/>

sem precisar de fios, o ACPAD é um dispositivo MIDI flexível que pode ser inserido no corpo de um violão e possui diversos sensores discretos e contínuos para aumentar a capacidade expressiva do instrumento acústico.



Figura 2.11: ACPAD.

Diversos controladores de DMIs são desenvolvidos para simular instrumentos (*instrument-like DMIs*) já existentes variando de instrumentos como os teclados e baterias eletrônicas (mais comumente vistos), a clarinetes, xilofones e guitarras (JORDA, 2005). Além disso, instrumentos como o Smomid¹⁴ (à direita na figura abaixo), criado por Nick Demopoulos, ou o Seaboard¹⁵ (à esquerda na figura abaixo) da empresa ROLI, que utilizam gestos herdados do violão e piano, respectivamente, para obter resultados completamente diferentes.



Figura 2.12: DMIs inspirados em instrumentos acústicos.

Na figura 2.14, podemos ver o ReacTable, criado por Sergi Jordà. O instrumento foi desenvolvido utilizando algoritmos de visão computacional para reconhecer objetos, que possuem marcadores, dispostos em uma mesa e determinam o tipo do objeto, a posição, a orientação, entre outros, e os movimentos das mãos realizados pelo usuário do sistema (JORDA et al., 2005). A partir deste

¹⁴ Detalhes e demonstrações disponíveis em: <http://www.nickdemopoulos.com/smomid.html>

¹⁵ Detalhes e demonstrações disponíveis em: <https://www.roli.com/products/seaboard-grand>

reconhecimento, o DMI utiliza estas informações para modificar determinados parâmetros musicais.



Figura 2.13: ReactTable.

Este é um exemplo de DMI classificado como *alternative* ou *alternate instruments*. Ou seja, instrumentos que não têm nenhuma relação nem com os instrumentos acústicos, com exceção da geração sonora, que pode ser o som de um instrumento já conhecido. Estes instrumentos podem possuir diversos formatos e funcionalidades, já que são o tipo que possui mais liberdade nos controles gestuais.

Certas propriedades como flexibilidade nas formas de interação, a ausência da dependência diversos tipos de *hardware*, interações mais naturais e principalmente a capacidade de expressar as nuances dos movimentos, não eram possíveis nas interfaces com botões (Saffer, 2008). Estas são algumas vantagens de se utilizar interfaces gestuais para produzir instrumentos musicais.

Todavia, este tipo de interface está longe de ser considerado perfeita para todos os tipos de interação. (SAFFER, 2008) ainda enfatiza uma preocupação expressada por Bill Buxton (BUXTON, 2007), ao desenvolver novas interfaces gestuais, na qual “quando se trata de tecnologia, tudo é melhor para certas coisas e pior para outras e gestos interativos não são exceções”¹⁶. É imprescindível investigar o problema a ser solucionado e quais tecnologias dispostas serão mais eficazes e podem ser utilizadas para produzir a melhor interação possível, sejam elas botões, reconhecimento por voz, gestos, etc. Cabe, então, ao designer e/ou pesquisador, a decisão de utilizar este tipo de interface para solucionar o problema encontrado ou realizar certa tarefa.

Particularmente, no caso dos DMIs, a investigação do uso destas tecnologias é apenas um dos desafios a serem enfrentados durante o seu desenvolvimento.

¹⁶ “when it comes to technology, everything is best for something and worse for something else and interactive gestures are no exception” (Saffer, 2008 p.16)

Diversos fatores devem ser considerados na concepção e criação deste tipo de instrumento.

2.3 Conclusões

Neste capítulo, foi mostrado como os avanços tecnológicos influenciaram as formas de interação homem-máquina. A criação de novas tecnologias e o aumento na capacidade de processamento dos computadores permitiram a exploração de novas interfaces que rompem com os paradigmas de interação já consolidados (que utilizam o modelo WIMP e interfaces compostas por mouse, teclado e monitor).

No contexto musical, a evolução das tecnologias e o surgimento dos computadores forneceram novas formas de interação com a música. O surgimento das interfaces gestuais proporcionou possibilidades de criação de Instrumentos Musicais Digitais (DMIs). Estes instrumentos utilizam gestos do usuário como entrada para controlar parâmetros sonoros, por meio de uma estratégia de mapeamento escolhida pelo *designer* do sistema.

3 DESAFIOS NA CRIAÇÃO DE DMIS

Este capítulo apresenta os desafios encontrados na construção de interfaces gestuais, mais especificamente dos DMIs, bem como demonstra algumas preocupações que foram levantadas na literatura a respeito destes novos instrumentos. Entretanto, é importante enfatizar que ao levantar estas dificuldades, deve-se explicitar o escopo, no qual estas estão inseridas.

Com estas infinitas possibilidades de criação de DMIs, é possível notar que a liberdade de processar as informações vinda dos instrumentos da maneira que for mais conveniente e a independência entre os parâmetros de controle e os módulos de geração sonora facilitam a “redução da sobrecarga de controle humano” (JORDA, 2005, p. 27).

A intenção deste capítulo é discutir as dificuldades de projetar DMIs que seriam realmente utilizados pelo público profissional com intuito de usufruir ao máximo as capacidades expressivas destes instrumentos. (MEDEIROS et al., 2014; SILVA, 2012).

Primeiramente, este capítulo apresenta um trabalho realizado pelo grupo de pesquisa que o autor desta dissertação faz parte (MusTIC), no qual DMIs que simulam instrumentos de percussão foram avaliados por músicos profissionais que levantaram problemas relativos à expressividade desses instrumentos (SILVA, 2012). Ademais, são apresentados os resultados de um questionário realizado com percussionistas profissionais de modo a investigar se controladores digitais de percussão são usados por esse público, além de levantar vantagens e desvantagens nesses sistemas.

Em seguida, são descritos os diversos desafios enfrentados no design de DMIs (MEDEIROS et al., 2014). Estes desafios podem ser separados em duas classes que estão relacionadas com os fatos desses DMIs serem **artefatos digitais** e serem **artefatos musicais**. Buscou-se relacionar estes desafios com os problemas levantados pelos músicos profissionais.

Por fim, os problemas investigados neste projeto são especificados, reduzindo o escopo dos desafios apresentados. Com isso, foi possível desenvolver a pergunta de pesquisa, além de propor objetivos para sua solução (apresentados no capítulo 1).

3.1 Apresentando DMIs ao público profissional

Em 2012, Silva (SILVA, 2012) apresentou uma pesquisa, na qual explorava a avaliação de tecnologias para construção de DMIs de percussão. Nesta pesquisa, duas tecnologias foram utilizadas para o desenvolvimento destes DMIs: o Nintendo Wii® *Remote* e o Microsoft Kinect® versão 1.

A construção destes instrumentos levou a um estudo detalhado sobre as capacidades de captura de movimentos destas duas tecnologias, baseadas nos movimentos primitivos do corpo, restritos pela anatomia humana. Para instrumentos de percussão, o foco na análise da captura dos membros superiores do corpo – ombros, braços, antebraços, punhos e mãos (separando a movimentação dos dedos) – é justificada pelo fato de que os gestos musicais, principalmente nos instrumentos de percussão, “são feitos em sua maioria com as mãos e braços, com algumas exceções” (SILVA, 2012, p. 40).

Gestos	Sensor Kinect	Wii remote
Movimentos com o polegar		x
Movimentos com os demais dedos		x
Movimentos com o punho	x	x
Flexão e extensão do antebraço	x	x
Movimentos de rotação com o antebraço		x
Movimentos com ombro e antebraço	x	

Tabela 1: Capacidade de rastreamento dos sensores Wii Remote e Kinect (Extraído de SILVA, 2012).

Os estudos destas capacidades mostraram primeiramente a limitação destas tecnologias no rastreamento de movimentos curtos e de precisão. Principalmente movimentos realizados com os dedos não são bem reconhecidos pelos sistemas utilizados. Portanto, a limitação da capacidade de rastreamento de nuances gestuais, relacionada com a precisão e resolução dos sensores, foi levantada previamente à avaliação junto ao público alvo – músicos profissionais.

Com o intuito de seguir uma abordagem de *design* centrado no usuário, foram escolhidos alguns instrumentos de percussão para que DMIs simuladores destes instrumentos fossem construídos e avaliados pelos músicos profissionais, de forma a avaliar as tecnologias estudadas e levantar os possíveis problemas destes instrumentos digitais quando comparados com os seus análogos acústicos. Para maiores detalhes sobre o protocolo experimental, recomenda-se a leitura de (SILVA, 2012).

Silva (SILVA, 2012) nos diz que os músicos levantaram diversos problemas com os instrumentos a ponto de um usuário se sentir “perdido” e sentir-se como “não houvesse tocado nada na vida”. Quatro problemas são destacados na pesquisa. Lembrando que estes experimentos foram realizados com músicos profissionais, o primeiro remete ao fato deles estarem realizando gestos no ar, diferentemente de tocarem no corpo do instrumento acústico, levantou preocupações com a **ausência de *feedback* háptico** do instrumento e as pistas visuais que o corpo do instrumento proporciona – ao delimitar o ponto de ataque que irá produzir o som, fazendo com que o músico tenha consciência de onde fará o instrumento

soar. Os outros problemas estão relacionados com a procura de **nuances gestuais** e imprecisão dos gestos, **variações de timbre** em um mesmo instrumento, controle da **intensidade das notas** e **latência** dos instrumentos em movimentos rápidos. Portanto, problemas relativos à expressividade do instrumento.

3.2 Questionário sobre as tecnologias de percussão digital existentes

Para entendermos melhor a relação do músico profissional com instrumentos de percussão digital, foi preciso realizar um questionário que primeiramente investigasse se estes instrumentos são realmente utilizados, em que contexto eles são utilizados e que funcionalidades e características destes instrumentos agradam ou não agradam os músicos profissionais.

Para isto, foi escolhida a plataforma *SurveyMonkey* para a condução deste questionário. O *SurveyMonkey* é uma plataforma *online* intuitiva desenvolvida para auxiliar os seus usuários a conduzirem questionários para pesquisas diversas, oferecendo diversos planos e serviços (gratuitos e pagos) para isto. O site não só oferece a possibilidade de criação de questionário, como já produz resultados à medida em que as respostas são obtidas. Além disso, ainda fornece dicas e modelos de questionários para pesquisadores iniciantes.

Assim, foi gerado um questionário que levantava questões sobre o uso de instrumentos de percussão digital. Exemplos desses sistemas são as baterias eletrônicas e pads de percussão, como o HandSonic¹⁷ (da Roland), o Wavedrum¹⁸ (da Korg).

O questionário (modelo apresentado no Apêndice I, página 116) envolveu oito perguntas (quatro objetivas e quatro discursivas) que são descritas abaixo:

- Qual é a sua idade?
- Há quanto tempo você trabalha no cenário musical?
- Você utiliza algum instrumento de percussão digital?
- Se sim, o que lhe levou a utilizar esse(s) instrumento(s)?
- Se não, por que não?
- Em que contexto você utiliza esse(s) instrumento(s)?
- O que você gosta nesse(s) instrumento(s)?
- O que você NÃO gosta nesse(s) instrumento(s)?

O link do questionário foi repassado por meio de fóruns a músicos profissionais, e também foi repassado manualmente a músicos ligados ao autor deste trabalho.

¹⁷ Informações disponíveis em: <http://roland.com.br/products/handsonic_hpd-20/>.

¹⁸ Informações disponíveis em: <<http://goo.gl/TkvBR1>>.

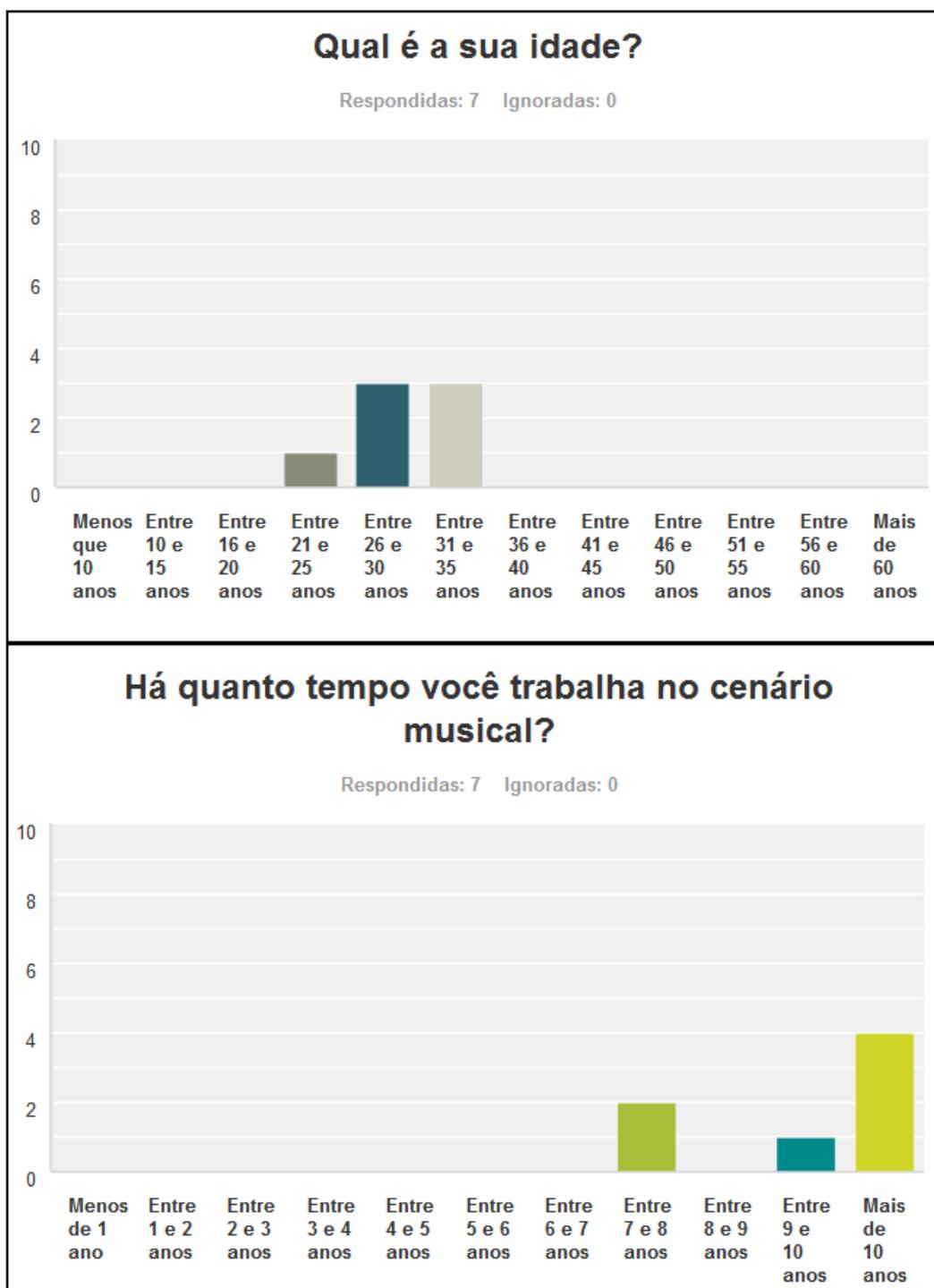


Figura 3.1: Perfil dos percussionistas no questionário.

Apesar de apenas 7 (entre 21 e 35 anos) músicos profissionais – dentre os quais 6 são percussionistas – terem respondido o questionário, todos possuem mais de 7 anos no meio musical (4 já trabalham a mais de 10 anos) e já conheciam este tipo de instrumento.

Um caso particular foi o de um violonista que utiliza este tipo de instrumento para produzir acompanhamento rítmico em seus trabalhos. Ele utiliza estes instrumentos para os mais variados contextos – ensaios, estudos, *performances*

e gravação. Entretanto, mesmo não sendo percussionista, as respostas aos questionamentos abertos e os comentários feitos, coincidem bastante com os músicos, já que utiliza este sistema de forma profissional e não há um instrumentista rítmico em seu grupo.

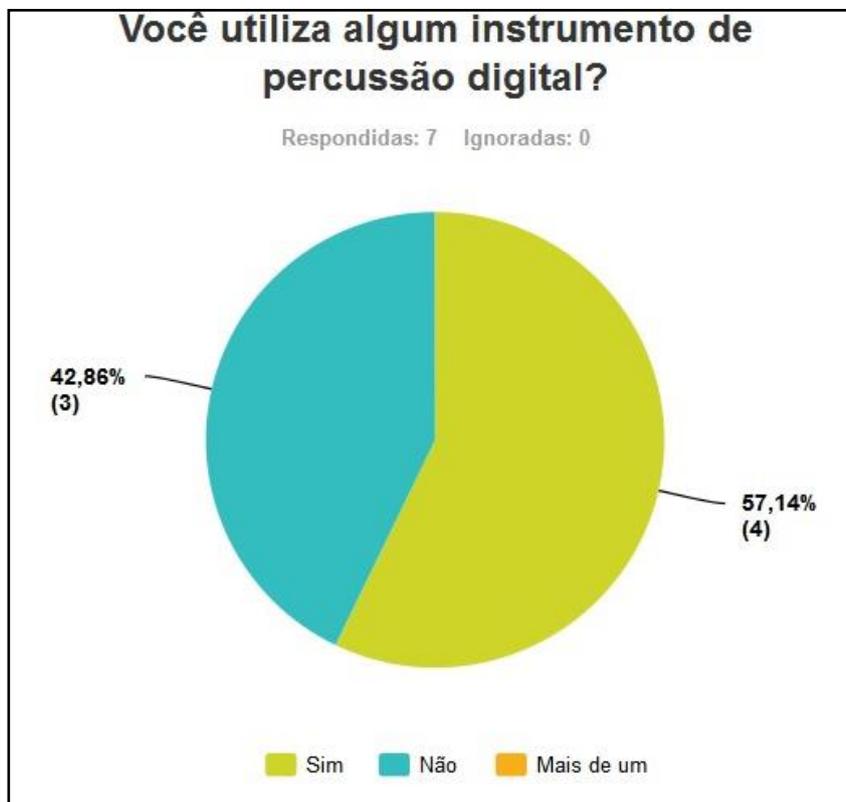


Figura 3.2: Gráfico representativo do uso de controladores digitais de percussão.

O gráfico mostra que um pouco menos da metade dos músicos que responderam o questionário não utilizam este tipo de instrumento, particularmente metade dos percussionistas (3 usuários).

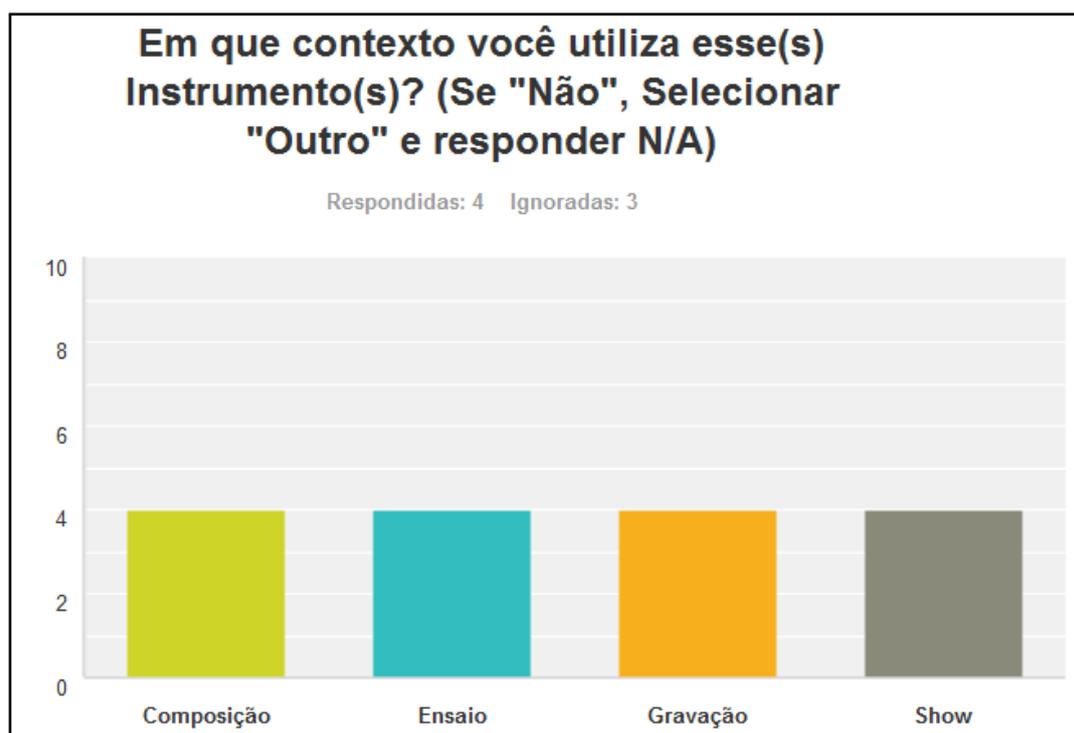


Figura 3.3: Contexto de uso de controladores digitais de percussão.

A partir do gráfico acima, é possível perceber uma variedade no contexto de uso (apenas um usuário não os utiliza em *performances* ao vivo) desses controladores. Isto pode estar atrelado às vantagens descritas por eles nas perguntas discursivas.

Os músicos que utilizam estes instrumentos (incluindo o caso particular) argumentam que a praticidade em relação ao transporte, o leque de possibilidades de timbres diferentes e “de instrumentos do mundo todo”, as vantagens de “não perder tempo passando o som, timbrando em momentos em que o tempo é curto... não precisa microfonar¹⁹, direcionar os canais na mesa, equalizar, etc.” são fatores que os agradam e que os levaram a aderir este tipo de instrumento.

Dentre o que levou estes músicos a não usarem esta tecnologia pode se destacar uma resposta em que o músico fala que “é interessante ser orgânico e ter a sensibilidade da pele, baqueta, palmada, podendo assim controlar a dinâmica”.

Ao serem questionados sobre o que eles não gostam nestes instrumentos, os músicos levantaram problemas como a memória interna pequena para se armazenar *samples*, o tempo de duração destes *samples* e a “falta de recursos para manipular os sons do banco de dados no computador ou nele mesmo”. Um

¹⁹ Técnica utilizada posicionando microfones em locais específicos dos instrumentos para capturar os sons e enviá-los à mesa de som, para que a regulação desses sons seja feita.

usuário ainda disse que acha “o som artificial, diferente do natural”. Outro especifica que “O timbre de peles realmente é algo interessante”.

Mesmo os percussionistas que utilizam estes instrumentos argumentaram que, apesar da variedade de possuir suas vantagens, a “pegada” e os próprios timbres não serem iguais aos dos instrumentos orgânicos. Um desses usuários menciona que “... em algumas ocasiões é muito válido, mas nunca vai se comparar a um instrumento percussivo acústico”, enquanto outro usuário diz que “os timbres e a pegada não são iguais aos instrumentos orgânicos”.

A partir destes resultados, é possível observar que mesmo quando há uma motivação por parte do músico em utilizar estes instrumentos, os fatores como o timbre e o material do instrumento proporcionam obstáculos a serem superados. Mesmo assim, pode-se notar vantagens quanto ao uso de tecnologias digitais para diversos usos de contexto, incluindo a praticidade e versatilidade nos timbres do instrumento.

3.3 Problemas do mapeamento

Para esclarecimento dos problemas de mapeamento enfrentados no desenvolvimento de um DMI é preciso entender as possibilidades e parâmetros de entrada (controles gestuais) e de saída (geração sonora) do mesmo.

Os gestos utilizados pela interface de controle podem ser classificados de acordo com várias abordagens. O trabalho de Calegario (2013) apresenta duas formas de classificação de controles gestuais que ajudam a entender a dimensão do problema em destaque. A primeira, apresentada na pesquisa de Wobbrock, Morris e Wilson (WOBBROCK; MORRIS; WILSON, 2009), classifica os gestos de acordo com o que está acontecendo com o sistema:

- **Gestos Discretos:** Pode ser dividido em gestos estáticos, e.g. uma pose ou postura no espaço X, Y, Z, e dinâmicos – gestos associados a algum tipo de movimento. Entretanto, a resposta do sistema só acontece após o gesto ser realizado.
- **Gestos Contínuos:** Neste caso, a resposta acontece durante a realização do gesto, um exemplo análogo é o caso do Thérémin, onde o movimento das mãos próximo às antenas do instrumento muda a frequência e/ou a amplitude do sinal.

A segunda, segue a linha de classificação utilizada na pesquisa de McGlynn (2011), que prioriza a maneira pela qual os dados do controle gestual serão utilizados:

- **Dados Brutos:** Estes dados podem ser discretos – como botões ou sensores digitais que apresentam valores quando estão ligados e

desligados – ou contínuos, como dados de rastreamento num espaço em duas ou três dimensões.

- **Dados Simbólicos:** Muitas vezes relacionados com combinações específicas dos dados brutos, e.g. uma posição determinada no espaço X, Y, Z.
- **Dados Gestuais:** Relacionados com o movimento predefinido.

O controle direto a partir dos **Dados Simbólicos** está, na maioria das vezes relacionados com os **Gestos Discretos**, enquanto que os **Dados Brutos** e **Dados Gestuais** podem estar relacionados tanto com os **Gestos Discretos** – e.g. apertar um botão e ligar um efeito ou fechar a mão e desligar o efeito – quanto com os **Gestos Contínuos** – e.g. aumentar o ganho de um efeito com um potenciômetro ou afastando as duas mãos.

Miranda e Wanderley (2006) ainda abordam o caso de a obtenção do controle gestual ser direta, indireta ou fisiológica:

- **Obtenção Direta:** Diversos sensores são utilizados para capturar os gestos que estão sendo realizados pelo usuário.
- **Obtenção Indireta:** Geralmente, neste caso, os sensores utilizados são microfones, os quais capturam informações de áudio que os gestos provocam e, a partir de técnicas de processamento de sinal em tempo real, é possível capturar informações sobre os gestos.
- **Obtenção Fisiológica:** Neste caso, os sensores obtêm informações provenientes do corpo do usuário, como, por exemplo, que músculos estão sendo contraídos para realizar aquele gesto.

No outro extremo do DMI, pode-se classificar o controle sonoro da seguinte maneira apresentada por Schloss (SCHLOSS, 1990):

- **Nível de Timbre:** Controle contínuo sobre parâmetros que irão afetar o timbre da saída sonora, e.g. mudar a frequência de uma nota ou adicionar um efeito.
- **Nível de Nota:** Controle, geralmente discreto, que muda diferentes notas de um instrumento, como em um teclado MIDI.
- **Nível de Processo:** Pode controlar parâmetros já predefinidos, como *samples*, redução de ruído, entre outros.

