

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**

**Escola de Música**

**Programa de Pós-graduação em Música**

Augusto Cesar Pereira Armondes

**ANÁLISE MULTIMODAL DA TÉCNICA DE *STRUMMING* NO VIOLÃO**

Belo Horizonte

2020

Augusto Cesar Pereira Armondos

ANÁLISE MULTIMODAL DA TÉCNICA DE *STRUMMING* NO VIOLÃO

Versão Final

Dissertação apresentada ao Programa de Pós-Graduação em Música da Universidade Federal de Minas Gerais como requisito parcial para a obtenção do título de Mestre em Música.

Orientador: Prof. Sérgio Freire

Linha de Pesquisa: Sonologia

Belo Horizonte  
2020

A733a

Armondes, Augusto Cesar Pereira.

Análise multimodal da técnica de strumming no violão [manuscrito] / Augusto Cesar Pereira  
Armondes. - 2020.  
110 f., enc.

Orientador: Sérgio Freire.

Linha de pesquisa: Sonologia.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Música.

Inclui bibliografia.

1. Música - Teses. 2. Performance musical. 3. Ritmo. 4. Violão - Instrução e estudo. I. Garcia, Sérgio Freire. II. Universidade Federal de Minas Gerais. Escola de Música. III. Título.

CDD: 787.607



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE MÚSICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM MÚSICA

### FOLHA DE APROVAÇÃO

Dissertação defendida pelo aluno **Augusto César Pereira Armondes**, em 03 de novembro de 2020, e aprovada pela Banca Examinadora constituída pelos Professores:

---

Prof. Dr. Sérgio Freire Garcia  
Universidade Federal de Minas Gerais  
(orientador)

---

Prof. Dr. Guilherme Augusto Soares de Castro  
Universidade do Estado de Minas Gerais

---

Prof. Dr. Maurício Alves Loureiro  
Universidade Federal de Minas Gerais

---

Prof. Dr. Davi Alves Mota  
CAPES-PRINT UFMG

Programa de Pós-Doutorado com Experiência no Exterior (PDEE)



Documento assinado eletronicamente por **Sergio Freire Garcia, Professor do Magistério Superior**, em 03/11/2020, às 14:31, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Guilherme Augusto Soares de Castro, Usuário Externo**, em 04/11/2020, às 09:24, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).

Documento assinado eletronicamente por **Davi Alves Mota, Usuário Externo**, em 04/11/2020, às 12:23, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de](#)



[outubro de 2015.](#)



Documento assinado eletronicamente por **Maurício Alves Loureiro, Professor do Magistério Superior**, em 04/11/2020, às 13:55, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufmg.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0345867** e o código CRC **637756E4**.

## **Agradecimentos**

O desenvolvimento de um trabalho como esse não seria possível sem o apoio de pessoas que acreditam na importância da educação e pesquisa como importante fator do crescimento humano. Com as dificuldades e incertezas que vieram com a pandemia da doença Covid-19 durante o processo, esse incentivo se mostra ainda mais importante.

Agradeço primeiramente aos meus pais, Orlando e Nádia, pelo apoio incondicional e por me mostrarem que o caminho certo é sempre aquele em que estamos seguindo.

À minha esposa, Camila, pelo amor e cuidados constantes, e por ser os pés ternos sobre os quais caminho rumo ao futuro.

Aos meus irmãos Vítor, Otávio e Larissa, por compartilhar as descobertas do mundo e pela troca de experiências.

À Dora, Nélio e Rodrigo, pelo acolhimento.

Ao Professor Sérgio Freire, pela orientação, conhecimentos e por confirmar a imensa importância do papel do docente na vida acadêmica.

Aos professores da UFMG, amigos do curso de Música e Popular e do LaPIS, com os quais aprendi muito além das paredes das salas.

Aos amigos de Ibiá, BH, Ouro Preto e República Aquarius, pelas conversas, incentivos e preocupações.

À CAPES, pelo investimento e por fomentar a pesquisa no Brasil, tão importante para o avanço do país.

Ao Ministério da Cultura e Secretaria Estadual de Cultura de Minas Gerais que, através da Lei Aldir Blanc, foram importante apoio para artistas em tempos difíceis.

Finalmente, à UFMG e à Escola de Música, pelo ensino gratuito e de qualidade.

A música é o tipo de arte mais perfeita: nunca revela o seu último segredo... (Oscar Wilde)

## Resumo

Novas tecnologias de áudio, vídeo e captura de movimento tem se desenvolvido expressivamente nas últimas décadas, trazendo novos métodos e perspectivas para os estudos sobre a performance. Equipamentos cada vez mais compactos, portáteis e com ótima relação custo–benefício contribuem para esse cenário, alavancando as possibilidades de pesquisa de diferentes técnicas utilizadas na prática musical.

Este trabalho realiza uma análise multimodal da técnica violonística de *strumming* através de áudio multicanal, vídeo em alta taxa de frames por segundo (substituído em um dos estudos realizados pelo sistema ótico de captação de movimentos Qualisys) e sensor de medida inercial com seis graus de liberdade (*IMU*), a fim de compreender os complexos gestos da mão direita utilizados na expressão rítmica. Com dados extraídos da execução de ritmos populares, situação de uso mais comum desta técnica, buscamos identificar e comparar as estratégias de cada músico, colaborando assim para uma melhor compreensão do *strumming*.

No primeiro capítulo, apresentamos os conceitos e ferramentas utilizadas nesta pesquisa. Iniciamos com a técnica de *strumming* e sua caracterização a partir de princípios psicoacústicos e do resultado sonoro gerado, seguido de uma breve revisão sobre multimodalidade e seu uso nos estudos de performance. Em seguida, descrevemos os equipamentos utilizados para extração das informações (*GuiaRT*, *MetaMotionR*, *Qualisys* e *GoPro*), o procedimento aproximativo de integração de curvas de aceleração e um estudo comparativo com os dados do *MetaMotionR* e *Qualisys*. A seguir, o método de análise dos componentes principais (*PCA*), finalizando com a estratégia para a avaliação conjunta dos dados.

O Capítulo 2 descreve os experimentos dedicados a dois ritmos selecionados, realizados em momentos distintos da pesquisa. Inicialmente faremos a comparação de performances do Ritmo 1, simples e regular, em duas situações: sem a fusão de dados do *IMU* executados por três músicos e com a fusão de dados executado por um músico. Em seguida, analisaremos as execuções do Ritmo 2, mais complexo com acordes em arpejo e separação de planos grave/agudo, com o sensor no modo de fusão de dados,



comparando três performances de um mesmo músico. Apresentamos ainda uma análise com os resultados prospectivos de todas as gravações.

As ferramentas utilizadas trouxeram contribuições importantes e complementares para a análise da técnica de *strumming*, propiciando não apenas a caracterização de diferentes aspectos das performances, mas também a diferenciação entre diferentes músicos. Como trabalho futuro, planejamos aumentar a abrangência da amostragem de músicos e de ritmos, e também aprofundar nas correlações entre os dados fornecidos pelas diferentes modalidades.

**Palavras-chave:** *Strumming*. Multimodalidade. Técnica violonística. Ritmo musical. Performance musical.

## Abstract

Audio, video, and motion capture technologies have developed significantly over the past few decades, bringing new methods and perspectives to performance studies. Equipment increasingly compact, portable, and with an excellent cost-benefit ratio contribute to this scenario, boosting the possibilities of researching different techniques used in musical practice.

This work performs a multimodal analysis of the strumming guitar technique using multichannel audio, high frame video (replaced in one of the studies carried out with Qualisys optical motion capture system), and a 6 DOF inertial sensor (IMU) to understand the complex gestures of the right hand used in the rhythmic expression. Using data extracted from the execution of popular rhythms - the most common use of this technique - we seek to identify and compare the strategies of different musicians, thus contributing to a better understanding of the strumming.

In the first chapter, we present the concepts and tools used in this research. We begin with the strumming technique and its characterization according to psychoacoustic achievements and the generated sound result, followed by a brief review on multimodality and its use in performance studies. Then, we describe the equipment used to extract information (GuiaRT, MetaMotionR, Qualisys, and GoPro), with the approximate procedure for integration acceleration curves and a comparative study with data from MetaMotionR and Qualisys. Next, we discuss the principal component analysis method (PCA), ending with the strategies for the joint data evaluation.

Chapter 2 describes the experiments dedicated to two selected rhythms carried out at different times in the research. Initially, we compare performances of Rhythm 1, a simple and regular one, in two situations: without IMU data fusion performed by three musicians and with data fusion performed by one musician. Then, we analyze the performances of the more complex Rhythm 2, bearing arpeggio chords and separation of bass/treble planes, with the sensor in the mode data fusion, comparing three performances by the same musician. We also present a prospective analysis mixing the results of all recordings.

The selected tools provided significant and complementary contributions to the analysis of the strumming technique, both in the characterization of different aspects of performances and in the differentiation between musicians.

In future work, we plan to increase the number of musicians and the variety of rhythms and deepen the study of correlations between data delivered by the different modalities.

**Keywords:** Strumming. Multimodality. Guitar technique. Musical rhythm. Musical performance.

Lista de artigos produzidos durante o mestrado, com conteúdo parcialmente utilizado, traduzido ou adaptado para esta dissertação

1 - Ferramentas para uma abordagem multimodal da técnica de strumming no violão - 5º NAS NUVENS... CONGRESSO DE MÚSICA. 2019

2 - Uma primeira aproximação multimodal da técnica de strumming. XXX CONGRESSO DA ANPPOM - MANAUS/AM. 2020

3 - Evaluation of Inertial Sensor Data by a Comparison with Optical Motion Capture Data of Guitar Strumming Gestures. *Sensors*. 20, 5722, 2020.

4 - GuiaRT, a Setup for Real-Time Transcription and Transformation of Music Played on a Hexaphonic Nylon Guitar. Under evaluation. 2020.

## Lista de Ilustrações

Figura 1- Exemplo de setas indicando a direção dos gestos na técnica de strumming com a cifra dos acordes.	21
Figura 2 – Sinal das 6 cordas no strumming em um intervalo de tempo de 500 ms, que compreende um gesto descendente (das cordas graves para as agudas) e um ascendente em um ritmo. O gesto ascendente utiliza apenas as cordas 1 e 2.	25
Figura 3 – Violão Yamaha equipado com captação hexafônica e conector multivias para aquisição independente dos sinais das cordas.	26
Figura 4 – Output do GuiaRT de um trecho com destaque para os onsets. Na ordem horizontal temos: evento, onset (ms), corda, nota MIDI, casa, duração (ms), amplitude (dB), flag de ligado, centroide espectral, indicadores de articulação (harmônico, pizzicato, vibrato) e microafinação.	27
Figura 5 – Fluxograma do algoritmo do GuiaRT. Ao lado esquerdo da figura podemos observar a estratégia de detecção de onsets, indicado como início em português.	29
Figura 6 – Patch do GuiaRT em Max/MSP com o pré-processamento do sinal e parâmetros reguláveis para cada corda.	30
Figura 7 – IMU MetaMotionR, da empresa MbientLab. À esquerda, as setas indicam a orientação dos três eixos no sentido positivo. À direita, vê-se este sensor bem compacto posicionado na mão do performer.	31
Figura 8 – Interface do drive escrito para MacOS 10.14 Mojave informando o status de conexão do sensor MetaMotionR pelo protocolo Bluetooth Low Energy	33
Figura 9 – Histograma do intervalo de tempo entre mensagens subsequentes com todas quantidades físicas (aceleração linear, velocidade angular e ângulos de Euler).	34
Figura 10 – Cálculo da latência do sensor com a preparação para a queda livre. O sinal de áudio é captado por um microfone de contato.	37
Figura 11 – Violão e IMU com os marcadores para mapeamento em corpo rígido pelo sistema de captura de movimento Qualisys.	38
Figura 12 – Partitura dos três ritmos utilizados neste experimento.	39
Figura 13 – Diferentes posições do IMU utilizadas pelos músicos nas gravações dos ritmos.	42
Figura 14 – Curvas em três dimensões do deslocamento de m2r1t2. Podemos observar que os gestos do strumming ocorre essencialmente em um plano.	43
Figura 15 – Gráfico com os quatro gestos de um ritmo regular e o início das três fases do gesto. O tamanho da seta indica a quantidade de notas tocadas e a flecha indica sua direção (descendente ou ascendente).	46

Figura 16 – Aplicação do PCA nas velocidades angulares em x, y e z de um take selecionado, com os três componentes principais resultantes do método. Note-se que o primeiro componente principal se aproxima da velocidade angular em z, indicando grande influência desse eixo na realização da técnica.	47
Figura 17 – Possibilidades de integração dos dados gerados pelo GuiaRT, IMU e vídeo. Os valores são normalizados na mesma escala, para uma boa visualização. Os pulsos estão subdivididos em semicolcheias. Os traços verticais (cinza e preto) indicam o início dos gestos (sentido descendente e ascendente).	49
Figura 18 – Partitura do trecho executado. As setas indicam o sentido do gesto.	51
Figura 19 – Posição do IMU na mão de um performer. Todos os músicos utilizaram a mesma orientação do sensor.	52
Figura 20 – Teste de medida do tempo de resposta do IMU. O instante em que o som é captado pelo microfone é comparado com a mudança de estado do sensor.	53
Figura 21 – Exemplo de dados extraídos pelo GuiaRT – 17 notas.	54
Figura 22 – Exemplo de dados extraídos do MetaMotionR – 170 ms.	55
Figura 23 – Lista dos pontos de segmentação de 17 gestos (ms) em três fases: preparação, sendo o momento de início dos gestos (troca de sentido); produção de som, considerado o momento do primeiro onset do acorde; finalização, tempo do último onset.	55
Figura 24 – PCA das médias das velocidades angulares dos músicos A, B e C. O primeiro componente tem peso significativo e está bem relacionado à regularidade dos gestos. As diferenças entre os músicos indicam as distintas estratégias de execução do trecho.	58
Figura 25 – Espectro dos componentes principais da velocidade angular dos três músicos. Considerando uma unidade de gesto como uma descida e uma subida, com metrônomo a 60 bpm, é esperada uma componente significativa na frequência de 2 Hz.	59
Figura 26 – Curvas de aceleração nos eixos x, y, z e seu módulo presentes na execução dos quatro primeiros acordes do primeiro take de cada músico.	61
Figura 27 – Sensor posicionado na mão do performer para registro do ritmo 1 programado para a fusão de dados.	62
Figura 28 – Gráfico com os ataques representados por setas, pulso subdividido em semicolcheias e direção dos gestos nos 3 takes do Ritmo 1.	65
Figura 29 – Velocidades lineares nos eixos x, y e z dos três takes do Ritmo 1.	66
Figura 30 – Velocidades angulares eixos x, y e z dos três takes do Ritmo 1.	67
Figura 31 – Partitura do Ritmo 2.	69
Figura 32 – Dois primeiros tempos dos ritmos 1 e 2, no take 1. As curvas de velocidade angular (gyro) em y e z e a aceleração linear em y possuem envelopes temporais semelhantes entre os ritmos, com menor intensidade nos gestos sem notas do ritmo 2, indicando movimentação constante da mão como estratégia de manutenção rítmica.	72
Figura 33 – Primeiro compasso dos takes do ritmo 2 com os acordes, representados por flechas, e o pulso subdividido em semicolcheias.	73

Figura 34 – Segundo compasso dos takes do ritmo 2 com os acordes e o pulso subdividido em semicolcheias. _____	74
Figura 35 – Segmentação dos gestos do strumming nos três takes do ritmo 2. _____	75
Figura 36 – Comparação das curvas de velocidade angular (gyro) dos dois compassos do ritmo 2. Apesar de mudança de acorde, as curvas apresentam comportamento semelhante. _____	77
Figura 37 – Curvas de velocidade angular, eixo x do sensor, combinadas com acordes, pontos de segmentação dos gestos e pulso no primeiro compasso de cada take. _____	78
Figura 38 – Curvas de velocidade angular nos eixos x e y nos três takes, primeiro compasso. _____	79
Figura 39 – Velocidade angular em z nos três takes. _____	80
Figura 40 – Velocidade linear do eixo z nos três takes do ritmo 2, primeiro compasso. Os pontos de velocidade nula apresentam bom alinhamento com os pontos de segmentação dos gestos. _____	81
Figura 41 – Velocidade linear em z e y nos três takes do ritmo 2. As curvas em y são constantes por todo o trecho. _____	82
Figura 42 – Velocidade linear no eixo x. A baixa amplitude das curvas indica que a referência inercial da posição do sensor em torno ao eixo z ("norte local") durante a gravação fez com que a movimentação no plano horizontal se alinhasse mais com o eixo y. _____	83
Figura 43 – Velocidade angular e velocidade linear no eixo y, nos takes 1 e 2 do ritmo 2. _____	85
Figura 44 – Acordes e pulso (em semicolcheias) dos músicos Exp2A e Exp3. _____	89
Figura 45 – Aceleração linear músicos Exp2A e Exp3 no take 2. Os músicos utilizam os eixos de maneira diferente na execução do ritmo 2. _____	90
Figura 46 – Velocidade angular músicos Exp2A e Exp3. Os músicos utilizam os eixos de maneira diferente na execução do ritmo 2. _____	91
Figura 47 – Exemplo de representação por setas do ritmo 2, em comparação com o modelo utilizado atualmente. _____	94

## Lista de Tabelas

Tabela 1 – Coordenadas dos marcadores de corpo rígido do IMU para alinhamento com a referência Qualisys L-shaped. _____	38
Tabela 2 – Quantidade de notas dos três músicos em cada take com a média e o desvio padrão. _____	56
Tabela 3 – Média e desvio padrão da duração total dos gestos, e do espalhamento dos acordes em ms, extraídos do take 1 de cada músico. _____	57
Tabela 4 – Comparativo entre as performances em relação a quantidade de notas. _____	63
Tabela 5 – Média e desvio padrão do espalhamento temporal e da quantidade de notas nos gestos descendentes e ascendentes. _____	64
Tabela 6 – Comparativo entre as performances do ritmo 2 em relação a quantidade de notas. _____	70
Tabela 7 – Spread temporal de acordes normais e arpejados nos três takes do ritmo 2, com média e desvio padrão geral. _____	71
Tabela 8 – Covariância entre velocidades angulares e lineares entre os eixos x, y e z, takes 1 e 2 do ritmo 2. _____	84
Tabela 9 – Valores médios da contribuição do primeiro componente principal das velocidades angulares para a explicação da variância do movimento e valores do autovetor correspondente. _____	87
Tabela 10 - Valores médios da contribuição do primeiro componente principal das acelerações lineares rotacionadas para a variância do movimento, e valores do eigenvector correspondente. _____	88



# Sumário

<b>Introdução.....</b>	<b>16</b>
<b>1. Conceitos e Ferramentas.....</b>	<b>20</b>
1.1. Strumming.....	20
1.2. Multimodalidade .....	23
1.3. Ferramentas .....	24
1.3.1. Captação hexafônica (GuiarT).....	24
1.3.1.1. Detecção de onsets .....	27
1.3.2. Sensor de movimento (IMU MetaMotionR).....	30
1.3.3. Qualisys .....	35
1.3.4. Comparação de dados gerados pelo MetaMotionR e pelo Qualisys .....	35
1.3.4.1. Medida do tempo de resposta do IMU .....	36
1.3.4.2. Setup .....	37
1.3.4.3. Gravações.....	39
1.3.4.4. Resultados.....	42
1.3.4.5. Discussão .....	44
1.3.5. Vídeo .....	45
1.3.6. PCA.....	46
1.3.7. Análise conjunta dos dados .....	48
<b>2. Experimentos.....</b>	<b>50</b>
2.1. Ritmo 1.....	51
2.1.1. Sensor sem a fusão de dados do <i>IMU</i> .....	51
2.1.1.1. Setup e gravações .....	51
2.1.1.2. Medida do tempo de resposta do <i>IMU</i> .....	52
2.1.1.3. Dados coletados .....	53
2.1.1.4. Resultados.....	56
2.1.2. Sensor com a fusão de dados do <i>IMU</i> .....	61
2.1.2.1. Setup e gravações .....	62
2.1.2.2. Resultados.....	63
2.1.3. Discussão .....	68
2.2. Ritmo 2.....	69
2.2.1. Setup e gravações.....	69
2.2.2. Resultados.....	69
2.2.3. Discussão .....	86
2.3. Análise prospectiva dos dados .....	87
2.4. Considerações .....	92
<b>Conclusão .....</b>	<b>93</b>
<b>REFERÊNCIAS.....</b>	<b>95</b>
<b>APÊNDICE A – Integração dos dados de aceleração linear gerados pelo IMU.....</b>	<b>99</b>
<b>APÊNDICE B – Resultados da comparação de dados gerados pelo MetaMotionR e pelo Qualisys.....</b>	<b>104</b>

## Introdução

O violão pode ser considerado um dos instrumentos mais populares do mundo. A partir da metade do Séc. XIX ganhou sua padronização atual na Europa, e nas décadas seguintes cresceu exponencialmente no gosto popular (TURNBULL; SPARKS, 2001). As técnicas desenvolvidas na prática musical variam de estilo e gênero em diversos períodos e locais, sendo algumas mais específicas, como o *rasgado* na música flamenca, e outras de domínio comum, como escalas e arpejos.

Dentre as diversas técnicas empregadas, tanto da mão que pressiona as cordas quanto da que realiza os ataques, o *strumming* se destaca por ser a mais básica interação com o instrumento, sendo um bom ponto de partida para iniciantes (LARSEN; OVERHOLT; MOESLUND, 2013). No entanto, sua grande importância didática para o domínio do violão é complicada pelo fato de ser uma das habilidades mais desafiadoras para serem desenvolvidas (PERSINGER, 2019). Além de carregar um aspecto bastante idiossincrático da performance, o *strumming* é de difícil ensino por ser baseado no ritmo, que é melhor compreendido quando absorvido intuitivamente (KITTLEBERGER, 2018).

Os estudos sobre performance musical podem colaborar na investigação desse complexo gesto presente na técnica violonística. Em seu livro “Musical performance: A guide to understanding” (2002), John Rink nos traz uma perspectiva geral da performance através da história, análise de som, escrita, aspectos neurológicos e gestuais, importantes não só para o estudo do *strumming*, mas para a prática musical em si. Nesse grande campo de estudos, as análises quantitativas e qualitativas aumentaram sua abrangência, partindo de relações entre som e partitura (ADORNO; PADDISON, 1982) e incorporando novas tecnologias para capturas de elementos pertinentes a sua realização, à medida que elas vão evoluindo (GLOWINSKI et al., 2012; MARIN; PERRY, 1999; MORLEY, 2002; RAHMAN et al., 2011).

Uma importante característica dos estudos relacionados à performance musical, no que tange a interpretação desta técnica específica, é a relativização do papel da partitura (ou tablatura, também muito comum) na sua aplicação. O conceito de Cook (2013) de que a escrita não carrega o sentido da performance, e sim o músico, encontra grande ressonância aqui, já que os elementos gráficos não conseguem traduzir os

elementos intrínsecos do *strumming*. É possível indicar a direção dos gestos e as notas utilizadas nessa estrutura, mas não os aspectos micro-temporais do som ou estratégias gestuais, que se tornam decisões pessoais do intérprete.

Com o recente avanço tecnológico na captura de movimento a relação gesto/som, que define o *strumming*, pode ser mais explorada. Tanto em situações com instrumentos, onde o gesto gera som (HSU, 2006; RASAMIMANANA; KAISER; BEVILACQUA, 2009), quanto em resposta gestual ao estímulo sonoro (CARAMIAUX; BEVILACQUA; SCHNELL, 2010) encontramos ferramentas que exploram esse diálogo muito importante da performance.

A trajetória que nos motivou a propor essa abordagem tem início na graduação em Música Popular, no programa de iniciação científica junto ao projeto *GuiaRT*. Através das extrações de *onsets* de trechos tocados em um violão adaptado com captação hexafônica, pudemos observar o comportamento micro-temporal do *strumming*. Essas análises trouxeram à luz questões pouco discutidas sobre regularidade e diferenças interpretativas dos performers no âmbito sonoro, e sua ligação com os gestos que realizam a técnica.

Para isso utilizamos diversas fontes de dados, caracterizando a multimodalidade da análise. Através de áudio, vídeo com alta taxa de frames por segundo e sensor de captura de movimento, com suporte de *softwares* específicos para registro e cálculos matemáticos, desenvolvemos ferramentas que possibilitaram tanto a extração das informações das performances quanto estratégias de interpretação dos dados.

Com diversas opções no mercado, o primeiro passo foi definir os equipamentos para essa abordagem. Em relação ao áudio, estávamos bem resolvidos com o *GuiaRT*. A primeira tentativa de capturar os gestos do *strumming* foi através do *Leap Motion*, um sensor *usb* que, com câmeras de infravermelho, estimava o deslocamento de diversos pontos da mão reconhecida por *software* próprio. Com um *sample rate* muito variável, este equipamento não se mostrou eficaz para a tarefa pretendida. Utilizamos então uma nova geração dos chamados *IMUs* (*Inertial Measurement Units*), que se mostrou mais compatível com o objetivo por ter conexão *wireless*, ser compacto e com a possibilidade de se prender em diferentes partes do corpo de maneira fácil.

Para o vídeo as câmeras de ação se apresentaram como uma boa opção por serem compactas e com alta taxa de *fps*. Em uma parceria com a Universidade McGill, tivemos a oportunidade de utilizar também um sistema ótico de captura de movimento, que enriqueceu nosso estudo sobre os gestos.

Após alguns estudos piloto, tínhamos bem desenvolvidas as ferramentas para a coleta final de dados, que contaria com gravações de ritmos simples e complexos na aplicação da técnica. Com performance de diversos músicos, experientes e iniciantes no *strumming*, teríamos um banco de dados satisfatório para a análise.

Infelizmente a pandemia do novo Corona Vírus não permitiu que essas gravações fossem feitas. A Covid19 nos fez mudar totalmente a rotina, seguindo protocolos rígidos de higiene e distanciamento social, afim de controlar a infecção e garantir a saúde da população mundial. A Escola de Música da UFMG foi fechada, com uma pequena abertura para professores, isoladamente, irem aos laboratórios e seus escritórios. A sala do Laboratório para Performance com Sistemas Interativos (LaPIS), local onde se concentram os principais *hardwares* utilizados, estava interditada, comprometendo a coleta final dos dados para este estudo. Nesse período de incertezas, não sabíamos se haveria tempo hábil para alcançar os objetivos pretendidos. Decidimos então usar os dados dos experimentos pilotos e conseguimos realizar uma única gravação adicional com um sistema portátil. Apesar dos contratemplos, acreditamos que os resultados apresentados são relevantes e sustentam as discussões realizadas. As ferramentas apresentadas neste trabalho não dizem respeito somente à extração de dados, mas também às estratégias interpretativas das informações da técnica de *strumming*. Portanto, devem ser consideradas nesses dois sentidos.