Entendidas as possibilidades dos extremos do modelo de um DMI (controles gestual e sonoro), o mapeamento será a relação entre eles. Entretanto, existem diversas formas de mapear os parâmetros de entrada e saída. Miranda e Wanderley (2006) classificam as formas de mapeamento como **Implícito** ou **Explícito**. O último ainda pode ser dividido em quatro categorias.

- **Mapeamento Implícito:** Utiliza algoritmos computacionais, como reconhecimento de padrões, redes neurais, entre outros, para mapear as entradas e saídas.
- **Mapeamento Explícito:** utiliza uma estratégia predefinida pelo designer do instrumento.

Este último ainda pode ser destrinchado em três categorias dependendo das escolhas de mapeamento (ROVAN et al., 1997):

- **Mapeamento um-para-um:** Cada parâmetro de entrada está associado com apenas um parâmetro de controle sonoro. Formato mais simples de mapeamento, porém, geralmente, o menos expressivo.
- **Mapeamento Divergente (um-para-muitos):** Cada gesto está relacionado com mais de um parâmetro de entrada. Muitas vezes essa é a escolha utilizada para redução da “sobrecarga de controle humano” (Jordà, 2005).
- **Mapeamento Convergente (muitos-para-um):** Diversos parâmetros de entrada são combinados para um controle mais preciso. Embora seja mais complexo, normalmente este tipo de mapeamento é mais expressivo.

Muitas vezes as interfaces de entrada ou saída estão relacionadas de acordo com uma metáfora gestual específica, sendo assim, a estratégia de mapeamento é muitas vezes escolhida para cada caso (FELS; GADD; MULDER, 2002). Portanto, a falta de um método bem estabelecido para construção de DMIs constitui um problema, já que as possibilidades são vastas (CALEGARIO, 2013).

3.4 Problema do entendimento e engajamento

Ligado ao problema de mapeamento, o entendimento sobre o que o músico que está tocando, por parte do público, pode ser afetado por uma estratégia de mapeamento descuidada. Como já dito anteriormente, nos DMIs a ligação causal entre o gesto e o som não existem como nos instrumentos acústicos, e estão relacionados pelas estratégias de mapeamento escolhidas pelo designer. Tendo isto em mente, diversos pesquisadores levantam a questão que esta liberdade influencia diretamente a compreensão da audiência sobre o que está acontecendo (O'MODHRAN, 2011; SCHLOSS, 2003). Sendo assim, o engajamento da plateia está diretamente relacionado com o entendimento do instrumento que resultará no entendimento da performance.

Isto aparece na literatura como a *transparência* do instrumento. Esta é o grau de entendimento das estratégias de mapeamento, que um instrumento possui, na visão da plateia e do usuário do mesmo. A pesquisa de Fels, Gadd e Mulder (FELS; GADD; MULDER, 2002) nos diz que, paralelamente, no caso dos instrumentos acústicos, dois fatores são fundamentais para a transparência. O primeiro, parte do princípio de que os gestos realizados para se tocar um

instrumento acústico são predeterminados por suas restrições físicas, fazendo com que exista um repertório de gestos que são ensinados ao longo de gerações. O segundo é exatamente a carga cultural que a audiência e o performer trazem em si.

Por exemplo, pelo fato de o piano ser um instrumento já bem estabelecido, ambos o performer e a audiência já esperam o resultado sonoro que o instrumento fará quando o *performer* realiza certos gestos. Com isso, ele fornece uma boa transparência. Ademais, informações visuais – sejam elas fornecidas por *feedback* visual ou pela escolha dos gestos do controle de entrada – colaboram na transparência do instrumento aos olhos da plateia (SCHLOSS, 2003).

Esta transparência é fundamental para aumento da expressividade do artista. Ou seja, se a audiência compreende o que o artista está fazendo, é possível que este consiga a expressividade necessária para a *performance* (FELS; GADD; MULDER, 2002).

O engajamento do músico com um DMI depende dentre diversos fatores. O'Modhrain (O'MODHRAIN, 2011, p. 34) argumenta que, primeiramente, o instrumento deve apresentar um “desafio a ser dominado, estimulando o desenvolvimento de uma virtuosidade”. Ademais, a pesquisa de Calegario (CALEGARIO, 2013) faz um comparativo do engajamento do músico com um grande problema no balanceamento de dificuldade em design de jogos digitais, no qual se o desafio é muito pequeno o usuário não se estimulará para desenvolver suas habilidades e se, por outro lado, for muito difícil, o sentimento de frustração suprime a vontade de melhorar.

Além disso, outras questões relativas ao engajamento são os fatos da ausência de repertório que inspire o músico a se motivar a aprender o instrumento que está ligado a questões de virtuosidade, que serão exploradas mais abaixo.

3.5 Latência

É importante lembrar que como “o tempo é característica central na música” (WANDERLEY; ORIO, 2002, p.72), solucionar os problemas relativos à latência e *jitter* são imprescindíveis para a construção de um instrumento que será utilizado por músicos profissionais.

Para se entender o problema, primeiramente deve-se estar em mente os conceitos de latência e *jitter*:

- **Latência:** atraso temporal que distancia a ação feita no controle de entrada e a resposta obtida no módulo de saída.
- **Jitter:** diferença de latência entre dois eventos.

A pesquisa de Calegario (CALEGARIO, 2013) destaca uma extensa análise feita na pesquisa de Jordà (JORDA, 2005), cuja qual nos diz que diferentes eventos sonoros são compreendidos pelo ouvido humano de formas diferentes, ou seja, os níveis de tolerância da latência dependem da complexidade do resultado sonoro. Enquanto sons simples, como dois cliques, podem ser percebidos em um intervalo de 2 milissegundos (ms), sons mais complexos ainda são percebidos simultaneamente mesmo quando separados de um intervalo de 20 - 30 ms.

Além disso, alguns autores relatam que percussionistas profissionais possuem um controle de notas com precisão de 1 ms em certas técnicas instrumentais, para isso delimitam a tolerância máxima de latência em 10 ms (WESSEL; WRIGHT, 2002).

Em contrapartida, algumas aplicações possuem maior tolerância em relação à latência. Instrumentos que relacionam o controle gestual com controles sonoros contínuos (e.g. aumento de um efeito em um fundo musical) podem não demandar uma baixa latência, contrariamente à sistemas que disparam eventos (e.g. disparo de samples e notas) (CALEGARIO, 2013).

Como discutido na seção 3.1, no caso de sistemas digitais de percussão (caso do sistema desenvolvido), a baixa latência foi identificada como elemento crucial para a experiência do usuário.

3.6 Feedback háptico e relação física entre o músico e o instrumento

Em toda interface – gestual ou não – *feedback* é importante para que o sistema seja responsivo ao seu usuário. i.e., “Toda ação realizada por um humano em uma interface gestual, não importa quão pequena, deve ser acompanhada de algum reconhecimento daquela ação sempre que possível e o mais rápido possível” (SAFFER, 2008, p. 20).

Para os instrumentos musicais, esta relação não é diferente. Estes instrumentos, por demandarem alta precisão espacial e temporal e por estarem inseridos em um meio de comunicação e interação com o seu usuário, também necessita de uma resposta à toda ação realizada pelo músico que o utiliza. Além disso, o uso de *feedback* pode “melhorar a precisão na hora de tocar um DMI” (CALEGARIO, 2013, p. 35).

Para melhor contextualização, seguindo a taxonomia de (Miranda e Wanderley, 2006), *feedback* pode ser classificado em:

- **Primário ou Secundário:** Em que o primário é fornecido ao usuário pelo controle gestual e o secundário pelos módulos de geração sonora.

- **Passivo e Ativo:** *Feedback* passivo é resultante das características físicas do instrumento (e.g. a sensação tátil ao tocar as cordas de um violão). Por outro lado, o *feedback* ativo é a resposta do instrumento à uma ação realizada pelo usuário (e.g. o som produzido pelo instrumento ou a vibração causada pelo som no corpo do instrumento).

O modelo de DMIs (Figura 2.10, página 19) utilizado nesse projeto facilita a distinção entre esses dois tipos de *feedback*.

Como já mencionado anteriormente, em um instrumento acústico, todo o processo de produção sonora é decorrente da mecânica do ar e das limitações físicas do próprio instrumento. Além disso, a relação física entre o instrumento está contida no material do instrumento e na mentalidade construída dominando suas técnicas por meio de um aprendizado motor (PAINE, 2013).

Tomando como exemplo um instrumento de percussão como o bongô, ao mesmo tempo em que o som é produzido pela vibração do ar decorrente do gesto aplicado em sua pele, tanto o *feedback* sonoro quanto a força que a pele faz em resposta ao gesto (*feedback* háptico) estão sendo proporcionados ao músico, já que a produção sonora é uma consequência mecânica da vibração do material do instrumento.

Além disso, esta relação cultural e motora estabelecida pelo músico ao aprender um instrumento permite com que ele possa prever como este irá funcionar antes de tocá-lo, resultante da experiência adquirida. Pesquisadores da área de música chamam estes de **gesto mental** (FREITAS, 2008), o qual incorpora ambas as ideias sonora e do movimento que são e serão realizados pelo músico. Medeiros et al. (MEDEIROS et al., 2014) destacam que esta relação é descrita por (PAINE, 2013) como uma **relação incorporada** (do inglês, *embodied relationship*).

Portanto, a ausência de *feedback* háptico e de uma relação entre o DMI, como artefato físico, e o músico são problemas a serem considerados no desenvolvimento deste DMI.

3.7 Expressividade do instrumento e virtuosidade do performer

Na área musical, a expressividade é indispensável para a comunicação da interpretação. De fato, alguns pesquisadores ainda afirmam que o estudo da música como uma comunicação expressiva é “Um dos temas fundamentais” (LINDSTRÖM et al., 2003, p. 1) nesta área de pesquisa. Os autores ainda levantam diversos achados de outros pesquisadores que descrevem a capacidade da música de “mover a plateia” ou de o performer “tocar com sentimento”.

A expressividade tratada neste projeto é a "expressividade na *performance*" descrita na pesquisa de Dobrian e Koppelman (DOBRIAN; KOPPELMAN, 2006), a qual envolve a forma com que o performer irá tocar a música, ou seja, a expressividade da interpretação daquela composição.

Pesquisas realizadas pelo grupo de pesquisa MustTIC, no qual o autor deste projeto está incluído, destacam que a limitação dos sensores (principalmente relativas à precisão e ausência da captura de nuances gestuais), a latência do sistema e a ausência de *feedback* háptico e visual influenciam na expressividade do performer (SILVA, 2012).

Pode-se considerar três parâmetros, utilizados na pesquisa de Wanderley (WANDERLEY, 2001), fundamentais na escolha de sensores: Sensitividade, Estabilidade e Repetitividade (SMITH, 1993).

- **Sensitividade:** a resolução de um sensor, ou seja, a capacidade de mudança de valores que estão sendo medidos.
- **Estabilidade:** capacidade de um sensor de manter a medida estável, sem que haja flutuações.
- **Repetitividade:** capacidade do sensor de medir os mesmos valores, dadas as mesmas condições do sistema, após ser avaliado diversas vezes (levando em consideração os erros de medida).

Para os controles de entrada, esta precisão irá depender "da resolução do conversor analógico-digital (ADC) e da taxa de amostragem da informação" (JORDA, 2005). Por exemplo, no caso de informações passadas via protocolo serial pelo Arduino, a resolução da informação é de 10-bit, i.e., os valores capturados pelo sensor variam entre 1024 valores – de 0 a 1023. O protocolo MIDI, por outro lado, possui uma resolução de 7-bit por canal, o que resulta em 128 valores diferentes para cada canal. Assim, é possível perceber que uma maior resolução dos sensores escolhidos proporcionará um controle mais refinado dos parâmetros de entrada, aumentando as possibilidades de captura de nuances de controle do instrumento.

Outra preocupação está relacionada com o fato de que alguns sensores não são capazes de capturar informações importantes para o sistema, causando limitação de na descrição dos controles gestuais e, conseqüentemente, na expressividade do sistema (MEDEIROS et al., 2014).

Com isso em mente, alguns autores destacam que sistemas mais complexos – e.g., eletrônicos de consumo como o iPhone ou Wii *Remote* – utilizam a combinação de sensores, de forma complementar, para a captura de nuances de movimento e combinação de gestos (SAFFER, 2008). Entretanto, os sistemas mencionados possuem um alto custo de implementação para serem desenvolvidos, demandando um alto nível técnico do *designer* (muitas vezes desenvolvidos por uma ou mais equipes dentro de uma empresa).

Embora esses sistemas possuam alto custo de implementação para serem desenvolvidos, as APIs disponíveis podem exigir um baixo custo de desenvolvimento para acesso das informações rastreadas. Entretanto, como visto anteriormente, uma avaliação conduzida pelo grupo de pesquisa no qual o autor deste projeto está inserido relata que alguns desses sistemas complexos ainda não solucionam as demandas de captura de nuances de expressividade na construção de DMIs (SILVA, 2012).

Com isto em mente, acredita-se que o uso de múltiplos sensores pode suprir esta carência, utilizando as informações provenientes de diferentes sensores de forma complementar. Desta forma, o sistema proporcionaria informação suficiente sobre o que está sendo capturado. Entretanto, busca-se explorar se essa captura de nuances pode ser atingida com baixo custo de implementação.

Outro fator importante é qualidade e variedade dos módulos de geração sonora. Independentemente do resultado sonoro desejado pelo *designer*, deve existir uma certa variação sonora para corresponder às diferentes nuances de controle gestual. A ausência disso prejudica o *feedback* secundário, mascarando as nuances de controle do instrumento e, conseqüentemente, afetando a sua expressividade.

Portanto, é importante que o controle gestual possua uma alta granularidade e a capacidade de capturar nuances, que os módulos de saída possuam uma boa qualidade sonora e variações de timbre correspondentes a essas nuances de controle e que o mapeamento seja desenvolvido de forma clara, tendo em mente as limitações dos sensores escolhidos.

Outro fator que contribui para a performance é o virtuosismo, ou seja, a capacidade técnica e musical que um músico possui com o instrumento utilizado. Não somente o domínio das técnicas, mas “todas as capacidades daquele instrumento com relativa facilidade” (DOBRIAN; KOPPELMAN, 2006, p. 3). Entretanto, para alcançar este nível de domínio, músicos devem estudar e praticar as tarefas específicas do instrumento durante vários anos. De fato, músicos novatos passam por diversas dificuldades tanto físicas quanto cognitivas ligadas ao aprendizado de um novo instrumento que são superadas com a prática (WESSEL; WRIGHT, 2002).

Como se o tempo necessário para se dominar um instrumento não fosse problemático o bastante, o engajamento de um músico com um instrumento (acústico ou digital) envolve, muitas vezes, a aproximação deste músico com um repertório que demonstre as capacidades expressivas e busque “amplificar as características daquele instrumento” (PAINE, 2013, p. 77) ou pelo simples fato do som produzido pelo instrumento ser agradável ao músico – meramente pelo timbre ou por uma *performance* de um virtuoso daquele instrumento – (MEDEIROS et al., 2014).

Ademais, pelo fato de os usuários necessitarem de tempo de prática ao engajar em um novo instrumento, é difícil distinguir os problemas que estão associados com o design do instrumento e as barreiras iniciais de aprendizado (MEDEIROS et al., 2014).

3.8 Contextos de uso

Outro critério importante a ser considerado no design de um DMI é o contexto no qual ele será utilizado. Diferentes contextos e perspectivas – perspectivas do performer, da plateia, do designer e do fabricante – irão necessitar que certos parâmetros sejam alterados e certos critérios de design sejam atingidos (MEDEIROS et al., 2014; O'MODHRAN, 2011). Por exemplo, um DMI desenvolvido para ser tocado em conjunto com outros instrumentos deve possuir algumas características como maiores amplitude e coesão sonoras que, se for desenvolvido para ser tocado solo, este mesmo DMI pode possuir uma amplitude um pouco menor e uma variação de timbres mais abrangente. Outros contextos envolvem a utilização do DMI para o estudo, para *performance* ao vivo, para tocar um tipo de música específico, entre outros. Com isto “o processo de *design* é mais difícil quando deve-se considerar múltiplos contextos que podem influenciar a experiência do usuário” (MEDEIROS et al., 2014, p. 7).

Os critérios de sucesso de um instrumento estão ligados a diversos fatores intrínsecos e extrínsecos. Os fatores mencionados anteriormente neste capítulo contribuem para a aceitação do instrumento, entretanto esta aceitação vai muito além das características do instrumento. Aspectos culturais, como a criação de um estilo musical próprio para aquele instrumento, comercialização e outros fatores também influenciam a adesão de novos usuários (MEDEIROS et al., 2014). Isto dificulta a avaliação destes DMIs, já que a adesão a um instrumento pode levar anos e, às vezes, não ser aceito pelo público geral.

Contrariamente, os instrumentos acústicos possuem milhares de anos de evolução e adesão cultural que determinaram o sucesso dos mesmos. Schloss (2003) ainda argumenta que apenas agora a música computacional atingiu um nível que pode ser considerado “maduro” o suficiente para ser explorada e utilizada em *performances* ao vivo.

3.9 Escopo da pesquisa

Neste capítulo, foram apresentados diversos desafios, relativos ao *design* de DMIs, levantados durante a revisão da literatura. Fatores como latência, transparência no mapeamento, limitações dos sensores, *feedback* háptico, entre outros, não somente afetam a experiência do usuário, mas estão diretamente relacionados com a capacidade de expressão musical do instrumento desenvolvido.

Este trabalho pode ser visto como uma continuação da pesquisa de Silva (SILVA, 2012), realizada no mesmo grupo de pesquisa (MusTIC), cujos resultados são apresentados na seção 3.1. Assim, os desafios abordados neste projeto focam nos problemas da **Latência** do instrumento, do **Feedback Háptico** e principalmente da **Expressividade** do instrumento e **captura de nuances gestuais**.

Particularmente, o problema do *feedback* abordado aqui é relativo apenas à necessidade levantada pelos percussionistas profissionais por uma superfície de contato que proporcionasse informações visuais e aspectos físicos do instrumento que está sendo tocado.

Sendo assim, este projeto pode ser considerado como uma nova iteração no processo de *design* de DMIs, buscando usar os princípios do processo de design utilizados em outra pesquisa do mesmo grupo (CALEGARIO, 2013) e visando o **baixo custo de implementação computacional** e a **integração de múltiplos sensores**.

4 ESTADO DA ARTE

Este capítulo descreve alguns avanços da área que buscam resolver os problemas descritos no capítulo anterior, bem como apresenta projetos similares na criação de interfaces artísticas (principalmente musicais), destacando o desenvolvimento de novos DMIs para percussão, qual foi a abordagem utilizada para isto e se houveram resultados relevantes às dificuldades encontradas.

4.1 Latência

Apesar do tempo ser característica central na música, poucos trabalhos científicos abordam, especificamente, métodos para solução dos problemas de latência na construção de instrumentos musicais. Entretanto, os estudos relativos à tolerância de latência em instrumentos musicais (JORDA, 2005; WESSEL; WRIGHT, 2002) e sistemas operacionais (WRIGHT; CASSIDY; ZBYSZYNSKI, 2004) fornecem parâmetros para *guidelines* na construção de DMIs.

Algumas pesquisas que abordam soluções para os problemas de latência se baseiam em um alto custo de implementação. Uma destas abordagens focou no desenvolvimento completo de um novo sintetizador digital na linguagem de programação Java (AUERBACH et al., 2007). Outro utiliza linguagens de descrição de *hardware* para programar um circuito integrado FPGA (Field Programmable Gate Array) que fornece módulos que transmitem informações de áudio e MIDI, módulos que possuem conversores analógico-digitais (ADC, do inglês *Analog-to-Digital Converter*) e módulos que transmitem informações gestuais por meio de pacotes User Datagram Protocol (UDP) – protocolo OSC que será descrito mais a frente – (WESSEL; WRIGHT, 2002).

A pesquisa feita por Bååth aponta outras possíveis soluções para os problemas de latência na construção de DMIs (BÅÅTH, 2011). Primeiramente, uma das abordagens é a utilização do protocolo MIDI para “disparar de sons pré-carregados ou registrar as batidas do usuário”, embora ainda haja problemas relativos a latências internas dos sistemas e APIs desenvolvidos (BÅÅTH, 2011, p. 2). A outra abordagem se baseia em utilizar tecnologias que operam em um nível próximo ao hardware especificamente para estas tarefas. Por exemplo, tecnologias como o Arduino operam sem a necessidade de sistemas operacionais e utilizam configurações de simples implementação.

4.2 Feedback Háptico

Pesquisas na área de DMIs buscam, frequentemente, descobrir novas formas de aprimorar controladores digitais já existentes e atribuir novas características aos seus conceitos para que o resultado final crie mais expressividade musical. Desta forma, a busca de implementação de *feedback* háptico em controladores musicais pode não só produzir informações visuais que melhorariam a relação

do músico com o instrumento (SILVA, 2012), mas também podem melhorar a precisão com a qual este é tocado (O'MODHRAIN; CHAFE, 2000).

Referente à esta última, Nichols buscou desenvolver um sistema que melhorasse o *feedback* de violinos MIDI (NICHOLS, 2002). Partindo de uma dificuldade pessoal ao utilizar este tipo de sistema, foi possível notar que as nuances gestuais, que são realizadas ao se tocar com o arco do violino, não eram devidamente capturadas utilizando as informações do protocolo MIDI – *pitch* e *velocity*. Para resolver este problema, foi desenvolvido o *vBow*: um sistema em acrílico que incorpora o arco do violino e possui 4 servo-motores para introduzir o *feedback*. Utilizando uma maneira similar de obtenção dos movimentos, outras pesquisas também implementaram *feedback* háptico em violinos (GROSSHAUSER; HERMANN, 2009).



Figura 4.1: o vBow (Extraído de Nichols, 2002).

O sistema simula cordas virtuais, num *software* de síntese musical, para que os servo-motores introduzidos capturem informações do quanto o arco está sendo movido – usando as rotações do eixo do motor, a rotação do arco paralelamente às cordas, a distância acima das cordas e a distância ao longo do corpo do violino, melhorando, assim, a capacidade de controle gestual do instrumento e uma maior expressividade.

Outro sistema que inclui *feedback* háptico para diversas aplicações (incluindo aplicações musicais) é o *MudPad*. O *Mudpad* é uma superfície multi-toque que possui quatro camadas sobrepostas para produzir diferenças de *feedback* háptico – variações de interações com objetos virtuais e texturas (JANSEN; KARRER; BORCHERS, 2010).

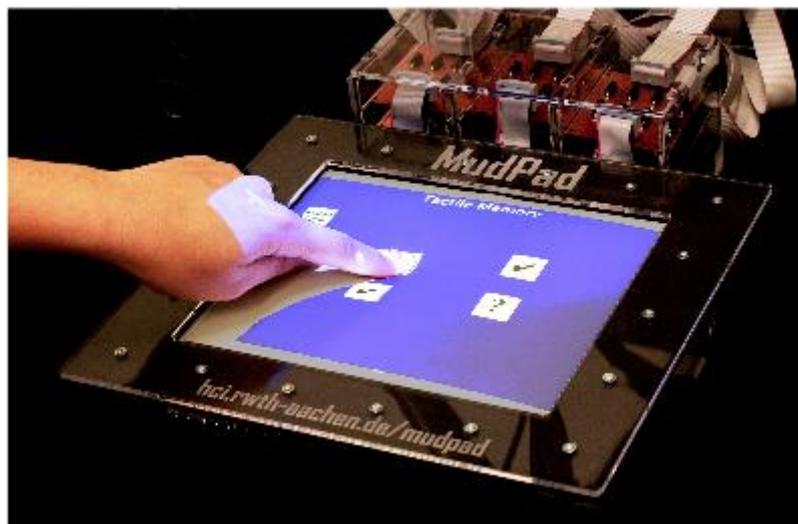


Figura 4.2: MudPad (Extraído de Jansen, Karrer e Borchers, 2010).

A camada mais baixa é composta por diversos eletroímãs que geram um campo magnético local. Acima desta, existe uma camada multi-toque resistiva, uma camada que possui um fluido magnético - *magnetorheological fluid* – e a última camada de látex branco para a projeção. Variando o campo magnético, é possível criar diferentes *feedbacks* e texturas para diferentes aplicações: rígida, pulso único, constantemente vibrando e macia. As aplicações podem variar de aplicações mais comuns de interação computacional a instrumentos musicais como pianos e sequenciadores virtuais mais expressivos. Além disso, como a tela multi-toque resistiva permite que o usuário toque levemente sobre a mesma, o sistema também pode utilizar o *feedback* da textura sem ativar o *trigger* da tela (JANSEN; KARRER; BORCHERS, 2010).

Outras pesquisas utilizaram *feedback* háptico a partir de um transdutor – dispositivo que transforma um tipo de energia em outro – (MONTAG et al., 2011) em mesas multi-toque. Nesta pesquisa, os autores ainda destacam outros sistemas que utilizam *feedback* háptico em mesas multi-toque. Ademais, outros sistemas apresentam este tipo de *feedback* em um sistema de joysticks e mouse (STEINER, 2004) para controle de parâmetros musicais.

Vale lembrar que, como visto no capítulo anterior, pesquisas descrevem a preocupação dos músicos com o *feedback* (SILVA, 2012). Entretanto, esta preocupação é, a priori, apenas para localização espacial do instrumento.

4.3 Expressividade musical

Esta seção busca demonstrar tecnologias recentes de interfaces digitais e seu impacto na expressividade musical. Aqui são apresentadas as tecnologias que foram desenvolvidas para solucionar problemas de expressividade musical. Além disso, alguns DMIs são apresentados de forma a exemplificar como estas

tecnologias são utilizadas em projetos acadêmicos, pessoais e/ou industriais e alguns resultados obtidos na avaliação de tecnologias existentes.

4.3.1 *Protocolos de comunicação*

Como já descrito na seção 2.2.2, a criação do protocolo MIDI proporcionou diversos avanços na criação de novas tecnologias musicais digitais. A capacidade de processar as informações enviadas (por meio desse protocolo) de qualquer maneira, utilizando *softwares* computacionais, proporciona diversos avanços para a expressividade musical.

Entretanto, algumas pesquisas acadêmicas levantam problemas relacionados com o protocolo MIDI (MCMILLEN, 1994; MOORE, 1988; WRIGHT, 1994). Segundo essas pesquisas, o protocolo MIDI não é preparado para lidar com variações de nuances de controle. Não apenas a largura de banda da comunicação desse protocolo é insuficiente para transmitir informações I/O de notas, mas também é insuficiente para reproduzir informações contínuas de múltiplas características desses instrumentos simultaneamente.

Segundo Wright, as mensagens MIDI podem ser divididas em duas categorias. A primeira categoria enquadra as mensagens I/O de notas e os chamados *aftertouches* (que especificam parâmetros como a pressão exercida em cada nota) são informações que atuam especificamente no nível da nota. A segunda categoria inclui modulações, controle de efeitos e outros parâmetros que são aplicados aos canais (não apenas às notas).