A dissertação está organizada em duas grandes seções, cada uma com um capítulo. A primeira parte trata do referencial teórico e metodológico, e a segunda trata dos experimentos realizados. No primeiro capítulo, dividido em três seções, discutimos os conceitos de *strumming* e multimodalidade, descrevemos as ferramentas de extração e análise, apresentamos um estudo comparativo dos dados gerados pelo *IMU* e *Qualisys* e o modo de integração dos dados. No capítulo dois apresentamos os experimentos com os dados que foram possíveis de captar, dividido em quatro seções: a) Ritmo 1, um ritmo simples executado por três músicos com o sensor sem o modo de fusão de dados e com

execuções de um músico com o sensor no modo de fusão de dados; b) Ritmo 2, um músico executando um ritmo complexo, com acordes ornamentados e separação de plano grave e agudo com o sensor no modo de fusão; c) resultados prospectivos dos estudos realizados; d) considerações. Uma conclusão e dois apêndices completam o texto.

# 1. Conceitos e Ferramentas

Com o objetivo de situar este estudo em seu recorte sobre a performance, iremos discutir os conceitos sobre a técnica de *strumming* e a multimodalidade, apresentar as ferramentas utilizadas na coleta dos dados das fontes escolhidas, os métodos de avaliação e as estratégias para sua análise conjunta. Como a bibliografia específica sobre o tema desta dissertação – a abordagem quantitativa da técnica de *strumming* – é bastante reduzida, preferimos não fazer uma seção específica de revisão de literatura; como alternativa, a cada tópico apresentado, discutiremos a bibliografia específica relacionada com os objetivos do estudo.

## 1.1. Strumming

A técnica de *strumming* é o gesto da mão que ataca as cordas do instrumento, produzindo som. É uma técnica bastante antiga e utilizada em diversos instrumentos de corda, apresentada com diferentes nomenclaturas de acordo com a região e época. Na Europa, a partir do século XVI, era chamada de *battuto*, *golpeado* ou *rasgueado* (STRIZICH; TYLER, 2001) e podemos ver o método padronizado para o alaúde em um tratado do século XVIII (ROLFHAMRE, 2010). Em sua aplicação na harpa e lira, uma mão amortece as cordas indesejadas enquanto a outra as ataca, com o uso de palheta ou unha, criando um som nítido e articulado. No *glissando*, podendo ser executado com as duas mãos, cria uma sonoridade bem característica da harpa, bastante utilizada como efeito sonoro (CLARK, 2006). Podemos ver seu uso também em violoncelo (BARKOSKIE, 2011), como técnica expandida para piano (ISHII, 2005) e no *koto* japonês (PORNPRASIT, 2010), instrumentos tradicionalmente não relacionados com o *strumming*.

Apesar de historicamente os nomes para essa técnica se modificarem, o termo *strumming* é o mais recorrente nas nossas pesquisas. No Brasil popularmente é chamado de levada ou batida, mas pode gerar ambiguidade, já que estes mesmos termos possuem conotação diferente dependendo do contexto. Assim a utilização do termo em inglês nos parece a mais adequada.

Existem diferentes modos de realizar a técnica, variando de acordo com o estilo em que está empregada. Na música flamenca espanhola os cinco dedos da mão que

realiza o ritmo atacam as cordas, gerando inúmeros *onsets* sucessivos em cada uma, propiciando uma forte identidade sonora e estética. Na vanera e guarânia brasileiras, utilizam-se ataques abafados, numa abordagem mais percussiva para o violão. No pop/rock, a palheta é utilizada também, conferindo um timbre diferente daquele gerado com o uso dos dedos. O *strumming* pode ser mesclado com outras técnicas, como o *plaquê*, em uma forma híbrida de performance, e aparece em situações de solo, com ataques abafados nas cordas que precedem a corda/nota alvo.

Apesar de sua grande utilização, a literatura sobre esta técnica é bastante escassa. Além da produção própria, encontramos apenas três estudos recentes ligados ao tema (BELLO; MAYOL, 2019; MATSUSHITA; IWASE, 2013; VISI; SCHRAMM; MIRANDA, 2014). Quanto ao *strumming* praticado na música popular, temos apenas métodos que propõem vários exercícios (KRENZ, 2010), além de diversas vídeo-aulas, baseadas principalmente na transmissão oral e com grafia simplista, utilizando setas que indicam a direção dos gestos (Figura 1).

Figura 1- Exemplo de setas indicando a direção dos gestos na técnica de *strumming* com a cifra dos acordes.



Frente a essa grafia, decisões sobre quais cordas tocar (separação de planos graves e agudos), velocidade do gesto (alterando o *Inter-onset Interval* entre as cordas), amplitude do movimento (deslocamento), posição da mão e ângulo da palheta em relação

as cordas cabem ao performer. A falta de uma descrição clara de como a técnica deve ser feita gera grandes diferenças interpretativas, visto que versões bem distintas são aceitáveis esteticamente. Esse fato abre uma discussão sobre como representar graficamente os eventos envolvidos no *strumming*, na ordem de ms, que não estão previstos na notação tradicional.

Em termos micro temporais, podemos posicionar um acorde tocado com a técnica de *strumming* aproximadamente entre um acorde *plaqué* e um arpejo. Aplicando os conceitos tipológicos de Schaeffer (1966), é possível dizer que com acordes *plaqué* o ataque ocorre de uma só vez, enquanto em acordes *strumming* acontece iterativamente, assim como em arpejos. Apesar de grande dificuldade em definir os limites dessas categorias, pois dependem de andamento, registro, dinâmica, etc., os resultados obtidos em nossas pesquisas anteriores permitem propor essa abordagem (FREIRE; ARMONDES; VIANA, 2017).

Na psicoacústica, a resolução temporal é, segundo Rossing (2007), "a capacidade de detectar alterações nos estímulos ao longo do tempo, por exemplo, para detectar um breve intervalo entre dois estímulos". Os resultados de diversas experiências indicam um intervalo entre 1 e 18 ms para essa habilidade, o que apresenta uma boa correlação com os valores por nós encontrados de, em média, 4 ms entre cordas vizinhas (FREIRE et al., 2018). Outra característica psicoacústica importante, com origem na análise da cena auditiva (BREGMAN, 1990), é a coerência dos acordes do violão (*plaqué*, *strummed* ou arpejado): nós os percebemos como uma unidade, como um único objeto sonoro, em vez de perceber cada nota separadamente.

Podemos descrever o *strumming*, em relação ao resultado sonoro gerado, pela quantidade de notas (densidade), espalhamento temporal (intervalo de tempo entre o primeiro e último *onset*) e direção (do grave para o agudo e vice-versa), e no âmbito gestual, pela duração total do gesto e deslocamento.

Em contextos musicais o estudo empírico desta técnica pode contribuir para compreensão não apenas de sua variabilidade, mas também de seu uso na expressão rítmica, principalmente em situações com pulsação clara.



## 1.2. Multimodalidade

Multimodalidade é um fenômeno em que diferentes modos sensoriais/semióticos são combinados e integrados.

A fusão de modalidades leva ao conceito de multimodalidade. Multimodalidade é baseada no uso de modalidades sensoriais com as quais humanos recebem informações, como toque, visão, audição etc., e requer o uso de ao menos duas modalidades de resposta em relação à apresentação da informação, como a atividade verbal e a manual. (MOURA, 2008, p. 37, t.n.)

Embora seja um fenômeno intrínseco a praticamente toda comunicação humana, os estudos de multimodalidade começaram a se desenvolver a partir dos anos 1990 com os escritores Gunther Kress e Theo van Leeuwen, principalmente na área de comunicação. Na área musical, observa-se uma grande atividade de pesquisa e criação a partir da década inicial do século XXI, como mostram os trabalhos de Camurri e o software *EyesWeb* (2000), a discussão sobre gestos iniciada por Cadoz e Wanderley (2000) e levada adiante por vários outros pesquisadores, dentre os quais Jensenius (2010), e a criação do conceito de *embodied music cognition* (LEMAN, 2008).

As tecnologias utilizadas nessas atividades estão em desenvolvimento constante, com o surgimento frequente de novas técnicas e equipamentos. As iniciativas de Camurri, por exemplo, basearam-se inicialmente em sinais de vídeo, incorporando paulatinamente novos recursos de captação gestual. Equipamentos dedicados à captação de movimentos (*mocap*, em inglês) tem se popularizado, com a utilização tanto de sensores passivos como ativos. Mais recentemente, as IMUs (*inertial measurement units*) tem ganhado atenção de fabricantes e pesquisadores. Embora a tecnologia de digitalização de áudio já tenha atingido há um bom tempo seu estado-da-arte, diferentes formas de captação sonora continuam a ser pesquisadas, seja pela utilização de microfones específicos e combinações, seja pelo uso de captadores de contato.

As performances musicais, objeto de estudo deste trabalho, são naturalmente multimodais (bem como praticamente todo processo de comunicação e expressão humana), tendo em vista que além do resultado sonoro temos também informações visuais sobre o gestual corporal do instrumentista, sua expressão facial, etc. No caso de

performances puramente instrumentais, os gestos do performer ao realizar uma determinada técnica influem diretamente em seu resultado sonoro.

Nesta abordagem multimodal, além do som e seus atributos, temos o gesto que realiza o *strumming* e sua relação com o pulso musical, indicando a expressividade do performer.

### 1.3. Ferramentas

Para esta abordagem multimodal da técnica de *strumming* lidaremos com diversas ferramentas para extração e análise dos dados, que serão apresentadas nesta seção, utilizando principalmente dois *softwares*: *Max/MSP* e *MATLAB*.

O *Max/MSP*<sup>1</sup>, da empresa *Cycling '74*, é uma linguagem de programação visual voltada para música e multimídia com processamento em tempo real, largamente utilizada em sistemas interativos. Os algoritmos para extração de *onsets* e leitura/registro dos dados do sensor de movimento estão implementadas nesta plataforma.

O *MATLAB*<sup>2</sup>, da empresa MathWorks, é um *software* interativo de alta performance voltado para o cálculo numérico, processamento de sinais e construção de gráficos, no qual importaremos os dados para manipulação, análises estatísticas e visualização.

Iniciaremos com a captação hexafônica e a estratégia para detecção de *onsets*.

#### 1.3.1. Captação hexafônica (GuiART)

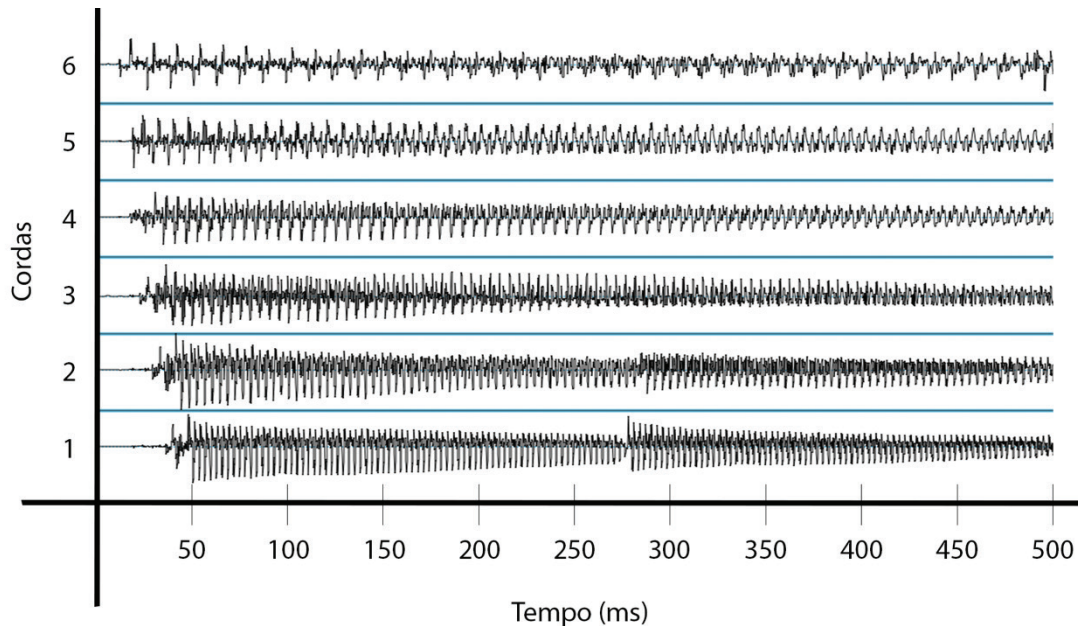
A captação hexafônica consiste em um captador individual para cada corda do violão, tanto nos formatos magnético ou piezo-elétrico, permitindo registro e análise do áudio de forma isolada. Desse modo ampliam-se consideravelmente as possibilidades para extração e caracterização dos diferentes tipos de sons tocados no instrumento e estudo de suas técnicas específicas, particularmente as que envolvem eventos iterativos com pequeno intervalo de tempo entre eles (ordem de grandeza em milissegundos) para a construção do objeto sonoro, como o *strumming* (Figura 2). No violão, a corda mais grave é chamada de corda 6, e a mais aguda corda 1.

---

<sup>1</sup> <https://cycling74.com/products/max/>, acessado em 20 de Setembro de 2020

<sup>2</sup> <https://www.mathworks.com/products/matlab.html>, acessado em 20 de Setembro de 2020

Figura 2 – Sinal das 6 cordas no strumming em um intervalo de tempo de 500 ms, que compreende um gesto descendente (das cordas graves para as agudas) e um ascendente em um ritmo. O gesto ascendente utiliza apenas as cordas 1 e 2.



Nos últimos anos desenvolvemos no nosso laboratório um sistema interativo de transcrição e expansão de trechos musicais tocados ao violão chamado *GuiaRT* (ARMONDES, 2017; FREIRE; ARMONDES; SILVA, 2020; MENESES; FREIRE; WANDERLEY, 2018), utilizando a captação hexafônica em um violão de nylon especialmente adaptado, que tem se mostrado muito útil para a composição interativa e análise de performances.

O *GuiaRT* é baseado em vários equipamentos de hardware e um software em sua implementação. O hardware conta com dois violões de nylon, um Yamaha modelo CG182 e um Alhambra modelo E-533, equipados com captadores RMC e conectores XLR fêmea de sete pinos (Figura 3), onde temos seis saídas não balanceadas de alta impedância. Para o casamento de impedância e balanceamento utilizamos *direct boxes* ativos Behringer alimentados por *phantom power* de uma interface Focusrite Saffire Pro40, a qual recebe o áudio e o digitaliza a 24 *bits* de resolução e 48 KHz de amostragem. O computador utilizado é um *iMac 27"* com sistema operacional *OSX Mojave 10.14.6*. Atualmente possuímos também uma configuração portátil para apresentações e coletas de dados *in loco*: um MacBook Pro com o mesmo sistema operacional e capacidade de

processamento similar, os mesmos *direct boxes* Behringer e uma interface Presonus 1810c. A programação é realizada em *Max/MSP* (versões 6 a 8), com objetos adicionais gratuitos desenvolvidos para processamento de sinais. Ao extrair um trecho podemos, ainda no *Max/MSP*, converter os dados em protocolo MIDI, reproduzindo-o com o banco de *samplers* de violão de nylon *Ilya Efimov* carregados no software *Kontakt*.

Figura 3 – Violão Yamaha equipado com captação hexafônica e conector multivias para aquisição independente dos sinais das cordas.



A partir desse sistema foram estudadas as técnicas de tremolo (FREIRE; NÉZIO, 2013) e *plaquê*, com suas possíveis combinações de *onsets* (FREIRE; NEZIO; REIS, 2014), e *strumming*, através da análise de *Inter-onset Intervals (IOI)*<sup>3</sup> e identificação de performers (FREIRE; ARMONDES; VIANA, 2017). No âmbito de composição e sistemas musicais interativos foi criada a peça *Nibanna* (SILVA; FREIRE, 2018), que utiliza

<sup>3</sup> *Inter-onset Interval* é o intervalo temporal entre os inícios de notas/eventos sucessivos, o intervalo entre os onsets, não incluindo a duração da nota/evento. É uma ferramenta importante para se estudar a expressividade rítmica (CLARKE, 2004). No estudo do ritmo no violão, podemos considerar os *IOIs* de diferentes maneiras: entre eventos produzidos em uma mesma corda, entre duas notas sucessivas de um acorde, entre notas ou acordes distintos.

identificação de *onsets*, variações criadas a partir de trechos tocados e processamento de áudio. Foi também criada e estreada em 2019 uma versão para GuiaRT do quarteto de violões *Música para 24 cordas*, de Sérgio Freire. No estudo de gestos, como recorte desta pesquisa, propusemos uma abordagem multimodal da técnica de *strumming* (ARMONDES; FREIRE, 2019) e a comparação de performers em um ritmo simples tocado com metrônomo (ARMONDES et al., 2020), que será apresentado como experimento nesta dissertação<sup>4</sup>.

Embora o GuiaRT nos possibilite extrair vários descritores de baixo nível, tais como nota tocada, duração, intensidade, centroide espectral, etc. (Figura 4), utilizaremos para este trabalho apenas a identificação de *onsets*. A estratégia de sua extração é apresentada a seguir.

Figura 4 – Output do GuiaRT de um trecho com destaque para os onsets. Na ordem horizontal temos: evento, onset (ms), corda, nota MIDI, casa, duração (ms), amplitude (dB), flag de ligado, centroide espectral, indicadores de articulação (harmônico, pizzicato, vibrato) e microafinação.

1,	1944	6	40	0	824	-18	0	7.696853	0	0	0	0.;
2,	1949	5	47	2	326	-15	0	7.265836	0	0	0	0.;
3,	1979	4	52	2	285	-11	0	5.659743	0	0	0	0.;
4,	2000	3	55	0	251	-13	0	5.688028	0	0	0	0.;
5,	2019	2	59	0	216	-13	0	5.178263	0	0	0	0.;
6,	2040	1	64	0	189	-12	0	4.27051	0	0	0	0.;
7,	2229	1	64	0	475	-12	0	3.381457	0	0	0	0.;
8,	2235	2	59	0	288	-16	0	3.758808	0	0	0	0.;
9,	2251	3	55	0	269	-16	0	3.981949	0	0	0	0.;

### 1.3.1.1. Detecção de onsets

O som típico do violão é caracterizado por um ataque rápido, com um considerável aumento de energia em um intervalo de tempo curto, seguido de ressonância (que pode ser bastante longa caso não seja abafada propositalmente). Neste caso, consideramos que o *onset* pode ser representado pelo final do ataque (ou pela clara aparição de um pico) uma vez que a inclinação dos ataques não varia muito entre os diferentes registros e articulações.

<sup>4</sup> Os artigos e peças podem ser acessados em <https://musica.ufmg.br/lapis/>



Mesmo com a separação do sinal de áudio por cordas existe um acoplamento mecânico entre os captadores e o corpo do instrumento que torna a tarefa de analisar os sinais de áudio mais complexa do que parece. Um toque em uma única corda pode gerar sinais de energia consideráveis nos captadores das demais cordas, fato que pode causar a identificação de falsos *onsets*. Não apenas a ressonância, mas também os transientes podem produzir níveis de energia mais elevados em uma corda em repouso do que aqueles produzidos por toques suaves na mesma corda.

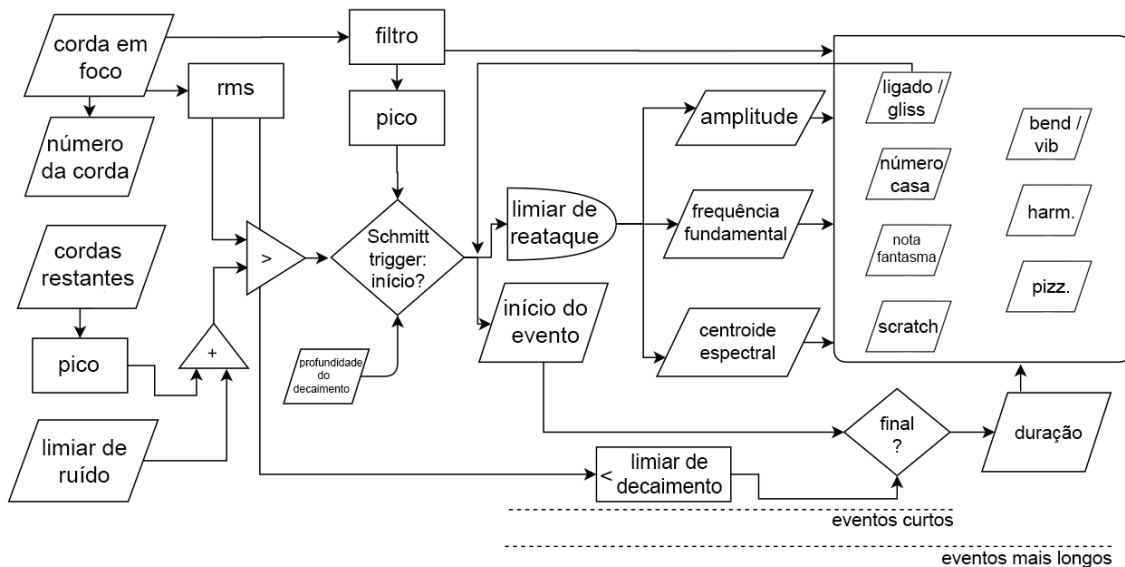
Utilizamos a escala dBFS (que usa o valor 1 como referência) para o cálculo de valores de pico e RMS do sinal de cada corda. No nosso setup observamos, com a utilização de uma *janela* de 1024 pontos e *hop size* de 256, um ruído de fundo de -80 dB em média, para cada corda. As amplitudes são estimadas como o valor RMS máximo encontrado entre o onset e um intervalo mínimo para ocorrência de um novo ataque (definido entre 50–80 ms), utilizando os mesmos parâmetros já mencionados. Diversos músicos já utilizaram o sistema em performances, com diferentes estilos e graus de experiência, e encontramos ataques entre -40 e -7 dB nessa amostragem.

O algoritmo de detecção de onsets possui uma fase de pré-processamento dos sinais captados. Para corda, é criado um submix das cordas restantes, e o pico deste sinal é calculado a cada 10 ms. Aos sinais de cada corda são aplicados dois filtros passa-banda de segunda ordem, cujas saídas são mixadas; em seguida calculamos os picos desses sinais filtrados em intervalos que variam de 1 a 3 ms (da corda mais grave para a mais aguda). O filtro passa-baixa foca nas frequências fundamentais presentes em cada corda, com frequência central na metade da escala (trítone na sexta casa), enquanto o filtro passa-alta destaca a energia dos transientes.

A estratégia básica para a detecção de *onset* é baseada na comparação da amplitude do pico do sinal filtrado de cada corda com um *threshold* adaptativo, usando um *Schmitt trigger* (um comparador com duplo *threshold*). O *threshold* adaptativo é o maior valor escolhido entre o ruído de fundo variável e o valor de RMS de cada corda. O ruído de fundo variável depende dos valores de pico das cordas restantes e de um parâmetro definido independentemente para cada corda. Este parâmetro está relacionado à inclinação de uma linha que associa o pico de amplitude das demais cordas entre -38 e 0 dB com o aumento do valor do ruído de fundo. Um outro parâmetro

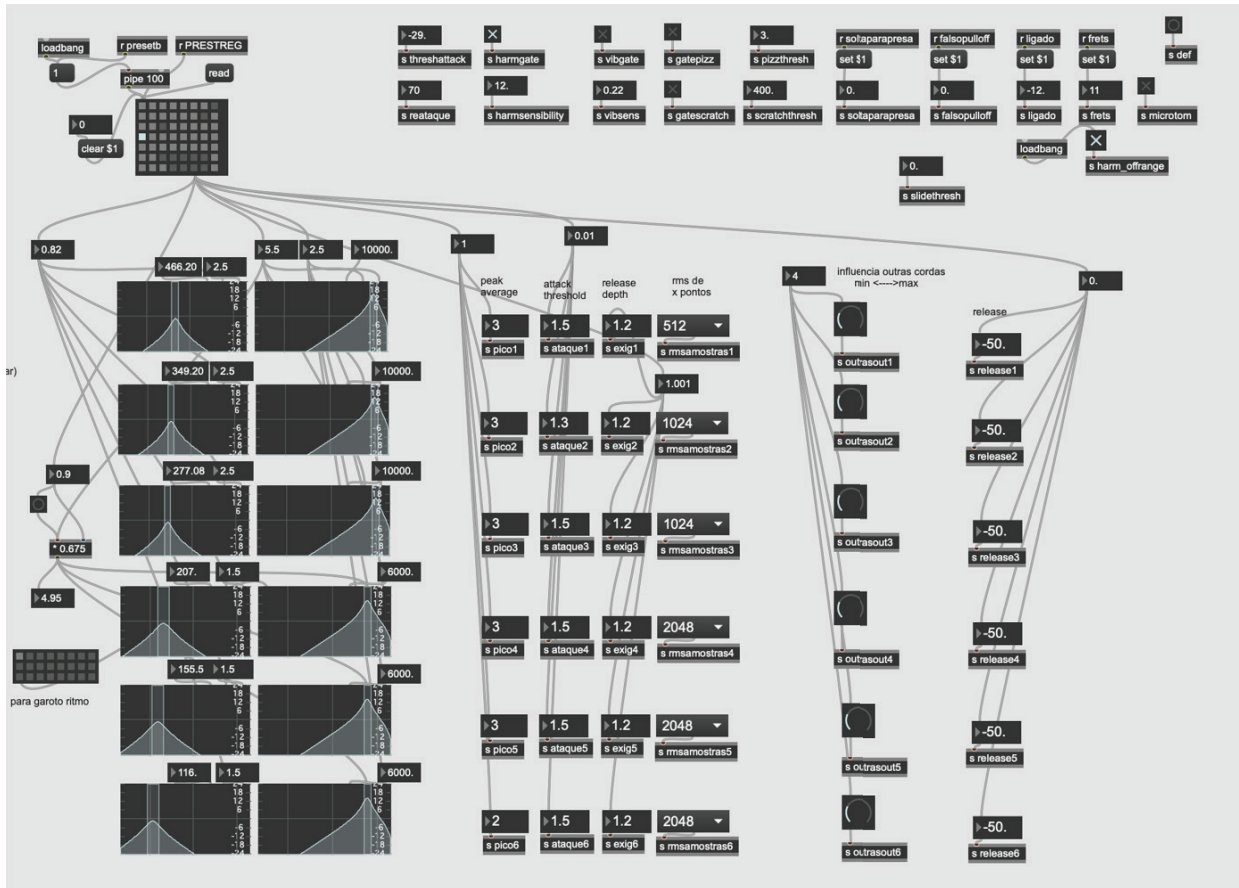
importante é a inclinação do decaimento, que controla a diferença entre os limiares máximo e mínimo do *Schmitt trigger*. Esse processo, que ocorre dentro de uma faixa de erro de 10 ms, pode ser visto no fluxograma da Figura 5.

Figura 5 – Fluxograma do algoritmo do GuiaRT. Ao lado esquerdo da figura podemos observar a estratégia de detecção de onsets, indicado como início em português.



No caso do *strumming* a energia do sinal não possui uma queda considerável entre os ataques na maioria das vezes, ocorrendo apenas um novo transiente de alta frequência. Por isto, o filtro focado nos transientes tem ganho maior do que aquele focado nas frequências fundamentais, o que permite uma modulação significativa do nível do sinal filtrado. Em certos casos pode acontecer até de uma corda não ser atacada e suas vizinhas sim, o que depende da direção do movimento, posicionamento da palheta e dedilhado da mão esquerda. Para estas especificidades temos outros parâmetros (valores para filtro de média móvel dos picos, ajustes de dinâmica para diferentes músicos, intervalo para validação de um novo *onset*) ajustáveis para cada situação, mostrado na Figura 6.

Figura 6 – Patch do GuiaRT em Max/MSP com o pré-processamento do sinal e parâmetros reguláveis para cada corda.



### 1.3.2. Sensor de movimento (IMU MetaMotionR)

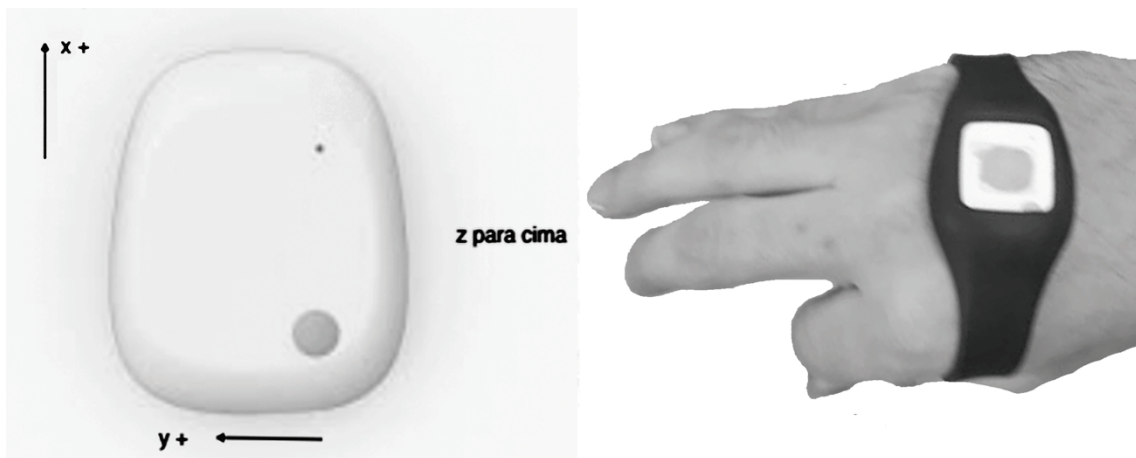
Recentemente houve um grande avanço tecnológico dos IMUs (*Inertial Measurement Units*) (PASSARO et al., 2017). Estes novos sensores são compactos, contam com bateria recarregável e transmissão de dados sem fio. Dentre os diversos produtos disponíveis no mercado utilizamos o *MetaMotionR*, da empresa *MbientLab* (Figura 7). Consiste em um pequeno sensor de nove graus de liberdade no formato de relógio que transmite, via protocolo *Bluetooth Low Energy (BLE)*, informações de aceleração linear, velocidade angular e magnetômetro em três eixos, além de temperatura, luminosidade ambiente e pressão barométrica. Com sua independência de infraestrutura, como câmeras e salas especialmente adaptadas, são uma alternativa para



estudo de gestos pela relação custo-benefício que oferecem quando comparados com os sistemas óticos consagrados.

O uso deste tipo de sensor ocorre nas mais diversas aplicações, desde controle de qualidade industrial, reabilitação médica, robótica, sistemas de navegação, aprendizado esportivo e sistemas de realidade aumentada (AHMAD et al., 2013). No estudo de performance Sawicki e Zieliński (2019) utilizaram um IMU no antebraço para detectar a expressividade musical dos gestos de violinistas. Schiesser e Schacher (2011) desenvolveram o SABRe, um clarinete baixo aumentado através da posição das chaves, pressão da boca e IMU acoplado ao instrumento.

Figura 7 – IMU MetaMotionR, da empresa MbiEntLab. À esquerda, as setas indicam a orientação dos três eixos no sentido positivo. À direita, vê-se este sensor bem compacto posicionado na mão do performer.



O sensor permite diversas configurações de uso. Para registro em memória interna, acessada através do aplicativo *Metabase App* (Windows, iOS e Android), opera com *sample rate* de até 800 Hz. No modo *wireless streaming* temos uma taxa de até 100 Hz para os acelerômetros e giroscópios, e de 25 Hz para o magnetômetro. Possui um algoritmo de fusão, desenvolvido pela empresa *Bosch*, permitindo estimativa de aceleração linear (sem a influencia da gravidade) e atitude espacial. Não podemos determinar o *sample rate* do modo fusão, estabelecido por suas rotinas internas, mas alcançamos a taxa de 100 Hz no nosso *setup*.

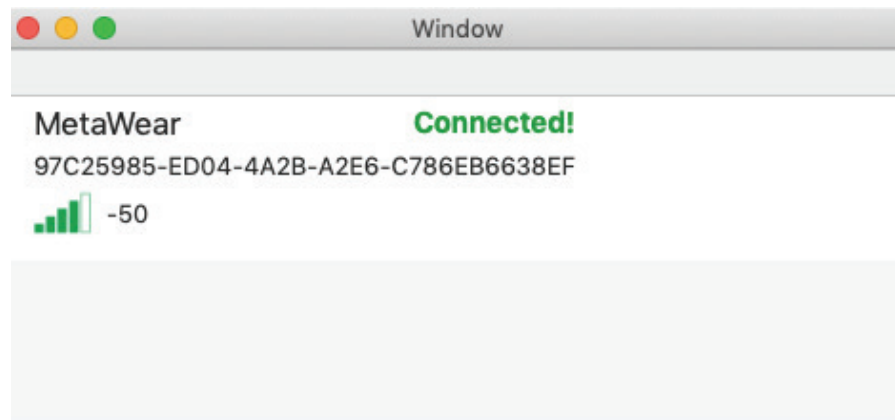
A fusão possui quatro modos de operação que combinam os sensores de formas diferentes: *NDoF* (nove graus de liberdade), que estima a orientação absoluta dos três sensores; *IMUPlus*, que estima a orientação relativa no espaço usando dados do acelerômetro e giroscópio; *Compass*, combinação de acelerômetro e magnetômetro; *M4G*, que usa o magnetômetro para detectar rotação. Para este estudo optamos pelo modo *IMUPlus*, que possui o maior *sample rate* e pelo fato de a orientação absoluta normalmente não oferecer problemas em performances musicais ao vivo. Além da aceleração linear e velocidade angular nos três eixos, recebemos os dados de atitude em Ângulos de Euler, por serem mais intuitivos para performance do que os quatérnios e por ser a unidade utilizada no Qualisys, um sistema ótico que utilizamos para comparação com o IMU, estudo que descreveremos na seção experimental.

O fabricante não oferece, até o momento, software para o uso desse sensor no modo *streaming* para sistemas operacionais *MacOS*. Tivemos, portanto, que escrever um drive para o sistema *MacOS 10.14 (Mojave)* na linguagem *Swift* a partir da adaptação de códigos fornecidos pela *MbientLab* e um *framework* dedicado ao *OSC (Open Sound Control* <sup>5</sup>), desenvolvido por Devin Roth (2016). Este drive prepara o sensor para envio de dados por *BLE* e, uma vez recebidos, são roteados para o *Max/MSP* através de um servidor local utilizando mensagens *OSC* via *UDP* (Figura 8).

---

<sup>5</sup> Open sound control é um protocolo de comunicação entre computadores, sintetizadores de som e outros dispositivos multimídia otimizado para tecnologia de redes moderna, como características principais ser Inter operável, preciso e flexível. <http://opensoundcontrol.org/introduction-osc>, acessado em 11/09/2020.

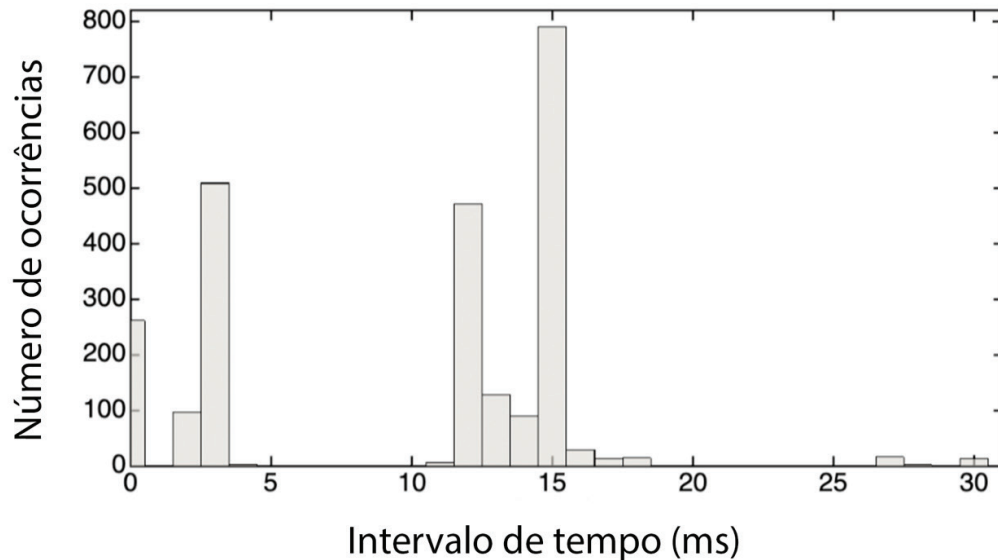
Figura 8 – Interface do drive escrito para MacOS 10.14 Mojave informando o status de conexão do sensor MetaMotionR pelo protocolo Bluetooth Low Energy



O protocolo *BLE* permite que os dispositivos em uso entrem em um modo de baixa energia ou inatividade quando a transmissão/recepção de um evento é concluída, operando com mais eficiência em largura de banda limitada. Diversos fatores podem afetar o desempenho nesse tipo de conexão, como hardwares, versão de software, distância entre os equipamentos, interferências e nível de bateria.

No *patch* do *Max/MSP* cada mensagem OSC é composta por 10 dados de 16 *bits*: três para aceleração linear (*g*), três para velocidade angular (graus/s) e quatro para os ângulos de Euler (graus); como um destes quatro valores é redundante, ele é excluído dos cálculos posteriores. Criamos neste *patch* um monitor para acompanhar a taxa de amostragem real da entrada de dados, e o número de mensagens recebidas por segundo é contado e apresentado como uma curva, que varia em torno a 100 Hz. Os dados de cada mensagem são gravados em um *buffer*, com resolução de 1 ms, e marcados com o seu tempo de chegada. Observando o intervalo de tempo entre duas mensagens subsequentes em um trecho, notamos a maior presença dos valores 0, 3, 12 e 15 ms (com média geral de 9,88 ms), como mostra a Figura 9. Estas irregularidades são intrínsecas ao protocolo *BLE*. Devido a isto, registramos também um *log* de erro contendo os instantes em que os intervalos de tempos são longos (> 30 ms).

Figura 9 – Histograma do intervalo de tempo entre mensagens subsequentes com todas quantidades físicas (aceleração linear, velocidade angular e ângulos de Euler).



A partir dessa observação percebemos a necessidade de reconstruir o *sample rate* de 100 Hz utilizando um *timer* e um filtro de média móvel. A cada 10 ms, um novo valor para cada uma das três grandezas é gerado na saída do filtro por um *clock* interno. Com a média geral do intervalo de tempo entre as mensagens está abaixo de 10 ms, e os valores ocorrendo em torno a 30 ms são raros, escolhemos utilizar um filtro com três pontos, alimentado com a entrada de um novo dado quando esta antes do próximo *tick*, ou com a repetição do dado anterior quando não chegam novas mensagens entre dois *ticks*. Esta ação impede a repetição de dois valores idênticos na saída do filtro. Embora a taxa de 100 Hz possa parecer insuficiente para dar conta de um fenômeno que conta com eventos ocorrendo (onsets sonoros) em um intervalo inferior a 10 ms, as curvas de aceleração e velocidade angular são suficientes para descrever todo o gesto envolvido na produção de um acorde.

Nas circunstâncias em que a pesquisa se desenvolveu, temos dados coletados com duas situações de programação do sensor. No primeiro experimento não utilizamos a fusão das grandezas, implicando na influência da gravidade sobre a aceleração linear; nas gravações seguintes, com a fusão implementada, eliminamos a gravidade, e obtivemos a aceleração linear de fato, além dos ângulos de Euler. Desse modo

realizamos dois testes de latência distintos, para cada situação, que serão descritos em momentos adequados.

Além dos dados gerados pelo *IMU* de aceleração linear e velocidade angular, podemos integrar essas informações para obter curvas de velocidade e deslocamento (com dupla integração), úteis tanto para a análise do *strumming* quanto para o estudo comparativo realizado. Apresentaremos no Apêndice A o método proposto para a integração da aceleração linear (no modo de fusão de dados) e sua aplicação em um caso de deslocamento com rotação.

### **1.3.3. Qualisys**

*Qualisys* é um sistema de captura ótica de movimento que consiste em um conjunto de câmeras de infravermelho conectadas via Ethernet. Cada câmera emite luz infravermelha, que é refletida por marcadores passivos gerando dados posicionais bidimensionais (2D). O software proprietário coleta esses dados 2D e calcula as posições tridimensionais (3D) desses marcadores no espaço.

É um equipamento já bem estabelecido no mercado, sendo usado para as mais diversas aplicações. Podemos destacar seu uso na área da saúde (CHEN et al., 2012; RIBEIRO et al., 2014) e robótica (ZHOU; BAI; LI, 2017). Em performance musical é utilizado em pesquisas que relacionam atividade cerebral e captura de movimento (MAIDHOF; KÄSTNER; MAKKONEN, 2014), movimento corporal com elementos estruturais da música (THOMPSON; LUCK, 2012) e comparação com outros sistemas de captura (BISHOP; JENSENIUS, 2020).

A infraestrutura utilizada para esse trabalho incluiu dois Oqus 300, quatro Oqus 400, quatro câmeras Oqus 700 e o software de gravação e edição *Qualisys Track Manager (QTM)* versão número 2019.1 (build 4420).

### **1.3.4. Comparação de dados gerados pelo MetaMotionR e pelo Qualisys**

No uso de novos equipamentos, é comum a análise comparativa com sistemas de referência, a fim de identificar sua performance e eficiência. Com a evolução dos *IMUs*, alguns estudos buscaram esse tipo de comparação com sistemas óticos consagrados

em diferentes situações (ABHAYASINGHE; MURRAY; SHARIF BIDABADI, 2019; ZÜGNER et al., 2019).

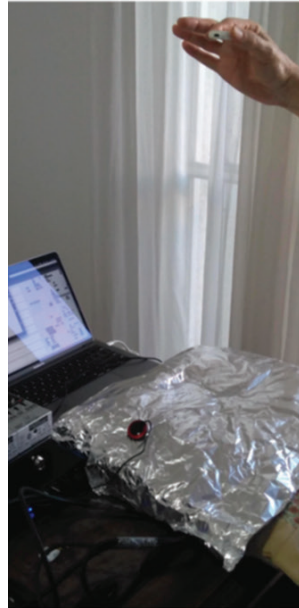
Em uma parceria acadêmica com a Universidade McGill, de Montreal, Canadá, tivemos a possibilidade de realizar um estudo utilizando o *MetaMotionR* e o *Qualisys*, um sistema ótico de captura de movimento, considerado uma referência nesta área. Com dois performers realizando dois ritmos com a técnica de *strumming* (além de alguns gestos percussivos), pudemos comparar os dados coletados pelos dois sistemas, compreendendo melhor a capacidade do *IMU* utilizado na pesquisa relatada nesta dissertação. Esse estudo nos rendeu um artigo em inglês (FREIRE et al., 2020), do qual retomaremos a seguir os experimentos e resultados.

Esta seção se organiza assim: com o sensor programado na função fusão de dados dos acelerômetros e giroscópios, realizamos um teste de latência, identificando o tempo de atraso de resposta nesta configuração. Em seguida, apresentaremos o *setup* e as posições dos marcadores. Descrevemos então o experimento e o processo de gravação. Finalmente, apresentamos os resultados do estudo, acompanhados dos condicionamentos dos dados necessários para cada comparação.

#### **1.3.4.1. Medida do tempo de resposta do IMU**

Esse experimento busca calcular de forma empírica o atraso de resposta do *IMU*. O sensor é colocado em queda livre na direção do eixo Z de uma altura de aproximadamente 60 cm, entrando em contato com uma almofada coberta com papel alumínio. Um microfone de contato está sobre o papel e ligado à interface de áudio. No *Max/MSP* são gravados dois canais, um com o som gerado pelo impacto com o papel alumínio e outro com a aceleração linear gerada pelo sensor. A Figura 10 mostra o procedimento para este experimento.

Figura 10 – Cálculo da latência do sensor com a preparação para a queda livre. O sinal de áudio é captado por um microfone de contato.



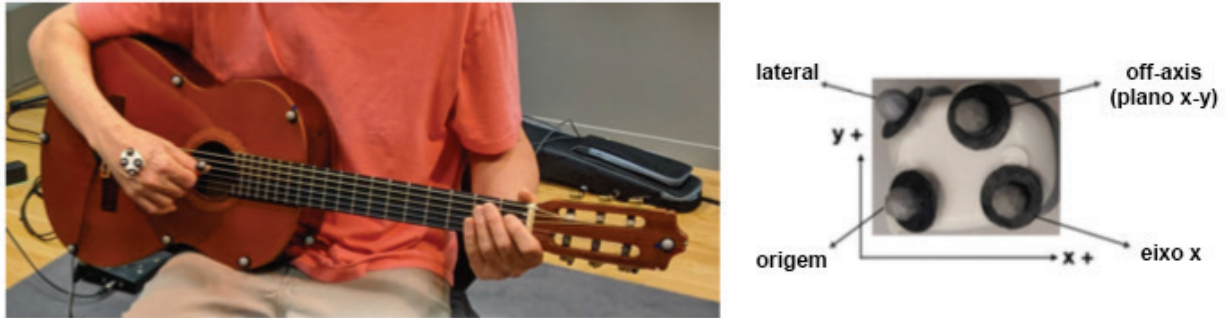
Manualmente medimos o intervalo entre o sinal de áudio e a mudança de estado da aceleração linear e calculamos, com 200 repetições, a média e desvio padrão desse processo. Encontramos o valor de  $24,4 \pm 6,4$  ms, que será utilizado na sincronização dos dados das três fontes: IMU, áudio multicanal e Qualisys.

#### 1.3.4.2. Setup

Colocamos marcadores passivos no IMU, no polegar (perto do ponto de contato entre a palheta - ou dedo - e cordas durante o *strumming*) e no violão. Marcadores passivos de 7 mm foram usados no IMU, enquanto um marcador de 10 mm foi colocado no polegar. A taxa de amostragem definida para a captura foi de 100 Hz, estando associada a uma gravação de vídeo (incluindo áudio) de 25 fps.

O violão foi mapeado como um corpo rígido com 14 marcadores de 13 mm, assim como o sensor, com 4 marcadores, como pode ser observado na Figura 11.

Figura 11 – Violão e IMU com os marcadores para mapeamento em corpo rígido pelo sistema de captura de movimento Qualisys.



Os eixos  $x$  e  $y$  do *IMU* foram alinhados com a estrutura de referência *Qualisys L-shaped* usada para a calibração, de forma que os quatro marcadores possuem as coordenadas que são mostradas na Tabela 1. A definição dos ângulos de Euler usou a opção: “Todos os três os eixos de rotação são iguais aos eixos de referência - o segundo e o terceiro eixos de rotação não são afetados pela primeira ou segunda rotações” (QUALISYS, 2011, p. 187), para torná-los compatíveis com os ângulos gerados internamente pelo *IMU*. Pelo mesmo motivo, o corpo rígido do *IMU* foi transladado ao seu ponto de origem e reorientado no software QTM, conforme Figura 11. O ponto de origem do *IMU* foi considerado a origem de seu referencial, em torno do qual são calculadas as rotações.

Tabela 1 – Coordenadas dos marcadores de corpo rígido do IMU para alinhamento com a referência Qualisys L-shaped.

Marcador	X (mm)	Y (mm)	Z (mm)
Origem	0	0	0
Eixo X	22	0	0
Off-axis	14,95	16,76	0
Lateral	-2,9	20,57	-3,28



### 1.3.4.3. Gravações

Com o objetivo de comparar os dois sistemas em situação de performance, escolhemos três ritmos: o primeiro, com gestos regulares descendentes e ascendentes; o segundo, um acompanhamento comum para balada, com acordes arpejados e variação na duração das figuras rítmicas; o terceiro, ataques percussivos no corpo do instrumento. Todos eles podem ser vistos na Figura 12.

Figura 12 – Partitura dos três ritmos utilizados neste experimento.

**Ritmo 1**

♩ = 60    Em                                  A                                  Em

**Ritmo 2**

♩ = 70    Am                                  G                                  Am

**Ritmo 3**

Gesto a	Gesto b	Gesto c
tambora (cavalete)	percussão com a palma da mão sobre a boca	percussão no tampo próximo ao braço

dedos restantes (m.d.) lado mais baixo

polegar (m.d.) lado mais alto

Dois músicos realizaram as gravações em dois *takes* cada. Identificaremos as gravações pela combinação de letras, relativas aos músicos ( $m1$  ou  $m2$ ) e o *take* ( $t1$  ou  $t2$ ). Por exemplo, o *take 2*, executado pelo Músico 2 será representado por  $m2t2$ . O  $m1$ , canhoto, executou os ritmos 1 e 2 com palheta, já o  $m2$ , destro, não utilizou palheta nos mesmos trechos.

As gravações nos sistemas *Qualisys* e *GuiaRT / IMU* foram feitas de forma independente. A gravação de áudio foi feita com um sistema portátil. A cadeira do músico foi posicionada na origem do quadro de referência, voltada para o eixo  $x$  positivo. Quando as câmeras de captura de movimento começavam a gravar, o músico pressionava um pedal, que iniciava a gravação de áudio/IMU no *Max/MSP*, junto com um clique tocado nos alto-falantes. Após quatro cliques, a performance começava. A duração da captura foi predefinida no software de captura de movimento *QTM*; para o violão e *IMU*, a liberação do pedal marcou o fim. O *IMU* foi posicionado de diferentes maneiras, com fita adesiva dupla-face, para testar se diferentes orientações em torno do eixo  $z$  poderiam influenciar nos resultados, conforme ilustrado na Figura 13. Observamos que algumas posições da mão direita dos violonistas favorecem a ocorrência de confusão entre os eixos (*gimbal lock*). Assim, este fator foi levado em consideração durante o posicionamento do sensor para evitar a criação de singularidades no cálculo dos ângulos de rotação.

Com dados vindo dos dois sistemas de forma independente, o alinhamento das informações se torna necessário para a correta interpretação, cujo processo apresentaremos a seguir.

A primeira etapa para alinhar os dados de ambos os sistemas é estimar o deslocamento entre as gravações de áudio, o atraso entre o áudio gravado em vídeo pelo *Qualisys* e o gravado no *GuiaRT*. Depois disso, também estimamos a distância entre o violão e a câmera em 4,5 m. Esta distância corresponde a um atraso de aproximadamente 13 ms a ser compensado do valor de deslocamento. O atraso de resposta do *IMU* foi fixado em 24,4 ms, de acordo com a análise empírica na seção 1.3.4.1. Por fim, observamos empiricamente que os melhores alinhamentos foram obtidos com um atraso de quadro de vídeo, que é de 40 ms.

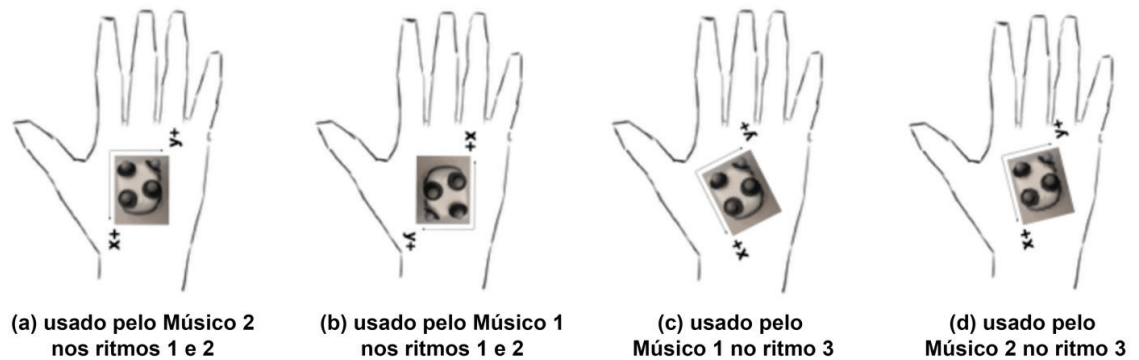
Portanto, o deslocamento entre as gravações de áudio deve ser compensado por esses três valores, a fim de atingir o alinhamento de tempo entre os dados do sistema de captura de movimento e do *IMU*. O resultado desta operação é arredondado para o múltiplo mais próximo de 10 ms (já que a taxa de amostragem é 100 Hz) e então aplicado para retornar os dados de captura de movimento.