Além disso, os eventos MIDI são passados por meio de sequências, o que influencia na latência da chegada dessas informações ao computador. McMillen exemplifica que um acorde de 10 notas em um teclado musical é passado ao computador com um atraso de cerca de 6,7 milissegundos (ms), uma latência que é próxima à tolerância proposta pela pesquisa de Wessel e Wright para apenas duas notas consecutivas (WESSEL; WRIGHT, 2002).

Com intuito de resolver essas restrições do protocolo MIDI, o protocolo *Open Sound Control* (OSC) foi desenvolvido em 1997, por Adrian Freed e Matthew Wright, para prover informações musicais entre computadores, sintetizadores sonoros e instrumentos, adaptando esse protocolo de informações às tecnologias de rede da época (e.g. comunicações Ethernet) (WRIGHT, 2005).

Além disso, esse protocolo envia informações através de pacotes User Datagram Protocol (UDP), por meio de um padrão de endereços similares às URLs, contendo argumentos, cujos tipos podem variar entre int32, float32, Osc-string (composto primariamente por caracteres ASCII), Osc-timetag (em números 64-bit), cores RGBA (em 32-bit), entre outros. Essas informações são enviadas por meio de mensagens OSC, que transmitem um endereço seguido dos argumentos (ou seguido de nenhum argumento) (WRIGHT, 2002). Ademais,

a capacidade do protocolo OSC de enviar e atualizar informações em múltiplos canais com uma maior resolução fornece uma maior possibilidade de processamento de nuances de controle.

Os protocolos MIDI e OSC são os mais comumente utilizados na construção de DMIs, devido à padronização na transmissão de informações e a grande variedade de sistemas que incorporam essas tecnologias, entre elas controladores, instrumentos e softwares digitais.

4.3.2 *Softwares de mapeamento*

Softwares de mapeamento foram criados para auxiliar no desenvolvimento de sistemas digitais, como, por exemplo, o MAX/MSP²⁰, Puredata²¹ (Pd), OSCulator²², libmapper, o Ableton Live (*software* de composição musical, mas que possui *features* de mapeamento), entre outros.

Diversas ferramentas de programação emergiram para aproximar leigos da possibilidade de criar sistemas de forma mais prática e rápida. Ferramentas como o MAX/MSP e o puredata ajudam iniciantes a romper as barreiras iniciais de aprendizado de sintaxe de linguagens, fornecendo uma interface de programação gráfica baseada em fluxo de informações e módulos lógicos – caixas que possuem objetos computacionais como operadores, estruturas de controle, entre outros.

²⁰ Informações disponíveis em: <<https://cycling74.com/products/max/>>.

²¹ Informações disponíveis em: <<http://puredata.info/>>.

²² Informações disponíveis em: <<http://www.osculator.net/>>.

que disponibiliza uma lista de parâmetros de entrada e controles de saída dos dispositivos e softwares conectados ao computador (CALEGARIO, 2013). Apesar da desvantagem de ser um sistema que envolve um nível técnico avançado para ser manipulado, o *libmapper* oferece a possibilidade de modificar o mapeamento do projeto sem a necessidade de compilar códigos e reiniciar outros sistemas. Além disso, esse sistema fornece integração com softwares de mapeamento como o MAX/MSP.

Recapitulando, esses softwares de mapeamento fornecem uma facilidade de visualização dos parâmetros de controle e de conexão entre esses parâmetros de entrada e controles sonoros e, conseqüentemente, facilitando as possibilidades de mapeamento (descritas no capítulo 3, seção 3.3), deixando a escolha das estratégias de mapeamento nas mãos do *designer*, e proporcionando maiores possibilidades para expressividade do sistema.

4.3.3 Síntese sonora

Como já mencionado anteriormente, a expressividade do instrumento não seria possível sem que os módulos de produção sonora possuam alta qualidade e variedade sonoras (independentemente do tipo de som que se deseja produzir). Para isso, existem diversas formas de síntese sonora e programas dedicados à composição musical.

No livro *The Computer Music Tutorial*, Curtis Roads faz um estudo das diversas formas de síntese sonora. Eles são compostos por uma combinação de elementos geradores de onda (e.g. osciladores) e modificadores (filtros, amplificadores, geradores de envoltória, etc.), além de existir casos específicos como o uso de *samples*.

Pode-se destacar cinco grandes formas de síntese: **Aditiva**, **Subtrativa**, **Síntese FM**, **Amostragem Digital**, **Modelagem Física** (ROADS, 1996).

- **Síntese Aditiva:** Neste caso, os sons são sintetizados seguindo o princípio das séries de Fourier, no qual pode-se descrever qualquer onda como uma soma de ondas cujas frequências são múltiplos inteiros da onda que se deseja formar. Assim, módulos de geradores de onda com frequências específicas são combinados para produzir o som desejado.
- **Síntese Subtrativa:** Este tipo de síntese utiliza sinais de áudio complexos (e.g. ruídos) como ponto de partida. A partir deste ponto, elementos modificadores são aplicados à amostra inicial para que o resultado sonoro seja alcançado. Este método era bastante utilizado em sintetizadores analógicos tradicionais.
- **Síntese FM:** Este caso envolve a manipulação de frequência de duas ou mais ondas relacionadas por um modelo matemático. Nesse modelo, a variação de uma certa onda (chamada de onda portadora) pode ser modificada por meio da variação de frequência de uma segunda onda

(chamada onda moduladora). Além disso, diferentes possibilidades se colocam: as ondas portadoras e moduladoras podem ser uma resultante de uma combinação de ondas; a onda moduladora pode ser uma outra onda portadora; e a portadora pode ser uma outra onda moduladora.

- **Síntese por Amostragem Digital:** Esta síntese possui duas vertentes. Um dos casos é a utilização de *samples* (ou seja, amostras sonoras previamente gravadas de instrumentos musicais), disparando-os conforme forem requisitados ou utilizando-os em laços (*loops*). A outra vertente é utilizar uma leitura de múltiplas tabelas. Essas tabelas contêm diversos módulos sonoros (envoltórias de onda, filtros, *samples*, entre outros) e combinam estas informações para gerar o resultado final.
- **Síntese por Modelagem Física:** Neste método, equações matemáticas são utilizadas para descrever a mecânica de produção sonora de um instrumento musical. Este tipo de modelo descreve os aspectos físicos do instrumento musical desejado (geração e acústica). Além disso, da mesma forma como a luz pode ser considerada como onda e partícula, a modelagem também pode considerar o som como partículas (síntese granular).

Estes métodos fornecem formas de geração sonora por intermédio de sistemas digitais. Entretanto, Roads ainda descreve alguns problemas na utilização destes métodos de síntese sonora. A maioria deles envolvem um alto custo computacional ou de implementação para produzirem resultados de alta fidelidade.

No primeiro caso, as sínteses aditiva, subtrativa e FM envolvem um grande número de módulos sonoros que acarretam numa alta demanda de processamento por parte do computador, tornando estes sistemas pouco viáveis. No caso da síntese por modelagem física, mesmo sendo o sistema que produz a maior fidelidade sonora, esse método envolve a elaboração de algoritmos complexos que demandam de um alto nível técnico por parte do desenvolvedor. Ademais, a utilização de *samples* possui problemas relativos a mudanças de timbre que são solucionados com transposição das amostras.

Sendo assim, a utilização de amostras pré-gravadas fornece alto nível de fidelidade musical com baixo nível de implementação. Para ampliar a expressividade sonora do instrumento utilizando esse método, é necessário registrar as diferentes nuances sonoras que corresponderão às nuances gestuais.

Diversos aplicativos digitais como o Ableton Live, GarageBand²³, FL Studio²⁴, Logic Pro²⁵, Sonic Visualizer²⁶, entre outros, incorporam parâmetros sonoros, sintetizadores virtuais, efeitos e amostrar MIDI ou de áudio, que podem ser utilizados para fornecer expressividade sonora ao sistema que será desenvolvido.

4.3.4 Sensores

Como visto no capítulo 2, os DMIs possuem controles gestuais que são desenvolvidos utilizando sensores. Avaliando os sensores utilizados pela comunidade científica na conferência *New Interfaces for Musical Expression* (NIME), cujo tema central também envolve a criação de novos DMIs, pode-se classificar os tipos de sensores da seguinte maneira (MEDEIROS; WANDERLEY, 2014a; MIRANDA; WANDERLEY, 2006):

- **Sensores analógicos:** Sensores que têm como parâmetros de saída informações contínuas de sinais elétricos.
- **Sensores digitais:** Contrariamente aos sensores analógicos, os parâmetros de saída de sensores digitais são informações discretas.
- **Eletrônicos de consumo²⁷:** Sensores desenvolvidos por empresas e comercializados para o público como smartphones, tablets, o Wii[®] e o Kinect[®].
- **Captura de movimentos:** Normalmente sensores deste tipo utilizam câmeras – RGB, infravermelho (IR), etc. Esta categoria inclui sensores como o Kinect[®], o Xtion[®], o Xsens[®], entre outros.

A pesquisa de Medeiros e Wanderley (MEDEIROS; WANDERLEY, 2014a) faz um levantamento dos sensores utilizados durante os anos de 2001 a 2013 da conferência *New Interfaces for Music Expression* (NIME) juntando os resultados realizados em pesquisas anteriores (MARSHALL, 2008).

Existe uma diversidade enorme de sensores que podem ser utilizados para capturar diferentes tipos de informações. Estes podem variar entre acelerômetros, giroscópios, sensores FSR (*Force Sensing Resistor*), discos piezoelétricos, sensores termoeletrônicos, sensores capacitivos, botões, potenciômetros, câmeras RGB, sensores infravermelho (IR, do inglês *Infrared*), sensores de luz, entre diversos outros. Esses sensores captam diferentes informações como orientação, aceleração, pressão, presença ou ausência de luz, posição, umidade, etc.

²³ Informações disponíveis em: <<https://www.apple.com/br/mac/garageband/>>.

²⁴ Informações disponíveis em: <<https://www.image-line.com/flstudio/>>.

²⁵ Informações disponíveis em: <<https://www.apple.com/br/logic-pro/>>.

²⁶ Informações disponíveis em: <<http://www.sonicvisualiser.org/>>.

²⁷ Traduzido de consumer eletrônicos.

Sensors	Occurrence (2009–2013)	Occurrence (2001–2008)
accelerometer	75 (30)	56
FSR™ (Force Sensing Resistors™)	38	68
buttons and potentiometers	29	110
gyroscope	30 (9)	
video/image	23	54
IR (infrared)	22	27
magnetometer	16 (4)	
capacitive	15	
biosensing	13	
piezoelectric disc	12	
non-definable	12	
microphone	11	29
textiles	11	
photo/light	10	
bend	9	21
Hall effect	7	
ultrasound	4	
pressure/flow	4	
fiber optic	2	

Tabela 2: Sensores que aparecem em artigos da conferência NIME (Extraído de MEDEIROS; WANDERLEY, 2014).

Os autores classificam separadamente eletrônicos como smartphones, tablets, Wii Remote®, Kinect®, Leap Motion® e sistemas de captura de movimentos (motion capture). Enquanto eletrônicos de consumo aparecem 71 vezes nas pesquisas, os sistemas de motion capture aparecem apenas 30 vezes.

Essa enorme quantidade de sensores fornece a possibilidade de captura de diferentes aspectos do ambiente para desenvolver o controle de entrada. Ademais, a combinação de sensores (ou integração de múltiplos sensores, como veremos no capítulo 5) proporciona um maior número de parâmetros de controle que podem ser utilizados para distinguir gestos. A pesquisa de Medeiros e Wanderley (MEDEIROS; WANDERLEY, 2014a) ainda oferece informações sobre a frequência com que estes sensores são utilizados.

Além disso, a comunidade adepta aos movimentos *Maker*, *Do-It-Yourself* (DIY) e tecnologias similares ao Arduino geralmente fornece módulos prontos e instruções de montagem dos circuitos eletrônicos que facilitam a experimentação e utilização desses sensores sem a necessidade de um alto conhecimento sobre instrumentação eletrônica.

É possível encontrar, na literatura, diversos usos desses sensores na construção de DMIs. Alguns sistemas envolvem o uso de sensores acelerômetros e giroscópios, acoplados às baquetas de uma bateria, para simular virtualmente partes de uma bateria (KANKE et al., 2012), a combinação de sensores flexíveis (*bend*) e piezoelétricos para criação de um sistema modular (WEINBERG; AIMI; JENNINGS, 2002) e o uso de uma combinação de sensores com algoritmos de aprendizagem de máquina para aumentar a capacidade expressiva de instrumentos convencionais (como o Sensory Percussion²⁸, desenvolvido pela empresa Sunhouse©).

Principalmente nos últimos anos, os números de sistemas musicais com eletrônicos de consumo cresceram nos trabalhos acadêmicos. Com exceção do Wii (que houve um decaimento de 2012 para 2013), vários DMIs estão sendo desenvolvidos com essas interfaces gestuais (MEDEIROS; WANDERLEY, 2014a).

No caso do Wii *Remote*, diversos sistemas musicais foram construídos utilizando o próprio Wii *Remote* (HEISE; LOVISCACH, 2008; WONG; YUEN; CHOY, 2008; TRALDI, AGUIAR, BARREIRO, 2011), a barra de captura IR (PENG; GERHARD, 2009), ou ambos (MILLER; HAMMOND, 2010). Entretanto, algumas dessas pesquisas mostram a ausência de controle fino dessas interfaces (WONG; YUEN; CHOY, 2008; TRALDI, AGUIAR, BARREIRO, 2011).

Diversos sistemas desenvolvidos com o Kinect abordam a simulação de instrumentos de percussão. Todas essas pesquisas demonstraram as limitações do sensor para ser aplicado profissionalmente. A latência (HSU et al., 2013; ODOVICHUK et al., 2011; ROSA-PUJAZÓN et al., 2013; TODOROFF; LEROY; PICARD-LIMPENS, 2011) e o erro no reconhecimento gestual (HSU et al., 2013; OKADA et al., 2014; TODOROFF; LEROY; PICARD-LIMPENS, 2011) foram as principais restrições levantadas.

O uso de interfaces multi-toque também é bastante explorado nos trabalhos acadêmicos. Trabalhos mencionados anteriormente como o *reactTable* (JORDA et al., 2005), o *MudPad* (JANSEN; KARRER; BORCHERS, 2010) e as pesquisas de Montag et. al. (MONTAG et al., 2011), Davidson e Han (DAVIDSON; HAN, 2006) e Calegário (CALEGÁRIO, 2013) exploram a capacidade de controle musical com estas interfaces gestuais.

²⁸ Informações em: <<http://sunhou.se/>>.

Poucos são os sistemas que utilizam o Leap Motion© na construção de interfaces de interação musical. O AirHarp²⁹ e o AirBeats³⁰ são DMIs desenvolvidos pela empresa Handwavy que visam simular instrumentos convencionais, respectivamente a harpa e instrumentos de percussão variados. Outros sistemas que buscam simular instrumentos acústicos são o Virtual Bowing³¹, o Theremix (BARBOSA et al., 2014), o aplicativo *Bongos!*³², o Air Piano (BARBOSA et al., 2014) e o trabalho de De Pra et. al. (DE PRA et al., 2014). Embora alguns desses trabalhos apresentem alguns resultados negativos relativos ao uso desse sistema, é possível que o uso de integração de sensores solucione os problemas encontrados (relativos à latência). Além disso, ferramentas de auxílio na comunicação do Leap Motion© com outros programas, como o GECO³³, possuem *features* interessantes, e.g. envio das informações do dispositivo por meio de protocolo OSC, para a construção de DMIs.

Além disso, empresas como a Roland©, Korg© e Alesis© também desenvolveram DMIs para comercialização. Baterias eletrônicas (como a DM10X Mesh Kit³⁴) e *pads* de percussão, como o HandSonic³⁵ e o Wavedrum³⁶, estão incluídos nesse conjunto de controladores MIDI.

²⁹ Vídeo disponível em: <<https://www.youtube.com/watch?v=U2HFcfOwDtl>>.

³⁰ Vídeo disponível em: <<https://www.youtube.com/watch?v=fU4ZWAPwvcM>>.

³¹ Vídeo disponível em: <<https://www.youtube.com/watch?v=kLkPvmr93K8>>.

³² Disponível em: <<https://apps.leapmotion.com/apps/bongos/windows>>.

³³ Disponível em: <<https://apps.leapmotion.com/apps/geco-midi/windows>>.

³⁴ Informações em: <<http://www.alesis.com/dm10-x-mesh>>.

³⁵ Informações em: <http://roland.com.br/products/handsonic_hpd-20/>.

³⁶ Informações em: <http://www.korg.com/us/products/drums/wavedrum_global_edition/>.

5 ABORDAGEM ADOTADA

Neste capítulo, são descritos os princípios utilizados durante a construção do DMI desenvolvido neste projeto, assim como é apresentado o processo de design seguido ao longo da pesquisa. Procura-se, aqui, deixar claro os métodos escolhidos para investigação do problema de pesquisa e desenvolvimento de uma possível solução.

5.1 Princípios utilizados

Esta seção apresenta os princípios seguidos na elaboração da solução para o problema de pesquisa e quais as diretrizes adotadas para guiar as escolhas realizadas durante o projeto.

5.1.1 Integração de múltiplos sensores

Hall e Llinas (HALL; LLINAS, 1997, p.1) argumentam que:

“Humans and animals have evolved the capability to use multiple senses to improve their ability to survive. For example, it may not be possible to assess the quality of an edible substance based solely on the sense of vision or touch, but evaluation of edibility may be achieved using a combination of sight, touch, smell, and taste. Similarly, while one is unable to see around corners or through vegetation, the sense of hearing can provide advanced warning of impending dangers. Thus multisensory data fusion is naturally performed by animals and humans to achieve more accurate assessment of the surrounding environment and identification of threats, thereby improving their chances of survival.” (HALL; LLINAS, 1997, p.1)

Embora o foco dos autores seja em utilizar este conceito para corroborar a técnica de *fusão de sensores* utilizada por eles, o princípio também pode ser utilizado para as técnicas de *integração de múltiplos sensores*.

A diferença entre estes termos técnicos é uma linha tênue, mas importante. Enquanto o último é utilizado para descrever um processo de combinar as informações de diversos sensores diretamente (de forma sinérgica), o termo técnico “ *fusão de sensores*” (*sensor fusion*) descreve uma técnica computacional, na qual informações capturadas por diversos sensores são combinadas para gerar novas informações aprimoradas (e.g. para reduzir erros de medida) que serão utilizadas pelo controle de entrada (MEDEIROS;

WANDERLEY, 2014a). A pesquisa de Medeiros e Wanderley ainda faz um levantamento da escolha do uso de sensores na conferência NIME (no período entre 2008-2013), juntando às informações do *survey* conduzido por Marshall (MARSHALL, 2008) e mostrando a frequência de se utilizar sensores combinados entre si (como visto na Figura 4.4, na página 49). Os resultados apresentados pelos autores mostram que, por um lado, projetos que utilizam sensores mais comuns – acelerômetros, sensores resistivos FSR (*Force Sensing Resistor*), giroscópios, sensores de luz – possuem um maior grau de utilização de múltiplos sensores. Entretanto, poucos são os trabalhos que utilizam eletrônicos como o Wii® e o Kinect® com outros sensores variados.

Mesmo que estes eletrônicos possuam diversas capacidades sensoriais – o Wii reconhece a cena a partir de informações obtidas por uma câmera IR e um acelerômetro e o Kinect possui um sensor de profundidade, uma câmera RGB e microfones –, presume-se que as dificuldades encontradas em diversos trabalhos acadêmicos serviriam de motivação para utilizar outros sensores que as sanassem. Vale lembrar a frase de Bill Buxton (BUXTON, 2007) que “tudo é melhor para certas coisas e pior para outras”.

Na literatura, algumas pesquisas relacionadas com integração de múltiplos sensores argumentam que existem vantagens – uso de redundância de informação para redução de erros, informações complementares, menor custo na obtenção de informações, entre outros – no uso deste tipo de sistema em comparação com sistemas que utilizam um único sensor (LUO; KAY, 1990). O uso de informações provenientes de dois ou mais sensores, de forma complementar, ajuda a descrever o ambiente com maior precisão. Algumas pesquisas utilizam essa abordagem para melhor rastreamento de um objeto na cena (MITICHE; AGGARWAL, 1986) ou para redução de erros de medida (HALL; LLINAS, 1997). Outras pesquisas seguem a mesma filosofia para o uso de “fusão de sensores” (MEDEIROS; WANDERLEY, 2014b), utilizando métodos que exigem um maior custo de implementação, já que envolvem a preparação e combinação de dados por meio de modelos matemáticos e estatísticos.

Poucas pesquisas relacionam a utilização de múltiplos sensores para captura de nuances de controle em interfaces gestuais. Enquanto alguns autores relacionam o uso de múltiplos sensores com sistemas mais complexos (como eletrônicos de consumo) (SAFFER, 2008), outras pesquisas acreditam que sistemas mais robustos são necessários, utilizando abordagens de fusão de sensores (MEDEIROS; WANDERLEY, 2014b). Entretanto, ambos os casos requerem um maior nível técnico e um maior custo de implementação.

Neste projeto, buscou-se aderir a filosofia de utilizar informações complementares provindas de diferentes sensores para resolver problemas de expressividade musical. Utiliza-se integração de sensores para capturar nuances gestuais que são essenciais na expressividade musical diretamente a partir dos

dados obtidos pelos sensores. Entretanto, buscou-se investigar se esta abordagem é válida em um sistema simples que requer baixo custo de implementação.

5.1.2 *Baixo custo de implementação*

A filosofia dos movimentos “faça você mesmo” (DIY, do inglês *Do It Yourself*) e *Maker* encorajam entusiastas a buscar conhecimento e técnicas para realizar projetos, reparar e/ou modificar objetos, “sem a ajuda de profissionais pagos” (KUZNETSOV; PAULOS, 2010). Kuznetsov e Paulos realizaram uma pesquisa com comunidades DIY, levantando suas motivações na realização de projetos e interação com a própria comunidade, as quais envolvem a divulgação de ideias, compartilhamento de técnicas e instruções de como realizar determinadas tarefas.

No contexto computacional, estas comunidades incorporam os princípios de software livre (Open Source) e buscam compartilhar técnicas e conhecimentos sobre diversos assuntos e, muitas vezes, disponibilizar os códigos de seus sistemas em plataformas desenvolvidas para tal. Essa comunidade oferece diversos módulos e códigos de softwares como o MAX/MSP e o puredata para auxiliar o engajamento de iniciante.

Como visto no capítulo 4, a linguagem de programação por fluxo de dados baseada em módulos oferece a possibilidade de o usuário interagir com parâmetros de controle mais próximos a ele, ao mesmo tempo que os “protege” de manipulações necessárias no contexto computacional que requerem um nível técnico avançado (como uma abordagem de “caixa preta”). Mesmo assim, fornece a possibilidade de o desenvolvedor buscar estas informações em camadas de abstração mais baixas, para que, à medida que forem aprendendo (ou se já possuírem um nível de programação mais avançado), possam entender como cada módulo funciona.

Além disso, empresas como o Arduino, Raspberry Pi, Intel, entre outras, criaram iniciativas para o desenvolvimento de microcontroladores que facilitam a criação de projetos eletrônicos, disponibilizando sistemas com interfaces intuitivas e práticas, facilitando a prototipação de projetos eletrônicos e computacionais, e, conseqüentemente, auxiliando na introdução de usuários leigos em um ambiente de programação.

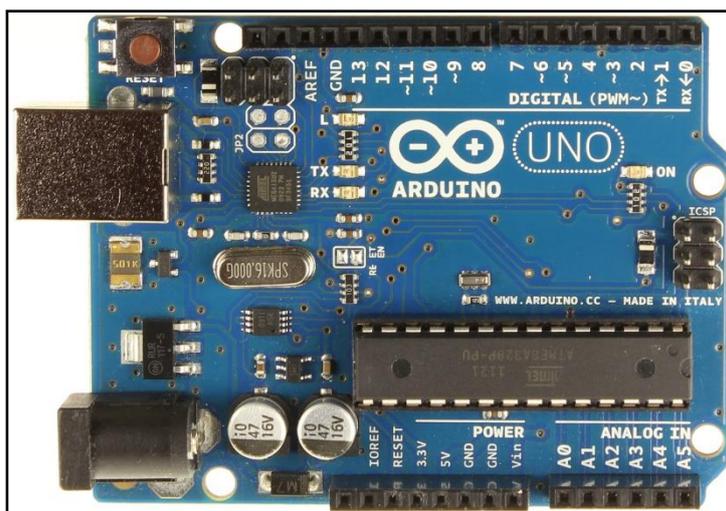


Figura 5.1: Microcontrolador arduino UNO.

Neste projeto, foi escolhida uma abordagem aproveitando estes princípios de facilidade na implementação visando uma possível extensão do projeto, na qual o músico seria capaz de personalizar o formato da superfície do instrumento, possivelmente modificando alguns parâmetros de mapeamento dos sensores.

Ademais, de forma a tornar o artefato mais acessível, foi escolhida uma abordagem de criação de um sistema, utilizando sensores que não gerassem um alto custo financeiro. Esta escolha parte do princípio de que se o artefato for altamente dispendioso e difícil de encontrar, o músico pode optar pelos instrumentos já consolidados no mercado que lhes trazem “segurança” na *performance* e que o seu funcionamento já é conhecido. Disponibilizar um sistema barato, aumenta as possibilidades de adesão de usuários, considerando que requisitos de expressividade sejam atingidos.

Músicos de percussão utilizam uma variedade enorme de instrumentos, em seus kits, para o acompanhamento de diversos ritmos musicais e buscando uma alta variedade de timbres, proporcionando um alto custo em instrumentos. Para que os profissionais de percussão possuam uma alternativa barata que lhes forneçam uma grande variedade de timbre e também para que o alto custo do produto não se torne um obstáculo na difusão desta tecnologia, foi feita a escolha de reduzir o custo e, ao mesmo, proporcionar uma experiência musical que possa ser usada profissionalmente.

5.1.3 Princípios de design

O conceito utilizado ao longo do projeto parte dos princípios de design centrado no usuário desenvolvidos e publicados por Donald Norman e Stephen Draper (NORMAN; DRAPER, 1986). Eles partem da ideia de buscar as soluções para os problemas de uma determinada área a partir das pessoas que fazem parte do contexto. Segundo Abras, Maloney e Preece (ABRAS; MALONEY-KRICHMAR; PREECE, 2004, p.1):

“For example, some types of UCS consult users about their needs and involve them at specific times during the design process; typically during requirements gathering and usability testing. At the opposite end of the spectrum there are UCD methods in which users have a deep impact on the design by being involved as partners with designers throughout the design process.”
(ABRAS; MALONEY-KRICHMAR; PREECE, 2004, p.1)

Outro conceito, utilizado por Calegario (CALEGARIO, 2013), é o de *Running Lean design*, desenvolvido por (MAURYA, 2012), cujos princípios fundamentais envolvem a rápida prototipação e a adaptação do projeto para suprir necessidades específicas do contexto no qual ele está inserido.