Os ângulos de atitude também precisavam de algum alinhamento. O primeiro passo foi inverter o sinal dos ângulos do *IMU*. Além disso, como não estávamos lidando com a orientação absoluta dos corpos rígidos, e também usando diferentes posições da *IMU* na mão, também foi necessário alinhar os ângulos de Euler com o eixo z para cada take. Esse procedimento foi feito observando-se os valores mínimo e máximo dessa grandeza física, a partir da qual foi estimado um deslocamento, e posteriormente refinado pela inspeção visual das curvas. Todas as rotações extrínsecas foram feitas na ordem xyz.

Para a conversão dos valores de aceleração entregues pela *IMU* em g (aceleração da gravidade), foi utilizado o valor  $9,81 \text{ m/s}^2$ . Os valores são expressos em graus para ângulos, em m / s para velocidades e em cm para deslocamentos. O ponto de origem da *IMU* (Figura 13b) foi considerado a origem de seu referencial, em torno do qual são calculadas as rotações.

Os dois sistemas se utilizam da referência “regra da mão direita” para orientação dos eixos e das rotações: os dedos indicador, médio e polegar da mão direita são posicionados ortogonalmente, indicando respectivamente a direção positiva dos eixos x, y e z. As rotações positivas se dão em sentido anti-horário, quando se olha em direção à origem de cada eixo. No entanto, os valores dos ângulos de atitude do *IMU* (e não as velocidades angulares) são recebidas com sinal contrário.

Figura 13 – Diferentes posições do *IMU* utilizadas pelos músicos nas gravações dos ritmos.



#### 1.3.4.4. Resultados

Os resultados obtidos estão divididos em três categorias: comparação dos dados de rotação (ângulos de atitude), comparação dos dados de translação (medidos diretamente e estimados por integração), resultados adicionais. Como as duas primeiras categorias possuem um caráter essencialmente técnico-estatístico, preferimos apresentá-los em um apêndice (Apêndice B) ao final desta dissertação. Deixamos no corpo do texto apenas os resultados adicionais e sua discussão geral.

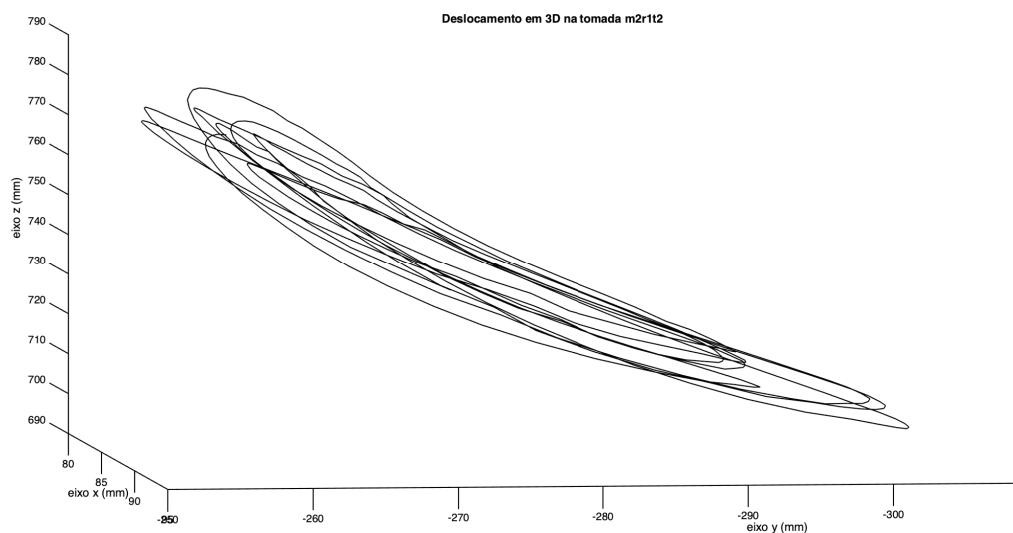
É importante notar que este estudo foi realizado entre os dois experimentos relatados na segunda parte da dissertação e, de certa forma, serviu para a validação do uso deste *IMU* na pesquisa.

Um resultado adicional deste estudo, voltado à comparação entre os dados gerados pelo *IMU* e pelo *Qualisys*, decorre da observação da correlação entre as curvas de deslocamento e de rotação calculadas por este último. Fizemos uma análise de covariância entre as curvas de deslocamento nos eixos  $x$ ,  $y$  e  $z$  e as curvas dos ângulos de Euler, em um total de nove comparações para cada gravação. Nos pares com maior correlação, tais como deslocamento no eixo  $y$  / rotação no eixo  $z$ , observa-se sempre uma defasagem de 1 a 6 quadros (10 a 60 ms) da última curva, indicando que os picos das curvas dos ângulos são atingidos pouco depois dos picos das curvas de deslocamento, Ou seja, é uma indicação de que as rotações se prolongam um pouco mais do que as translações, que pode ser explicada pela necessidade de reposicionamento da mão antes de fazer o deslocamento na direção inversa.

Uma outra observação, que pode ser útil para comparações futuras realizadas sem um sistema ótico de captação de movimentos, é que nas comparações entre as curvas de deslocamento medidas pelo *Qualisys* e aquelas estimadas pelo método de integração desenvolvido, existe a tendência de um pequeno atraso dos picos destas últimas. Estamos nos referindo aqui apenas às curvas com coeficiente de covariância alto, geradas por movimentos oscilatórios bem evidentes. Tal resultado pode decorrer das filtragens aplicadas (filtros de média móvel) ou do próprio método de integração. Um efeito colateral dessa observação é que ao se lidar apenas com os dados do *IMU*, pode-se obter uma defasagem menor entre as curvas de translação e rotação do que as observadas neste experimento,

Finalmente, os dados posicionais do *Qualisys* permitem inferir que o movimento da mão durante o *strumming* ocorre essencialmente em um plano, ao se observar as curvas desenhadas em três dimensões (Figura 14), e também os *PCAs* dessas grandezas, cujos dois primeiros componentes explicam mais de 98% da variação em todas as gravações.

Figura 14 – Curvas em três dimensões do deslocamento de m2r1t2. Podemos observar que os gestos do *strumming* ocorre essencialmente em um plano.



### 1.3.4.5. Discussão

De forma geral, os dados gerados pelo *IMU* apresentaram boa correlação com os dados do *Qualisys*, dadas as diferenças entre eles: cinemática versus dinâmica, preço, e instalação complexa versus equipamento portátil/vestível.

As três comparações diferentes feitas aqui (entre rotações, acelerações e deslocamentos) têm fontes de erros diferentes e um tanto cumulativas. A comparação mais direta, que ocorreu entre ângulos de atitude, é afetada pelos *yaw offsets*<sup>6</sup> e pelos eixos mais explorados pelos gestos. No entanto, obtivemos resultados entre aceitáveis e toleráveis, quando comparados aos descritos na literatura consultada (BEANGE, 2019; BESSONE et al., 2019).

A comparação entre acelerações, obtida através da derivação de dados posicionais, é também afetada pela acurácia dos dados de rotação, pelo uso de matrizes de rotação e possivelmente também pelo posicionamento do violão/violonista. A comparação dos deslocamentos, obtida por meio da integração da aceleração, é a mais complexa, pois depende de vários fatores: acurácia dos ângulos de rotação, efetividade da detecção de *zero crossings*, tipo de gesto a ser analisado, limitações do método de integração. Alguns resultados poderiam ser melhorados através de segmentação.

Com base nisso, acreditamos que este sensor pode ser uma alternativa viável para análises de movimento fora de ambientes especializados e controlados, como podemos observar, por exemplo, na área de esportes (DAHL et al., 2020). Dentre suas vantagens podemos citar o funcionamento sem fio, tamanho e peso pequenos, fácil de prender em diferentes partes do corpo, ou até mesmo de fazer malabarismos. Uma desvantagem significativa está relacionada ao *jitter* na transmissão *BLE* e à largura de banda restrita, que limita o número de sensores a serem usados. O atraso de resposta, que pode inviabilizar algumas aplicações interativas, pode ser compensado posteriormente na análise dos dados.

---

<sup>6</sup> O eixo *yaw* diz respeito à atitude de rotação do corpo rígido em torno do eixo *z*. Com o *IMU* atribuindo um norte arbitrário, utilizamos um ajuste (*offset*) para a correção dessa orientação em relação ao *Qualisys*.

### 1.3.5. Vídeo

Para a captação das performances, os equipamentos utilizados foram as câmeras *GoPro HERO8 Black*, com resolução de 1080p (1920 x 1080 *pixels*) e *fps* de 240, e *GoPro HERO3 Black*, a 480p (848 x 480 *pixels*) e 240 *fps*. Este registro permite segmentar os gestos de *strumming* em sua totalidade, já que podemos identificar os momentos de início e fim (repouso da mão) que acontecem antes e depois do ataque efetivo das cordas. É possível observar também a maioria dos onsets nos acordes tocados com *strumming*, uma vez que 240 frames por segundo correspondem a um intervalo de 4,16 ms, condizente com os valores observados entre os ataques nas cordas. Além disso podemos analisar a posição dos dedos, o modo como o músico segura a palheta e demais detalhes visuais da performance.

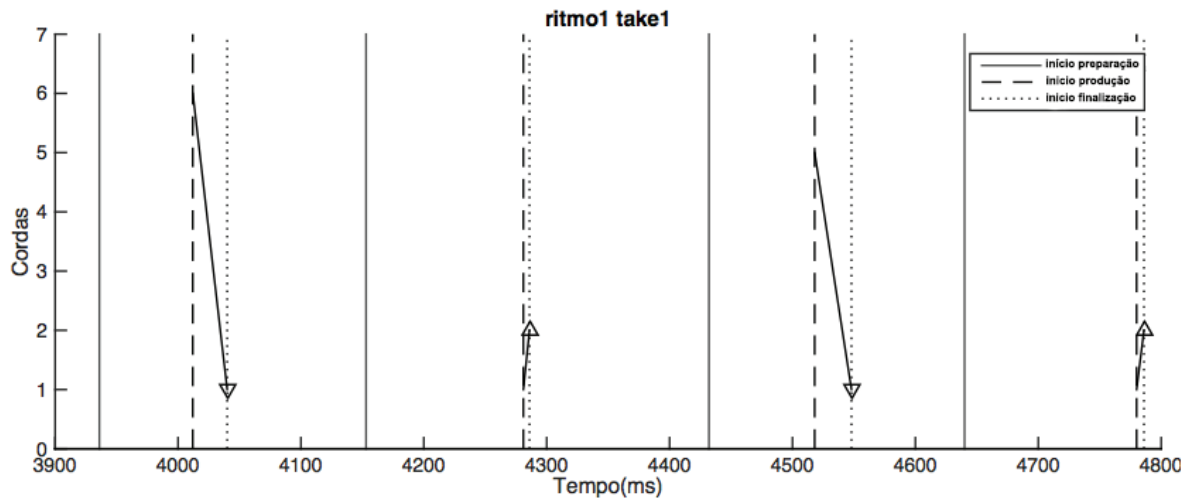
Dividimos o gesto da execução de um acorde *strummed* em 3 fases. A fase inicial, chamada de preparação, consiste no intervalo de tempo entre o início do gesto, o momento em que a mão sai do repouso, descendente ou ascendente, e o ataque da primeira nota do acorde. É esperado que nessa fase exista grande variação da aceleração linear no plano xy do sensor (paralelo ao plano do violão) e da velocidade angular no(s) eixo(s) referente(s) ao ajuste de posição da mão antes do impacto com as cordas. Importante salientar a presença da aceleração da gravidade, que pode estar a favor ou contra o sentido do movimento, sendo importante para a regularidade na adaptação rápida do performer.

A fase seguinte é a de produção do som, indo dos ataques entre a primeira e última corda do acorde, e não da distância entre todas as cordas do violão. A definição de *onset* aqui, como discutido anteriormente, ocorre no instante em que a palheta libera a corda para vibração. Por características sonoras dessa técnica, com a presença constante de sons e pouco *decay*, o número de cordas utilizadas varia constantemente, mesmo em situações densas com acordes de seis notas. Esse fato não invalida subjetivamente a performance, indicando diferenças interpretativas de cada músico. Percebemos assim que acordes onde se esperam ataques nas seis cordas possam aparecer com apenas três, por exemplo.

A fase de finalização vai do ataque na última nota do acorde até o repouso da mão, onde o ciclo se reinicia. É esperado que haja grande desaceleração nessa fase. As

três fases e a quantidade de notas por gesto, com indicação de direção, podem ser vistas em um exemplo na Figura 15. Em ritmos com grande intervalo de tempo entre os gestos, é comum, como estratégia para manutenção do pulso, que apareça um gesto sem ataque nas cordas, como veremos nos ritmos propostos nos experimentos.

Figura 15 – Gráfico com os quatro gestos de um ritmo regular e o início das três fases do gesto. O tamanho da seta indica a quantidade de notas tocadas e a flecha indica sua direção (descendente ou ascendente).



### 1.3.6. PCA

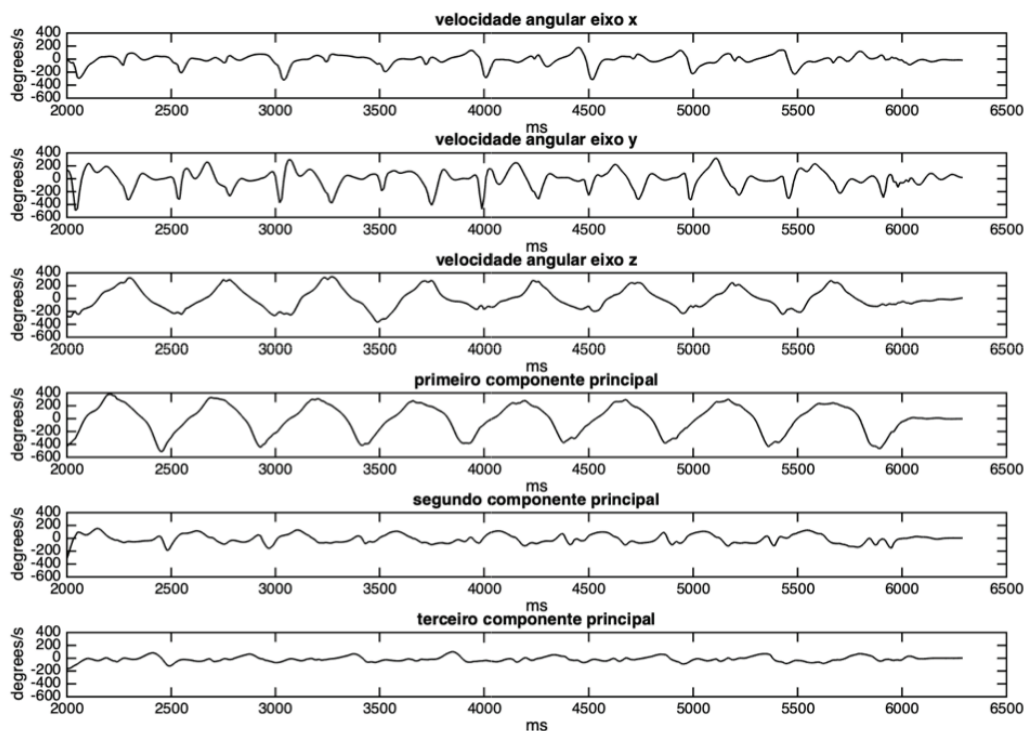
A análise de componentes principais (ACP), muitas vezes designada por sua sigla em inglês PCA, “é um procedimento matemático que utiliza uma transformação ortogonal (ortogonalização de vetores) para converter um conjunto de observações de variáveis possivelmente correlacionadas num conjunto de valores de variáveis linearmente não correlacionadas chamadas de componentes principais” (“Análise de Componentes Principais”, 2019). A partir do cálculo da covariância entre as variáveis analisadas, esse procedimento determina eixos ortogonais nos quais ocorrem variações conjuntas máximas, representando assim uma rotação das variáveis para um novo sistema de coordenadas, com dimensões menores (ou no máximo iguais) aos da distribuição original. Cada componente principal é calculado como uma combinação linear das variáveis observadas, à qual pode ser atribuído um peso de sua contribuição para a variação total.



Em música observamos o uso do PCA em estudos relacionados à ansiedade da performance (RAE; MCCAMBRIDGE, 2004) e principalmente em análise de áudio para classificação de gênero (PANAGAKIS; KOTROPOULOS; ARCE, 2010; PANAGAKIS; BENETOS; KOTROPOULOS, 2008) e decomposição de sinais (P. HUANG et al., 2012).

Como exemplo, usaremos os dados de velocidade angular nos três eixos cartesianos (x, y e z) medidos pelo giroscópio do MetaMotionR em uma das gravações realizadas. A Figura 16 representa em sua parte superior as variações da velocidade angular em cada um dos eixos, e na parte inferior os componentes principais calculados para este conjunto de dados. Pode-se notar que o primeiro componente agrupa as variações simultâneas das velocidades nos eixos x e z, que neste caso ocorrem no mesmo sentido. A título de ilustração: cada ponto do primeiro componente é assim calculado:  $c1(t) = 0,38x(t) - 0,03y(t) + 0,93z(t)$ , o que revela o grande peso do eixo z na componente mais importante dos recursos técnicos empregados nesse trecho.

Figura 16 – Aplicação do PCA nas velocidades angulares em x, y e z de um take selecionado, com os três componentes principais resultantes do método. Note-se que o primeiro componente principal se aproxima da velocidade angular em z, indicando grande influência desse eixo na realização da técnica.



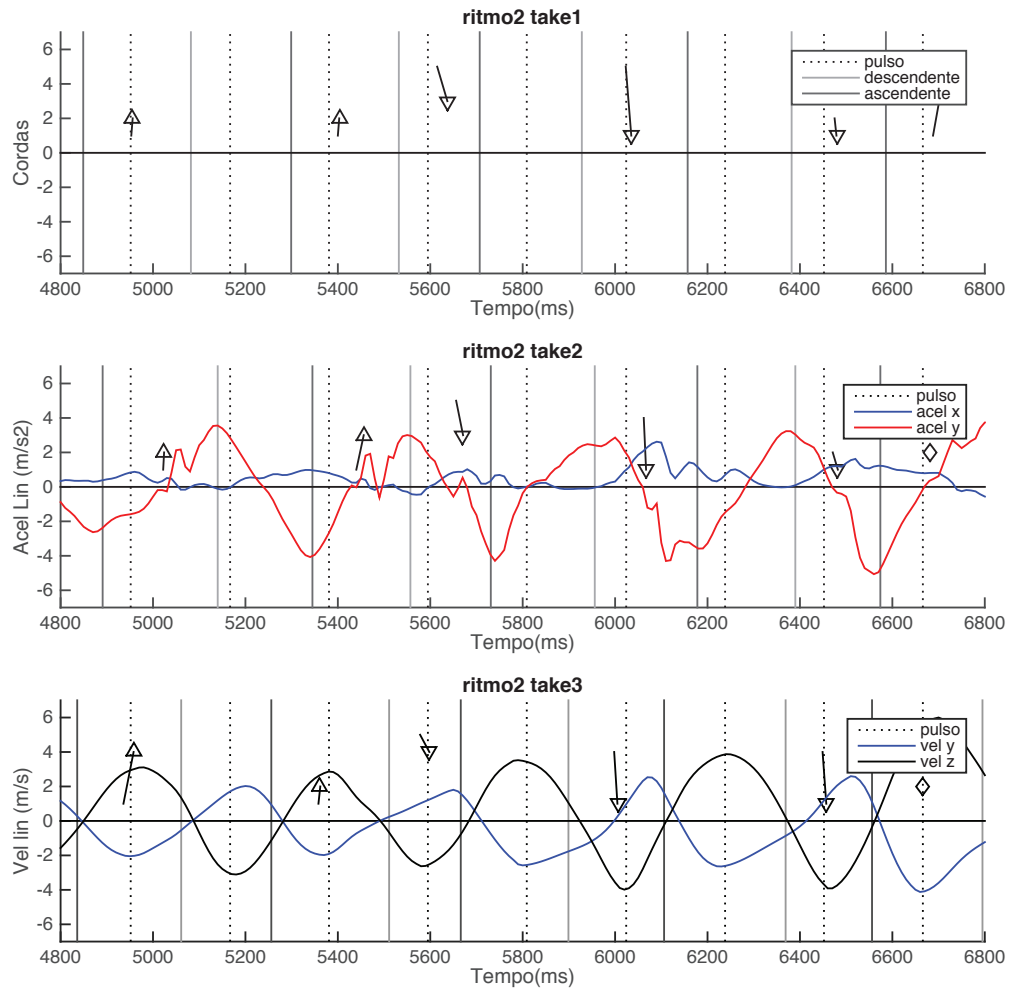
### 1.3.7. Análise conjunta dos dados

Nessa abordagem multimodal da técnica de *strumming* a análise dos diferentes tipos de dados, vindos de fontes distintas, precisam de um ambiente onde possam ser processados, analisados e transformados em gráficos e dados estatísticos. Para isso, a formatação das informações deve seguir um padrão, facilitando o processo.

Outro fator importante para a correta interpretação dos dados é a sincronia. Ao lidar com dados gerados por equipamentos que não estão sincronizados entre si, alguns cuidados se impõem, para que se possa realizar sua análise conjunta. Entre os problemas principais se encontram o *jitter* (variações na transmissão), a latência (atraso entre causa física e seu registro) e a sincronização (determinação de um tempo inicial comum).

O áudio, dados do GuiaRT e do *IMU* são gravados no *Max/MSP* através do mesmo comando, garantindo sua sincronização ou uma defasagem fixa. O vídeo, diferentemente das outras três fontes, possui um comando de gravação independente. A sincronização é feita através dos áudios (gravado pelo *GuiaRT* e câmera) no software *After Effects 2020*, da *Adobe*, que também é o ambiente que nos permite avançar *frame a frame* e localizar os pontos de mudança de direção e ataques nas cordas. Desse modo alinhamos temporalmente os dados para o estudo e podemos criar gráficos com arranjos diversos das informações registradas (Figura 17).

Figura 17 – Possibilidades de integração dos dados gerados pelo GuiaRT, IMU e vídeo. Os valores são normalizados na mesma escala, para uma boa visualização. Os pulsos estão subdivididos em semicolcheias. Os traços verticais (cinza e preto) indicam o início dos gestos (sentido descendente e ascendente).



## 2. Experimentos

Os experimentos realizados nesta pesquisa se deram em momentos distintos. A programação e desenvolvimento das ferramentas foram se aperfeiçoando e, com análise de alguns pilotos, alcançamos uma estratégia para estudar o *strumming* à luz dos dados captados em diferentes modalidades. Assim chegamos ao ponto de uma coleta de dados confiável, fundamentada nos estudos produzidos, que seria feita com músicos experientes na técnica, afim de contribuir para seu entendimento e discussões. No entanto, com a chegada da pandemia, acompanhada dos protocolos de distanciamento social, fomos obrigados a mudar a estratégia e os objetivos dos experimentos programados.

Dividiremos este capítulo em quatro seções. Inicialmente apresentaremos os estudos do Ritmo 1 em duas situações: (1) sem a implementação da fusão no *IMU* e, portanto, com a influência da gravidade na aceleração linear, com gravações realizadas por três músicos; (2) gravação de três *takes* por um mesmo músico com a fusão de dados do *IMU* implementada. Na seção seguinte, os estudos do Ritmo 2, com gravações realizadas com o *IMU* em modo de fusão de dados, por um único músico. Finalmente faremos uma análise prospectiva dos dados de todas as gravações, incluindo alguns dados produzidos no experimento de comparação entre o *IMU* e o sistema Qualysis, seguido de considerações mais gerais sobre os resultados.

## 2.1. Ritmo 1

Para essa análise inicial utilizamos um ritmo simples, com gestos descendentes e ascendentes consecutivos, com uma troca de acorde, mostrado na partitura da Figura 18.

Figura 18 – Partitura do trecho executado. As setas indicam o sentido do gesto.



### 2.1.1. Sensor sem a fusão de dados do *IMU*

Apresentaremos a seguir o setup utilizado para as gravações nesse modo de configuração do sensor, o teste de latência do *IMU*, os dados coletados e os resultados.

#### 2.1.1.1. Setup e gravações

Os registros feitos para esse experimento ocorreram na disciplina “Técnica violonística de strumming: estudo, aplicações e análise de ritmos da música popular”, ofertada no primeiro semestre de 2019 para a graduação da Escola de Música da UFMG, na qual foi realizado um estágio docência. Para a captação de áudio, utilizamos o equipamento descrito na seção 1.3.1. Com alunos de nossa escola, discutimos a técnica de *strumming* e sua aplicação em diversos ritmos e, com o suporte do GuiaRT, observamos as singularidades e semelhanças dos performers, ressaltando as diferenças interpretativas de cada um em um mesmo trecho.<sup>7</sup> Três executantes realizaram os registros em três *takes* cada.<sup>8</sup>

<sup>7</sup> Essas análises geraram um artigo (ARMONDES et al., 2020)

<sup>8</sup> Uma gravação em vídeo de um *take* de cada músico pode ser acessada em: <https://vimeo.com/420043911>

As gravações foram realizadas com metrônomo, tendo dois pulsos de espera antes do início, com a utilização de palheta. O sensor de movimento foi posicionado centralizado nas costas da mão com mesma orientação para todos os músicos, utilizando uma pulseira de borracha neste experimento (Figura 19). Este é o local de maior proximidade da palheta que, comparado com o pulso do músico, é bem mais sensível aos movimentos de translação e rotação que fazem parte da técnica. Todos performers utilizaram a mesma palheta: *Dunlop Tortex* de nylon com .87 mm de espessura. Os músicos não relataram desconforto em relação à palheta nem à fixação do *IMU*.

Figura 19 – Posição do IMU na mão de um performer. Todos os músicos utilizaram a mesma orientação do sensor.



Para a análise conjunta dos dados, iniciaremos com a investigação da latência global do *IMU*. Em seguida, apresentaremos os resultados dos registros para cada *take* e análises.

#### **2.1.1.2. Medida do tempo de resposta do IMU<sup>9</sup>**

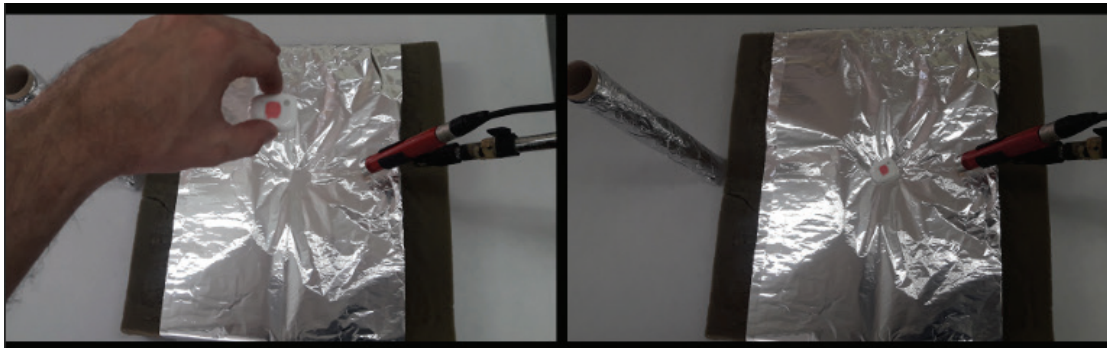
O teste procura medir o tempo de reação do sensor *IMU* em uma mudança de estado. Na configuração deste experimento, o sensor envia aceleração linear, com a influência da gravidade, e a velocidade angular dos três eixos, com uma frequência de amostragem de 100 Hz.

---

<sup>9</sup> Esta medição foi realizada anteriormente à relatada na seção 1.3.4.1.

O sensor é suspenso no ar a cerca de 30 cm, é largado e entra em queda livre até bater numa superfície com papel alumínio. O som da batida é captado por um microfone posicionado bem próximo ao papel (Figura 20) e comparado com a curva de aceleração linear no eixo z que está perpendicular à superfície. Observações preliminares mostraram que o sensor indica situação de queda livre (aceleração total igual a zero) em alturas dessa magnitude. Assim, uma mudança brusca neste valor indica o momento do choque por parte do sensor. Foram realizadas 200 repetições, obtendo-se uma média de 19,43 ms para o tempo de reação, com desvio padrão de 9,99 ms. Essas medidas levam em conta toda a cadeia de transmissão, iniciando-se nos cálculos internos do sensor, conexão *bluetooth*, programação em Swift, conversão em OSC, envio em UDP do sinal, recepção dos dados no *Max/MSP*.

Figura 20 – Teste de medida do tempo de resposta do IMU. O instante em que o som é captado pelo microfone é comparado com a mudança de estado do sensor.



Considerando-se a relação desse valor médio de resposta - 19,43 ms - com sua taxa de amostragem - 10 ms -, decidimos utilizar um valor de 20 ms como a latência geral do IMU a ser levada em conta na sincronização dos demais dados.

### 2.1.1.3. Dados coletados

Como resultado temos, para cada registro:

- a) uma lista gerada pelo GuiaRT (Figura 21), com número do evento, momento de ataque, corda, frequência fundamental em MIDI, casa, duração, amplitude, flag de nota ligada, centroide espectral e indicadores de articulação. Esta lista passa por uma conferência dos tempos iniciais das notas de cada

acorde, e é feito um realinhamento dos ataques, quando necessário. O áudio mixado é também usado para a sincronia com o vídeo;

b) uma lista gerada pelo sensor de movimento (Figura 22), com número do evento, aceleração linear nos eixos x, y e z, e velocidade angular também nos eixos x, y e z. Esta lista é resultado da reconstrução da frequência de amostragem do *IMU*, conforme tratado no capítulo anterior (seção 1.3.2).

c) uma lista com os pontos de segmentação dos gestos, oriunda da análise do vídeo da performance (Figura 23);

Todos estes dados são importados para o Matlab, para análise conjunta.

Figura 21 – Exemplo de dados extraídos pelo GuiaRT – 17 notas.

```

1, 2032 6 40 0 958 -20 0 5.776393 0 0 0 0.;
2, 2037 5 47 2 272 -17 0 4.639483 0 0 0 0.;
3, 2040 4 52 2 278 -14 0 3.689055 0 0 0 0.;
4, 2040 3 55 0 272 -17 0 6.359123 0 0 0 0.;
5, 2051 2 59 0 240 -13 0 4.05957 0 0 0 0.;
6, 2051 1 64 0 237 -11 0 2.888298 0 0 0 0.;
7, 2288 1 64 0 251 -13 0 2.947878 0 0 0 0.;
8, 2291 2 59 0 240 -15 0 3.44004 0 0 0 0.;
9, 2309 5 47 2 467 -20 0 2.990937 0 0 0 0.;
10, 2318 4 52 2 210 -17 0 3.749005 0 0 0 0.;
11, 2531 3 55 0 251 -18 0 4.752764 0 0 0 0.;
12, 2531 2 59 0 245 -16 0 3.033165 0 0 0 0.;
13, 2539 1 64 0 232 -13 0 2.612999 0 0 0 0.;
14, 2771 1 64 0 261 -11 0 3.178079 0 0 0 0.;
15, 2776 2 59 0 245 -14 0 4.52068 0 0 0 0.;
16, 2782 3 55 0 237 -16 0 4.503349 0 0 0 0.;
17, 2798 4 52 2 210 -18 0 4.17829 0 0 0 0.;

```



Figura 22 – Exemplo de dados extraídos do MetaMotionR – 170 ms.

	Aceleração Linear			Velocidade Angular		
	x	y	z	x	y	z
0,	0.742188	0.484131	0.321045	-16.646341	6.280488	-8.597561;
1,	0.731934	0.48291	0.326904	-15.853659	3.536585	-10.97561;
2,	0.714233	0.497192	0.33374	-16.158537	2.5	-12.987805;
3,	0.718384	0.498657	0.342285	-16.890244	1.402439	-13.963415;
4,	0.737915	0.499023	0.324585	-19.146341	-1.097561	-15.060976;
5,	0.752441	0.534668	0.29126	-21.463415	-3.170732	-17.012196;
6,	0.758911	0.563843	0.281738	-22.134148	-3.902439	-18.963415;
7,	0.775635	0.569214	0.293701	-21.768293	-5.853659	-23.475611;
8,	0.807007	0.557739	0.309814	-21.280489	-8.963415	-27.682928;
9,	0.832764	0.531128	0.306519	-21.097561	-12.378049	-27.012196;
10,	0.838867	0.525757	0.315186	-20.609756	-13.780488	-23.353659;
11,	0.832153	0.536377	0.334595	-16.402439	-13.719513	-19.268293;
12,	0.804932	0.526855	0.356689	-10.	-11.402439	-13.902439;
13,	0.750122	0.532227	0.363159	-4.207317	-6.95122	-10.243902;
14,	0.741211	0.529541	0.373657	1.219512	-3.719512	-9.878049;
15,	0.722412	0.496582	0.412354	6.463415	-3.353659	-8.170732;
16,	0.680176	0.511597	0.42981	8.414635	-9.146341	-6.829268;
17,	0.653931	0.528687	0.443726	7.682927	-16.829269	-4.939024;

Figura 23 – Lista dos pontos de segmentação de 17 gestos (ms) em três fases: preparação, sendo o momento de início dos gestos (troca de sentido); produção de som, considerado o momento do primeiro onset do acorde; finalização, tempo do último onset.

Preparação	Produção	Finalização
1937	2032	2051
2141	2288	2318
2400	2528	2539
2625	2771	2798
2887	2990	3032
3112	3248	3261
3362	3480	3512
3587	3747	3760
3858	3960	3992
4083	4248	4262
4362	4472	4501
4591	4744	4760
4858	4960	4992
5087	5237	5262
5370	5480	5491
5583	5739	5779
5854	5963	5987
6116		

#### 2.1.1.4. Resultados

Mesmo em um ritmo muito simples observamos diferenças globais entre os performers. A Tabela 2 mostra os dados relacionados à quantidade de notas de cada performer nos 3 *takes*. De acordo com a partitura eram esperadas 94 notas; no entanto deve-se relativizar este número, já que nessa técnica as partituras normalmente não prescrevem como se deve dar a execução. É comum que os acordes sejam indicados por cifras ou tablatura, e que o ritmo seja indicado por flechas para cima e para baixo. O número de cordas, a velocidade do gesto e a dinâmica na realização de cada acorde são decisões pessoais de cada músico.

Tabela 2 – Quantidade de notas dos três músicos em cada take com a média e o desvio padrão.

	Quantidade de notas		
	Músico A	Músico B	Músico C
<i>Take 1</i>	82	87	53
<i>Take 2</i>	87	82	44
<i>Take 3</i>	88	86	39
<b>Média e desvio padrão</b>	85,67 ± 3,21	85 ± 2,65	45 ± 7,1

Os músicos A e B se aproximaram na quantidade de notas tocadas, com médias de 85,67 e 85 no trecho, respectivamente. Já o músico C possui média de 45 notas. Isso ocorre pela clara diferenciação entre os planos graves e agudos utilizada por este performer, no qual o gesto descendente mira as cordas graves e o ascendente as agudas. Com desvio padrão de 2,65 o músico B se apresenta como o mais regular nos 3 *takes*, seguido do músico A, com 3,21, e do músico C, com 7,1, indicando que um dos motivos para menor regularidade do número de notas pode ser a intenção da separação dos planos.

O espalhamento temporal dos acordes e a duração total dos gestos também indicam diferenças entre os músicos. Os músicos estavam sentados na gravação, com o violão posicionado aproximadamente perpendicular ao plano do chão, de modo que os gestos estão diretamente ligados à influência da gravidade: no sentido descendente a

seu favor e no ascendente oposto a ela. Essa adaptação constante em relação a gravidade é também um fator relevante para o músico na regularidade do seu gesto. A Tabela 3 mostra as informações extraídas do *take* 1 de cada performer.

Tabela 3 – Média e desvio padrão da duração total dos gestos, e do espalhamento dos acordes em ms, extraídos do take 1 de cada músico.

	Média e desvio padrão da duração total e do espalhamento temporal (ms) – Take 1			
	Gesto total		Produção de som	
	Descendente	Ascendente	Descendente	Ascendente
<b>Músico A</b>	226,33 ± 15,7	267,75 ± 10,73	25,78 ± 10,46	22,25 ± 9,88
<b>Músico B</b>	295,78 ± 102,58	249 ± 3,54	53,44 ± 31,3	48,25 ± 19,45
<b>Músico C</b>	266,78 ± 68	248,75 ± 12,6	35,78 ± 32,82	18,12 ± 13,6

Notamos que a fase de produção de som ocupa uma fração pequena do gesto total nessa amostra, variando de 7% a 19%. Cada músico possui uma característica distinta em relação à duração do gesto. Na descendente: A é mais regular e possui intervalo geral menor; B é consideravelmente mais irregular e mais lento; C está no meio termo entre os outros músicos. Essa natureza se reflete no período de produção de som no mesmo sentido, mantendo aproximadamente o mesmo aspecto. Já na ascendente, B e C estão próximos e mais rápidos que A, invertendo a condição anterior, porém seu reflexo na produção de som não se mantém. Como discutiremos a seguir, a rotação da mão e o modo como os dedos seguram a palheta podem justificar essas diferenças.

O gesto do *strumming* é composto principalmente por 2 movimentos consideravelmente dependentes entre si: o de subida e descida da mão (translação), e o da rotação da mão na junção com o pulso (rotação). Ambos são captados pelo sensor, por meio da aceleração linear em cada eixo (embora neste caso esses valores estejam influenciados pela aceleração da gravidade) e velocidade angular, respectivamente. Concentremos na velocidade angular do gesto. Aplicando o PCA sobre as médias das curvas de cada músico (Figura 24), o primeiro componente representa de modo mais

intuitivo o gesto. Analisaremos também o espectro dos componentes (Figura 25) para observar a regularidade e suavidade desses gestos. As inclinações ascendentes e descendentes do primeiro componente principal são coerentes com a duração total dos gestos apresentadas na Tabela 3, indicando forte correlação entre essas grandezas.

A coerência entre as rotações nos diferentes eixos ortogonais também pode ser indicada pelas amplitudes do primeiro componente, observáveis na Figura 24. Para o músico A, o peso deste componente é de 84,4%, para o músico B é de 73% e para o músico C 56,9%, o que por si só já indica a presença de estratégias gestuais bastante distintas entre os participantes.

Figura 24 – PCA das médias das velocidades angulares dos músicos A, B e C. O primeiro componente tem peso significativo e está bem relacionado à regularidade dos gestos. As diferenças entre os músicos indicam as distintas estratégias de execução do trecho.

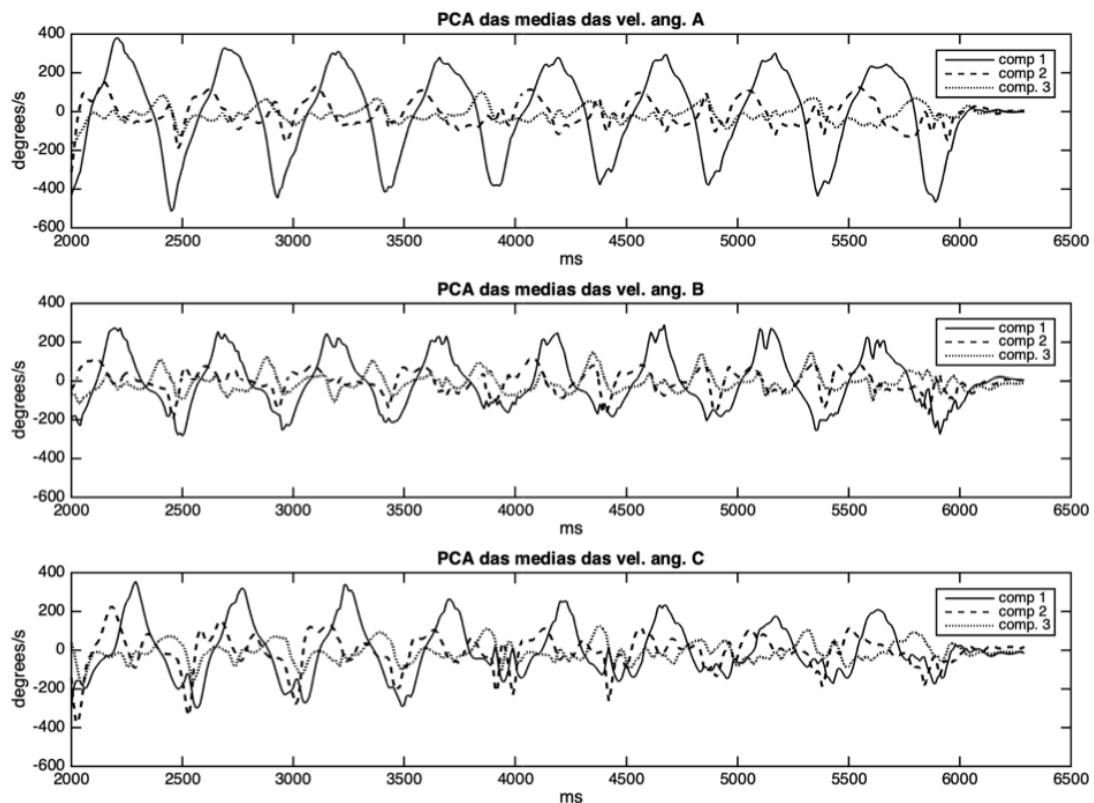
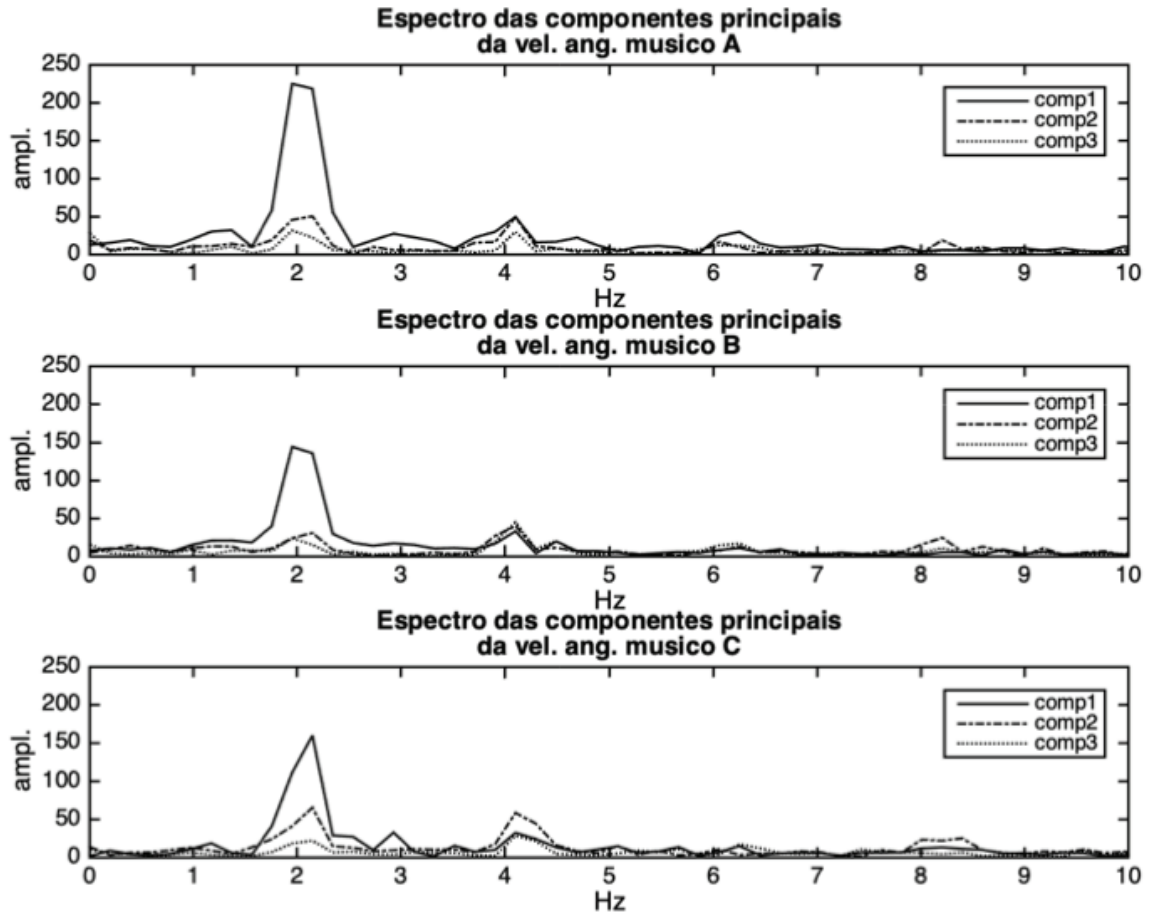


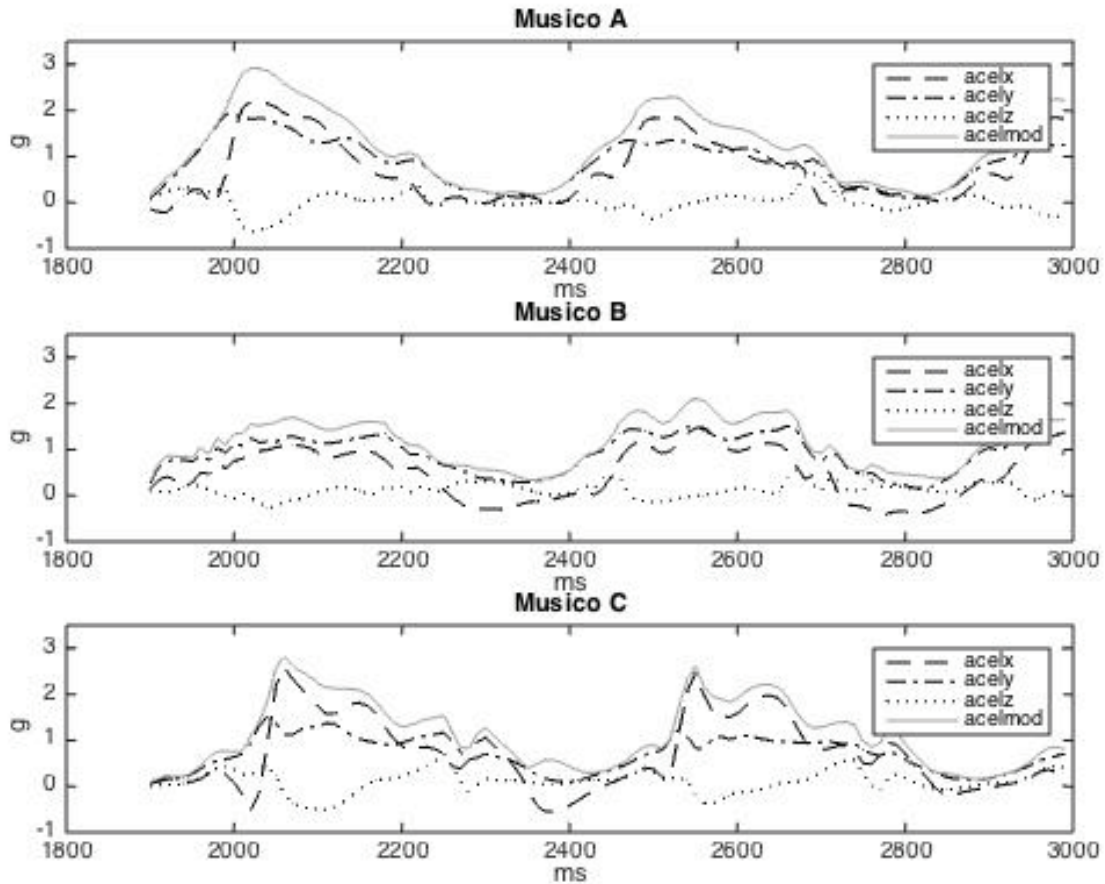
Figura 25 – Espectro dos componentes principais da velocidade angular dos três músicos. Considerando uma unidade de gesto como uma descida e uma subida, com metrônomo a 60 bpm, é esperada uma componente significativa na frequência de 2 Hz.



O andamento do trecho executado é de 60 bpm. Com o ritmo em semicolcheias, contendo 4 acordes/gestos por tempo, e considerando que cada unidade de movimento possui uma descida e uma subida da mão (ou duas semicolcheias), é esperado que a frequência do gesto seja de 2 Hz. Todos os músicos apresentam essa frequência bem proeminente no primeiro componente, destacando-se a maior amplitude do músico A. Os demais harmônicos se derivam da já comentada assimetria das curvas ascendentes e descendentes (a título de exemplo, uma onda dente-de-serra, com seus harmônicos com intensidades decrescentes, seria um exemplo extremo dessa assimetria). A forte presença de 4 Hz nos demais componentes indicam a existência de movimentos menores e mais rápidos acompanhando o gesto principal.

Mesmo com os resultados afetados pela influência dinâmica da gravidade, podemos também comparar as curvas de aceleração. A Figura 26 mostra as curvas de aceleração nos três eixos (e também seu módulo) presentes na execução dos quatro primeiros acordes do primeiro *take* de cada músico. Observando os módulos das acelerações, pode-se notar uma curva com menor amplitude do músico B, que certamente se deve à posição de sua mão ao segurar a palheta: enquanto o polegar e o indicador a seguram (como fazem os demais músicos), os demais dedos permanecem esticados e afastados, como se não pertencessem ao órgão responsável pelo toque. Já o músico A, que apresenta as curvas de aceleração  $x$  e  $y$  bem similares, usa toda a mão (com os demais dedos acompanhando a curvatura do indicador) para executar o gesto. A similaridade dessas curvas, que estão associados aos eixos  $z$  e  $x$  de rotação, respectivamente, certamente está correlacionada à grande amplitude do primeiro componente principal das velocidades angulares desse músico (ver Figuras 24 e 25). O músico C segura a palheta de modo semelhante ao músico B, embora integre bem mais os dedos restantes em seu gesto, o que explica uma amplitude mais acentuada do módulo de sua curva de aceleração. Devido à diferenciação que ele faz entre os planos grave e agudo, suas rotações são menos coerentes entre si, e do mesmo modo as acelerações nos eixos  $x$  e  $y$ . As ondulações presentes em alguns momentos das curvas do músico B (e em menor grau também dos demais músicos), podem estar relacionadas à reação da mão ao choque da palheta com as cordas, que são destacadas com uma “pegada” mais leve.

Figura 26 – Curvas de aceleração nos eixos x, y, z e seu módulo presentes na execução dos quatro primeiros acordes do primeiro take de cada música.



A seguir apresentaremos o estudo do mesmo ritmo com o sensor programado para o modo de fusão dos dados, que elimina a influência da gravidade nos dados de aceleração linear.

### 2.1.2. Sensor com a fusão de dados do *IMU*

Este experimento foi realizado com uma série de limitações impostas pela pandemia do Covid-19. Sem a possibilidade de gravação com diversos músicos, que nos traria uma amostragem maior sobre a utilização do *strumming*, conseguimos com o sistema portátil realizar a coleta de dados com um performer experiente no uso da técnica.



A metodologia utilizada aqui segue os padrões relatados no experimento anterior. Nessa seção apresentaremos rapidamente a configuração dos equipamentos para o experimento e os resultados obtidos.

### 2.1.2.1. Setup e gravações

O *setup* para essa coleta de dados contou com o *IMU* em modo fusão de dados de aceleração e giroscópio, *GuiarT* (utilizando um sistema portátil) e registro de vídeo em *FULL HD 1920 x 1080 pixels 240 fps*. Utilizamos fita dupla face para posicionamento do sensor na mão do performer, como mostrado na Figura 27. Apesar dessa posição ser diferente da usada no Ritmo 1 sem a fusão de dados do *IMU*, isto não interfere nas análises, pois o plano *xy* se mantém na mesma orientação em relação à mão nos dois casos, alterando apenas a origem e, portanto, as direções das rotações. Como anteriormente, um pedal marca o início do registro dos dados gerados pelo *MetaMotionR* e *GuiarT* no *Max/MSP*, enquanto o vídeo possui um controle independente, sincronizado posteriormente através do áudio. Foram realizados três *takes* por um mesmo performer.<sup>10</sup>

Figura 27 – Sensor posicionado na mão do performer para registro do ritmo 1 programado para a fusão de dados.



Os registros coletados seguem o padrão descrito na seção 2.1.1.3 e o teste de latência para o sensor na configuração de fusão de dados já foi relatado na seção 1.3.4.1. Com os dados de apenas um músico, buscaremos analisar a regularidade e estratégia

---

<sup>10</sup> As gravações podem ser vistas em <https://vimeo.com/463810986>



gestual, baseadas na interpretação dos *onsets* e sua relação com o pulso, curvas de aceleração linear e velocidades angular e linear (integrada da aceleração).

### 2.1.2.2. Resultados

Iniciaremos esta análise com uma visão geral sobre a performance do ritmo proposto, identificando as estratégias em três takes diferentes em relação às informações sonoras e gestuais.