O processo adotado neste projeto segue a linha da abordagem utilizada em (CALEGARIO, 2013), a qual utiliza um processo iterativo que pode ser dividido nos seguintes passos: Inspiração, Investigação, Ideação, Prototipação e Avaliação.

Os métodos de **Inspiração** envolvem primeiramente entender o que se sabe sobre o assunto, e atentar para oportunidades que possam gerar motivação para o desenvolvimento do projeto.

A **Investigação** envolve buscar na literatura, formal e informal, o que já foi pesquisado sobre o assunto. É importante que na primeira fase de iteração do processo, se elabore uma pergunta de pesquisa (descrita no capítulo 1), para guiar esta exploração.

A **Ideação** envolve reunir as informações coletadas na etapa de investigação e elaborar formas de tentar solucionar o problema de pesquisa.

Prototipação é fundamental para materializar suas ideias geradas e refiná-las, eliminando características desnecessárias que não contribuirão no processo de *design* e experimentando novas ideias.

A **Avaliação** do sistema deve ser feito ao final de cada processo iterativo. A partir dela é possível analisar as contribuições apresentadas pelo protótipo criado e como ele pode ser refinado. Neste passo, pode-se apresentar o protótipo ao usuário final, para que este possa apontar possíveis problemas que passem despercebidos pelos *designers* ou nuances que são essenciais para o projeto.

5.1.4 Conexão com tecnologias já existentes

A difusão dos protocolos MIDI e OSC e as possibilidades de conexão com diversos sistemas incentivou a aceitação destes protocolos por empresas

desenvolvedoras de sistemas musicais (virtuais ou reais), se tornando “indispensável na produção de música eletrônica” (CHAGAS, 1992, p. 16–17).

As possibilidades de desenvolvimento de interfaces MIDI/OSC com baixo custo de implementação – facilidade na programação e baixo custo financeiro – junto à flexibilidade do padrão de comunicação criadas por estes protocolos facilitam a integração dos mesmos com diversos sistemas computacionais e, particularmente, musicais (virtuais ou reais).

Portanto, este projeto busca a integração do artefato desenvolvido com sistemas musicais já existentes e familiares para o público alvo, procurando um maior grau de adesão e facilidade de uso por parte dos músicos profissionais.

5.1.5 *Mobilidade*

Como já descrito anteriormente, a capacidade dos protocolos MIDI e OSC de proporcionar um imenso leque de timbres, facilita a aceitação desta tecnologia para um público que busca diversas variações de timbre para acompanhamento musical. Ademais, o fato de os DMIs serem artefatos digitais faz com que seja possível construir sistemas portáteis.

Além disso, é importante notar a dificuldade no transporte dos *kits* de percussão de um músico profissional. Diversos destes instrumentos, principalmente àqueles projetados para performance ao vivo, normalmente possuem dimensões que se tornam empecilhos para o transporte, principalmente se levarmos em consideração a quantidade de instrumentos contidos nestes *kits*.

Portanto, a abordagem adotada neste projeto ainda leva em consideração a mobilidade do artefato final.

5.2 Métodos de design adotados

Nesta seção, são apresentadas as escolhas que foram feitas baseadas nos princípios apresentados anteriormente. São apresentados os métodos de design, as escolhas de tecnologia, o tipo de instrumento escolhido, bem como as justificativas para tais escolhas.

5.2.1 *Método de inspiração*

Este projeto partiu da necessidade de responder questionamentos levantados em pesquisas anteriores. Músicos profissionais demonstraram interesse em possuir este tipo de instrumento, porém levantaram diversos problemas encontrados nos DMIs desenvolvidos com as tecnologias de interfaces gestuais Wii e Kinect (SILVA, 2012).

Com tantas tecnologias gestuais emergindo, há uma certa necessidade de explorar as capacidades destas tecnologias para o desenvolvimento de DMIs para investigarmos há possibilidade de resolver estes problemas encontrados.

Ao mesmo tempo se torna uma contribuição para a avaliação das tecnologias em si.

5.2.2 *Método de ideação*

Desde o início do projeto, buscou-se definir diretrizes para seguir um método de desenvolvimento estruturado. A definição destes princípios auxilia o desenvolvedor a se guiar. Estes princípios, descritos na seção 5.1, são imprescindíveis para facilitar as escolhas a serem feitas, tendo em vista os objetivos que se deseja alcançar. Além disso, no desenvolvimento de um artefato, é importante procurar saber que valores – e.g. baixa latência, alta precisão, etc. – são realmente necessários para responder a pergunta de pesquisa e quais são opções almejados pelo designer.

Ademais, é essencial que também seja definido o perfil dos usuários, já que, partindo dos princípios definidos para a realização deste projeto (descritos na seção anterior), esta definição ajuda a definir os valores a se entregar e quais os problemas devem ser investigados.

Como mencionado no método de inspiração, ao explorarem instrumentos desenvolvidos com tecnologias de interfaces gestuais, percussionistas profissionais encontraram problemas que afetam a *performance* ao vivo de maneira drástica. Isto acontece porque a percussão é altamente demandante de precisão temporal – alguns autores argumentam que para se realizar certas técnicas de percussão são controladas pelos músicos com precisão de 1 milissegundo (WESSEL; WRIGHT, 2002) – e precisão espacial para explorar o material dos instrumentos, buscando as nuances de expressividade.

Portanto, a escolha de percussionistas profissionais como sendo o público-alvo deste projeto parte de que estes usuários irão demandar que a tecnologia seja explorada para atingir as altas expectativas e necessidades para uma *performance* expressiva.

5.2.3 *Métodos de investigação*

Os métodos de investigação utilizados para a construção deste projeto partiram da análise da literatura da área, à procura de artefatos que utilizam interfaces gestuais para criação de instrumentos musicais. Esta busca de referências na literatura formal – a partir de trabalhos acadêmicos – e informal – demonstrações de DMIs disponíveis em sites de compartilhamento de vídeos – introduzem novos conceitos e possibilidades de técnicas a serem trabalhados e suportam as tomadas de decisão que deverão ser feitas ao longo do projeto.

Em seguida, após a elaboração de uma primeira hipótese sobre uma possível solução, a análise das tecnologias existentes e as possibilidades de utilização de tecnologias emergentes foram estudadas para a construção de um primeiro protótipo a ser avaliado.

Feita a redução do escopo para instrumentos de percussão, foi preciso compreender melhor os problemas e questionamentos elencados por Silva (2012) – particularmente se os DMIs são realmente utilizados pelo público profissional – e buscar informações mais detalhadas sobre controladores MIDI de percussão já existentes no mercado. Para isto, foi conduzido um questionário online utilizando a plataforma *SurveyMonkey* dirigida ao público alvo do projeto – percussionistas profissionais.

Ademais, também foi realizada uma análise dos gestos do instrumento de percussão escolhido a ser simulado. Com o levantamento destes gestos, é possível avaliar as capacidades da tecnologia.

5.2.4 Método de prototipação

O método de prototipação escolhido para este projeto pode ser descrito em duas etapas: Uma primeira etapa, na qual o desenvolvedor parte apenas de informações hipotéticas e análises da literatura, e uma segunda etapa que envolve o aprimoramento das tecnologias escolhidas a partir de novos conhecimentos adquiridos pela familiarização com estas tecnologias, descobertas de novas técnicas e avaliação do protótipo construído na primeira etapa.

Durante a primeira iteração, protótipos preliminares visando resultados similares devem ser elaborados para se estudar as limitações das tecnologias utilizadas, para se familiarizar com a arquitetura do sistema e com as possibilidades de desenvolvimento. Além disso, é nesta fase em que o desenvolvedor procura descobrir potenciais erros na ideação do projeto, bem como adicionar novas funcionalidades que havia negligenciado ou necessidades que não sabia que existiam.

Após esta primeira etapa, tendo um maior conhecimento sobre as tecnologias escolhidas, os protótipos desenvolvidos começam a ganhar forma e os resultados das avaliações internas e/ou avaliações com os potenciais usuários guiarão o *designer* a explorar novas técnicas, descartar e consolidar técnicas utilizadas, buscar novas funcionalidades do sistema e procurar solucionar problemas que não foram resolvidos e/ou problemas que emergirem.

5.3 Escolha do instrumento

Como já mencionado, a escolha de simular instrumentos de percussão parte do princípio de que estes requerem um alto grau de precisão sonora e espacial. Entretanto, de forma a reduzir o escopo para uma análise mais efetiva, foi necessário focar um instrumento de percussão que possuísse diversas nuances de expressividade.

As tumbadoras (Figura 5.2), também conhecidas como congas, são instrumentos da classe dos *membranofones* – classe de instrumentos de

percussão, em que o som é produzido pela vibração de uma membrana – que possuem um grande repertório de gestos e nuances gestuais para serem tocadas profissionalmente. O formato semelhante a diversos instrumentos de percussão da mesma classe e o vocabulário sonoro variado permitiram que elas se tornassem bastante difundidas em todo o globo e utilizadas em ritmos variados, como, por exemplo, diversos ritmos latino-americanos (conga, rumba, samba, entre outros), rock, jazz, R&B e até utilização em arranjos de música clássica (WARDEN, 2005).



Figura 5.2: Tumbadoras (ou Congas).

Portanto, a escolha de simular uma conga parte da versatilidade na utilização deste instrumento em diversos ritmos musicais, bem como na variedade do repertório gestual que pode ser utilizado para avaliação das tecnologias utilizadas neste projeto. Entretanto, as limitações do FOV do Leap Motion© fizeram com que apenas um tambor pudesse ser simulado.

No caso dos gestos específicos das tumbadoras, diversos vídeos e websites instrucionais foram analisados a fim de levantar os gestos e as nuances gestuais utilizadas pelos percussionistas para tocar as congas. Estes vídeos estão dispostos no site de compartilhamento Youtube e são fornecidos por empresas e estúdios musicais que produzem materiais instrucionais para percussionistas iniciantes e avançados.

Diversas nuances e gestos foram encontradas analisando estes vídeos. A maior parte destas utilizam apenas as duas mãos para serem realizadas, enquanto algumas são realizadas com uma ferramenta de madeira (uma espécie de baqueta) que é utilizada para tocar as congas (na pele e no corpo do instrumento) e abafá-las de diversas maneiras de acordo com a distância entre a baqueta e a mão.

Foram analisados os gestos que produzem as notas **Open Tone**, **Muffle Tone** (*Open Tone* abafado), **Slap Tone**, **Closed Slap Tone** (*Slap Tone* abafado), **Pressed Slap Tone** (*Slap Tone* abafado com a outra mão), **Bass Tone**, **Heel Tone**, **Tip Tone**, **Gliss** (ou *Glissando*) e as chamadas **Grace Notes**. As figuras aqui apresentadas foram extraídas de um site instrucional de uma empresa chamada *Pulsewave*³⁷, o qual possui uma seção ensinando os gestos para tocar congas.

As **Grace Notes** são “notas que não recebem valor musical, tocadas o mais próximo possível da nota subsequente, que irá receber o valor musical completo” (EVANS, 1966). Estas notas chegam ao instrumento para produzir uma nuance de expressividade que não é possível com apenas uma nota. A *Grace Note* mais comum em instrumentos de percussão é chamada de “*Flam*”. Estas são, na maioria das vezes, realizadas com duas notas iguais subsequentes. Entretanto, em alguns casos, outros gestos são utilizados para produzir *Grace Notes*, como usar a mão com o punho (a qual foi atribuído o nome **Fist Note**) fechado antes de uma nota.

Extrair *samples* destas notas é uma tarefa bastante complicada, já que elas são notas que são produzidas quase simultaneamente às notas que as sucedem, além de suas intensidades serem mais baixas.

No caso de a *Grace Note* ser feita com a mão em formato de punho – que também é utilizada com uma nota independente pelos percussionistas –, apesar de o Leap Motion© conseguir reconhecer o gesto realizado, a nota não foi incorporada no sistema devido à ausência do *sample* sonoro produzido pelo mesmo. Entretanto, para avaliação deste gesto, um *sample* sonoro de outro instrumento foi utilizado para que este movimento pudesse ser distinguido dos demais.

A nota **Open Note** é tocada com os dedos nas extremidades do instrumento. A técnica é realizada golpeando a borda do instrumento com a junção entre a palma e os dedos, permitindo que estes atinjam a pele e produzam o som. Geralmente, o músico golpeia o instrumento com os dedos separados.

³⁷ <http://pulsewave.com/classes-workshops/conga-hand-positions-sounds/>



Figura 5.3: Open Tone.

O **Muff** ou **Muffled Tone** utiliza a mesma técnica anterior, porém o músico pressiona a pele do instrumento para que esta não vibre completamente, causando um tom abafado. Muitas vezes, nesta técnica, o músico junta os dedos da mão (com exceção do polegar), diferentemente do que acontece no *Open Tone*.



Figura 5.4: Muffled Tone.

O **Slap Tone** ou **Slap** é feito um pouco mais ao centro do instrumento, utilizando os dedos e parte da palma da mão para fazer o som. No *Slap*, os dedos da mão do músico ficam separados, igualmente ao *Open Tone*. Entretanto, o som produzido no *Slap* é mais agudo.



Figura 5.5: Slap Tone.

No **Closed Slap**, os dedos ficam mais próximos e, muitas vezes, a mão faz um formato curvado. Além disso, pressiona-se a pele do instrumento para abafar a nota, produzindo um som mais agudo e mais “seco” que o *Slap*.



Figura 5.6: Closed Slap.

Outro tom derivado do *Slap* é o chamado **Pressed Slap** ou **Muted Slap**. Neste caso, a outra mão do músico é posicionada mais ao centro do instrumento e é usada para pressionar a pele. A outra mão irá realizar um gesto similar ao *Slap* ou ao *Closed Slap* (mais frequentemente este último) para produzir um som mais agudo e mais “seco” do que os outros dois.



Figura 5.7: Pressed Slap.

O **Bass Tone**, o tom mais grave dentre estes, golpeando o instrumento com toda a mão (palma e dedos) paralelamente à sua pele e deixam ela bater e voltar. Este gesto é feito mais ao centro do instrumento. Muitas vezes os músicos levantam os dedos para que apenas a palma atinja o instrumento.



Figura 5.8: Bass Tone.

No **Heel Tone**, o músico utiliza a base da mão com os dedos levantados para produzir um som grave mais suave e distinto. Este é frequentemente usado com o **Tip Tone** em diversos ritmos musicais.



Figura 5.9: Heel Tone.

O **Tip Tone**, é realizado com os dedos que estão levantados. Pode-se levantar um pouco mais a palma para iniciar o movimento e golpear a pele com os dedos da mão, saindo um som mais agudo, porém similarmente suave ao do *Heel Tone*.



Figura 5.10: Tip Tone.

Por último temos o **Gliss** ou **Glissando**, no qual o músico produz arrastando o dedo na pele do instrumento. Com isto, pode-se produzir som bastante peculiar e distinto dos outros sons percussivos.



Figura 5.11: Gliss Tone.

Algumas outras nuances, como utilizar as unhas dos dedos para produzir um som diferente, ou golpear o instrumento com apenas um dos dedos, ainda aparecem em diversos vídeos. Entretanto, vários destes são técnicas particulares que variam para diferentes músicos. A falta do controle sobre a qualidade do som destas nuances e a particularidade do uso por apenas um pequeno número de músicos (diversas vezes apenas um), fez com que este projeto se limitasse a focar nos gestos acima. Outra particularidade mais comum, é utilizar o cotovelo para abafar as notas. Deste jeito, pode-se obter outra variação dos tons acima.

5.4 Método de aquisição dos samples

Durante a aquisição dos gestos, a análise dos vídeos proporcionou a aquisição dos *samples* sonoros, gerados pelas congas, relativas a estes gestos. Vale ressaltar que os vídeos, a partir dos quais os *samples* foram coletados, são vídeos de instituições profissionais que fornecem vídeo-aulas para músicos de diversos níveis. Sendo assim, lembrando que o intuito deste projeto é avaliar as tecnologias e a interface de interação para a construção futura de DMIs alternativos (e não simular instrumentos acústicos já existentes), os sons coletados aqui serão utilizados para aplicações puramente acadêmicas e para este trabalho apenas. Gravações futuras de *samples* deverão ser feitas para outras aplicações.

6 CONSTRUÇÃO DO PROTÓTIPO

Este capítulo descreve a evolução dos protótipos desenvolvidos ao longo deste projeto de maneira cronológica. Aqui, são detalhadas as tecnologias e sensores que foram utilizados no design do DMI que foi denominado V-Conga. Em seguida, todas as versões são explicitadas, discutindo as mudanças de *design* efetuadas e suas justificativas. Os resultados discutidos para cada versão são posteriormente condensados e apresentados no próximo capítulo. Esta apresentação cronológica das versões busca mostrar o processo iterativo de *design* do DMI.

O primeiro protótipo consiste em um estudo realizado para familiarização com o dispositivo de captura de gestos (o Leap Motion©) e a construção de um primeiro instrumento musical com o sistema: o Crystal Piano. Apesar de apresentar diversos problemas, o piano virtual forneceu uma visão geral sobre o sistema e as possibilidades de interação musical com este sensor.

Após mudanças de abordagem, outros protótipos foram criados e avaliados com a tentativa de resolver os problemas apresentados no Crystal Piano, nos trabalhos acadêmicos revisados na literatura, especialmente o trabalho de Silva (2012) e os problemas encontrados no questionário realizado com percussionistas profissionais (apresentados no capítulo 3, seção 3.2).

Todos os protótipos foram implementados e avaliados (com exceção do Crystal Piano que foi apenas avaliado) em um computador com processador intel i5 (2,6 – 3,1 GHz) e 8 Gb de memória RAM, rodando Windows 8/8.1.

As versões “alfa” são versões que foram avaliadas internamente, pelo autor deste projeto, sem apresentá-las à músicos profissionais. A versão “beta” foi apresentada ao público alvo e utilizada para avaliação quantitativa do projeto. Por fim, a versão final foi implementada de forma a aprimorar a versão apresentada ao músico e com intuito de construir um artefato físico com um melhor acabamento.

6.1 Sensores Utilizados

Como era esperado, de acordo com o método de *design* escolhido, mudanças ocorreram ao longo do processo de *design*, ocasionando na construção de algumas versões diferentes do protótipo. Estas mudanças e suas justificativas serão apresentadas mais à frente neste capítulo.

Nesta seção, são apresentados os principais sensores que serviram como base para a concepção do artefato, bem como as justificativas para seu uso.

6.1.1 Leap Motion©

O Leap Motion Controller® utiliza duas câmeras infra-vermelho (IR, do inglês *Infrared*) e três LEDs IR para realizar o mapeamento dentro de um campo de visão (FOV, do inglês *Field Of View*) limitado de aproximadamente 60 centímetros cúbicos com um ângulo de visão de 150 graus.

Apesar de seu campo de visão ser limitado, diferentemente do Kinect que consegue capturar todo o corpo, desde a primeira versão do SDK, o algoritmo de reconhecimento do Leap Motion® conseguia extrair mais características das mãos com um tempo de processamento menor.

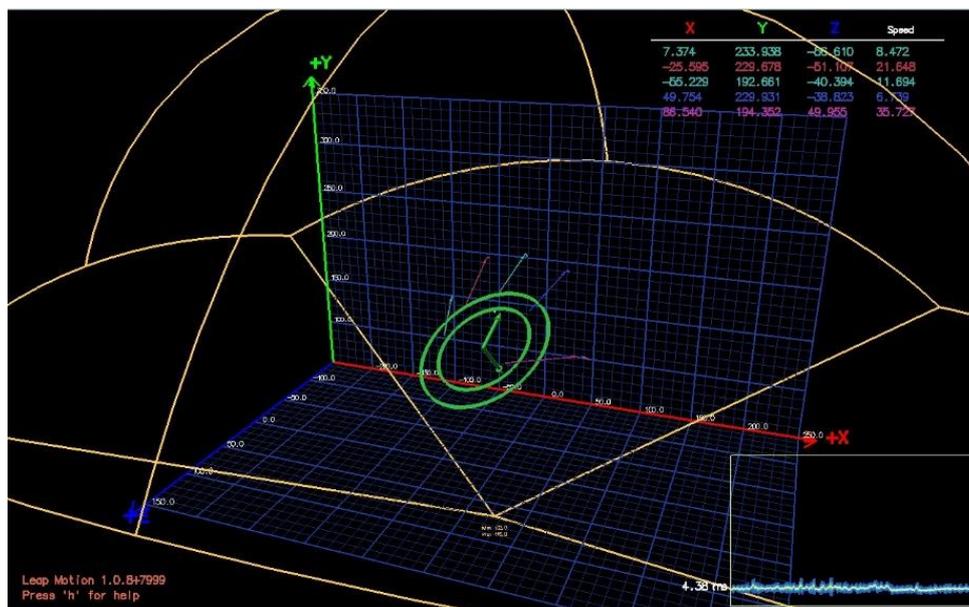


Figura 6.1: Campo de Visão do Leap Motion® SDK (v1.0).

O dispositivo opera enviando *frames* que contém diversas informações a respeito da cena em questão. Informações como posição e velocidade de movimento das mãos, ponta dos dedos e juntas de cada osso das mãos, diâmetro das mãos, o comprimento e largura do polegar e de cada osso do mesmo e dos demais dedos e seus ossos, vetor que determina para onde os dedos estão apontando, além de possuir a capacidade de reconhecer alguns tipos de objetos (com a limitação de serem finos e longos), denominados *tools*, disponibilizar a imagem bruta das câmeras e reconhecer alguns gestos programados previamente pela equipe de desenvolvimento.

Como descrito no capítulo 3, estudos anteriores levantaram o repertório de gestos que outras interfaces de reconhecimento, mais especificamente o Kinect® (versão 1.0) da Microsoft e o *Wii Remote*® da Nintendo, conseguem identificar e disponibilizar para o desenvolvimento de novos DMIs (SILVA, 2012). Como visto na Tabela 1 (Seção 3.1, página 24), enquanto o sensor Kinect® apresentou falhas na captura de movimentos com os dedos, o *Wii Remote*® não conseguiu capturar movimentos com os ombros.

Elaborando uma tabela comparativa da capacidade de reconhecimento do Leap Motion© com as interfaces previamente estudadas, pode-se reconhecer o nível de aperfeiçoamento e precisão do dispositivo. Esta tabela segue a nomenclatura utilizada na pesquisa de (SILVA et al., 2013).

Gestos	Kinect© v1.0	Wii Remote©	Leap Motion©
Movimentos com o ombro	X		X*
Movimentos com o cotovelo	X	X	X*
Rotação do antebraço		X	X*
Movimentos com o pulso	X	X	X
Movimentos com o polegar (todas as articulações)		X	X
Movimentos com os demais dedos		X	X
Movimento das falanges proximais (demais dedos)		N/A	X
Movimento das falanges intermediárias (demais dedos)		N/A	X
Movimento das falanges distais (demais dedos)		N/A	X

Tabela 3: Movimentos capturados pelas interfaces.

***Capturados por translação ou rotação das mãos.**

Este tipo de granularidade na captura de gestos é extremamente importante para possibilitar o desenvolvimento de novos aplicativos e interfaces digitais que sejam capazes de expressar diferentes nuances de movimento, as quais são imprescindíveis para a expressividade na *performance* de diversos instrumentos musicais, incluindo os instrumentos percussivos.

Este resultado apresenta um grande potencial de captura das nuances. Entretanto, deve ser feita uma análise mais rigorosa da captura dos gestos do instrumento escolhido, bem como investigar a captura desses gestos durante uma performance musical, ou seja, enquanto o músico varia os gestos que está fazendo. Deste modo, pode-se avaliar a capacidade de captura no contexto de uso e a estabilidade do sistema.

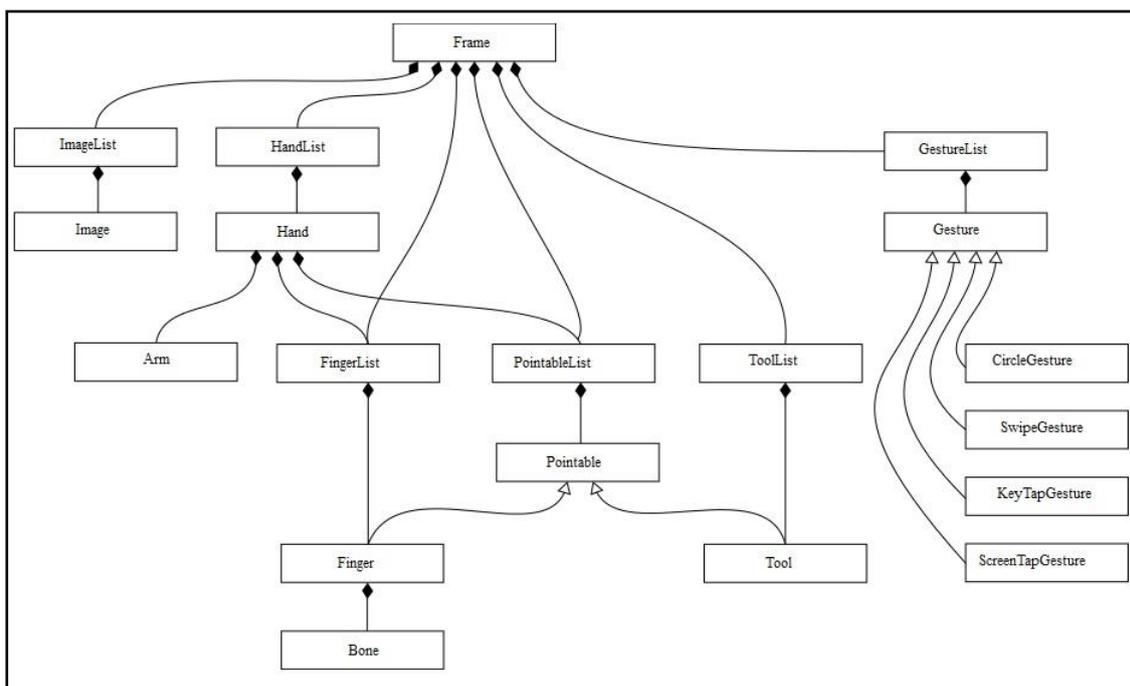


Figura 6.2: Arquitetura do frame de reconhecimento.

Este sensor possui a capacidade de reconhecer a cena em três dimensões, na área limitada pelo seu FOV, e foi projetado especificamente para realizar o rastreamento das mãos do usuário. Essas informações são fornecidas aos desenvolvedores a partir de *frames* (sendo capaz de armazenar os últimos 30 frames na memória). Além disso, a arquitetura da API funciona em um sistema de múltiplos processos de execução (*threads*) que facilita a construção de um sistema que utiliza integração de sensores.

A primeira versão do SDK liberada aos desenvolvedores possuía três modos de rastreamento e, conseqüentemente, de processamento de informações. Que capturava algumas informações. Entretanto possuía limitações no rastreamento que foram identificadas pelo nosso grupo de pesquisa (SILVA et al., 2013).

Em 2014, a empresa lançou a versão 2.0 do SDK, aprimorando radicalmente o algoritmo de rastreamento. A versão 2.0 denominada “*Skeletal Tracking*” (ou Rastreamento de Esqueleto), trouxe a capacidade de reconhecer todos os ossos dos dedos e a posição das juntas dos dedos. Alguns meses mais tarde a empresa lançou o programa de suporte a desenvolvimento específico em realidade virtual, fornecendo a imagem bruta capturada pelas câmeras e aprimorando o algoritmo de rastreamento com o sensor em posições não triviais, como, por exemplo de cabeça para baixo ou conectado a um display de cabeça HMD (Head Mounted Display).

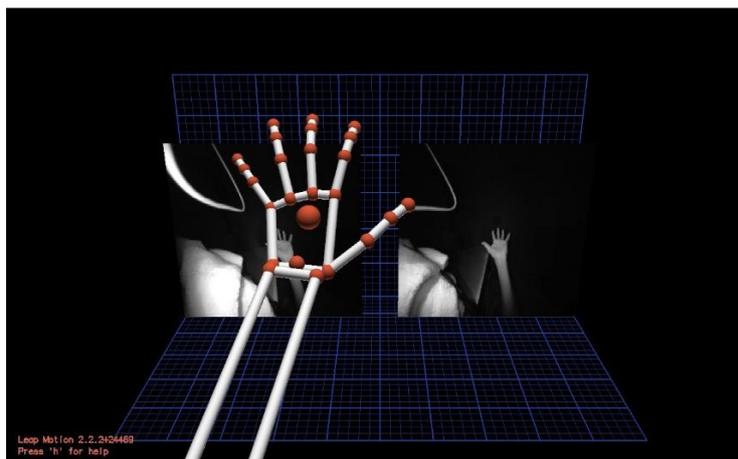


Figura 6.3: Visualizador do Leap Motion© (SDK v2.0).

Nesse novo SDK, as informações de captura eram mais estáveis comparadas à versão anterior. A captura das informações dos ossos e suas junções permitiram a captura de maiores nuances – e.g. identificar a base das mãos -, além de melhorar o algoritmo dos métodos.

Portanto, a escolha da utilização deste sensor se baseia na promessa de que sua capacidade de captura de gestos mostrou grande potencial para o reconhecimento das diversas nuances gestuais realizadas para a *performance* de um instrumento de percussão.

6.1.2 *Sensor piezoeléctrico*

O sensor piezoeléctrico (mais conhecido como piezo ou microfone de contato) é um sensor que capta vibrações de uma determinada superfície pela vibração consequente do cristal piezoeléctrico contido nele. A vibração do cristal gera um sinal elétrico entre 0 e 5 Volts (ou valores entre 0 e 1023 no Arduino).

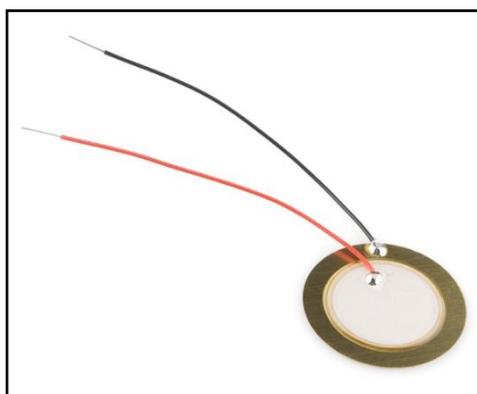


Figura 6.4: Sensor piezoeléctrico.

Como o Arduino processa os dados em um nível muito próximo ao hardware, a latência de captura das informações do piezo é muito baixa (ou nula). Deste

modo, pode-se utilizar o piezo como um *trigger* para reduzir a latência do sistema.

O esquema de montagem do piezo é o seguinte:

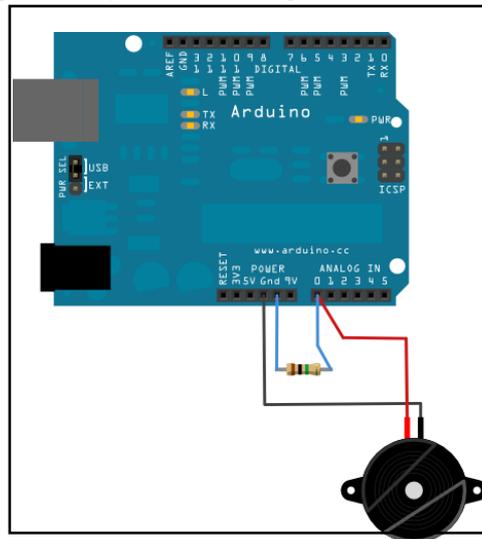


Figura 6.5: Esquema de conexão do piezo com o arduino.

É importante que um resistor seja conectado ao circuito para proteger a placa Arduino de picos de voltagem inesperados que podem danificá-la. Para isso, foi utilizado um resistor de 10KΩ conectado em paralelo.

6.1.3 Sensor de pressão

Os sensores de pressão FSR (*Force Sensing Resistor*) são compostos por polímeros condutivos que variam suas resistências de acordo com a pressão exercida sobre eles (MEDEIROS; WANDERLEY, 2014a). Para obter essa variação de resistência, o sensor é conectado ao pino de 5 Volts do Arduino, gerando valores entre 0 e 1023 (similarmente ao piezo).

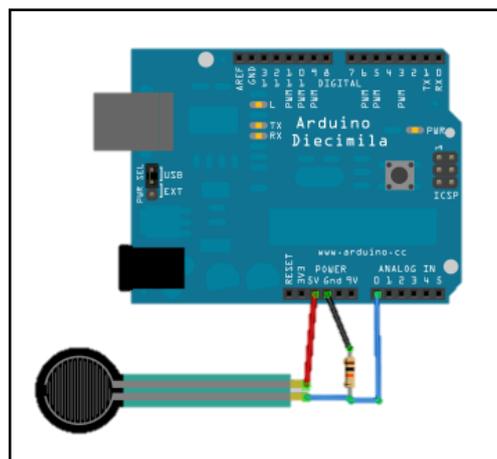


Figura 6.6: Esquema de conexão do FSR com o arduino.

A integração do FSR ao sistema foi feita de forma a capturar as nuances de abafamento das notas da conga. O sistema proporciona uma nova forma de abafamento das notas que possui maior similaridade com o gesto realizado no instrumento acústico. Neste caso, as informações são enviadas

6.2 Tecnologias utilizadas

Nesta seção, são apresentadas as tecnologias que foram escolhidas para a construção do artefato. O Arduino serve para conectar os sensores piezo e FSR ao sistema, enviando os dados coletados por eles para computador. Já o Ableton Live foi a tecnologia escolhida para produção sonora.

6.2.1 Integração dos sensores

A versão final do protótipo construído possui a seguinte arquitetura:

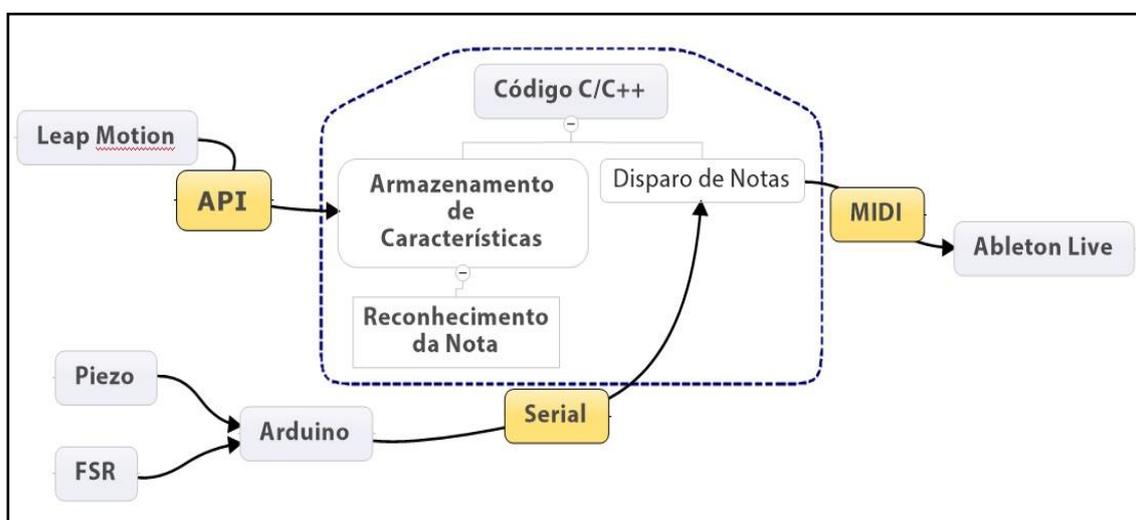


Figura 6.7: Arquitetura do protótipo final.

Como veremos mais detalhadamente na descrição das versões do sistema, essa integração permitiu que problemas encontrados nas primeiras versões fossem solucionados. Primeiramente, a captura das características das mãos já apresentava o potencial de captura das nuances gestuais. Entretanto, o sistema apresentava latência e algumas nuances não eram reconhecidas.

A descrição das versões do artefato mostra em que etapa de iteração foram integrados os sensores piezo e FSR.

6.2.2 Arduino

O Arduino (“Arduino”, 2015) é uma plataforma de prototipação de projetos eletrônicos lançada em 2005 e desenvolvida pelos fundadores Massimo Banzi, David Cuartielles, Tom Igoe, Gianluca Martino e David Mellis. Foram desenvolvidos diversos microcontroladores (mais chamadas de placas) que

possuem uma interface intuitiva e seu próprio ambiente integrado de desenvolvimento (IDE, do inglês *Integrated Development Environment*).

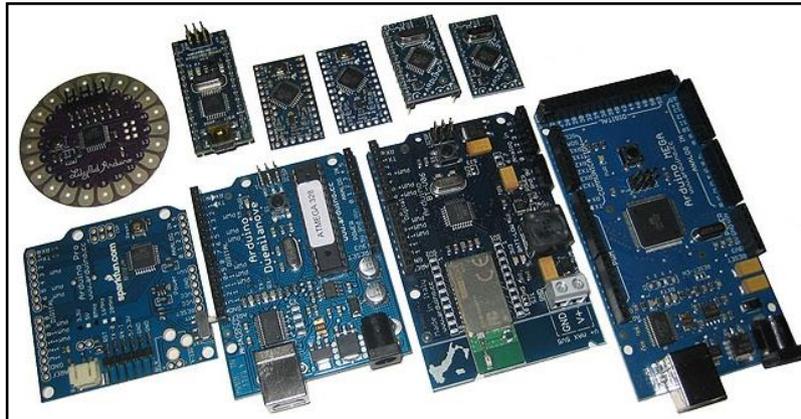


Figura 6.8: Diferentes microcontroladores arduino.

A iniciativa foi criada com intuito de auxiliar entusiastas que não possuem um nível técnico avançado de instrumentação eletrônica e/ou programação. Assim, artistas e *designers* podem se munir de diversas ferramentas tecnológicas para aprimorar seus projetos.

Possuindo a capacidade de integração com mais diversos sensores existentes – sensores de luz, capacitivo, resistivo, flexíveis, microfones de contato, entre outros – os projetos desenvolvidos com estes microcontroladores podem tanto funcionar conectados a um computador – enviando e recebendo informações – ou independentemente (projetos *stand-alone*).

Esta possibilidade de integração com computadores foi possível com o desenvolvimento do IDE para os sistemas operacionais Windows, Mac e Linux, implementando o *framework* Wiring: um framework desenvolvido para programação de microcontroladores, baseado na plataforma de programação em Java “Processing”.

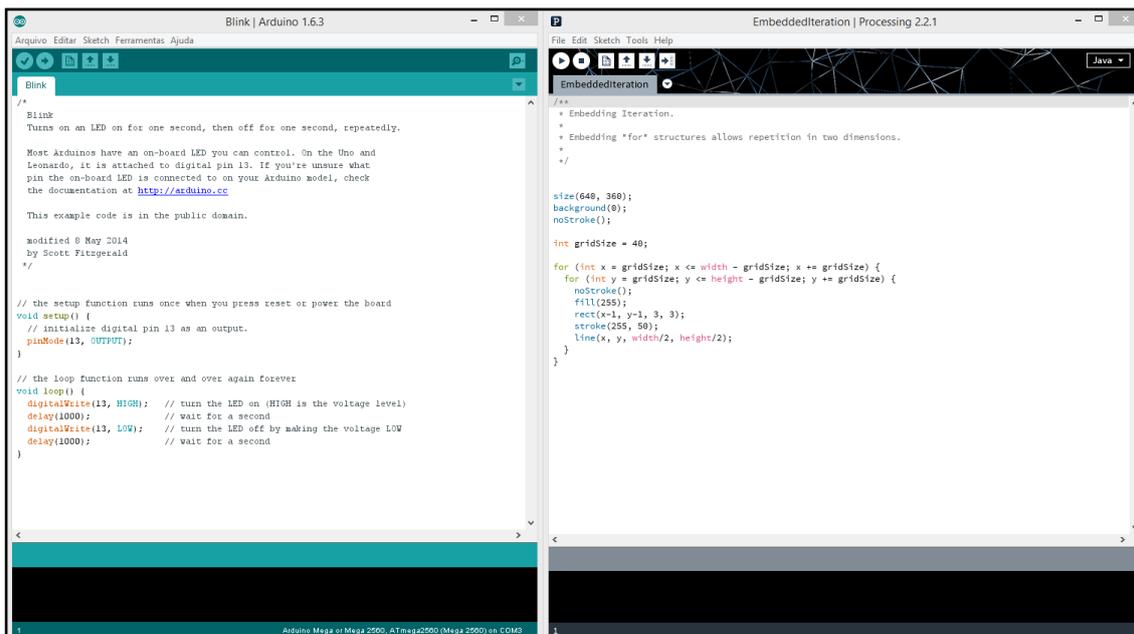


Figura 6.9: Ambientes de programação do Arduino (à esquerda) e Processing (à direita).

As placas foram construídas partindo dos princípios da comunidade *Open Source* de baixo custo de implementação de projetos, uma interface de programação clara e simples, possuir integração entre múltiplas plataformas, e que seus *software* e *hardware* pudessem ser expansíveis. Para este último, foram desenvolvidos os chamados *Shields*, com os quais novas funcionalidades podem ser incorporadas às placas originais (GOMES; TAVARES, 2013).

No contexto de programação, a comunidade aderiu esta iniciativa, elaborando diversas bibliotecas para serem integradas ao IDE do arduino ou de forma a incorporar a comunicação destas placas em seus ambientes, e.g. a plataforma de desenvolvimento OpenFrameworks.

Estas placas possuem comunicação utilizando o protocolo serial que podem ser utilizadas para transmitir as informações dos sensores conectados a qualquer ambiente de programação que seja capaz de receber este tipo de informação.

Sendo assim, a capacidade de incorporação de diversos sensores diferentes junto aos baixos custos de implementação e financeiro e a facilidade de integração desta tecnologia com diversos ambientes de programação, fazem com que a incorporação da plataforma arduino seja ideal para o desenvolvimento do artefato deste projeto.

6.2.3 Ableton Live

Ableton Live (“Ableton”, 2015) é um *software* de composição musical com o qual usuários podem explorar diferentes maneiras de interagir com o som. O Live é uma ferramenta poderosa, possuindo diversas características, variando entre

gravação de vários canais de áudio, composição utilizando o protocolo MIDI, incorporação de efeitos sonoros em tempo real e possuindo a capacidade de integração com diversos instrumentos musicais e módulos de sintetização sonora físicos ou virtuais.

Como o nome sugere, o software foi desenvolvido para ser utilizado em composições em tempo real, possuindo uma alta resposta ao se utilizar suas funcionalidades e baixa latência. Além disso, possui uma imensa diversidade de *samples* dos mais variados sons, efeitos e síntese de diversos instrumentos.

A interface do sistema segue um formato modular, na qual cada *sample* de áudio ou MIDI que é inserido no programa pode ser processado de maneiras diferentes: adicionando efeitos independentes, modificando o formato da envoltória da onda, etc.



Figura 6.10: Interface do Ableton Live 9.

Uma das funcionalidades que se pode destacar neste programa é o *Drum Rack*. Este é uma funcionalidade que utiliza o protocolo MIDI, com a qual é possível mapear os mais variados controles de entrada em *samples* que são escolhidos pelo usuário.



Figura 6.11: Drum Rack do Ableton Live (Extraído de http://en.wikiaudio.org/Ableton_Live:Drum_rack).

O *Drum Rack* possui 128 módulos que podem funcionar de forma independente. Cada módulo é ativado recebendo um sinal de nota MIDI (de 0 a 127, possibilitando os 128 módulos), enviado por um (ou mais de um) dos 16 canais de entrada disponíveis, e possui a capacidade de armazenar um *sample* musical que é escolhido pelo usuário. Além disso, ao inserir um *sample* no módulo, o usuário pode expandir o módulo para que diversos processamentos de som possam ser realizados com o *sample* contido nele.

O *Drum Rack* ainda possui a capacidade de os *samples* estarem interligados ou não. Por exemplo, ao se tocar um determinado som, pode ser desejável que outro som toque ao mesmo tempo, ou que não seja possível tocar outro som específico, ou que a envoltória de dois *samples* sejam modificadas a partir de um certo parâmetro, entre diversas outras possibilidades.

Além disso, a facilidade na conexão de dispositivos de controle, com o uso do protocolo MIDI, torna a configuração de sistemas musicais mais acessível para o público profissional.



Figura 6.12: Configuração do mapeamento MIDI de parâmetros sonoros.

Na Figura 5.10, cada funcionalidade destacada em roxo pode ser mapeada em um parâmetro de controle diferente. Para sua configuração, o usuário precisa apenas clicar na funcionalidade e depois utilizar o controle específico do controlador MIDI, ao qual ele deseja mapear naquela funcionalidade. A maioria dos controladores MIDI são reconhecidos automaticamente pelo sistema – em alguns casos é necessária a utilização de drivers, principalmente para controladores virtuais –, tornando mais prática a integração dos mesmos.

Todas estas ferramentas disponíveis no Live, fizeram com que diversos artista utilizem o Ableton Live para seus projetos, já que as possibilidades de interação com áudio para *performances* em tempo real são imensas.

Portanto, o Ableton Live foi utilizado com o intuito de buscar integração com sistemas que já são familiares aos músicos profissionais e também pelo fato de o Live possuir uma vasta gama de possibilidades de interação sonora em tempo real e de fácil conexão.

6.3 Protótipo preliminar com o Leap Motion®: Crystal Piano

De forma a entender melhor o funcionamento do sensor e se familiarizar com a arquitetura do sistema, foi desenvolvido um primeiro instrumento musical simulando um piano³⁸. Para o protótipo, também foi desenvolvida uma interface virtual que produz *feedback* visual.

³⁸ Vídeo disponível em: <<https://youtu.be/W4lkCbcG12Y>>.

6.3.1 Tecnologias utilizadas

Para o desenvolvimento deste protótipo, uma abordagem exploratória foi adotada. A implementação do *software* desta versão, desenvolvida pelo pesquisador Jáder Anderson Oliveira de Abreu, foi feita utilizando o ambiente de programação Eclipse e a linguagem de programação Java, a partir do SDK (versão v0.8.1.6221) disponibilizada no site do dispositivo. Além disso, pela falta de conhecimento, por parte dos autores do trabalho, sobre o desenvolvimento de sistemas musicais digitais, o *software* foi construído baseado em exemplos também disponibilizados junto ao SDK.

As bibliotecas utilizadas para o desenvolvimento utilizam foram as seguintes:

- **LeapMotion:** Biblioteca oficial fornecida no SDK, a qual permite que o programa capture informações de estruturas pré-estabelecidas reconhecidas pelo dispositivo.
- **LWJGL:** Biblioteca Java. Fornece o acesso a bibliotecas multi-plataforma de alta *performance*, como OpenGL, OpenCL, OpenAL.
- **SLICK:** Biblioteca compatível com a *LWJGL* para facilitar a construção de jogos 2D.

6.3.2 Configuração

Nesta configuração, apenas o Leap Motion® era utilizado para capturar as características das mãos do usuário. Levando em conta as limitações do campo de visão (FOV) do sensor, o sistema simulava uma oitava de um piano (da nota Dó C4 à nota Dó C5).

O sistema utilizava uma folha de vidro de 10 milímetros de espessura, a aproximadamente 12 centímetros do sensor, proporcionando *feedback* háptico (tátil). A folha possuía marcações simulando um piano, oferecendo *feedback* visual na própria superfície. As marcações em preto delimitavam as teclas brancas do piano, enquanto as marcações em vermelho limitavam as teclas pretas.



Figura 6.13: Folha de vidro com marcações.

O mapeamento do sistema foi configurado para que o plano de ativação das notas estava exatamente na parte de cima da folha de vidro, fazendo com que o usuário só ativasse as notas quando tocasse a tela, criando um tipo de “tela multi-toque”.

Além disso, o software possuía uma interface visual que mostrava os dedos do usuário, constituindo outro tipo de *feedback* visual que podia ser projetado. Portanto, o usuário poderia escolher entre tocar olhando apenas para a superfície de contato ou olhando para a interface projetada.

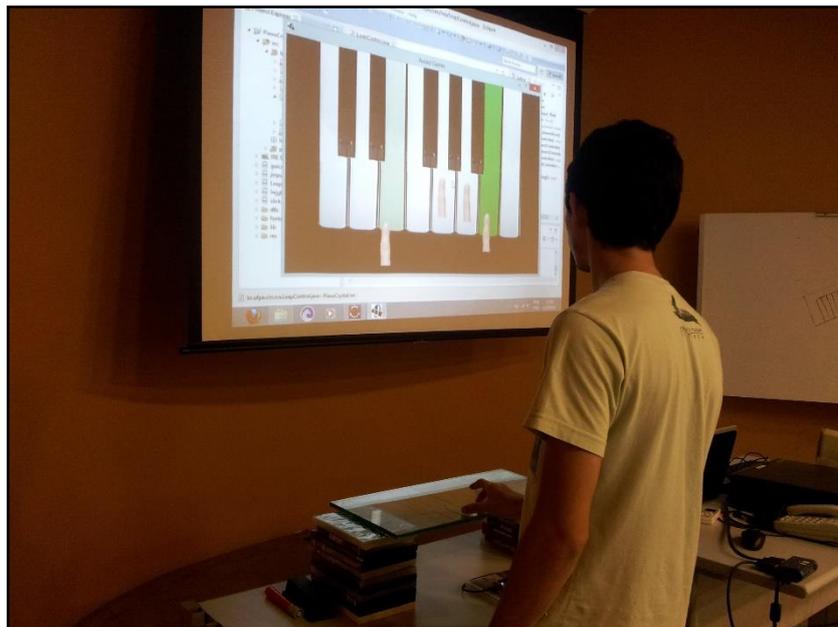


Figura 6.14: Perspectiva do usuário.

6.3.3 Discussões

Para avaliação da latência, foram medidos os sons do usuário golpeando o vidro e o som do piano, gerado pelo computador. Excedendo 100 milissegundos (aproximadamente 150 milissegundos), a latência apresentada pelo sistema causava sérios problemas para o seu uso. O usuário só era capaz de tocar melodias numa velocidade lenta. Mesmo ao tentar tocar melodias numa velocidade natural, a enorme latência causava desconforto e “quebra” no fluxo da música e perda do ritmo da melodia.

Além disso, o sistema apresentava problemas com o rastreamento das mãos. Neste estado do SDK, o algoritmo não estava desenvolvido o suficiente para contornar problemas frequentes em rastreamento com câmeras, como a oclusão de objetos rastreados no espaço. Por exemplo, ao tocar um acorde, os dedos usados para tocar poderiam obstruir o reconhecimento dos outros dedos da mão, o que ocasionava em erros de reconhecimento e poderia causar notas falso positivas ou falso negativas. Outro problema de rastreamento que causava notas falso positivas ou falso negativas ocorria quando dois ou mais dedos estavam muito próximos – característica de rastreamento do algoritmo da versão SDK.

Notas falso positivas são notas que o usuário não pretendia tocar, mas o sistema reconhece a nota e gera o som. Já notas falso negativas são as notas que o usuário queria tocar, mas, por algum motivo, o sistema não reconhece a nota, causando frustração no usuário. Solucionar estes problemas é essencial para que a experiência musical seja possível.

Apesar dos problemas apresentados acima, as capacidades de reconhecimento dos gestos da mão foram bastante promissoras, rastreando todos os movimentos apresentados na Tabela 3 (Seção 6.1.1, página 68). Entretanto, este sistema não possuía controles da dinâmica do instrumento – e.g. controle na intensidade da nota.

6.4 Versão alfa 1

Após a escolha de simular instrumentos de percussão, particularmente a conga, foi necessária a busca de uma nova abordagem de desenvolvimento do sistema. Esta versão³⁹ usava uma nova técnica de implementação para simular uma conga e produzir o som disparando *samples* sonoros previamente armazenados na memória do computador.

6.4.1 Tecnologias utilizadas

A grande latência apresentada pelo Crystal Piano fez com que fosse cogitada a mudança de linguagem de programação e o método de produção sonora. Para

³⁹ Vídeo disponível em: <<https://youtu.be/O-Dalk-vt6Y>>.

isto a implementação foi realizada na linguagem C/C++ na plataforma de programação Visual Studio 2012 tentando aprimorar a *performance* do sistema.

O sistema foi desenvolvido utilizando apenas as bibliotecas fornecidas pelo SDK do Leap Motion© e bibliotecas inseridas no próprio sistema operacional do Windows.

6.4.2 Configuração

Para esta versão, o sistema funcionou praticamente da mesma maneira que o Crystal Piano, entretanto a refração causada pelo vidro fez com que o mesmo fosse desconsiderado na configuração final desta versão.

Como na conga, diversos gestos realizados para produzir o som são feitos na borda do instrumento. Neste caso, a refração dos raios emitidos pelo dispositivo e as dimensões da folha de vidro (medida na versão anterior para simular uma oitava do piano) causavam problemas de rastreamento das mãos. Ainda assim, a folha de vidro foi utilizada para medição da latência, já que era possível medi-la tocando logo acima do sensor.

O sistema utilizava a mesma forma de mapeamento do Crystal Piano. A conga era simulada virtualmente para o usuário que tocava atravessando as mãos em um plano imaginário virtual.