O ritmo é constituído de gestos sucessivos que se alternam nas direções descendente e ascendente, em semicolcheias, ocorrendo ataque nas cordas em todos eles. A Tabela 4 mostra a quantidade de notas em cada *take*.

Tabela 4 – Comparativo entre as performances em relação a quantidade de notas.

	Quantidade de notas
<b>Take 1</b>	69
<b>Take 2</b>	64
<b>Take 3</b>	72
<b>Média e desvio padrão</b>	68,33 ± 4,04

De acordo com a partitura, eram esperadas 94 notas, mas, como já dito em relação ao experimento anterior, esse número deve ser relativizado pela falta de orientação exata para a realização da técnica. No ritmo estudado, que não apresenta acordes com ornamentos diferentes, não ocorre o decaimento considerável do som (podendo assim ocorrer algum mascaramento ou a ilusão da ocorrência de novos ataques), colaborando para essa variação de quantidade de notas observada.

Iremos comparar o espalhamento temporal (*spread*) entre *onsets* (primeiro e último) de cada gesto, uma vez que representam a mesma figura rítmica (semicolcheia). Com os dados da Tabela 5, percebemos uma considerável diferença entre os sentidos dos *takes* 1 e 2: *spread* maior na descendente e menor na ascendente. Isso não se relaciona diretamente com a velocidade do gesto, como veremos adiante, mas com a

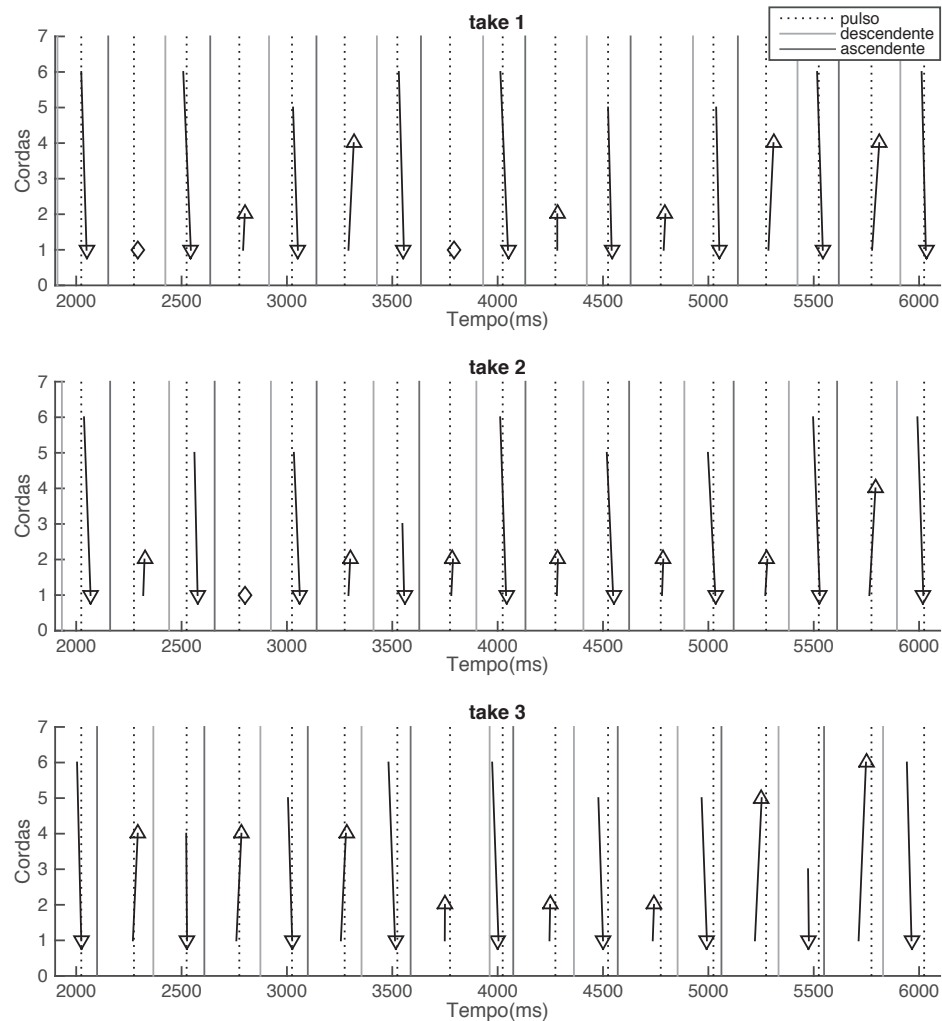
quantidade de cordas tocadas em cada um. Já o *take 3* apresenta um equilíbrio maior, tanto no *spread* quanto na quantidade de notas.

Tabela 5 – Média e desvio padrão do espalhamento temporal e da quantidade de notas nos gestos descendentes e ascendentes.

	Descendente		Ascendente	
	<i>Spread</i> (ms)	Quantidade de Notas	<i>Spread</i> (ms)	Quantidade de Notas
<b>Take 1</b>	25 ± 7,42	5,67 ± 0,5	13,37 ± 13,17	2,25 ± 1,04
<b>Take 2</b>	25,89 ± 8,74	5,22 ± 0,97	9 ± 8,6	2,12 ± 0,83
<b>Take 3</b>	18,11 ± 10,95	4,78 ± 1,48	21,12 ± 16,28	3,62 ± 1,41

Em complemento à análise da Tabela 5, a Figura 28 apresenta os três takes com os ataques representados por setas que indicam a quantidade de cordas utilizadas e a direção de cada gesto, o pulso subdividido em semicolcheias e pontos de segmentação dos gestos. Podemos observar que os números de notas nos gestos ascendentes em geral são menores do que no sentido contrário e que os inícios dos ataques tendem a ocorrer após o pulso, indicando uma estratégia interpretativa do performer. Vale destacar também o aumento de notas na ascendente no final do trecho, que indica uma variação de dinâmica, um *crescendo* para a resolução da frase, outra estratégia interpretativa. No geral o músico apresenta uma boa regularidade na quantidade de notas entre os takes, salvo no primeiro compasso do take 3, onde ocorrem mais notas no sentido ascendente do que nos outros takes. É clara a diferença entre os gestos descendentes e ascendentes, o primeiro com mais notas e o segundo menos denso, às vezes com uma única nota. Esse fato pode estar relacionado com a ação da gravidade, que no caso está a favor do gesto descendente, e ao ângulo da palheta e da mão do performer em relação as cordas do violão.

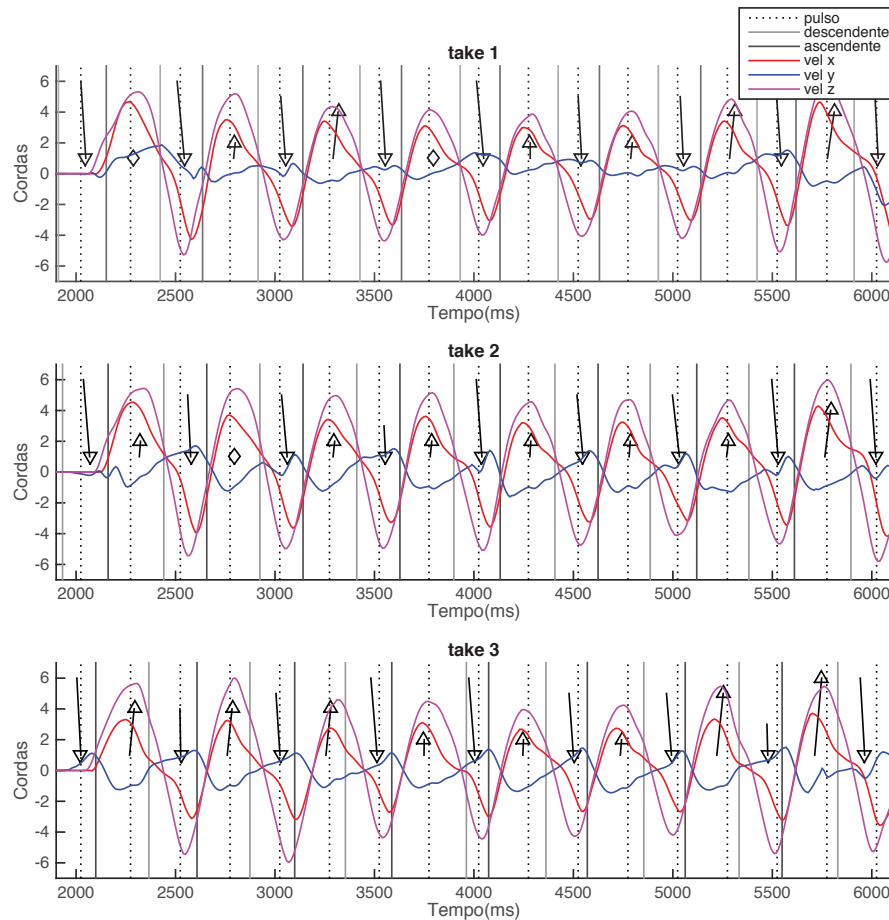
Figura 28 – Gráfico com os ataques representados por setas, pulso subdividido em semicolcheias e direção dos gestos nos 3 takes do Ritmo 1. As barras verticais cinza e preta indicam o início dos gestos (descendente e ascendente).



Focaremos agora nas curvas de velocidade linear - sem a influência da gravidade e calculadas pela integração das acelerações rotacionadas - e de velocidade angular. A Figura 29 mostra as velocidades lineares nos eixos x, y e z dos três takes do Ritmo 1. Os dados extraídos do sensor de movimento são normalizados<sup>11</sup> para melhor visualização das curvas.

<sup>11</sup> As curvas dos 3 eixos são normalizadas conjuntamente, para não se perder a proporção entre elas. Em seguida, realizamos uma multiplicação por 6, para facilitar a visualização quanto plotadas juntamente com as flechas das cordas utilizadas em cada acorde.

Figura 29 – Velocidades lineares nos eixos x, y e z dos três takes do Ritmo 1.

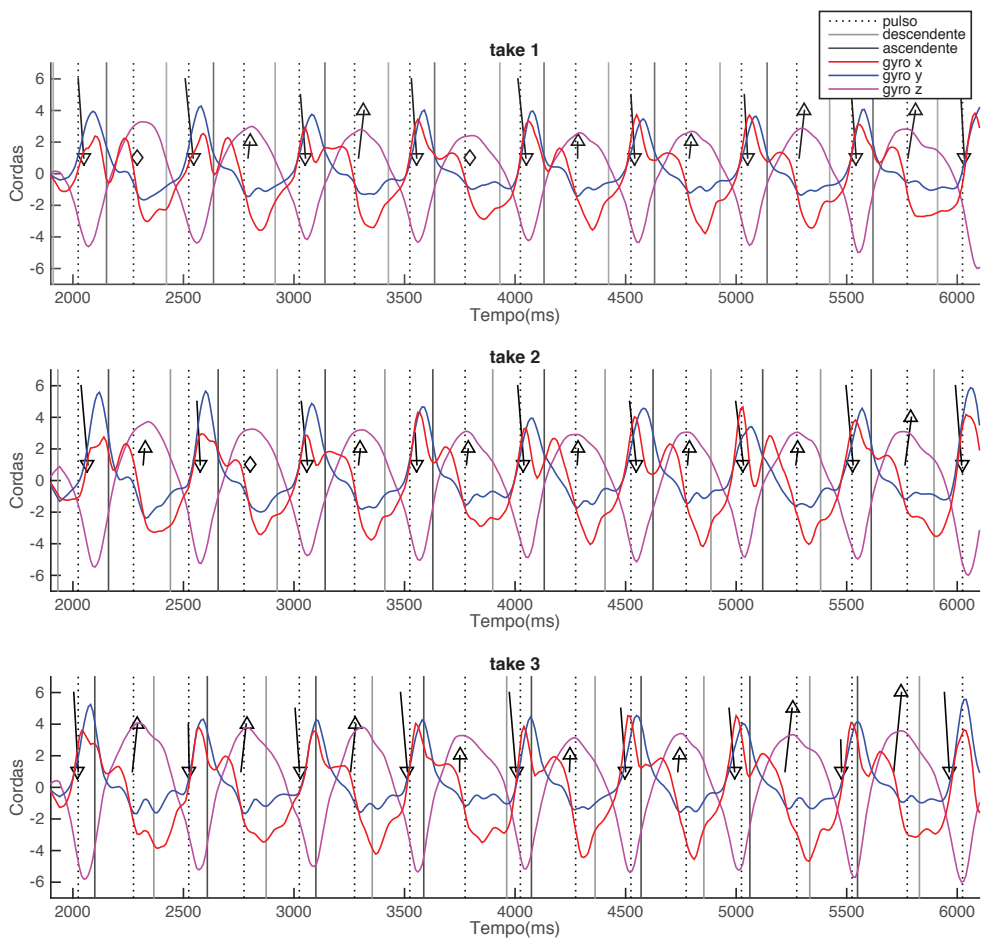


De modo geral, as curvas de velocidade linear nos três eixos possuem grande regularidade nos três takes. As velocidades em X e Z apresentam maiores variações de intensidade do que no eixo Y, indicando que os gestos envolvidos no *strumming* ocorrem principalmente em um plano gerado pelos dois eixos. Retomaremos esse raciocínio mais à frente. Outro fato importante a se destacar é que a quantidade de cordas/notas utilizadas em cada acorde não está diretamente ligada ao gesto, pois em curvas muito parecidas a variação de notas é significativa. Tomemos como exemplo o quarto e sétimo acordes do trecho. O quarto acorde, na ascendente, apresenta curvas bem próximas, porém com duas notas no take 1, uma nota no take 2 e quatro notas no take 3. De modo similar, o sétimo acorde, na descendente, contém seis notas nos takes 1 e 3 e três notas

no take 2. No âmbito gestual, as curvas nos três eixos se cruzam próximas ou coincidentes aos pontos de mudança de direção do gesto (da ascendente para descendente) e tanto X e Z se cruzam com Y no sentido contrário (descendente para ascendente). Esse fato nos indica uma possibilidade de segmentação dos gestos desse tipo de *strumming* através dos dados gerados pelo *IMU*, hipótese que deve ser verificada com uma amostragem maior.

A Figura 30 mostra as velocidades angulares.

Figura 30 – Velocidades angulares eixos x, y e z dos três takes do Ritmo 1.



Diferente da velocidade linear, a velocidade angular apresenta valores maiores e equilibrados entre si em todos os eixos nos lóbulos positivos, e intensidade menor no eixo Y nos lóbulos negativos. O destaque ocorre em Z, que mantém um envelope quase senoidal no trecho, indicando um uso bem regular de giro da mão nesse eixo. Pode-se

também notar que os máximos e mínimos da curva Z se localizam próximos ao momento de toque nas cordas, o que corrobora a regularidade da rotação neste eixo.

### **2.1.3. Discussão**

O primeiro estudo sobre o Ritmo 1 nos mostrou que os dados gerados pelas diferentes fontes utilizadas (número de notas, spread temporal dos acordes, duração total dos gestos, PCA das velocidades angulares, curvas de aceleração e velocidades, video) foram capazes de apontar diferenças significativas entres os músicos, além de permitir apontar algumas correlações entre as diferentes modalidades. O segundo estudo nos permitiu uma análise mais aprofundada sobre as curvas de aceleração linear e angular, e em conjunto com os demais dados foi possível apontar as estratégias gestuais do performer nos três *takes* analisados, percebendo-se uma interessante regularidade entre eles.

Na sequência, apresentaremos o estudo do Ritmo 2, mais complexo, com acordes ornamentados e divisão de planos grave e agudo.

## 2.2. Ritmo 2

Essa próxima análise segue as mesmas especificações técnicas do estudo anterior. Apresentaremos sucintamente o ritmo a ser estudado e em seguida os resultados.

O segundo ritmo desse estudo pode ser observado na Figura 31, contendo duas trocas de acorde, com um arpejo no segundo tempo de cada compasso.

Figura 31 – Partitura do Ritmo 2.

The musical score for Ritmo 2 is written in 4/4 time with a tempo marking of ♩ = 70. The key signature is one flat (Am). The score consists of 12 measures. The first measure is an Am chord. The second measure features an arpeggiated accompaniment pattern with a downward arrow below it. The third measure is a G chord. The fourth measure features an arpeggiated accompaniment pattern with a downward arrow below it. The fifth measure is an Am chord. The sixth measure features an arpeggiated accompaniment pattern with a downward arrow below it. The seventh measure is a G chord. The eighth measure features an arpeggiated accompaniment pattern with a downward arrow below it. The ninth measure is an Am chord. The tenth measure features an arpeggiated accompaniment pattern with a downward arrow below it. The eleventh measure is a G chord. The twelfth measure features an arpeggiated accompaniment pattern with a downward arrow below it. The score ends with a double bar line.

### 2.2.1. Setup e gravações

O *setup* para as gravações do ritmo 2 seguem as mesmas especificações apresentadas na seção 2.1.2.1. Foram realizados três *takes* desse ritmo por um mesmo músico.

### 2.2.2. Resultados

Iniciaremos as análises com os dados extraídos do áudio e vídeo e acrescentando as informações do *IMU* posteriormente. O ritmo 2 possui notas com durações mais longas do que o ritmo 1, com um acorde arpejado. Apesar não indicado na partitura, nos *takes* analisados observamos uma clara separação dos planos grave/agudo, como será mostrado na sequência. A Tabela 6 mostra a quantidade de notas em cada *take*.

Tabela 6 – Comparativo entre as performances do ritmo 2 em relação a quantidade de notas.

	<b>Quantidade de Notas</b>
<b><i>Take 1</i></b>	64
<b><i>Take 2</i></b>	64
<b><i>Take 3</i></b>	63
<b>Média e desvio padrão</b>	63,66 ± 0,58

De acordo com a partitura, eram esperadas 96 notas para o ritmo 2. Apesar de ser um ritmo mais complexo, o músico obteve uma regularidade maior no ritmo 2 que no ritmo 1 (ver Tabela 4). Isso pode ser relacionado com o fato de esse ritmo possuir separação de planos grave/agudo, como interpretação do performer, e um acorde arpejado. Nessas situações, a precisão (todas as notas do arpejo e a diferenciação de planos grave e agudo, ambos de forma clara) se torna um fator importante para a validação estética da performance. No primeiro ritmo, que não apresenta acordes com ornamentos diferentes, não ocorre o decaimento considerável do som (podendo assim ocorrer algum mascaramento ou a ilusão da ocorrência de novos ataques) como no segundo, colaborando para essa variação de quantidade de notas observada.

Iremos analisar o *spread* de gestos que utilizam todas as cordas no acorde (cinco para Am e seis para G). Aqui temos duas situações: acordes normais, presentes no primeiro tempo de cada compasso, e arpejados, no segundo tempo de cada compasso. A Tabela 7 mostra os intervalos e média total desses acordes nos três *takes*.



Tabela 7 – Spread temporal de acordes normais e arpejados nos três takes do ritmo 2, com média e desvio padrão geral.

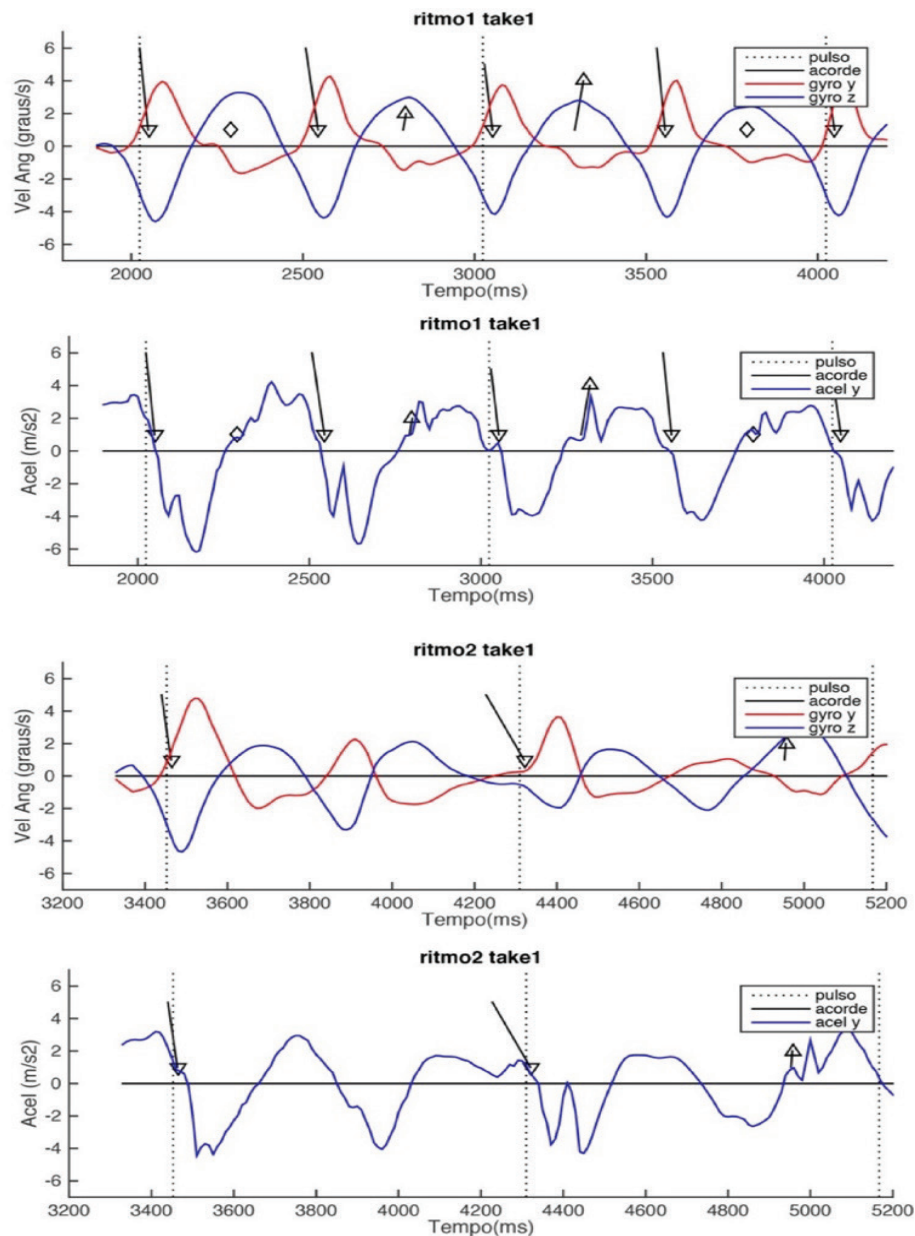
		Acordes normais (ms)	Acordes arpejados (ms)
<b>Take 1</b>	<b>Am</b>	23	92
	<b>G</b>	22	88
<b>Take 2</b>	<b>Am</b>	18	59
	<b>G</b>	24	89
<b>Take 3</b>	<b>Am</b>	19	72
	<b>G</b>	22	58
<b>Média geral</b>		21,33 ± 2,34	76,33 ± 15,47

Podemos observar a relação de duração entre estes dois acordes. O acorde arpejado, que exige uma técnica mais apurada para execução, é menos regular nos *takes* gravados e dura em média aproximadamente o triplo de tempo de um acorde normal, este já bem regular em relação à duração. A quantidade de cordas utilizadas nos arpejos não possui relação com o *spread*, explicitado no *take 2*: o arpejo que utiliza menos cordas dura mais tempo. Ressaltamos também a janela temporal em que esses acordes acontecem, de 58 ms a 92 ms, permitindo uma variedade de interpretação dentro deste intervalo.

Apesar do Ritmo 2 ser bastante diferente do Ritmo 1 no âmbito sonoro, podemos observar semelhanças no aspecto gestual, como mostrado na Figura 32, compreendendo os dois primeiros tempos do *take 1*, que apresentam curvas similares aos *takes* restantes. Mesmo em momentos de notas longas (no caso, uma semínima no primeiro compasso do Ritmo 2), o performer mantém o gesto subdividido em semicolcheias, como estratégia de manutenção rítmica. As curvas de velocidade angular (y e z) e a aceleração linear em y (eixo mais relacionado com o gesto) possuem 2 Hz por pulso aproximadamente nos dois *takes*. Podemos observar nos dois casos fases semelhantes entre a velocidade angular z e a aceleração linear y, e fases de 180 graus entre a velocidade angulares y e z, além de amplitudes menores nos momentos em que não acontecem ataques, o que aponta para alguma semelhança gestual entre as performances. Apesar de uma

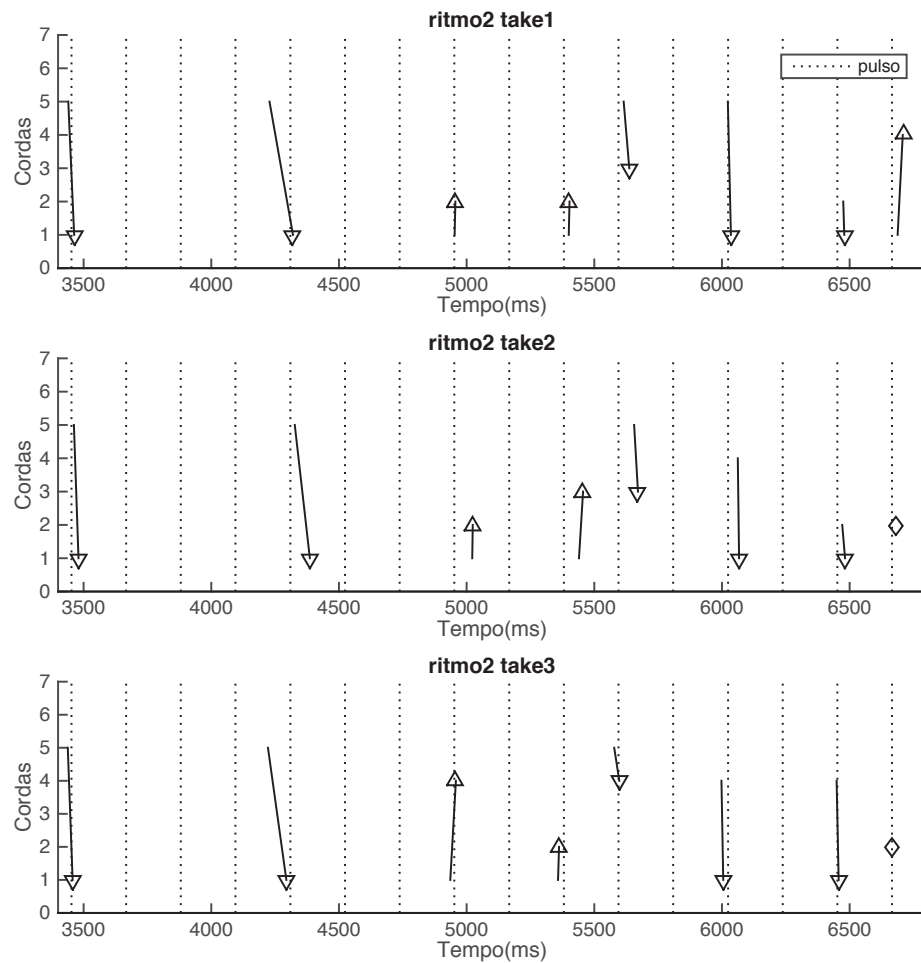
amostragem muito pequena, levantamos a hipótese de que esse gesto contínuo contribua para a precisão rítmica dos ataques. Com a comparação prospectiva dos dados, em seção própria a frente, retomaremos essa discussão.