Tendo em mente os mesmos problemas encontrados por Silva (SILVA, 2012), esta versão não foi desenvolvida com o intuito de ser a versão final, a mesma foi desenvolvida para investigar se a nova abordagem adotada produziria melhores resultados que a implementação do Crystal Piano.

6.4.3 Discussões

Neste caso foi notada uma redução da latência para aproximadamente 71 milissegundos, o que ainda atrapalhava bastante na *performance* de músicas em velocidades mais rápidas, entretanto já representava um grande avanço na direção de um sistema musical aceitável.

Esta versão já possuía a capacidade de reconhecer seis gestos específicos da conga: os gestos associados aos sons *Open Tone*, *Muffled* (ou *Muted Open Tone*), *Slap*, *Muted Slap*, *Bass Tone* e *Muted Bass Tone*. O sistema realizava um mapeamento espacial com a variação de a mão do usuário estar aberta com os dedos espaçados (denominada mão aberta) e mão aberta com os dedos juntos (denominada mão fechada). Vale ressaltar que, a esta altura, diversas nuances gestuais ainda não tinham sido percebidas na análise dos gestos da conga.

Assim, o sistema mostrava grande potencial para o desenvolvimento de um sistema virtual, entretanto o problema da latência ainda constituía um grande problema que precisava ser resolvido

6.5 Versão alfa 2

Como os resultados obtidos anteriormente, apesar de serem promissores, não solucionavam os problemas encontrados na literatura, outra abordagem foi avaliada. Utilizando o mesmo conceito da versão alfa 1, foi necessária a exploração de outros métodos de mapeamento e foi cogitada, pela primeira vez, a integração de outros sensores para complementar o sistema.

6.5.1 Tecnologias utilizadas

A partir dos resultados obtidos nas versões anteriores, foi necessário buscar métodos de reduzir a latência do sistema. A escolha feita, neste momento, foi incorporar um sensor piezoelétrico. A integração do piezo com o sistema se deu por intermédio do Arduino, que fornece uma grande facilidade de integração com o sistema já existente. Outra vantagem é o fato de que no próprio site do Arduino já existem alguns esquemas de montagem de sensores (para algumas aplicações) e um exemplo de código que implementa a conexão do arduino com outros programas (por meio do protocolo serial).

As informações do piezo são mandadas usando um protocolo serial para o computador que irá processá-las de acordo com as estratégias de mapeamento escolhidas. O piezo, nesta configuração, serve como um *trigger* para o sistema. Desta maneira, quando o sistema receber as informações do piezo (armazenando-as em um buffer), ele ativa a nota que é rastreada pelo Leap Motion© (que está armazenada). Desta forma, a latência é dada pelo tempo que leva para a informação do piezo chegar ao Ableton Live.

Buscando novas formas de implementação, foi escolhida a linguagem de programação do puredata. Este foi escolhido por suas facilidades de desenvolvimento e integração com *softwares* de composição musical, mais especificamente com o Ableton Live.

Ainda foram utilizados dois aplicativos para passar as informações de rastreamento das mãos: o GECO e o ManoOSC.

6.5.2 Configuração

As informações do Leap Motion© eram enviadas ao puredata utilizando o protocolo OSC por meio dos aplicativos GECO e ManoOSC. A utilização dos dois aplicativos se deve ao fato de que as informações das características de mão aberta/mão fechada extraídas pelo GECO não eram bem processadas pelo programa desenvolvido com o puredata. Então, esta era a única informação processada pelo ManoOSC.

O mapeamento era feito todo pelo *puredata*, que adquiria as informações do arduino e enviava notas MIDI para o *Drum Rack Ableton Live*, o qual continha os 6 *samples* da conga.

Nesta configuração, a folha de vidro foi testada novamente, a fim de fornecer o *feedback* háptico desejado pelos músicos profissionais e testar se a robustez do sistema havia melhorado, devido à recentes mudanças no algoritmo de rastreamento do SDK (versão v.1.2.2.11000) do dispositivo.

6.5.3 *Discussões*

A latência deste sistema foi reduzida consideravelmente para aproximadamente 17 milissegundos graças a incorporação do piezo no sistema. Esta latência já era aceitável segundo Jordà (JORDA, 2005), porém, como alguns autores identificavam que certos controles percussivos necessitam uma latência abaixo de 10 milissegundos (WESSEL; WRIGHT, 2002), era necessário aproximar a latência do sistema ainda mais deste último valor.

Além disso, o sistema também era capaz de reconhecer os gestos da conga reconhecidos na versão anterior. Entretanto, apesar de o *puredata* ser uma plataforma interessante para controle de fluxos de informação, encontrava alguns problemas no reconhecimento dos gestos, ocasionando em notas falso negativas.

Com isto, foi confirmado que a integração com o Ableton Live auxiliaria bastante na configuração do sistema por parte do usuário. Ademais, facilita a construção de novos DMIs simplesmente pelo fato de que o músico pode utilizar seus próprios *samples*, sejam eles de instrumentos acústicos ou não.

Todavia, o excesso de programas integrados pode causar adições desnecessárias de latência, levando ao retorno de uma implementação em C/C++, porém utilizando o protocolo MIDI para enviar as informações ao Live.

6.6 **Versão beta 1**

Esta versão⁴⁰ incorpora o princípio de integração de sensores da abordagem da versão anterior, com a tentativa de alta performance da implementação em C/C++ buscada na versão alfa 1. Além disso, para reduzir ainda mais a latência encontrada na versão anterior, esta configuração envolve uma nova abordagem de captura de gestos.

6.6.1 *Tecnologias Utilizadas*

Nesta época a empresa Leap Motion© lançava a segunda versão de rastreamento (o *Skeletal Tracking*) que envolvia capacidade de reconhecer os ossos da mão e diversas outras características, e os algoritmos de rastreamento

⁴⁰ Vídeo disponível em: <<https://youtu.be/wpA24R6fFpc>>.

havia sido melhorados. Além disso, o suporte da companhia para aplicações em realidade virtual (VR), fez com que o dispositivo tivesse sua robustez aprimorada para ser utilizado em outras configurações (anexado a óculos de realidade virtual, rastreando de cima para baixo, entre outros), contanto que o sensor esteja estável.

Com isto, a ideia de utilizar este sensor rastreando a cena de cima para baixo facilitava seu uso em qualquer superfície e também o transporte do sistema, já que não seria preciso carregar a folha de vidro. Outra vantagem de não utilizar a folha de vidro é que a problemas de rastreamento, causados por uma possível refração dos feixes de luz do sensor, não existiriam.

O sistema ainda utilizava o piezo conectado ao arduino, enviando as informações de *trigger* para o programa desenvolvido em C/C++ no Visual Studio. O piezo ainda captura informações relativas à intensidade do golpe do músico na superfície de contato, à qual ele está anexado, fornecendo assim, um controle de dinâmica do volume (parâmetro *velocity* do MIDI) do instrumento. Esta funcionalidade foi implementada demanda encontrada após a análise dos resultados do questionário, referente aos controladores digitais de percussão, realizado com os percussionistas profissionais, os quais mencionaram que a ausência desta funcionalidade é um fator desagradável neste tipo de instrumento.

As bibliotecas utilizadas foram a do próprio dispositivo Leap Motion®, a biblioteca MIDI *RtMidi* e também as bibliotecas do sistema operacional, as quais se encarregam de diversas funções, entre elas a comunicação serial com o arduino.

A biblioteca MIDI utilizada foi desenvolvida na universidade de McGill, no Canadá, por Gary P. Scavone e modificada por diversos outros pesquisadores. Esta biblioteca multi-plataforma – Mac, Windows e Linux – fornece uma interface de programação de aplicativo (API, do inglês *Application Programming Interface*) que simplifica o protocolo MIDI para os desenvolvedores.

É importante ressaltar que para reduzir a latência do sistema e para que o protocolo MIDI seja passado para o Ableton Live, dois drivers foram usados: o ASIO4ALL e um driver MIDI.

ASIO (Audio Stream Input/Output) é um protocolo de som desenvolvido para fornecer baixa latência e alta performance entre as aplicações do computador e sua placa de som. O ASIO4ALL é um *software* gratuito que simula este driver e disponibiliza o protocolo para usuários que não possuem o driver.

Além disso, um driver MIDI é necessário já que o sistema operacional Window 8/8.1 não possuem um acesso fácil ao mapeamento MIDI da placa de som do computador.

6.6.2 Configuração

A configuração deste sistema contava com o Leap Motion© anexado a um pedestal improvisado, rastreando a cena de cima para baixo. Deste jeito, o usuário golpeia uma superfície qualquer, por exemplo uma mesa, na qual o microfone de contato está colado.

Entretanto, a utilização do sensor apontado diretamente para a mesa faz com que os feixes IR emitidos pelos LEDs do sensor sejam retrorefletidos pela superfície, ou seja, refletidos de volta para as câmeras do sensor. Isto “cega” o dispositivo, que não consegue distinguir os objetos próximos à superfície de contato.

Para solucionar este problema, uma folha preta fosca de Etil Vinil Acetato (EVA), mais conhecida como emborrachado, foi posta em cima da superfície. Assim, os raios IR não são retrorefletidos e o sensor consegue capturar as informações.

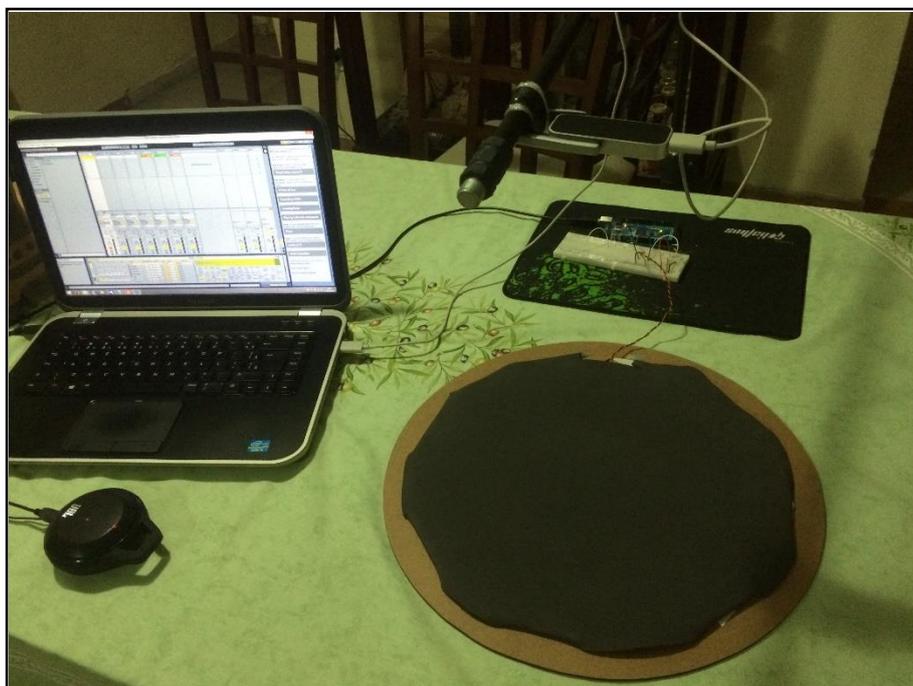


Figura 6.15: Configuração do protótipo versão beta 1.

Para reduzir ainda mais a latência do sistema e se aproximar aos valores descritos por Wessel and Wright (Wessel and Wright, 2002), a abordagem de rastreamento adotada envolve a delimitação de uma área em três dimensões um pouco acima da superfície, na qual as características capturadas pelo sensor são armazenadas em variáveis que serão utilizadas para distinguir os diferentes gestos. Isto é possível tendo em mente que, quando o músico vai realizar um determinado gesto, existe uma certa distância da superfície na qual o gesto não muda – variando apenas outras características do sistema, como intensidade da nota e pressão na pele do instrumento – e o movimento é realizado pela inércia. Assim, o sistema identifica o gesto realizado antes de receber a informação do piezo e gera uma nota MIDI que é mandada para o Ableton Live. Portanto, o

rastreamento pode ser feito um pouco acima da superfície, já que a intensidade é capturada pelo piezo. Ao receber as informações do piezo, o sistema envia a nota MIDI previamente calculada para o Live.

6.6.3 Discussões

Com este sistema, os requisitos de latência foram alcançados e as medidas de latência resultaram em aproximadamente 7 milissegundos, mostrando que a implementação do piezo e a abordagem de rastreamento permitem esta redução.

Nesta versão, a análise de gestos mostrou a capacidade de captura da maioria dos gestos da conga. Entretanto certos gestos como o *Glissando* (ou *Gliss*), o abafamento das notas, pressionando a pele com a mão, e o gesto de se tocar com a mão com o punho fechado não eram reconhecidos pelo sistema.

No primeiro caso, o gesto é realizado pela vibração da pele do material ao se arrastar o dedo. A tensão aplicada na pele esticada e o atrito com o dedo do músico causam som distinto. Ou seja, o gesto está diretamente relacionado com a síntese sonora, o material e o formato do instrumento acústico.

No sistema digital, a diferença de material e a ausência da tensão no mesmo não provocam a mesma vibração e o microfone o sistema não consegue captar gesto. Como o controle de entrada gestual e a síntese sonora estão separados nos DMIs, a captura de alguns gestos não são possíveis, mesmo que possam ser incorporados de outra forma. Além disso, a escolha de utilizar o piezo para servir como *trigger* do sistema e a utilização de *samples* prejudicam no reconhecimento deste gesto.

O caso do abafamento das notas também leva em conta o material do instrumento e a relação direta entre o gesto e a produção sonora. O abafamento ocorre quando o músico pressiona a pele do instrumento, impedindo que a mesma vibre como faria normalmente, isto causa a geração de outra sonoridade do instrumento. Novamente, como a produção sonora numa conga acústica é causada pela vibração da pele, abafar a mesma produzirá um som diferente, o que não é possível nesta versão do sistema DMIs.

No último caso, a nota tocada com a mão em formato de punho fechado é utilizada pelos percussionistas nas chamadas *grace notes*, são notas muito rápidas que precedem outras notas com a finalidade de produzir um som diferenciado. Neste caso, capturar o áudio gerado por estas notas é bastante complicado. Por isto, mesmo que o sistema seja capaz de capturar o gesto realizado, o mesmo não possui esta nota pela dificuldade na captura do *sample* sonoro.

6.7 Versão final

Esta versão segue a mesma linha do último protótipo, entretanto procura resolver os problemas encontrados no mesmo, i.e., os problemas da falha de captura de alguns gestos.

6.7.1 *Tecnologias Utilizadas*

Esta versão utiliza as mesmas tecnologias da versão anterior, com a única exceção da incorporação de um novo sensor para reconhecimento de novas nuances gestuais.

A solução do problema de captura das nuances gestuais do abafamento de notas foi realizada incorporando um sensor resistivo FSR (do inglês, *Force Sensitive Resistor*). O sensor é integrado ao sistema utilizando o arduino e seus valores também são enviados por meio de um protocolo serial.

Neste caso, o valor da resistência gerada pelo resistor varia à medida em que o usuário o pressiona. Deste modo, é o sensor ideal para simular a pressão exercida pelo músico na pele do instrumento acústico.

Uma melhoria feita a partir dos resultados da entrevista com o usuário foi a sensibilidade do sensor piezo. Incorporando um novo algoritmo de detecção do sensor piezo, modificado de um algoritmo disponível *online*⁴¹, foi possível aprimorar a captura do mesmo.

Um sistema físico, que está sendo desenvolvido em colaboração com o luthier Paulo Sérgio Nunes⁴², ainda incorpora mais microfones de contato para melhorar a sensibilidade da superfície de contato.

6.7.2 *Configuração*

Nesta configuração, o EVA é colocado em uma nova camada de madeira que fica em cima do sensor de pressão. Ao pressionar a superfície, o sensor manda o valor do FSR para o *software* que irá processar o valor do abafamento. Este valor pode ser usado para a distinção de um novo *sample* ou enviado diretamente ao Ableton Live, controlando algum efeito de abafamento.

Além disso, o instrumento físico pretende incorporar uma haste para que o Leap Motion© seja fixo, sem precisar de um pedestal externo. Esta escolha está relacionada com o aumento da portabilidade do sistema.

6.7.3 *Discussões*

⁴¹ Disponível no site: <<http://www.instructables.com/id/Arduino-Xylophone/step7/Connecting-to-the-Arduino/>>.

⁴² Site com informações sobre o luthier: <www.psluthier.com>.

Os resultados deste sistema são extremamente similares aos da versão anterior, com a exceção de que há a capacidade de reconhecimento das nuances de abafamento das notas.

Além disso, mesmo que o gesto do *Glissando* não tenha sido reconhecido, observou-se que ele pode ser incorporado ao sistema utilizando o gesto de *swipe*, intrínseco ao Leap Motion®. Entretanto, uma avaliação da usabilidade deste gesto não convencional integrado aos gestos do instrumento deve ser feita.

Outro gesto que não é possível de ser capturado, pela configuração do sistema é o abafamento que alguns percussionistas fazem com o cotovelo. Neste sistema não é possível distinguir este tipo de abafamento por limitações de rastreamento do Leap Motion®, já que o mesmo não captura o antebraço nem o ponto do cotovelo do usuário (como é possível no Kinect®) com uma boa estabilidade, devido às limitações espaciais do sensor.

Todavia, nesta configuração, a nuance gestual de tocar o instrumento com a mão fechada em formato de punho foi implementada utilizando outro *sample* sonoro para que a avaliação do reconhecimento deste gesto fosse testada.

É importante ressaltar que algumas nuances encontradas *online* não foram capturadas pelo fato de serem usadas por indivíduos específicos e pela baixa qualidade sonora dos vídeos, nos quais elas aparecem.

7 RESULTADOS OBTIDOS

Neste capítulo, os resultados obtidos ao longo da pesquisa são explicitados. Primeiramente, são apresentados os métodos de avaliação quantitativas e qualitativas. Em seguida, são apresentados os resultados para a latência de cada versão do protótipo e os resultados relativos às capacidades de rastreamento do dispositivo, particularmente os gestos do instrumento escolhido (Congas). Por fim, são apresentados os resultados qualitativos do experimento realizado com o músico profissional.

7.1 Métodos de avaliação

Após a escolha do instrumento a ser simulado, foi feito um levantamento dos gestos necessários para realização das nuances musicais possíveis de serem alcançadas com o instrumento. Assim, avaliando a capacidade do sistema de reconhecer estas nuances, pode-se desenvolver um sistema musical expressivo. Esta abordagem parte de uma análise quantitativa do reconhecimento de gestos por parte das tecnologias utilizadas.

As tecnologias foram testadas seguindo os mesmos princípios de baixo custo de implementação e, tendo em vista a integração de sensores, a apresentação da avaliação segue uma cronologia de integração de cada sensor no sistema desenvolvido e quais nuances podem ser capturadas em cada etapa. Esta cronologia segue as iterações do processo de *design* e os protótipos desenvolvidos ao longo do projeto.

Diferentes métodos foram utilizados, de forma a obter informações quantitativas e qualitativas complementares que proporcionem uma visão geral sobre as capacidades de reconhecimento dentro e fora do laboratório. Ou seja, uma avaliação quantitativa sobre as capacidades de reconhecimento e a apresentação do protótipo do sistema ao público alvo.

Com isso, busca-se entender o quão longe as tecnologias atuais estão de permitir a elaboração de sistemas musicais expressivos, como são os instrumentos acústicos, identificar os problemas apresentados pelo sistema desenvolvido e responder à pergunta de pesquisa elaborada durante o projeto.

7.1.1 Método de avaliação da latência

Um dos problemas relatados na avaliação das tecnologias analisadas por Silva (SILVA, 2012) foi a latência dos DMIs construídos.

Com isto, a análise da latência foi considerada essencial, tendo em vista também os estudos de tolerância de latências dispostos na literatura (JORDA, 2005; WESSEL; WRIGHT, 2002).

A medição foi realizada utilizando um microfone posicionado para capturar o som das mãos do usuário golpeando a superfície de contato (que fornece o *feedback* tátil) e a produção sonora gerada pelo computador. Utilizando o *software* gratuito *Kristal Audio Engine*, foi possível gravar os áudios e enviar para um programa de análise de áudio para comparar a diferença temporal na saída sonora.

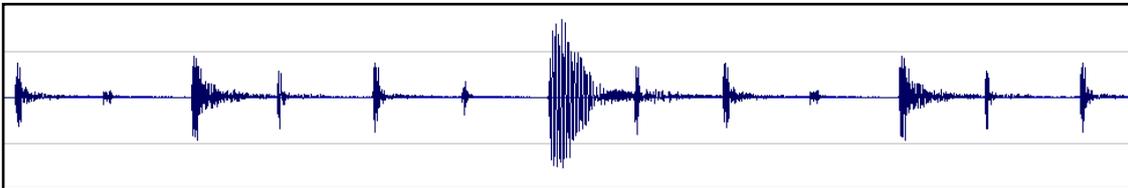


Figura 7.1: Formato de onda de percussão.

Geralmente, o som de um instrumento de percussão possui uma forma similar ao mostrado na Figura 5.3. Os sons percussivos consistem, na maioria das vezes, de um ataque súbito seguido de um decaimento, com muito pouca ou quase nenhuma sustentação. Este formato de onda é propício para distinção de dois sons no *software* de análise. Reconhecendo os dois picos do som causado pelo golpe na superfície e do som gerado pelo computador, pode-se calcular a latência do instrumento. Apesar de não ser exatamente o momento em que o áudio realmente começa – primeira amostra com amplitude acima do nível de ruído –, existe uma dificuldade para medi-lo, principalmente ao comparar dois sons muito próximos. Como os máximos locais das amostras sonoras percussivas são muito próximos ao início do som, é possível utilizar este momento para a medição. Além disso, a diferença temporal entre dois sons percussivos consecutivos será a mesma medindo da primeira amostra sonora ou do máximo local.

7.1.2 Método de avaliação do reconhecimento de gestos

Para avaliar a capacidade do sistema em reconhecer os gestos e nuances de expressividade do instrumento escolhido (congas), foi utilizado um método de duas etapas, no qual os gestos são realizados um determinado número de vezes e os erros de reconhecimento são computados para que seja possível avaliar a confiabilidade no reconhecimento destes gestos.

Na primeira etapa, o usuário realiza os gestos com apenas uma das mãos sendo reconhecida pelo sistema. O usuário repete este gesto até completar a quantidade pré-estabelecida e os erros de reconhecimento são computados para elaboração de uma tabela.

Na segunda etapa, os gestos são realizados com as duas mãos sendo reconhecidas pelo sistema. Primeiramente, com uma das mãos e, posteriormente, com a outra mão. Da mesma forma, os gestos que necessitam das duas mãos para serem reconhecidos são feitos com cada uma das mãos.

Desta maneira, é possível identificar a taxa de confiabilidade na captura dos gestos, e assim determinar se a mesma é válida em diferentes contextos de uso da interface.

Além disso, partindo de uma abordagem similar à utilizada por Wanderley e Orio (WANDERLEY; ORIO, 2002), durante a apresentação do protótipo ao público profissional, foram apresentadas diferentes tarefas musicais foram realizadas. Com isto pode-se avaliar a capacidade de reconhecimento gestual no contexto de uso do instrumento, no qual há uma variação dinâmica dos gestos realizados para produção do acompanhamento musical.

7.1.3 Método de avaliação qualitativa

Após a elaboração do último protótipo, o mesmo foi apresentado ao público profissional de forma a coletar *feedback* sobre o sistema para descobrir se o sistema satisfaz as demandas de contexto de uso, quais funcionalidades atendem ao público alvo e quais destas não são desejadas por eles.

Foi seguido um método de avaliação qualitativa de DMIs utilizado por alguns pesquisadores da área (BARBOSA et al., 2011; SILVA, 2012). Este tipo de avaliação qualitativa pode ser dividido em três etapas:

- **Exploração livre:** Nesta etapa, o instrumento é apresentado ao usuário. Os gestos são explicitados e o usuário pode testar o instrumento de maneira mais informal, sem nenhuma restrição.
- **Exploração guiada:** Após a primeira etapa, os usuários devem realizar certas tarefas rítmicas comuns ao instrumento escolhido, para que uma análise da captura dos gestos seja feita enquanto o usuário está utilizando o instrumento normalmente.
- **Entrevista semiestruturada:** Nesta etapa, uma entrevista é conduzida para coletar informações sobre a experiência do usuário que podem não ficar claras durante a avaliação quantitativa.

Segundo Stowell et al. (STOWELL et al., 2009), a utilização de métodos qualitativos proporciona a descoberta de informações que não podem ser obtidas na avaliação quantitativa. A subjetividade do contexto de uma *performance* musical faz com que certos aspectos do sistema, como *affordance* – capacidade de um artefato de prover informações suficientes ao usuário de como ser manipulado –, confiabilidade para um determinado contexto de uso, confiança do músico ao tocar o instrumento, entre outros, possam não ser enxergados por avaliações quantitativas.

Desta forma, a realização de entrevistas e a utilização do método de análise de discurso podem beneficiar a avaliação, complementando os fatos obtidos quantitativamente com informações mais subjetivas. Assim, pode-se obter resultados mais realistas.

A análise de discurso segue um processo que pode ser dividido da seguinte maneira (SILVA, 2012, p. 46):

- **Transcrição:** Etapa em que o áudio do material obtido (gravações de áudio e vídeo) são traduzidos em linguagem textual.
- **Livre associação:** Etapa em que se busca anotar impressões para serem analisadas posteriormente.
- **Discriminação:** Etapa em que se tenta extrair informações, da transcrição, sobre os temas relativos à pergunta de pesquisa. O texto é quebrado em diversos pedaços e informações comuns são reconhecidas, tendo em mente as terminologias utilizadas durante a avaliação.
- **Reconstrução do mundo do usuário:** Etapa na qual busca-se reconstruir a experiência do usuário a partir dos elementos emergidos da etapa anterior, tendo em mente o contexto no qual ele está inserido.
- **História Central:** Nesta etapa, todas as análises são combinadas para construir a experiência geral e buscar responder à pergunta de pesquisa. Assim, as informações relativas às experiências particulares podem ser generalizadas para o contexto de uso do sistema que está sendo avaliado.

Silva ainda mostra que apesar de este método de avaliação possuir 5 etapas, a “Livre associação” não é estritamente necessária para a análise completa do sistema. Pode-se, assim, extrair as informações relevantes para a pergunta de pesquisa utilizando apenas 4 etapas.

7.2 Resultados da avaliação da latência

Como mencionado no capítulo de construção do protótipo, a latência foi aprimorada durante a construção do artefato, escolhendo novas abordagens e implementando novos sensores. A tabela abaixo apresenta a latência de cada versão desenvolvida. Os resultados apresentados já são aproximados.

	Crystal Piano	Versão alfa 1	Versão alfa 2	Versão beta 1	Versão final
Latência (ms)	150	71	17	7	7

Tabela 4: Latência do artefato desenvolvido.

Entre a versão “alfa 1” e a versão “alfa 2” nota-se uma enorme redução da latência que se deve ao fato da integração do microfone piezoelétrico no sistema. Da versão “alfa 2” para as versões “beta 1” e “final”, a redução da latência parte da mudança de ambiente de programação e incorporação do driver ASIO ao sistema.

7.3 Reconhecimento de gestos

Nesta seção, são apresentados os resultados da avaliação da captura dos gestos do protótipo “beta 1” em relação aos gestos específicos das congas.

O sistema desenvolvido neste projeto, consegue capturar a maioria dos gestos específicos das tumbadoras. Como foi explicado mais detalhadamente no capítulo 6, a versão beta1 da V-Conga (Seção 6.6) só não é capaz de capturar os gestos que produzem o *Gliss*, a nuance de abafamento com o cotovelo e os abafamentos pressionando a pele do instrumento. Entretanto, as notas que utilizam a mão pressionando o instrumento puderam ser diferenciadas utilizando um parâmetro de controle que identifica se os dedos da mão (com exceção do polegar) estão juntos ou separados – descritos no sistema como “mão fechada” e “mão aberta, respectivamente.

Já na versão final da V-Conga (Seção 6.7) a integração de um sensor de “pressão” FSR (*Force Sensitive Resistor*) permite que outro nível de abafamento seja implementado. Com isto, um abafamento mais “natural” pode ser simulado, já que o usuário pode pressionar a superfície do instrumento analogamente ao instrumento acústico.

Para avaliação da captura dos gestos deste sensor, 60 batidas foram realizadas e computadas para detecção de possíveis erros de captura. Foram realizadas batidas com apenas uma mão e batidas com as duas mãos presentes. Neste último caso, a detecção de cada uma das mãos foi analisada independentemente.

Como cada gesto está associado a um *sample* sonoro distinto (a geração sonora é discreta), pode-se definir que um gesto “correto”, ou a captura “correta” do gesto, é quando o usuário executa o gesto e o *sample* correspondente é disparado. Se o *sample* disparado estiver associado a outro gesto, ou se o sistema não disparar nenhum som, é identificado o erro de captura.

Os gestos foram realizados pelo autor desse projeto sem acompanhamento musical, porém em uma velocidade de 140 batidas por minuto (bpm). Para essa análise, cada gesto foi feito independentemente, seguindo a sequência de gestos da tabela abaixo.

Notas	Porcentagem de Captura
Bass	93,33%
Open	98,33%
Muff	100%
Slap	98,33%
Closed Slap	100%

Heel	100%
Tip	96,67%
Fist	100%

Tabela 5: Reconhecimento dos gestos da conga com apenas uma mão.

É possível notar que, com apenas uma mão sendo reconhecida pelo sistema, todos os gestos são capturados com um alto grau de estabilidade do sistema, apresentando poucos erros.

No sistema, as notas *Tip* e *Bass* são diferenciadas a partir da movimentação da base da palma da mão em relação à superfície. Possíveis flutuações no reconhecimento dos gestos podem causar esses erros de reconhecimento. Uma maior tolerância no rastreamento pode auxiliar o controle do sistema.

A avaliação com ambas as mãos seguiu um protocolo experimental similar. Neste caso, os gestos foram realizados com uma das mãos (igualmente ao protocolo anterior), porém o mesmo gesto era avaliado com a segunda mão antes de passar para o próximo.

Notas	Percentagem de captura da mão 1	Percentagem de captura da mão 2
Bass	0%	0%
Open	96,67%	5%
Muff	100%	8,33%
Slap	96,67%	6,67%
Closed Slap	91,67%	13,33%
Pressed Slap	100%	8,33%
Heel	100%	5%
Tip	100%	0%
Fist	100%	1,67%

Tabela 6: Reconhecimento dos gestos da conga com ambas as mãos.

Neste caso, foi possível notar a instabilidade do sistema no reconhecimento da segunda mão. Com exceção da *Bass Note*, todos os gestos foram reconhecidos com grande estabilidade na mão 1, enquanto na mão 2 o reconhecimento apresentou diversas falhas. Elas estão relacionadas com problemas relacionados com **oclusão de uma das mãos** e **perdas de rastreamento**.

A partir desses resultados, um aplicativo disponível na galeria de desenvolvedores do Leap Motion®, chamado VR Intro⁴³, foi testado para verificar se os problemas de rastreamento eram uma particularidade do sistema desenvolvido neste projeto. Da mesma forma, o aplicativo apresentou diversos problemas de perda de rastreamento de uma das mãos ao se aproximar da superfície do instrumento.

A falta de robustez apresentada pelo Leap Motion® ao ser utilizado na configuração de cima para baixo, quando o usuário está usando as duas mãos para tocar o instrumento, faz com que o sensor altere os parâmetros de descrição da cena. Como o mapeamento das nuances requer uma boa precisão para que os gestos sejam reconhecidos, uma robustez baixa faz com que notas falso positivas e falso negativas ainda sejam apresentadas pelo sistema.

Além disso, na apresentação do protótipo ao público profissional, quatro tarefas foram selecionadas para serem realizadas para também serem computados os erros. Novamente, nos movimentos dinâmicos, mais rápidos e nos quais as mãos ficavam muito próximas e se sobrepunham constantemente, os músicos não conseguiram realizar as tarefas. Uma das tarefas que pôde ser realizada com apenas uma das mãos, foi computada e ainda apresentou alguns erros. Esses erros não estão relativos a falhas de reconhecimento do sistema, mas relativos ao usuário não realizar o gesto na posição espacial correta. Neste caso, o usuário pretendia tocar a *Slap Note*, mas não realizava o gesto mais ao centro do instrumento, o que proporcionava a produção da *Open Note*.

A implementação do sistema utilizando as mãos do autor deste projeto fez com que algumas notas não fossem reconhecidas quando o usuário tentava tocar. Devido às diferentes características de inclinação, tamanho da mão, largura dos dedos, entre outros, a implementação dos gestos passa a ser mais complicada para atender à diferentes músicos, que certamente possuirão outras medidas de mão. Em particular, o *Heel Tone*, por ser capturado a partir da inclinação da mão em relação ao plano da superfície, foi um dos gestos que não foram reconhecidos quando realizado pelo músico.

7.4 Resultados qualitativos

A avaliação do reconhecimento de gestos por meio de uma abordagem quantitativa proporciona fatos sobre as capacidades do reconhecimento de gestos do sistema. Entretanto, de forma a levantar vantagens e desvantagens e analisar a experiência do usuário com este sistema, foi necessário apresentar o sistema ao público alvo – músicos profissionais. Desta maneira, pode-se coletar informações que não podem ser obtidas por meio da análise quantitativa⁴⁴.

⁴³ Aplicativo disponível em: <<https://developer.leapmotion.com/gallery/leap-motion-vr-intro>>.

⁴⁴ Vídeo demonstrativo em: <https://youtu.be/4lw_Zfc74Xc>.

Devido aos diversos problemas e instabilidade do sistema observados durante o primeiro experimento, considerou-se prudente realizar a entrevista com apenas um usuário, já que o sistema não estava suficientemente estável para ser apresentado aos percussionistas profissionais. Sendo assim, foi feita a escolha de solucionar estes problemas do sistema antes de apresentá-lo ao público alvo, deixando uma maior análise qualitativa para trabalhos futuros. Ainda assim, os resultados do experimento qualitativo são apresentados abaixo.

Mesmo que esses resultados não possam ser generalizados, decidiu-se seguir rigorosamente o método de avaliação qualitativa.

7.4.1 *Protocolo Experimental*

O protocolo experimental da avaliação qualitativa que foi desenvolvido possui três etapas: Exploração Livre, Exploração Guiada e Entrevista semiestruturada.

Na primeira etapa, o usuário é apresentado ao sistema. O usuário explora o instrumento sem nenhuma instrução, apenas seguindo os conhecimentos musicais prévios ao experimento. Assim, todos os gestos são apresentados e o usuário busca testá-los antes de explorar a dinâmica do instrumento.

Já etapa seguinte, quatro tarefas musicais foram escolhidas para testar o reconhecimento dos gestos em rudimentos de percussão – ritmos utilizados para treinos e desenvolvimento de técnicas de percussão. Os ritmos escolhidos tentam explorar os gestos implementados e avaliar a capacidade do sistema de rastrear as mudanças de gestos. Uma maior variedade de gestos está diretamente relacionada com um maior repertório sonoro do instrumento que é essencial na expressividade musical.

Dois dos ritmos selecionados foram retirados do livro “*The Book of Contemporary Conda Techniques*” (NECIOSUP; ROSA, 2007), enquanto a técnica do ritmo 3 e o ritmo 4 são muito comuns no aprendizado de congas e, particularmente o ritmo 4, serve de base para diversas variações rítmicas. Nas notações dos ritmos, as letras **R** e **L** determinam a mão com a qual o músico deve executar determinada nota: direita e esquerda, respectivamente. Algumas notações ainda mostram onomatopeias dos sons que são produzidos ao tocar cada nota.

Os ritmos escolhidos foram os seguintes:

- **“The Four Stroke Ruff”**: Ritmo simples de 4 tempos alternando as mãos. Neste caso, as três primeiras notas são *Open Tones* (notação **O**) e a quarta é um *Slap* (notação **S**). Este ritmo foi escolhido por ser simples e possível de ser executado com apenas uma das mãos ou ambas.

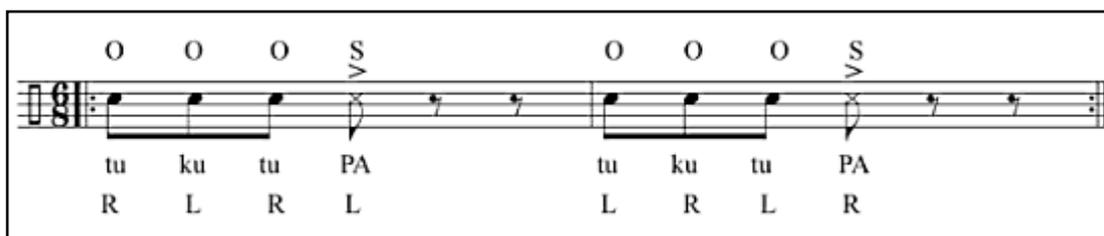


Figura 7.2: Ritmo 1 – Four Stroke Ruff.

- **“The Flam Tap”**: Neste ritmo, o usuário executa um *Flam* e logo em seguida, executa uma nota qualquer. Escolhendo este gesto pode-se avaliar a velocidade na detecção das notas do *Flam*, e variar as notas subsequentes. Neste ritmo, a notação temporal da velocidade não é mostrada, já que os autores do livro utilizam este ritmo com o intuito de o leitor variá-la.



Figura 7.3: Ritmo 2 – Flam Tap.

- **“The Heel-Tip Technique”**: Esta técnica segue a mesma linha do ritmo anterior. É uma técnica bastante comum em diversos ritmos musicais e utilizadas em diversos rudimentos. Nesta tarefa, o usuário deve executar a técnica e aumentar a velocidade para testar a confiabilidade do reconhecimento deste gesto. Na figura, as notações **P** e **T**, denominam as notas *Heel* (ou *Palm*) e *Tip*, respectivamente.



Figura 7.4: Ritmo 3 – Heel-Tip Technique.

- **“The Tumbao”**: Este ritmo é bastante comum e possui diversas variações para diferentes estilos musicais. É um exercício básico que utiliza algumas notas da conga. Neste caso, um ritmo comum, porém mais complexo, pode testar a robustez do sistema durante a permutação de gestos.



Figura 7.5: Ritmo 4 – Tumbao.

Por fim, foi coletada uma entrevista semiestruturada de forma a coletar informações sobre a experiência do usuário ao utilizar o sistema em questão.

7.4.2 *Entrevista*

A entrevista foi feita para extrair informações referentes aos seguintes temas:

- Como foi a sua experiência com o instrumento?
- Você sentiu um algum *delay*?
- O que você achou do material do instrumento?
- Quais são as vantagens que você percebeu neste sistema?
- Quais são as desvantagens que você percebeu neste sistema?
- Quais melhorias podem ser feitas ao sistema?

Esta última pergunta foi respondida ao longo das observações durante o experimento. O usuário notificava constantemente quando um erro era notado por ele.

Durante a entrevista, foi utilizado o termo “instrumento” para identificar o sistema, de forma a facilitar as comparações com instrumento acústicos, o processo cognitivo do usuário e a análise do discurso.

A transcrição da entrevista está disponível *online*⁴⁵.

7.4.3 *Reconstrução do mundo do usuário e observações durante o experimento*

O experimento foi realizado com um músico profissional formado pelo Conservatório Pernambucano de música e que trabalha como *sound designer*, seguindo o protocolo experimental descrito na seção 7.4.1 e durando aproximadamente 50 minutos. Durante sua realização, o músico relatou diversos problemas e possíveis melhorias que podem ser feitas ao mesmo.

Foi observado que mesmo após a explicação dos gestos do sistema – posição espacial das notas e as características das mãos usadas para o rastreamento –, o usuário ainda realizava os gestos de acordo com sua formação musical. Os gestos feitos com apenas uma das mãos foram capturados pelo sistema. Entretanto, como o mapeamento foi modelado usando a mão do autor deste projeto, gestos feitos por vezes não correspondiam com os desejados pelo usuário.

Pelo fato de cada músico executar o gesto de maneira um pouco diferente, mesmo usando a mesma técnica (e.g. curva da mão no *Closed Slap* pode ser maior ou menor ou a inclinação na *Heel Note* pode variar), observou-se que o

⁴⁵ Disponível em: <http://www.4shared.com/rar/82yBI9Kcba/Experimento_Conga_Virtual_-_Ed.html>. Acessado em: 04/08/2015.

sistema precisa ser ter maior tolerância a variações que podem ocorrer por essas diferenças ou por questões anatômicas (e.g. tamanho da mão de cada músico).

Durante a primeira parte do experimento (exploração livre), já foi possível notar a instabilidade do sistema no reconhecimento das duas mãos. Essa instabilidade causa problemas relativos ao reconhecimento incorreto das características das mãos e perda de rastreamento que, conseqüentemente, acarretam em notas falso negativas e disparo de *samples* incorretos. Assim, o usuário não pôde completar a maioria das tarefas musicais escolhidas para a segunda etapa do instrumento.

Embora o ritmo 1 possua a alternância das mãos, o usuário conseguiu completá-lo utilizando apenas uma das mãos. Isto foi feito para avaliar se o comportamento do sistema seria mais estável nesta forma de rastreamento. No caso do ritmo 3, que também pôde ser feito com uma mão, a diferença anatômica das mãos do usuário em comparação com as mãos usadas para implementar o sistema fez com que apenas a *Tip Note* fosse reconhecida.

Devido a esses problemas observados durante a utilização da conga, outro instrumento foi apresentado ao usuário para investigar se a direção de rastreamento do sensor influencia a estabilidade do sistema. O instrumento escolhido foi um “Cajón” com apenas 4 notas. Este instrumento também é bastante versátil e utilizado em diversos estilos musicais.



Figura 7.6: Cajón virtual (esquerda) e cajón acústico (direita).

O cajón utiliza a mesma técnica de rastreamento das congas. As diferenças entre estes dois instrumentos são percebidas na orientação de captura do sensor – de cima para baixo no caso das congas e de lado no caso do cajón – e no mapeamento do instrumento (devido à diferença dos gestos).

Durante a exploração livre, o usuário percebeu que o sistema apresentou uma maior estabilidade relatando que o instrumento estava “bem melhor” em comparação a conga e permitindo que o usuário pudesse tocar um ritmo da sua escolha. Entretanto, este foi um caso particular, já que foi o único ritmo completado com as duas mãos durante toda a primeira etapa do protocolo. Já na segunda etapa, apenas os ritmos 1 e 2 podem ser avaliados no cajón. Mesmo assim, o sistema ainda apresentou problemas de instabilidade, durante essa exploração guiada, ao usar as duas mãos.

No caso do cajón, ainda foi possível perceber a presença de notas falso negativas. Neste caso, isso está relacionado com a sensibilidade do microfone piezo de capturar os golpes no instrumento. A vibração mínima para o disparo de uma nota é afetada pela espessura do material no qual o cajón foi desenvolvido. O músico ainda levantou uma possível melhoria do sistema relacionada com a mudança para um material mais fino, para que melhorar a captura do microfone.

7.4.4 Codificação e Mapa mental

Nesta seção, são apresentados os elementos comuns (ou códigos comuns), emergidos da análise da entrevista, utilizados para construir o mapa mental (que demonstra as relações entre estes elementos). A definição destes códigos facilita o entendimento do mapa mental da seção 8.5.

1. **Instrumento Virtual** – Instrumento, desenvolvido no projeto, que simula um instrumento de percussão.
2. **Instrumento Acústico** – Instrumentos de percussão acústicos.
3. **Características do Instrumento** – Propriedades intrínsecas e extrínsecas do instrumento.
 - 3.1. **Material do instrumento** – Material (ou materiais) que fazem parte do instrumento (madeira, pele, corda, etc.).
 - 3.1.1. **Espessura do material** – Uma das dimensões de um objeto tridimensional.
 - 3.2. **Vibração do Instrumento** – Vibração física do material do instrumento (corpo e pele).
 - 3.3. **Portátil** – Capacidade de um objeto de ser transportado facilmente.
 - 3.4. **Tolerância para com erros do usuário** – Possuir a capacidade de permitir certos erros que podem acontecer, sem prejudicar a experiência.
 - 3.5. **Latência** – Atraso relativo ao gesto realizado pelo músico e a produção sonora.
4. **Conhecimento Musical** – Conhecimento inerente ao músico.
 - 4.1. **Gesto Mental** – Gestos formados cognitivamente pelo músico a partir do seu conhecimento musical. Gestos que são antecipados pelo usuário durante a *performance*.

5. **Novo Aprendizado** – Processo de passar a ter conhecimentos sobre algo novo. Re-aprender.
6. **Tempo** – Medida de progressão de eventos.
7. **Costume** – Hábito comum aos membros de um certo grupo social.
8. **Estudo** – Ato de procurar adquirir o conhecimento sobre algo.
9. **Correspondência Gesto e Som** – Relação entre o gesto que é feito no instrumento e o som que é produzido.
10. **Estranhamento** – Sentimento de achar algo fora do comum.
11. **Mobilidade** – Relativo ao transporte de certos objetos.
12. **Gravação** – Processo de registrar composições em estúdios musicais.
13. **Sensores** – Dispositivos eletrônicos que capturam informações da cena.
 - 13.1. **Precisão** – Rigor no registro de valores. Exatidão nas medidas.
 - 13.2. **Sensibilidade** – Qualidade de identificar mínimas variações de valores de medida de um determinado sensor.
14. **Experiência do usuário** – Vivência no contexto do uso de determinado objeto.
15. **Vantagem** – Pontos positivos do sistema.
16. **Desvantagem** – Pontos negativos do sistema.
17. **Melhorias** – Aprimoramentos no sistema.

7.4.6 *História Central*

Essa seção descreve a história central obtida por meio da análise das relações entre os códigos emergentes na análise da entrevista. Essas relações foram extraídas a partir da análise da entrevista. Entretanto, as possíveis melhorias foram extraídas nos relatos do usuário durante o experimento.

A história segue abaixo:

Diversos aspectos afetam a experiência do usuário. As primeiras características do instrumento notadas são relativas aos aspectos físicos do instrumento. Muitas vezes, nos DMLs, a ausência do corpo e diferenças dos materiais do instrumento (em comparação com os instrumentos acústicos) provocam estranhamento por parte do músico que sente falta da pele do instrumento e da vibração do corpo do instrumento. Este estranhamento, porém, é superado pelo usuário que se acostuma com o material do sistema depois de um certo tempo.

Mesmo que a análise quantitativa da latência tenha mostrado resultados abaixo de 10 milissegundos, o usuário ainda percebeu um pouco de latência vinda do instrumento. Isto pode ser causado pela perda do reconhecimento da cena. Esse problema, proporciona notas falso positivas que podem atrapalhar o processo cognitivo do usuário que interpreta como uma latência.

As principais vantagens desse tipo de instrumento estão relacionadas com os fatos de o instrumento ser portátil e facilitar a gravação e o estudo. No primeiro caso, por ser construído a partir de sensores conectados a uma superfície de madeira e a um computador, o sistema permite uma maior mobilidade, solucionando um problema que é encontrado na vivência do percussionista profissional. No contexto de estudo do instrumento, o fato de o som ser gerado pelo computador se torna uma característica interessante por possibilitar o uso de fones de ouvido para um estudo privado. O sistema ainda possibilita uma facilidade no contexto de gravação, já que a saída sonora é uma informação digital (MIDI).

Entretanto, o sistema ainda possui diversas desvantagens quando comparado aos instrumentos acústicos. Além da latência sentida pelo músico, problemas relativos à precisão e sensibilidade dos sensores afetam bastante a experiência do usuário.

A necessidade de realizar os gestos com precisão faz com que muitas vezes não haja correspondência entre o movimento do usuário e o som que sai do instrumento – diferentemente do instrumento acústico que possui maior tolerância para com erros do usuário. Essa ausência na correspondência atrapalha o gesto mental do usuário, gerando confusão e afetando a experiência.

No outro caso, a sensibilidade do microfone de contato está diretamente relacionada com a espessura do material utilizado no instrumento e à captura

dos disparos. Problemas com a sensibilidade podem causar notas falso negativas e disparo de mais de uma nota ao mesmo tempo.

A solução desses problemas é essencial para o progresso do sistema e melhoria na experiência do usuário.

8 CONCLUSÕES E TRABALHOS FUTUROS

Neste projeto, foi apresentada uma abordagem para desenvolver sistemas digitais, mais especificamente para desenvolver Instrumentos Digitais Musicais Expressivos. Com o sistema desenvolvido, buscou-se simular um instrumento acústico de percussão e avaliar as capacidades de tecnologias existentes para desenvolvimento de instrumentos musicais digitais, em particular para simular uma conga. Foi dada ênfase em se obter baixa latência, grande capacidade de reconhecimento de nuances gestuais e no provimento de algum feedback tátil. Essa abordagem partiu da análise de trabalhos anteriores (principalmente trabalhos realizados dentro do próprio grupo de pesquisa do autor deste projeto), nos quais músicos profissionais levantaram os problemas associados com a expressão musical do instrumento. Ademais, a abordagem focou na integração de múltiplos sensores e visou baixo custo de implementação, utilizando um processo de desenvolvimento iterativo que envolve a evolução do instrumento por meio da avaliação do protótipo ao final de cada iteração.

Tecnologias como o Arduino permitiram utilizar uma abordagem que envolve a integração de múltiplos sensores com baixo custo de implementação, desenvolvendo um sistema que combina a alta granularidade de captura de características das mãos do sensor Leap Motion©, a baixa latência de captura do sensor piezoelétrico e um sensor de pressão. Desta maneira foi possível capturar a maioria das nuances gestuais utilizadas para se tocar uma conga, ainda oferecendo uma baixa latência e proporcionando *feedback* háptico ao usuário. Entretanto, apesar do grande potencial para captura de nuances gestuais, limitações relativas à robustez do sensor Leap Motion© apresentaram grandes problemas de instabilidade do sistema. Enquanto os gestos realizados com apenas uma das mãos era reconhecido pelo sistema, a taxa de reconhecimento da outra mão do usuário era mínima. Isso fez com que o sistema apresentasse restrições que afetaram a experiência do músico profissional.

Assim, foi possível desenvolver uma abordagem de criação de DMIs expressivos, utilizando integração de múltiplos sensores, para solucionar problemas relacionados com captura de nuances de expressividade, mesmo utilizando algoritmos de baixo custo de implementação. Para isso, utiliza-se uma estratégia de design do sistema, na qual cada um dos sensores é utilizado para capturar características específicas e a complementaridade dessas informações é usada para solucionar os problemas encontrados durante cada iteração do processo. Com isso, pode-se construir Instrumentos Musicais Digitais que possuem controles gestuais bastante refinados.

Ainda assim, embora haja um bom reconhecimento com “uma mão”, a instabilidade na captura das duas mãos mostra que novos avanços devem ser realizados para a construção de um sistema que possa ser utilizado por músicos profissionais. Algoritmos de visão computacional ou aprendizagem de máquina

devem ser investigados na tentativa de solucionar os problemas encontrados. Outra possível melhoria é a implementação de síntese por modelagem física que resultaria em uma maior fidelidade sonora, ainda incorporando elementos da acústica do instrumento. Entretanto, a utilização de algoritmos mais complexos pode trazer um maior custo computacional ao sistema, influenciando na latência do instrumento.

Além disso, é importante a constante avaliação e aprimoramento do sistema desenvolvido. Dito isto, mais testes com músicos profissionais e a incorporação de novas funcionalidades podem levantar novos problemas e oferecer valores que não foram cogitados ao longo do projeto.

Ademais, é almejada a integração com sistemas de prototipação e experimentação para construção de DMIs. Sistemas como o Sketchument (Calegario, 2013) e o Illusio (Barbosa et. al., 2013) fornecem a capacidade de configuração da interface do instrumento que é de grande ajuda na construção de novos instrumentos musicais.

Como já discutido anteriormente, a escolha de simular um instrumento de percussão se deu pelo fato de que estes demandam bastante de uma precisão temporal e espacial. O intuito deste projeto foi desenvolver um sistema expressivo que possa ser utilizado para criar DMIs alternativos.

8.1 Lições aprendidas com o Leap Motion©

Nesta seção, busca-se apresentar as conclusões específicas relacionadas com o Leap Motion Controller©. Visamos relatar, aqui, as vantagens e desvantagens de uso do dispositivo.

8.1.1 Vantagens

O sensor mostrou a capacidade de captura de diversas características que a maioria dos sensores disponíveis no mercado não conseguem rastrear. Conseguir capturar características como a posição, o comprimento, a direção do vetor, a velocidade, o ângulo (em relação aos eixos do sensor), grau de extensão para cada articulação e dedos das mãos; grau de abertura, ângulos de rotação, dimensões e velocidade da palma de cada mão; entre outras características, mostram o potencial de uso desse sensor para aplicações interativas com interfaces gestuais.

O sensor foi essencial na captura dos gestos e nuances da conga. O reconhecimento do ângulo da palma e a abertura das mãos permitiu que o ângulo de ataque do músico e o abafamento fossem capturados, mostrando a possibilidade de implementar timbres distintos para cada situação e, conseqüentemente, possibilitando a criação de instrumentos mais expressivos.