Figura 32 – Dois primeiros tempos dos ritmos 1 e 2, no take 1. As curvas de velocidade angular (gyro) em y e z e a aceleração linear em y possuem envelopes temporais semelhantes entre os ritmos, com menor intensidade nos gestos sem notas do ritmo 2, indicando movimentação constante da mão como estratégia de manutenção rítmica.



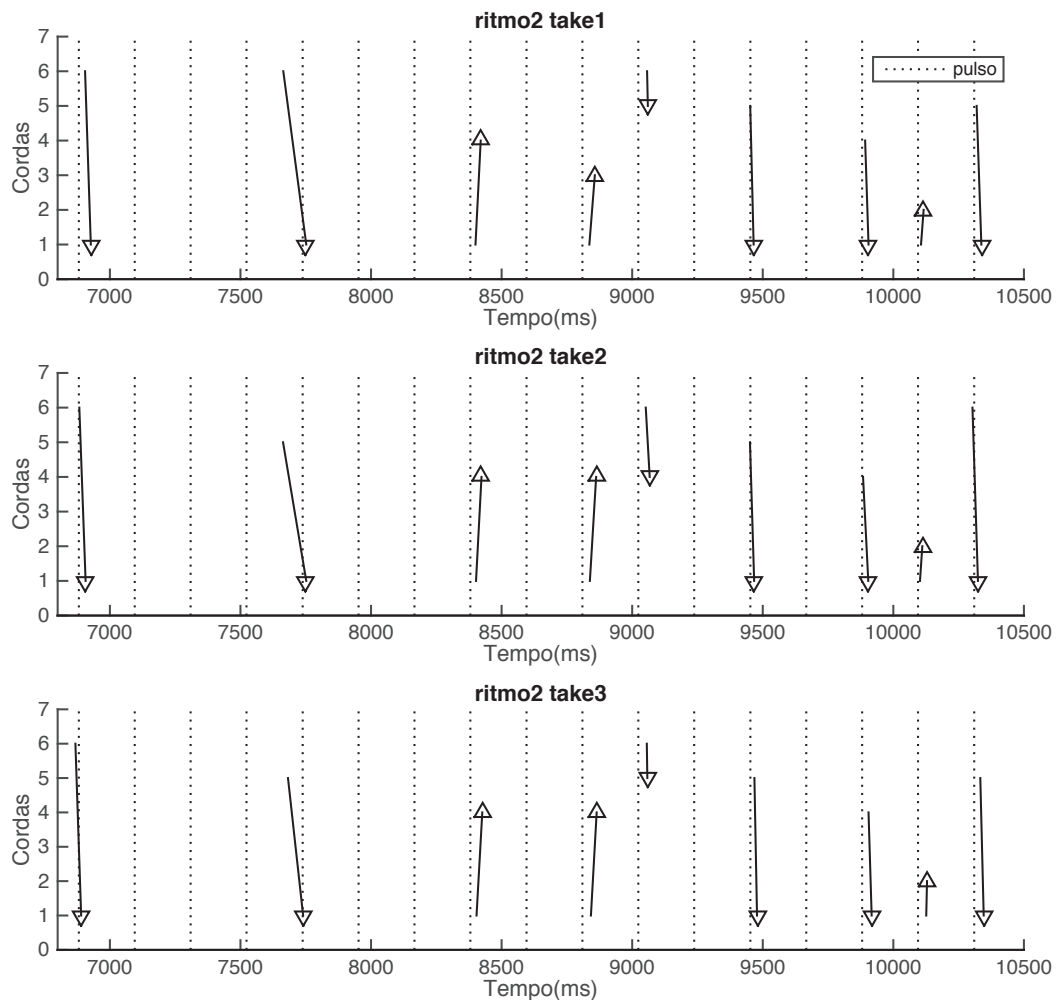
Podemos observar a relação dos ataques com o pulso (subdividido em semicolcheias), no primeiro compasso do trecho, na Figura 33. O *take 2* apresenta um pequeno atraso em relação ao pulso, enquanto os restantes estão mais alinhados. Nos *takes 1* e *3* o arpejo (segundo acorde) tem início praticamente uma fusa antecipada, com o pulso caindo nas cordas mais agudas. Já no *take 2* ocorre o oposto: os últimos ataques ocorrem uma fusa à frente do pulso. Outro ponto interessante de ser observado é a diferenciação dos planos grave/agudo. Os acordes 3 e 4 estão no plano agudo e o acorde 5 no plano grave, bem distintos no *take 1*. Nos demais, apesar da mudança na quantidade de cordas tocadas, essa diferenciação se repete, indicando a estratégia interpretativa do performer para esse ritmo.

Figura 33 – Primeiro compasso dos takes do ritmo 2 com os acordes, representados por flechas, e o pulso subdividido em semicolcheias.



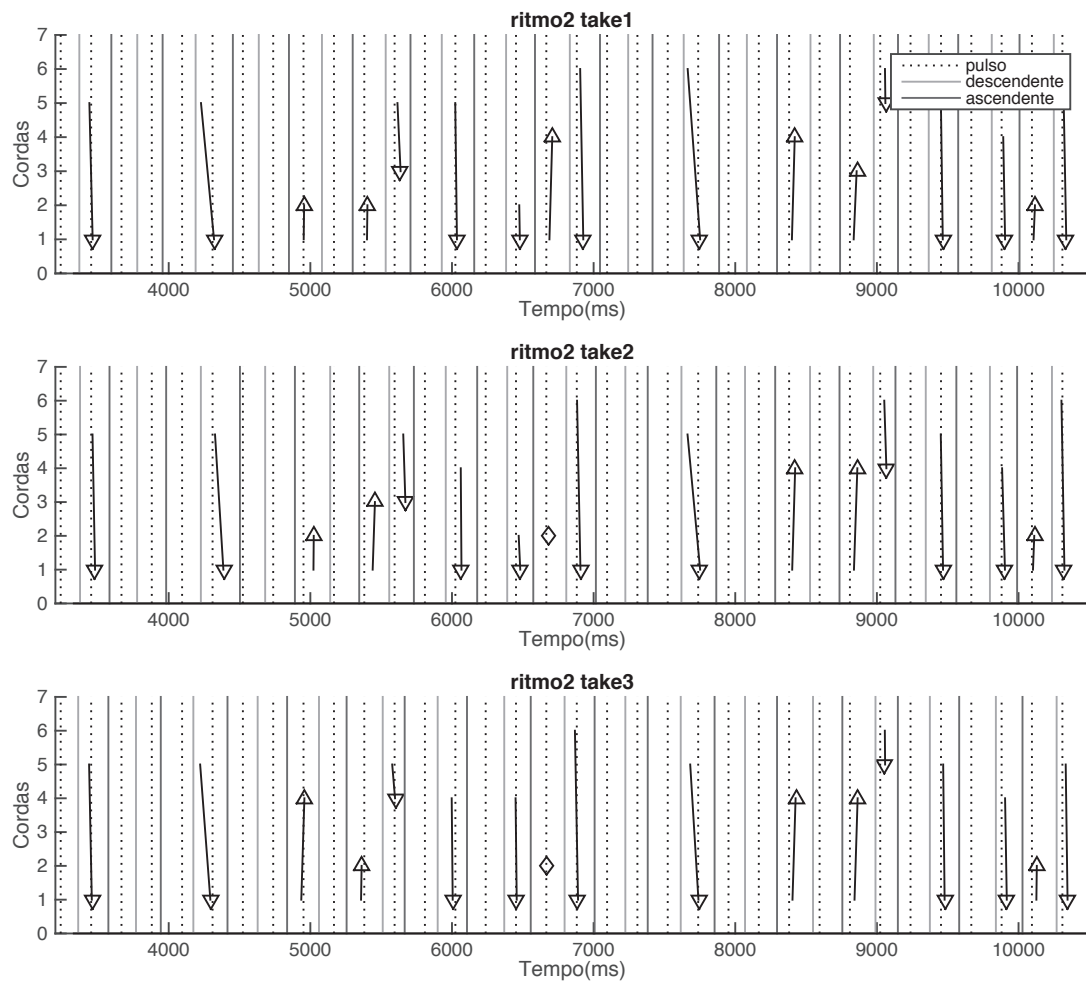
A Figura 34 mostra o segundo compasso. Os arpejos são bem semelhantes em sua relação com o pulso, diferenciando-se apenas na quantidade de cordas. O plano agudo agora utiliza mais cordas do que no compasso anterior, podendo se relacionar com o fato de o acorde de Sol maior compreender todas as cordas do violão e indicando que essa pequena mudança em relação a quais cordas tocar interfere na maneira como o performer abordou essa separação. Tanto no primeiro como no segundo compasso os ataques na descendente (exceto o arpejo) apresentam, de modo geral, um *spread* menor que no sentido oposto em acordes a princípio iguais. Esse gesto é a favor da força de gravidade, considerando a posição do performer, o que pode explicar tal fato.

Figura 34 – Segundo compasso dos takes do ritmo 2 com os acordes e o pulso subdividido em semicolcheias.



A segmentação dos gestos, através dos pontos observados nos vídeos, nos traz uma visão geral de seu comportamento temporal (Figura 35). Existe uma regularidade considerável no intervalo de tempo entre os gestos descendentes e ascendentes, mesmo onde não há acordes, com pulso se mantendo em uma região mediana na duração, exceto no primeiro compasso do *take 2*, onde ocorre um atraso dos ataques. Essa análise indica que o performer mantém a duração dos gestos regular independentemente do ornamento do acorde: normal, arpejado ou na separação dos planos.

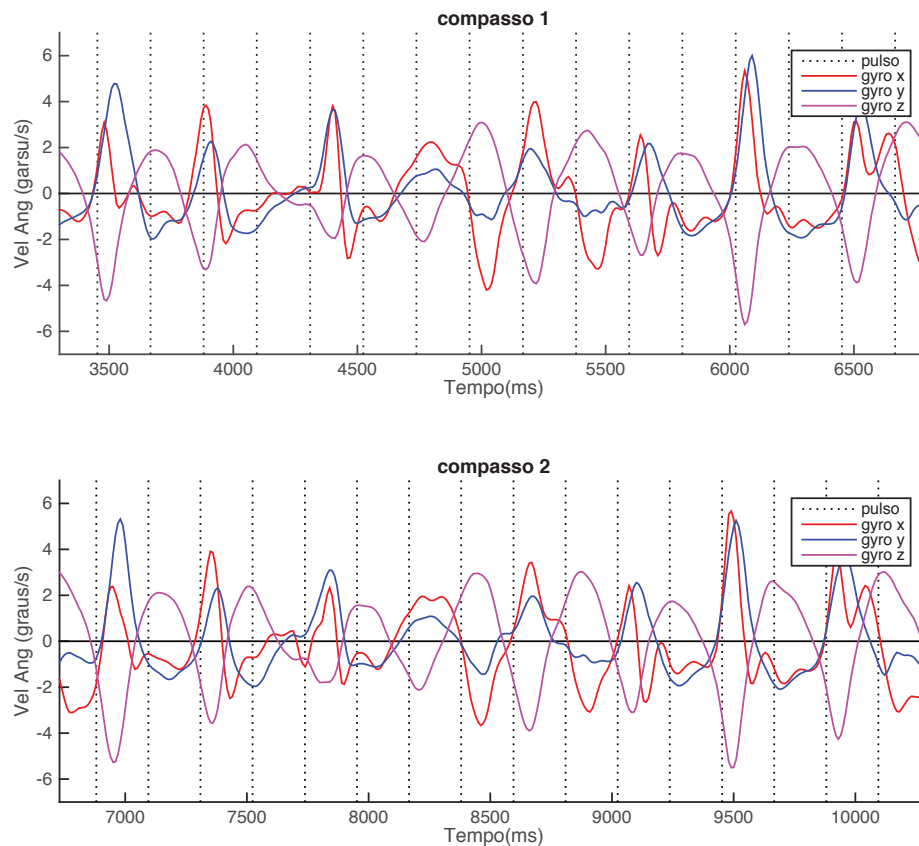
Figura 35 – Segmentação dos gestos do strumming nos três takes do ritmo 2.



Após essa análise com os dados extraídos do *GuiaRT* e do vídeo, incluiremos nas discussões as informações do sensor de movimento, com os valores normalizados para melhor visualização das curvas. O gesto do *strumming* contém uma grande complexidade, com movimentos de translação e rotação que não estão dissociados entre si em uma performance. Uma rotação no eixo z, por exemplo, implica uma aceleração linear considerável em x, já que o eixo de rotação não é centralizado no sensor, e sim no pulso da mão do performer, movimentando-o também em outras direções (ARMONDES; FREIRE, 2019). Apesar desse fato, buscaremos relacionar as diferentes curvas de forma independente nas mudanças de direção e evolução do gesto.

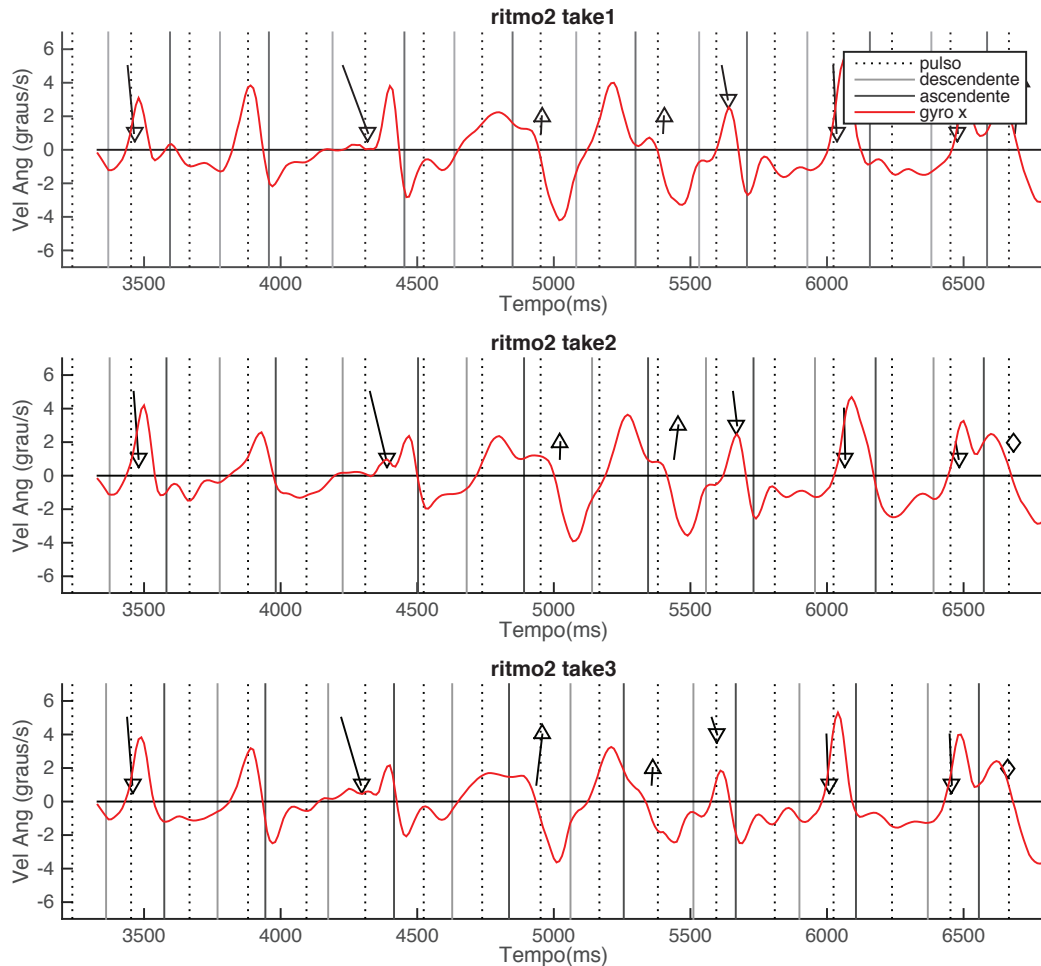
Para uma melhor visualização, apresentaremos as curvas apenas do primeiro compasso de cada *take*, visto que o ritmo se repete no segundo compasso e elas são similares, com uma leve diminuição dos valores no acorde menos denso. Esse fato demonstra grande regularidade gestual do performer, independente do acorde em questão. A Figura 36 mostra um exemplo da similaridade dos compassos com as curvas de velocidade angular em x, y e z.

Figura 36 – Comparação das curvas de velocidade angular (gyro) dos dois compassos do ritmo 2. Apesar de mudança de acorde, as curvas apresentam comportamento semelhante.



Como podemos observar na Figura 37, a velocidade angular no eixo x apresenta valores maiores nos gestos descendentes (lóbulo positivo), indicando uma contribuição maior dessa rotação do que no sentido oposto. Esse cenário se modifica nos acordes do plano agudo (gestos ascendentes), quando há um aumento de módulo no lóbulo negativo, ou seja, para mirar nas cordas mais agudas o performer utiliza um giro maior nesse eixo. Os picos de velocidade ocorrem, no geral, após o último ataque nas cordas, e durante o arpejo praticamente não há velocidade, apontando que a estratégia gestual desse tipo de ataque não utiliza rotação no eixo. Nos gestos sem acordes o comportamento geral se mantém.

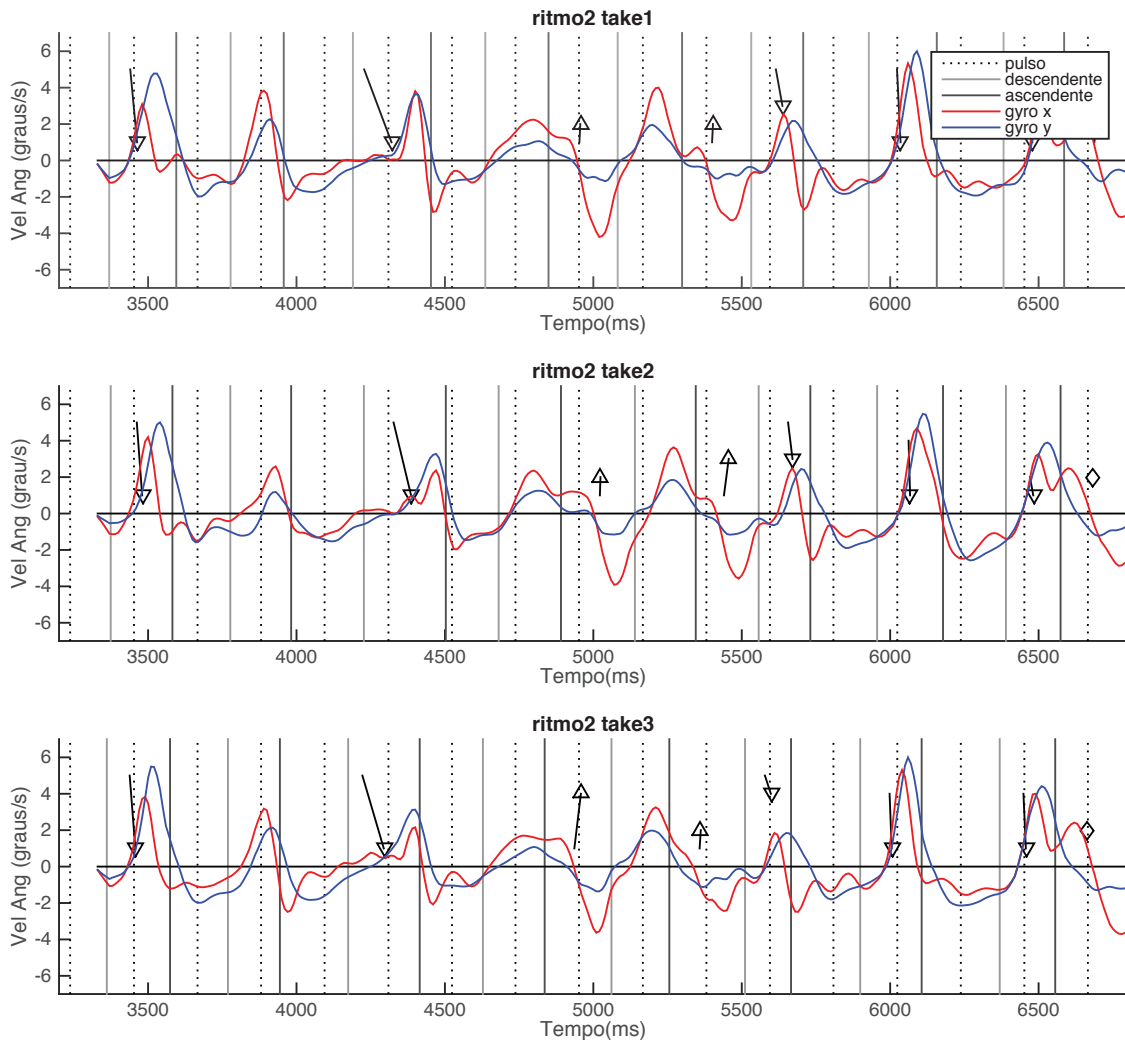
Figura 37 – Curvas de velocidade angular, eixo x do sensor, combinadas com acordes, pontos de segmentação dos gestos e pulso no primeiro compasso de cada take.



A velocidade angular no eixo y tem semelhanças com a o eixo x, sugerindo que as observações anteriores valem também para esse eixo. A principal diferença ocorre na amplitude dos lóbulos negativos nos gestos ascendentes, onde o giro em y não apresenta contribuição como em x, visto na Figura 38. As relações dessas duas curvas com os pontos de mudança de direção não indicam claramente um comportamento regular, dificultando seu uso na segmentação dos gestos.

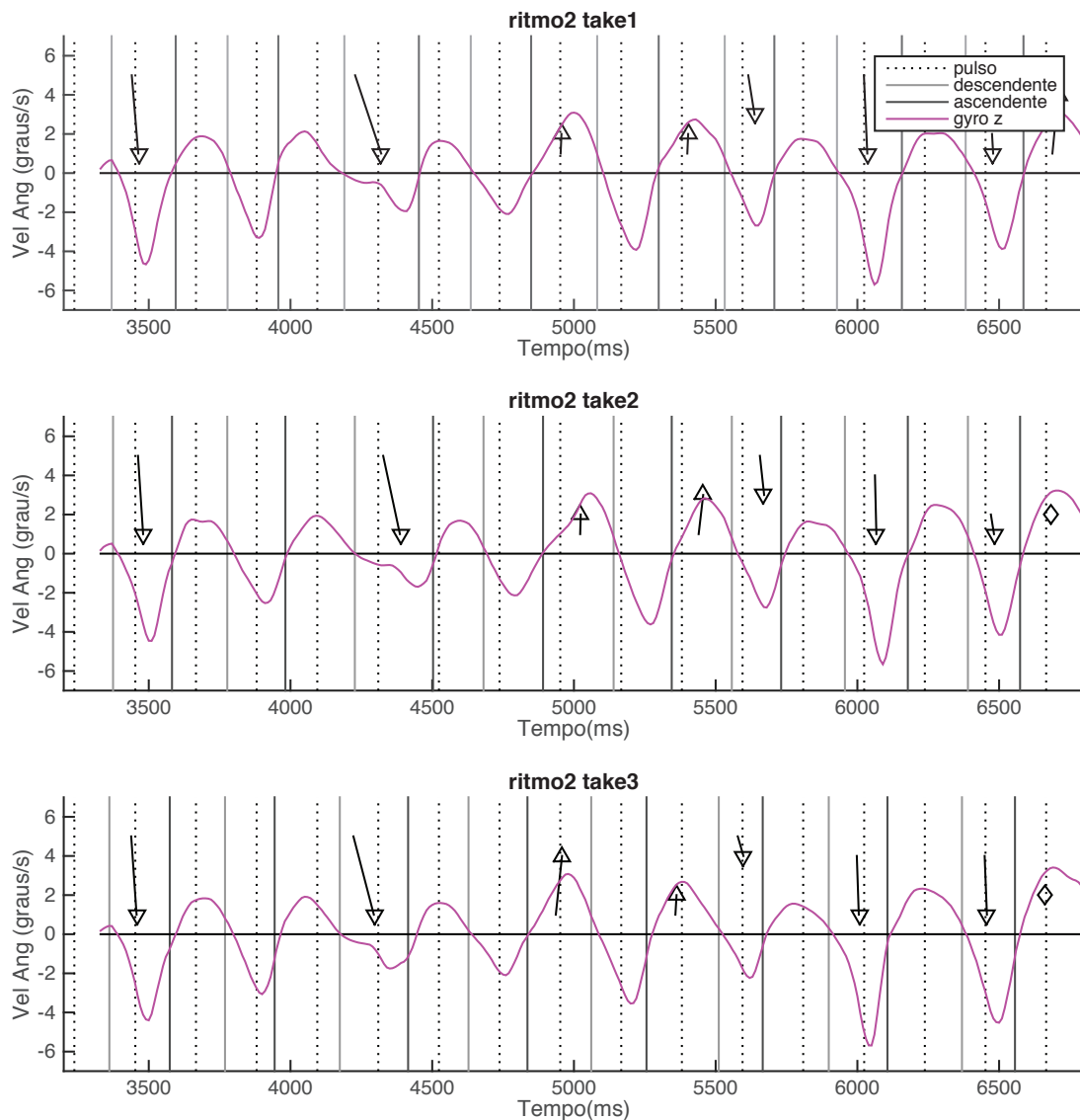


Figura 38 – Curvas de velocidade angular nos eixos x e y nos três takes, primeiro compasso.