Além disso, a arquitetura do sistema é bastante favorável à construção de interfaces interativas, particularmente DMIs. A API disponível é altamente baseada em orientação a objetos, fornecendo uma linguagem simples de entender e, conseqüentemente, de implementar. O SDK fornece suporte para diversas linguagens de programação (JavaScript, Unity, C#, C++, Java, Python, Objective-C) e um plugin para a *engine* Unreal. Outro fator positivo é a arquitetura do sistema, que é naturalmente desenvolvida em múltiplos processos (*threads*) simultâneos, o que facilita o desenvolvimento da aplicação e a integração com outros sensores. Ademais, possui o suporte para aplicações em realidade virtual (VR), com modos de rastreamento específicos para isso. Assim, pode-se utilizar o sensor em diversas orientações de rastreamento – existem guidelines disponíveis para o uso com *Head-Mounted Displays* (HMDs) (e.g. o Oculus Rift®), de “cima para baixo”, entre outras.

8.1.2 *Desvantagens*

O sensor mostrou instabilidade (na captura da segunda mão) na configuração apresentada, mesmo em aplicativos disponíveis na *App Store*. Aplicações que envolvem uma superfície de contato devem ser desenhadas com maior precaução. Mesmo utilizando uma superfície não refletora (utilizando um emborrachado de cor preta fosca), o sistema teve problemas com rastreamento da segunda mão. Isso mostra que mudanças no algoritmo de rastreamento do dispositivo deve ser aprimorado para que possa ser utilizado nesse tipo de aplicação.

A arquitetura do sistema não permite que mais de um sensor seja utilizado no mesmo computador, o que limita o campo de visão da sua aplicação. Ademais, encarece projetos que utilizam mais de um sensor, já que será necessário outro computador.

Como consequência de ter sido desenvolvido para aplicações em que o usuário utiliza apenas as mãos e está na frente do computador, o campo de visão do sensor ainda é limitado, principalmente quando comparado com sensores como o Kinect®. Isso restringe as possibilidades de aplicação. Além disso, o dispositivo deve estar “fixo” para funcionar corretamente em outras orientações. O sensor apresentou instabilidade ao ser testado de “cima para baixo” e de “lado” sem estar fixo ao pedestal, mesmo sem a superfície de contato. Ademais, o uso de aplicativos que usam gestos livre no ar causa fadiga nos braços e mãos.

Ademais, o sensor apresentou uma latência considerável (71 milissegundos) ao ser avaliado na versão alfa 1 (utilizando C/C++ e disparando *samples*). Com isso, foi necessária a integração de um sensor piezoelétrico para solucionar esse problema.

8.1.3 *Considerações Finais*

Por ser um dispositivo que captura gestos em três dimensões, o *feedback* visual/háptico deve ser implementado para fornecer pistas que facilitem a interação. Como já foi discutido, o uso de *feedback* háptico em interfaces gestuais facilita a interação do usuário. O uso de superfícies transparentes contribuiu bastante para fornecer um feedback nas primeiras versões do DMI desenvolvido. Desta forma, pôde-se usar o Leap Motion® na configuração “natural”. Mesmo assim, o uso de superfícies entre o dispositivo e as mãos pode causar refração nos raios IR disparados pelo dispositivo, o que pode causar problemas no rastreamento das mãos.

Pode-se utilizar este sensor para desenvolver DMIs que não necessitem de uma latência tão baixa quanto instrumentos de percussão que disparem notas ou utilizá-lo na sua configuração de rastreamento natural, e.g. desenvolver um ganzá ou um instrumento que controle efeitos no ar. Deste modo, os problemas levantados podem ser contornados e o *designer* pode fazer uso da precisão gestual do sensor.

Referências

- Ableton**. Disponível em: <<https://www.ableton.com/>>. Acesso em: 17 jul. 2015.
- ABRAS, C.; MALONEY-KRICHMAR, D.; PREECE, J. User-centered design. **Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications**, v. 37, n. 4, p. 445–456, 2004.
- Arduino**. Disponível em: <<https://www.arduino.cc/>>. Acesso em: 17 jul. 2015.
- AUERBACH, J. et al. **Real-time music synthesis in Java using the Metronome garbage collector**. Proceedings of the International Computer Music Conference, Copenhagen, Denmark. **Anais...2007**. Disponível em: <<http://goo.gl/ScXYlu>>. Acesso em: 31 jul. 2015.
- BÅÅTH, R. Construction of a Low Latency Tapping Board. **LUCS minor**, v. 17, p. 362–366, 2011.
- BARBOSA, A. L. W. et al. Theremix: Uma aplicação musical com Interação Natural. 2014. Disponível em: <<http://goo.gl/2BpiHx>>. Acesso em 04 ago. 2015.
- BARBOSA, J. et al. **Towards an evaluation methodology for digital music instruments considering performer’s view: a case study**. Proceedings of 13th Brazilian Symposium on Computer Music. **Anais...2011**. Disponível em: <<http://batebit.cc/wp-content/uploads/2014/07/verse3o-final-sbcm.pdf>>. Acesso em: 17 jul. 2015.
- BLADES, J. **Percussion instruments and their history**. London, United Kingdom: Bold Strummer Limited, 1992.
- BUCHMANN, V. et al. **FingARtips: gesture based direct manipulation in Augmented Reality**. Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia. **Anais...ACM**, 2004. Disponível em: <<http://goo.gl/oJ6UO2>>. Acesso em: 17 jul. 2015
- BUSH, V. As we may think. **The atlantic monthly: As we may think**, v. 176, n. 1, p. 101–108, 1945.
- BUXTON, B. Multi-touch systems that I have known and loved. **Microsoft Research**, v. 56, p. 1–11, 2007.
- BUXTON, W.; HILL, R.; ROWLEY, P. Issues and techniques in touch-sensitive tablet input. **ACM SIGGRAPH Computer Graphics**, v. 19, n. 3, p. 215–224, 1985.
- CALEGARIO, F. C. DE A. Sketchument: ambiente de experimentação para criação de instrumentos musicais digitais. Dissertação de mestrado, Universidade Federal de Pernambuco, 2013.

CHAGAS, P. Le MIDI et la musique électronique: Quelques remarques esthétiques et techniques. **Revue informatique et statistique dans les sciences humaines**, v. 28, p. 15, 1992.

CHIAVENATO, I. **Introdução à teoria geral da administração**. [s.l.] Elsevier Brasil, 2003.

DAVIDSON, P. L.; HAN, J. Y. **Synthesis and control on large scale multi-touch sensing displays**. Proceedings of the 2006 conference on New interfaces for musical expression. **Anais...IRCAM—Centre Pompidou**, 2006. Disponível em: <<http://dl.acm.org/citation.cfm?id=1142269>>. Acesso em: 17 jul. 2015.

DE PRA, Y. et al. Infrared vs. Ultrasonic Finger Detection on a Virtual Piano Keyboard. 2014.

DIETZ, P.; LEIGH, D. **DiamondTouch: a multi-user touch technology**. Proceedings of the 14th annual ACM symposium on User interface software and technology. **Anais...ACM**, 2001. Disponível em: <<http://dl.acm.org/citation.cfm?id=502389>>. Acesso em: 17 jul. 2015

DOBRIAN, C.; KOPPELMAN, D. **The'E'in NIME: musical expression with new computer interfaces**. Proceedings of the 2006 conference on New interfaces for musical expression. **Anais...IRCAM—Centre Pompidou**, 2006. Disponível em: <<http://dl.acm.org/citation.cfm?id=1142283>>. Acesso em: 17 jul. 2015.

ENGELBART, D. C. **Augmenting Human Intellect: A Conceptual Framework**. Disponível em: <<http://www.doungengelbart.org/pubs/augment-3906.html>>. Acesso em: 17 jul. 2015.

ENGELBART, D. C.; ENGLISH, W. K. **A Research Center for Augmenting Human Intellect**. Disponível em: <<http://goo.gl/EMySah>>. Acesso em: 17 jul. 2015.

EVANS, B. **Authentic Conga Rhythms (Revised): A Complete Study: Contains Illustrations Showing the Current Method of Playing the Conga Drums and All the Latin Rhythms**. [s.l.] Alfred Music, 1966.

FELS, S.; GADD, A.; MULDER, A. Mapping transparency through metaphor: towards more expressive musical instruments. **Organised Sound**, v. 7, n. 02, p. 109–126, 2002.

FREITAS, D. et al. Development and evaluation of a Kinect based motor rehabilitation game. **Proceedings of SBGames**, p. 144–153, 2012.

FREITAS, E. M. C. DE. **O gesto musical nos métodos de percussão afro-brasileira**. Dissertação de mestrado. Universidade Federal de Minas Gerais, 2008.

GOLDSTINE, H. H.; GOLDSTINE, A. The electronic numerical integrator and computer (ENIAC). **Mathematical Tables and Other Aids to Computation**, p. 97–110, 1946.

GOMES, E.; TAVARES, L. Uma Solução Com Arduino Para Controlar E Monitorar Processos Industriais. **Trabalho de Conclusão de Curso Apresentado ao Instituto Nacional de Telecomunicações**, 2013.

GROSSHAUSER, T.; HERMANN, T. Augmented haptics—an interactive feedback system for musicians. In: **Haptic and Audio Interaction Design**. [s.l.] Springer, 2009. p. 100–108.

HADJAKOS, A. **Analysis of Piano Playing Movements Spanning Multiple Touches**. Proceedings of the International Conference on New Interfaces for Musical Expression. **Anais...2010**. Disponível em: <<http://goo.gl/y58AUS>>. Acesso em: 17 jul. 2015.

HALL, D. L.; LLINAS, J. An introduction to multisensor data fusion. **Proceedings of the IEEE**, v. 85, n. 1, p. 6–23, 1997.

HAN, J. Y. **Low-cost multi-touch sensing through frustrated total internal reflection**. Proceedings of the 18th annual ACM symposium on User interface software and technology. **Anais...ACM**, 2005. Disponível em: <<http://dl.acm.org/citation.cfm?id=1095054>>. Acesso em: 17 jul. 2015.

HARTREE, D. R. The ENIAC, an electronic computing machine. **Nature**, v. 158, n. 4015, p. 500–506, 1946.

HEISE, S.; LOVISCACH, J. **A versatile expressive percussion instrument with game technology**. Multimedia and Expo, 2008 IEEE International Conference on. **Anais...IEEE**, 2008. Disponível em: <<http://goo.gl/rSqrSL>>. Acesso em: 17 jul. 2015.

HSU, M.-H. et al. **Spider King: Virtual musical instruments based on microsoft Kinect**. Awareness Science and Technology and Ubi-Media Computing (iCAST-UMEDIA), 2013 International Joint Conference on. **Anais...IEEE**, 2013. Disponível em: <<http://goo.gl/X3WQiT>>. Acesso em: 17 jul. 2015.

JANSEN, Y.; KARRER, T.; BORCHERS, J. **MudPad: tactile feedback and haptic texture overlay for touch surfaces**. ACM International Conference on Interactive Tabletops and Surfaces. **Anais...ACM**, 2010. Disponível em: <<http://dl.acm.org/citation.cfm?id=1936655>>. Acesso em: 17 jul. 2015.

JOHNSON, J. et al. The xerox star: A retrospective. **Computer**, v. 22, n. 9, p. 11–26, 1989.

JORDA, S. Digital Lutherie: Crafting musical computers for new musics performance and improvisation. **PhD diss., Universitat Pompeu Fabra, Departament de Tecnologia**, 2005.

KANKE, H. et al. Airstic drum: a drumstick for integration of real and virtual drums. In: **Advances in Computer Entertainment**. Springer, 2012. p. 57–69. Disponível em: <http://ubi.eeddept.kobe-u.ac.jp/papers/kanke_ace2012.pdf>. Acesso em: 04 ago. 2015.

KERRIS, N.; DOWLING, S. **Apple - Press Info - Apple Reinvents the Phone with iPhone**. Disponível em: <<https://goo.gl/8ZyPGd>>. Acesso em: 17 jul. 2015.

KIEFER, C.; COLLINS, N.; FITZPATRICK, G. **HCI methodology for evaluating musical controllers: A case study**. Proceedings of the 2008 conference on New interfaces for musical expression. **Anais...** University of Genova, Italy 2008. Disponível em: <<http://goo.gl/j7AQAO>>. Acesso em: 17 jul. 2015.

KUZNETSOV, S.; PAULOS, E. **Rise of the expert amateur: DIY projects, communities, and cultures**. Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries. **Anais...ACM**, 2010. Disponível em: <<http://dl.acm.org/citation.cfm?id=1868950>>. Acesso em: 17 jul. 2015.

LAMPSON, B. W.; TAFT, E. Alto User's Handbook. **Xerox Palo Alto Research Center**, v. 3333, p. 31–62, 1976.

Leap Motion. Disponível em: <<https://www.leapmotion.com/>>. Acesso em: 17 jul. 2015.

LEE, S. K.; BUXTON, W.; SMITH, K. C. **A multi-touch three dimensional touch-sensitive tablet**. ACM SIGCHI Bulletin. **Anais...ACM**, 1985. Disponível em: <<http://dl.acm.org/citation.cfm?id=317461>>. Acesso em: 17 jul. 2015

LINDSTRÖM, E. et al. “Expressivity comes from within your soul”: A questionnaire study of music students' perspectives on expressivity. **Research Studies in Music Education**, v. 20, n. 1, p. 23–47, 2003.

LUO, R. C.; KAY, M. G. **A tutorial on multisensor integration and fusion**. Industrial Electronics Society, 1990. IECON'90., 16th Annual Conference of IEEE. **Anais...IEEE**, 1990. Disponível em: <<http://goo.gl/JQuF3c>>. Acesso em: 3 ago. 2015

MANNING, P. **Electronic and computer music**. New York, United States of America: Oxford University Press, 2013.

MARSHALL, M. T. **Physical interface design for digital musical instruments**. Tese de Doutorado, McGill University, Canadá, Citeseer, 2008.

MAURYA, A. **Running lean: iterate from plan A to a plan that works**. Califórnia, United States of America: O'Reilly Media, Inc., 2012.

MCGLYNN, P. Towards more effective mapping strategies for digital musical instruments. 2011.

MCMILLEN, K. Zipi: Origins and motivations. **Computer Music Journal**, p. 47–51, 1994.

MEDEIROS, C. B.; WANDERLEY, M. M. A comprehensive review of sensors and instrumentation methods in devices for musical expression. **Sensors**, v. 14, n. 8, p. 13556–13591, a 2014a.

MEDEIROS, C. B.; WANDERLEY, M. M. Multiple-model linear Kalman filter framework for unpredictable signals. **Sensors Journal, IEEE**, v. 14, n. 4, p. 979–991, b 2014b.

MEDEIROS, R. et al. Challenges in Designing New Interfaces for Musical Expression. In: **Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience**. Springer, 2014. p. 643–652.

MILLER, J.; HAMMOND, T. **Wiiolin: a virtual instrument using the Wii remote**. Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME 2010), University of Technology, Sydney, Australia. **Anais...2010**. Disponível em: <<http://goo.gl/trh4IT>>. Acesso em: 10 jul. 2015.

MIRANDA, E. R.; WANDERLEY, M. M. **New digital musical instruments: control and interaction beyond the keyboard**. [s.l.] AR Editions, Inc., 2006. v. 21. Disponível em: <<https://goo.gl/4IWq0K>>. Acesso em: 04 ago. 2015.

MITICHE, A.; AGGARWAL, J. K. Multiple sensor integration/fusion through image processing: a review. **Optical Engineering**, v. 25, n. 3, p. 253380–253380, 1986.

MITRA, S.; ACHARYA, T. Gesture recognition: A survey. **Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on**, v. 37, n. 3, p. 311–324, 2007.

MONTAG, M. et al. **A Low-Cost, Low-Latency Multi-Touch Table with Haptic Feedback for Musical Applications**. Proceedings of the 2011 conference on New interfaces for musical expression. **Anais...** University of Oslo, Norway, 2011. Disponível em: <<http://goo.gl/W5ldNd>>. Acesso em: 17 jul. 2015.

MOORE, F. R. The Dysfunctions of MIDI. **Computer music journal**, p. 19–28, 1988. Disponível em: <<http://goo.gl/8SUguP>>. Acesso em: 04 ago. 2015.

NECIOSUP, H. P.; ROSA, J. **The Book of Contemporary Conga Techniques**. Outskirts Press, United States of America, Contemporary Percussion, 2007.

NICHOLS, C. **The vBow: development of a virtual violin bow haptic human-computer interface**. Proceedings of the 2002 conference on New interfaces for musical expression. **Anais...** University of Singapore, 2002. Disponível em: <<http://dl.acm.org/citation.cfm?id=1085179>>. Acesso em: 17 jul. 2015

NORMAN, D. A.; DRAPER, S. W. **User Centered System Design: New Perspectives on Human-Computer Interaction**. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers, 1986.

ODOWICHUK, G. et al. **Sensor fusion: Towards a fully expressive 3d music control interface**. Communications, Computers and Signal Processing (PacRim), 2011 IEEE Pacific Rim Conference on. **Anais...IEEE**, 2011. Disponível em: <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6033003>. Acesso em: 17 jul. 2015

OKADA, K. et al. **Virtual Drum: Ubiquitous and playful drum playing.** Consumer Electronics (GCCE), 2014 IEEE 3rd Global Conference on. **Anais...IEEE**, 2014. Disponível em: <<http://goo.gl/Ks5DdB>>. Acesso em: 17 jul. 2015

O'MODHRAIN, S. A framework for the evaluation of digital musical instruments. **Computer Music Journal**, v. 35, n. 1, p. 28–42, 2011.

O'MODHRAIN, S.; CHAFE, C. **Incorporating haptic feedback into interfaces for music applications.** Proceedings of the International Symposium on Robotics with Applications, World Automation Conference. **Anais...2000**. Disponível em: <<http://goo.gl/aafE9n>>. Acesso em: 04 ago. 2015.

PAINE, G. New Musical Instrument Design Considerations. **MultiMedia, IEEE**, v. 20, n. 4, p. 76–84, 2013.

PENG, L.; GERHARD, D. **A wii-based gestural interface for computer-based conducting systems.** Proceedings of the 2009 conference on new interfaces for musical expression. **Anais...2009**. Disponível em: <<http://goo.gl/Dv4LF3>>. Acesso em: 17 jul. 2015.

ROADS, C. **The computer music tutorial.** United States of America: MIT press, 1996.

ROSA-PUJAZÓN, A. et al. **Drum-hitting gesture recognition and prediction system using kinect.** I simposio espanol de entrenamiento digital seed. **Anais...2013**. Disponível em: <<http://goo.gl/38VcmP>>. Acesso em: 17 jul. 2015.

ROVAN, J. B. et al. **Instrumental gestural mapping strategies as expressivity determinants in computer music performance.** Kansei, The Technology of Emotion. Proceedings of the AIMI International Workshop. **Anais...Citeseer**, 1997. Disponível em: <<http://goo.gl/G6tO9k>>. Acesso em: 17 jul. 2015.

SAFFER, D. **Designing gestural interfaces: Touchscreens and interactive devices.** [s.l.] O'Reilly Media, Inc., 2008. Disponível em: <<https://goo.gl/F172E7>>. Acesso em: 06 ago. 2015.

SATO, Y.; SAITO, M.; KOIKE, H. **Real-time input of 3D pose and gestures of a user's hand and its applications for HCI.** Virtual Reality, 2001. Proceedings. IEEE. **Anais...IEEE**, 2001. Disponível em: <<http://goo.gl/V5wO98>>. Acesso em: 17 jul. 2015.

SCHLOSS, W. A. **Recent advances in the coupling of the language Max with the Mathews/Boie Radio Drum.** Proceedings of the International Computer Music Conference. **Anais...1990**. Disponível em: <<http://goo.gl/rPykQs>>. Acesso em: 17 jul. 2015.

SCHLOSS, W. A. Using contemporary technology in live performance: The dilemma of the performer. **Journal of New Music Research**, v. 32, n. 3, p. 239–242, 2003.

SHEPARD, B. K. **Refining Sound: A Practical Guide to Synthesis and Synthesizers**. [s.l.] Oxford University Press, 2013.

SILVA, E. S. et al. **A preliminary evaluation of the leap motion sensor as controller of new digital musical instruments**. Anais do SBCM 2013. **Anais...** In: 14º SIMPÓSIO BRASILEIRO DE COMPUTAÇÃO MUSICAL. 2013. Disponível em: <http://compmus.ime.usp.br/sbcm/2013/pt/docs/art_tec_1.pdf>. Acesso em: 3 ago. 2015

SILVA, J. V. S. Avaliando Interfaces Gestuais Para a Prática de Instrumentos virtuais de Percussão. **Dissertação de Mestrado**, p. 89, 2012.

SINGER, E.; LARKE, K.; BIANCIARDI, D. **LEMUR GuitarBot: MIDI robotic string instrument**. Proceedings of the 2003 conference on New interfaces for musical expression. **Anais...**National University of Singapore, 2003. Disponível em: <<http://dl.acm.org/citation.cfm?id=1085759>>. Acesso em: 10 jul. 2015

SMITH, C.; EVANS, B. **Apple - Press Info - Apple Launches iPad**. Disponível em: <<https://www.apple.com/pr/library/2010/01/27Apple-Launches-iPad.html>>. Acesso em: 17 jul. 2015.

STEINER, H.-C. **Stickmusic: Using haptic feedback with a phase vocoder**. Proceedings of the 2004 conference on New interfaces for musical expression. **Anais...**National University of Singapore, 2004. Disponível em: <<http://dl.acm.org/citation.cfm?id=1085932>>. Acesso em: 17 jul. 2015

STOWELL, D. et al. Evaluation of live human-computer music-making: Quantitative and qualitative approaches. **International Journal of Human-Computer Studies**, v. 67, n. 11, p. 960–975, 2009.

TODOROFF, T.; LEROY, J.; PICARD-LIMPENS, C. Orchestra: Wireless sensor system for augmented performances & fusion with kinect. **QPSR of the numediart research program**, v. 4, n. 2, p. 29–36, 2011.

WACHS, J. P. et al. Vision-based hand-gesture applications. **Communications of the ACM**, v. 54, n. 2, p. 60–71, 2011.

WADLOW, T. A. The xerox alto computer. **Byte Magazine**, v. 6, n. 9, p. 58–68, 1981.

WANDERLEY, M. M. **Gestural control of music**. International Workshop Human Supervision and Control in Engineering and Music. **Anais...**2001. Disponível em: <<http://goo.gl/p5gqtS>>. Acesso em: 17 jul. 2015.

WANDERLEY, M. M.; ORIO, N. Evaluation of input devices for musical expression: Borrowing tools from hci. **Computer Music Journal**, v. 26, n. 3, p. 62–76, 2002.

WARDEN, N. A History of the Conga Drum. **Percussive Notes**, v. 43, p. 8–15, 2005.

WEINBERG, G.; AIMI, R.; JENNINGS, K. **The Beatbug network: a rhythmic system for interdependent group collaboration.** Proceedings of the 2002 conference on New interfaces for musical expression. **Anais...**National University of Singapore, 2002. Disponível em: <<http://dl.acm.org/citation.cfm?id=1085197>>. Acesso em: 17 jul. 2015

WESSEL, D.; WRIGHT, M. Problems and prospects for intimate musical control of computers. **Computer Music Journal**, v. 26, n. 3, p. 11–22, 2002.

WESTERMAN, W.; ELIAS, J. G.; HEDGE, A. **Multi-touch: A new tactile 2-d gesture interface for human-computer interaction.** Proceedings of the Human Factors and Ergonomics Society Annual Meeting. **Anais...**SAGE Publications, 2001. Disponível em: <<http://pro.sagepub.com/content/45/6/632.short>>. Acesso em: 17 jul. 2015

WOBBROCK, J. O.; MORRIS, M. R.; WILSON, A. D. **User-defined gestures for surface computing.** Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. **Anais...**ACM, 2009. Disponível em: <<http://dl.acm.org/citation.cfm?id=1518866>>. Acesso em: 17 jul. 2015

WONG, E. L.; YUEN, W. Y.; CHOY, C. S. **Designing Wii controller: a powerful musical instrument in an interactive music performance system.** Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia. **Anais...**ACM, 2008. Disponível em: <<http://dl.acm.org/citation.cfm?id=1497205>>. Acesso em: 10 jul. 2015

WRIGHT, M. A comparison of MIDI and ZIPI. **Computer Music Journal**, p. 86–91, 1994.

WRIGHT, M. **The Open Sound Control 1.0 Specification.** Disponível em: <http://opensoundcontrol.org/spec-1_0>. Acesso em: 31 jul. 2015.

WRIGHT, M. Open Sound Control: an enabling technology for musical networking. **Organised Sound**, v. 10, n. 03, p. 193–200, 2005.

WRIGHT, M.; CASSIDY, R. J.; ZBYSZYNSKI, M. F. **Audio and gesture latency measurements on linux and osx.** Proceedings of the ICMC. **Anais...**2004. Disponível em: <<http://goo.gl/XHB4rT>>. Acesso em: 31 jul. 2015.

Apêndice A – Modelo do questionário online

*** 1. Qual é a sua idade?**

- | | |
|--|--|
| <input type="radio"/> Menos que 10 anos | <input type="radio"/> Entre 36 e 40 anos |
| <input type="radio"/> Entre 10 e 15 anos | <input type="radio"/> Entre 41 e 45 anos |
| <input type="radio"/> Entre 16 e 20 anos | <input type="radio"/> Entre 46 e 50 anos |
| <input type="radio"/> Entre 21 e 25 anos | <input type="radio"/> Entre 51 e 55 anos |
| <input type="radio"/> Entre 26 e 30 anos | <input type="radio"/> Entre 56 e 60 anos |
| <input type="radio"/> Entre 31 e 35 anos | <input type="radio"/> Mais de 60 anos |

*** 2. Há quanto tempo você trabalha no cenário musical?**

- | | |
|--|---|
| <input type="radio"/> Menos de 1 ano | <input type="radio"/> Entre 6 e 7 anos |
| <input type="radio"/> Entre 1 e 2 anos | <input type="radio"/> Entre 7 e 8 anos |
| <input type="radio"/> Entre 2 e 3 anos | <input type="radio"/> Entre 8 e 9 anos |
| <input type="radio"/> Entre 3 e 4 anos | <input type="radio"/> Entre 9 e 10 anos |
| <input type="radio"/> Entre 4 e 5 anos | <input type="radio"/> Mais de 10 anos |
| <input type="radio"/> Entre 5 e 6 anos | |

*** 3. Você utiliza algum instrumento de percussão digital?**

- Sim Não Mais de um

*** 4. Se sim, o que lhe levou a utilizar esse(s) Instrumento(s)? (Se "Não", Responder N/A)**

*** 5. Se não, por que não? (Se "Sim" ou "Mais de um", Responder N/A)**

*** 6. Em que contexto você utiliza esse(s) Instrumento(s)? (Se "Não", Selecionar "Outro" e responder N/A)**

Composição

Ensaio

Gravação

Show

Outro (especifique)

7. O que você gosta nesse(s) instrumento(s)?

8. O que você NÃO gosta nesse(s) instrumento(s)?