A velocidade angular no eixo z possui uma relação direta com a segmentação dos gestos: onde ocorre a mudança de sentido do giro seu valor é zero ou próximo de zero (Figura 39). Esse alinhamento indica o uso dessas informações como estratégia para segmentação dos gestos sem o uso do vídeo, ao menos no presente caso. As amplitudes no sentido descendente são maiores que na ascendente, assim como foi apontado nos outros eixos.

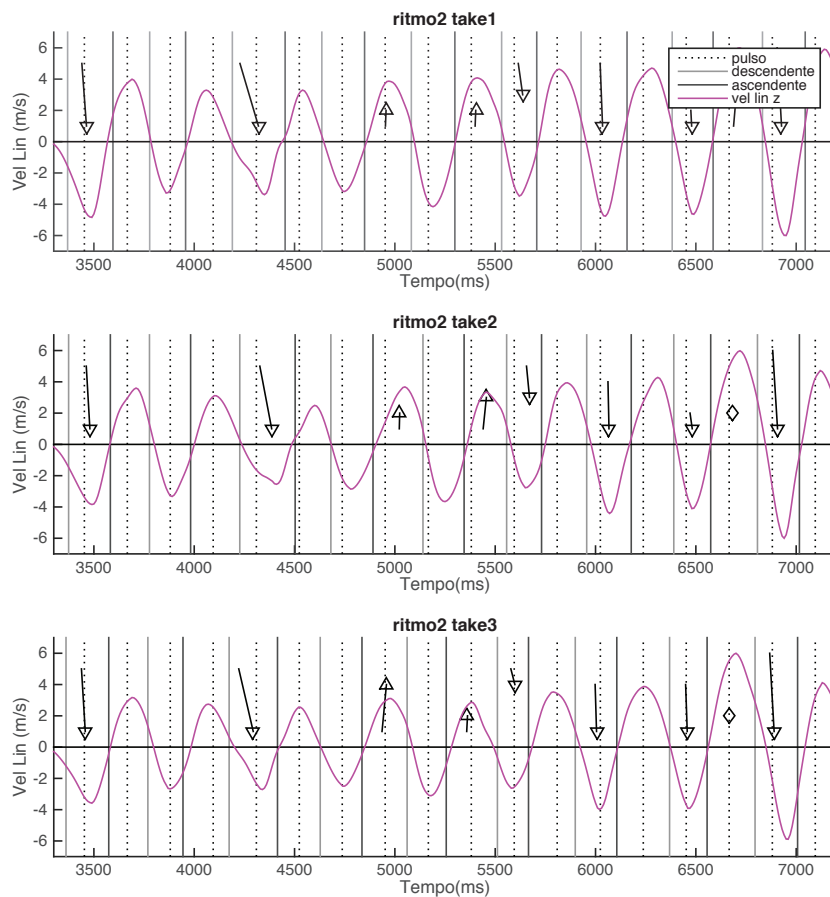
Figura 39 – Velocidade angular em z nos três takes.



Desse modo podemos observar uma característica gestual marcante do performer: nos ataques descendentes utiliza bastante o giro da mão, nos três eixos, em relação aos ascendentes, exceto na diferenciação dos planos e no arpejo, quando o foco em um grupo de cordas menor ou um espalhamento temporal maior traz mudanças à regularidade dos gestos. Para complementar a análise iremos incluir as velocidades lineares, obtidas através da integração dos dados de aceleração linear.

Entre as velocidades lineares<sup>12</sup>, calculadas com a rotação prévia das curvas, a do eixo z (direção vertical, neste caso) é a mais forte e possui alinhamento com os pontos de segmentação dos gestos, assim como na velocidade angular, mais uma vez indicando esse eixo como foco para segmentação dos gestos através dos dados gerados pelo *IMU*. Como podemos ver na Figura 40, as curvas mostram um quase aspecto senoidal, com lóbulos bem equilibrados, indicando o uso desse eixo tanto no gesto descendente quanto no ascendente. As amplitudes aumentam gradativamente até o ponto de troca de acorde, trecho com gestos que realizam três ataques sucessivos em semicolcheias, mostrado uma relação com a cadência rítmica. Nas outras regiões, onde acontecem gestos sem ataque, esses valores são menores.

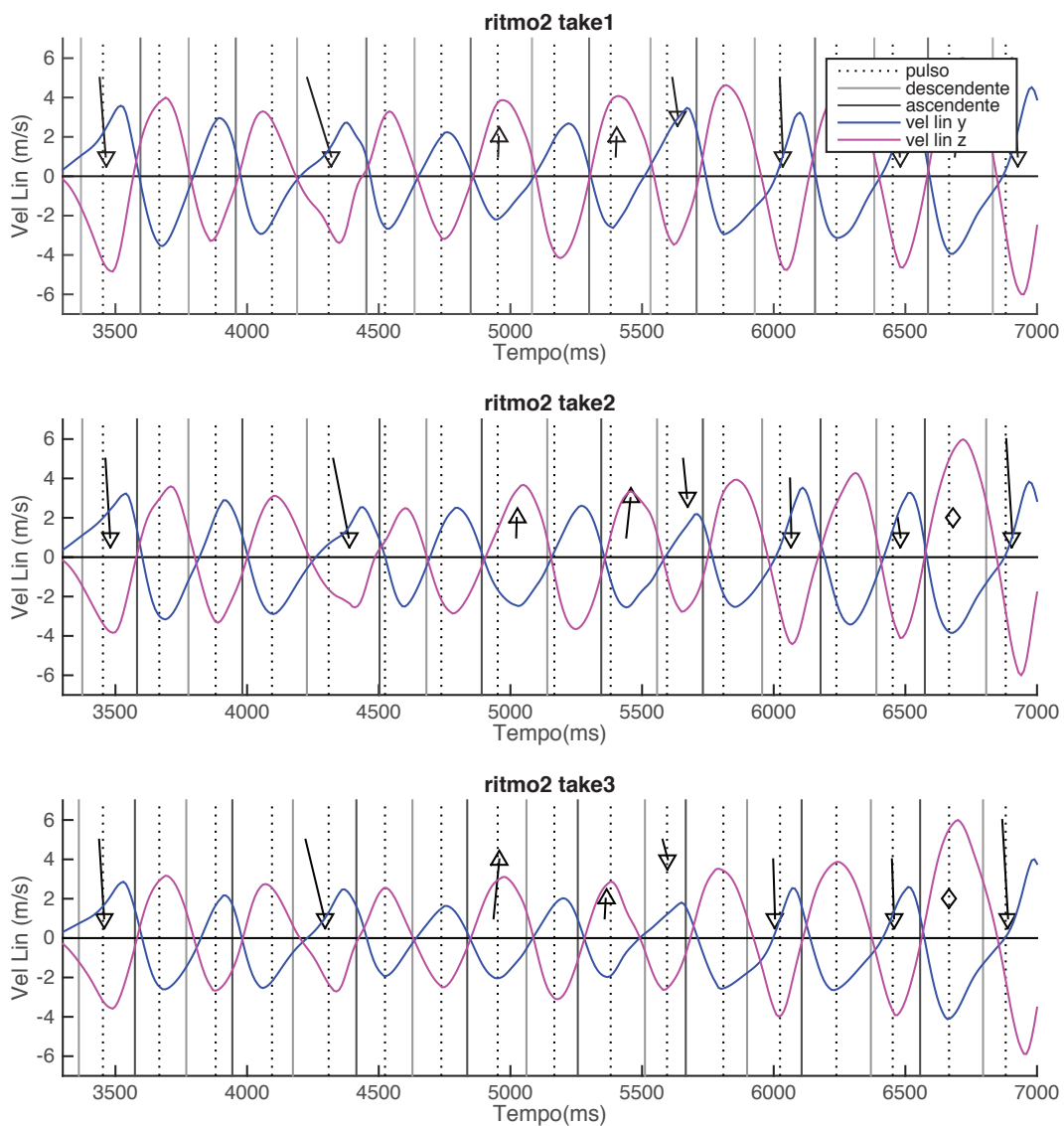
Figura 40 – Velocidade linear do eixo z nos três takes do ritmo 2, primeiro compasso. Os pontos de velocidade nula apresentam bom alinhamento com os pontos de segmentação dos gestos.



<sup>12</sup> Calculadas pela integração das acelerações rotacionadas segundo os ângulos de Euler do próprio *IMU*.

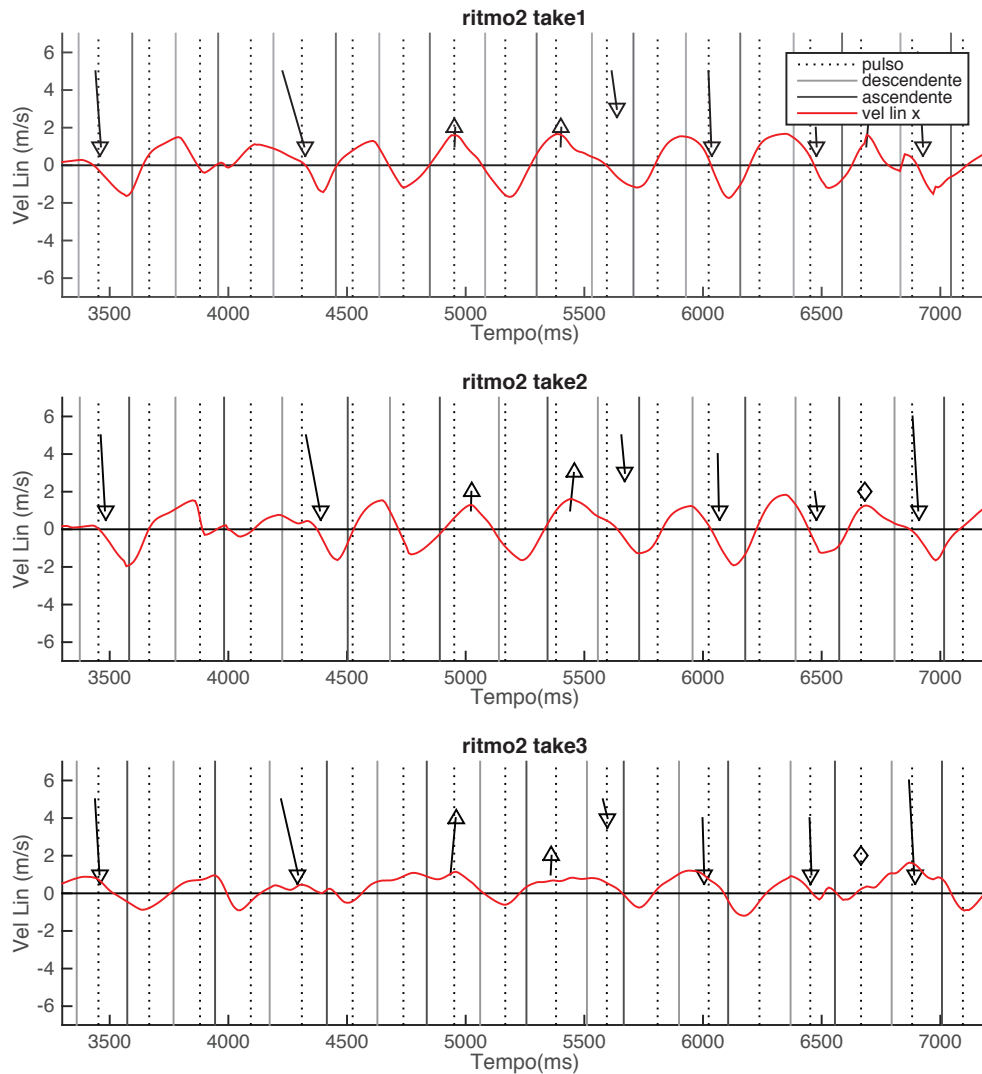
A velocidade linear em y apresenta diversas semelhanças com z, dentre elas a coincidência com os pontos de segmentação do gesto (diferente da velocidade angular em y), a regularidade dos lóbulos em um período e o aumento gradativo até o ponto de mudança de acorde. Ou seja, o performer utiliza os dois eixos de forma parecida nas diversas situações do ritmo proposto (Figura 41).

Figura 41 – Velocidade linear em z e y nos três takes do ritmo 2. As curvas em y são constantes por todo o trecho.



Diferente dos outros dois eixos, a velocidade linear em x é bastante irregular (Figura 42), relacionada um pouco com o eixo z em alguns momentos, com uma defasagem que se modifica na evolução do trecho. As curvas de velocidades angulares e lineares observadas anteriormente sem mantém semelhantes nos três *takes*, fato que aqui não ocorre: o *take 3* apresenta valores bem menores que nos outros dois. Este fato pode ser alterado com uma defasagem (com um valor constante) dos ângulos z calculados pelo IMU, o que será feito a seguir.

Figura 42 – Velocidade linear no eixo x. A baixa amplitude das curvas indica que a referência inercial da posição do sensor em torno ao eixo z ("norte local") durante a gravação fez com que a movimentação no plano horizontal se alinhasse mais com o eixo y.



Por último, fizemos também uma análise da covariância entre os movimentos de translação e rotação. Antes da integração das acelerações, aplicamos uma defasagem de 65 graus nos ângulos z, de modo que as acelerações rotacionadas nos eixos x e y fiquem na mesma faixa de variação. Isto não altera a translação, que continua a ser planar (a contribuição dos dois primeiros componentes principais é maior que 97%), mas torna a integração nos três eixos mais efetiva.

A tabela 8 mostra os resultados da covariância entre as velocidades angulares e as velocidades lineares entre todos os eixos para duas gravações do Ritmo 2.

Tabela 8 – Covariância entre velocidades angulares e lineares entre os eixos x, y e z, takes 1 e 2 do ritmo 2.

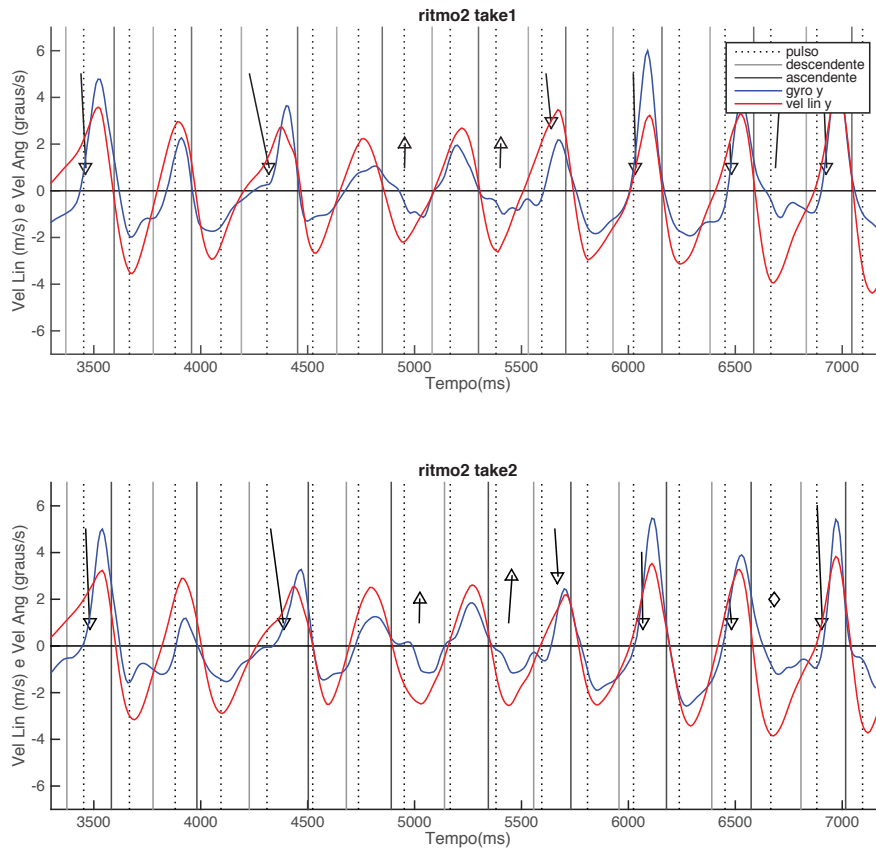
<i>Take 1</i>						
vel lin \ vel ang	x		y		z	
	covar	lag	covar	lag	covar	lag
x	0,508	0	0,549	0	0,468	-12
y	0,745	2	0,807	0	0,568	-12
z	0,276	-12	0,295	12	0,907	1
<i>Take 2</i>						
vel lin \ vel ang	x		y		z	
	covar	lag	covar	lag	covar	lag
x	0,642	1	0,648	0	0,434	-12
y	0,756	2	0,807	0	0,540	-12
z	0,245	-12	0,310	12	0,913	1

Iremos analisar as maiores covariâncias, com valores acima de 0,7. Na prática, essas covariâncias mostram o grau de integração do performer entre o giro do pulso (velocidade angular) e o gesto linear (ou quase) do conjunto pulso e antebraço

(velocidade linear). A maior covariância se dá entre a velocidade angular em z e velocidade linear em z, com valores bem aproximados nos dois *takes* (0,907 no *take* 1 e 0,913 no *take* 2). O *lag* de valor 1 corresponde a um deslocamento de um *frame* pela integração da aceleração. Esses valores confirmam a similaridade que pode ser observada nas Figuras 43 e 44.

A velocidade angular em y e velocidade linear em y também possuem grande covariância, com valor igual de 0,807 nos dois *takes*, sem *lag*. A diferença aqui se destaca no gesto ascendente, onde a velocidade linear é mais “lisa”, similar a uma onda senoidal, enquanto a velocidade angular demonstra pequenas variações, principalmente na fase de finalização do gesto, depois dos ataques (Figura 43). Essa variação do giro pode estar relacionada com o fato de após o contato com as cordas ocorrer um ajuste do ângulo da palheta, tanto pelo término do atrito com as cordas (alguma forma de energia acumulada liberada) quanto para a preparação do gesto seguinte.

Figura 43 – Velocidade angular e velocidade linear no eixo y, nos takes 1 e 2 do ritmo 2.



Por último temos a velocidade angular em y e linear em x, com 0,745 e 0,756 de covariância para os *takes* 1 e 2, respectivamente, e com lag de 2 *frames* nos dois casos. Como já discutido anteriormente, a velocidade angular por si só gera velocidade linear nos outros eixos, pois o giro não é centralizado no sensor, e sim no pulso do performer. Esse fato pode indicar essa covariância observada.

### **2.2.3. Discussão**

Com essa análise sobre os dados percebemos uma característica marcante do performer: ele mantém o movimento da mão constante, mesmo em situações sem previsão de ataque, com estratégias distintas para os gestos descendentes e ascendentes. A regularidade gestual se traduziu em uma boa regularidade rítmica, com as performances sendo bem similares.

Conseguimos avaliar quais grandezas e eixos contribuem mais para os diferentes gestos utilizados na execução desse ritmo. Embora a translação seja essencialmente coplanar, a rotação da mão se dá nos três eixos e é um fator de grande importância na aplicação da técnica, ao menos para o performer analisado.

Essa investigação, além de avaliar quantitativamente a performance, mostra uma potencial estratégia para segmentação do gesto a partir dos dados gerados pelo *IMU*. Uma futura coleta de dados, com outros performers, indicará a validade ou não dessa abordagem.



### 2.3. Análise prospectiva dos dados

O Ritmo 1 foi explorado tanto na comparação dos dados gerados pelos diferentes sistemas de captura de movimento quanto no experimento próprio. Temos assim, performances de cinco músicos diferentes; um deles fez gravações com o sensor nos dois modos de programação. De todas estas performances, podemos comparar diretamente as velocidades angulares, que não foram afetadas pelas programações distintas usadas para o *IMU*. Iremos, aqui, estender a comparação feita no primeiro experimento com esse ritmo para as demais execuções.

Chamaremos de Exp1 os dados extraídos das execuções do Ritmo 1 com o sensor sem o modo de fusão de dados, realizados por 3 músicos (seção 2.1.1). O Exp2 se refere às execuções do mesmo ritmo realizadas por 2 músicos no estudo comparativo entre MetaMotionR e Qualisys (seção 1.3.4). E finalmente Exp3 indicando as performances de um mesmo músico com o sensor programado no modo de fusão de dados apresentadas na seção 2.1.2. A tabela 9 mostra as médias dos PCAs (e dos autovetores) de cada músico.

Tabela 9 – Valores médios da contribuição do primeiro componente principal das velocidades angulares para a explicação da variância do movimento e valores do autovetor correspondente.

músico		Exp1A	Exp1B	Exp1C	Exp2A	Exp2B	Exp3
%		84,4	73	56,9	87,5	94,2	83,2
autovetor	x	0,35	0,42	0,18	0,93	0,72	-0,53
	y	-0,43	-0,44	-0,46	0,25	0,43	-0,48
	z	0,83	0,79	0,87	0,28	-0,28	0,69

O músico Exp1A é também o músico Exp3. A importância do primeiro componente principal manteve o mesmo peso, embora as contribuições das rotações em cada eixo tenham se modificado. Os valores com sinais trocados para o eixo x se referem ao posicionamento do sensor na mão.

Observando-se apenas a contribuição do primeiro componente, há uma boa separação entre os cinco músicos, havendo apenas alguma aproximação entre o músico

Exp2A e o mesmo músico que fez as gravações do Exp1A do Exp3. Neste caso, a distribuição das contribuições de cada eixo serve como diferencial, já o Exp2A músico apresenta um peso bastante significativo do eixo x, e pouca contribuição dos outros.

É possível também comparar as acelerações lineares (já rotacionadas com base nos ângulos do próprio *IMU*) produzidas pelos músicos dos experimentos 2 e 3 neste mesmo trecho de ritmo simples, e também no ritmo mais complexo (Tabela 10).

Tabela 10 - Valores médios da contribuição do primeiro componente principal das acelerações lineares rotacionadas para a variância do movimento, e valores do eigenvector correspondente.

		ritmo 1			ritmo 2		
músico		Exp2A	Exp2B	Exp3	Exp2A	Exp2B	Exp3
%		80,1	90,5	85,9	88,1	74,8	90,3
autovetor	x	-0,12	-0,37	0,56	0,49	-0,52	0,17
	y	-0,39	-0,43	-0,10	-0,17	-0,02	-0,58
	z	0,91	0,81	0,82	0,85	0,85	0,80

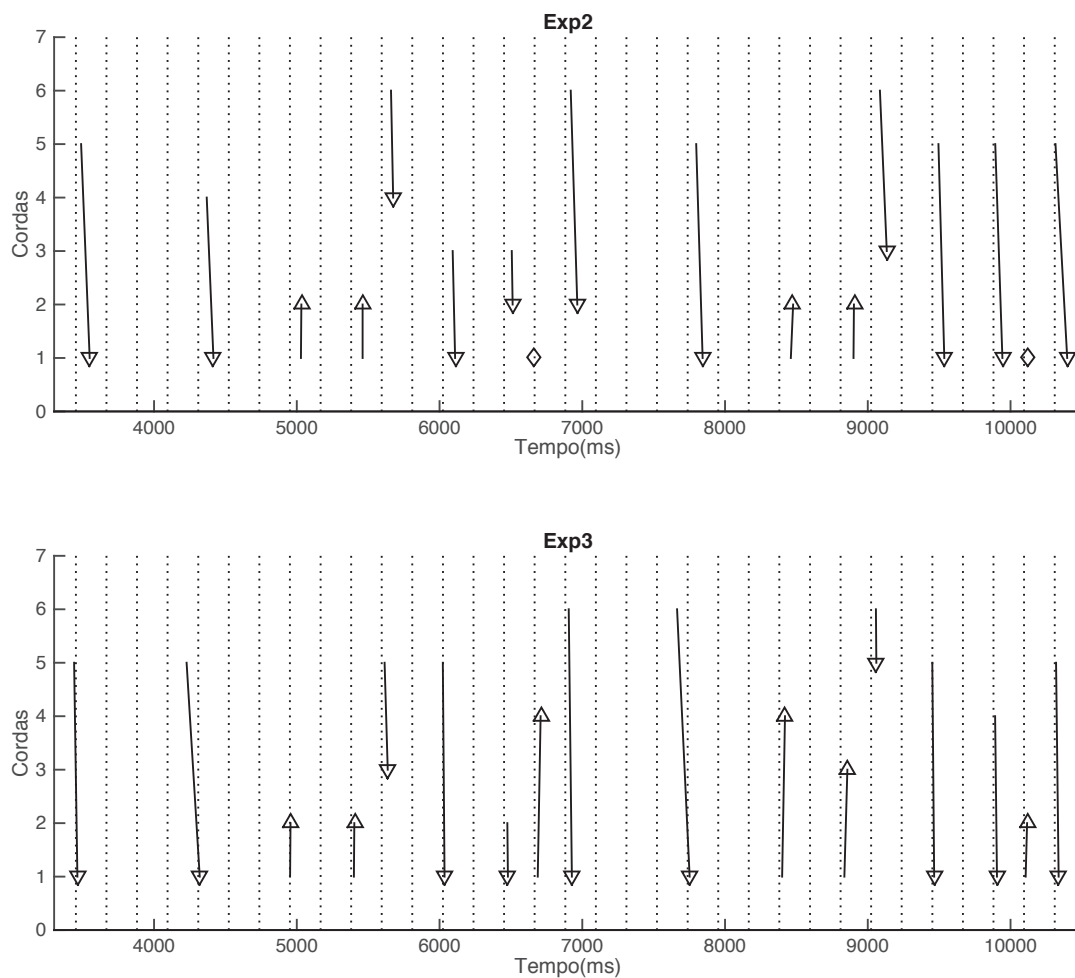
Para o ritmo simples, o peso do primeiro componente principal é suficiente para uma diferenciação (5% de diferença entre os músicos). Já para o ritmo 2, a diferenciação entre o músico Exp2A e o Exp3 é mais difícil, tanto pela observação da porcentagem quanto da contribuição de cada eixo. Note que como não estamos lidando com um valor de *yaw* fixo, a distribuição das contribuições dos eixos x e y pode variar. Mesmo observando-se os resultados de *PCA* para as velocidades angulares desses músicos neste trecho, não se nota uma diferença marcante. No entanto, pode-se observar estratégias totalmente distintas de execução, devido à expertise, como será discutido em seguida. Estes resultados nos alertam para o fato de que as comparações realizadas devem se basear em algum teste de similaridade prévio entre as estratégias de performance utilizadas no *strumming*.

Finalmente comparamos uma versão do ritmo 2 (Exp3) com uma versão do mesmo ritmo tocada pelo músico A no Exp2. Como este último não tem muita experiência com este tipo de acompanhamento, além de ser canhoto (porém toca da forma tradicional,

com a mão direita atacando as cordas), a comparação de alguns aspectos já será suficiente para qualificar e diferenciar as performances.

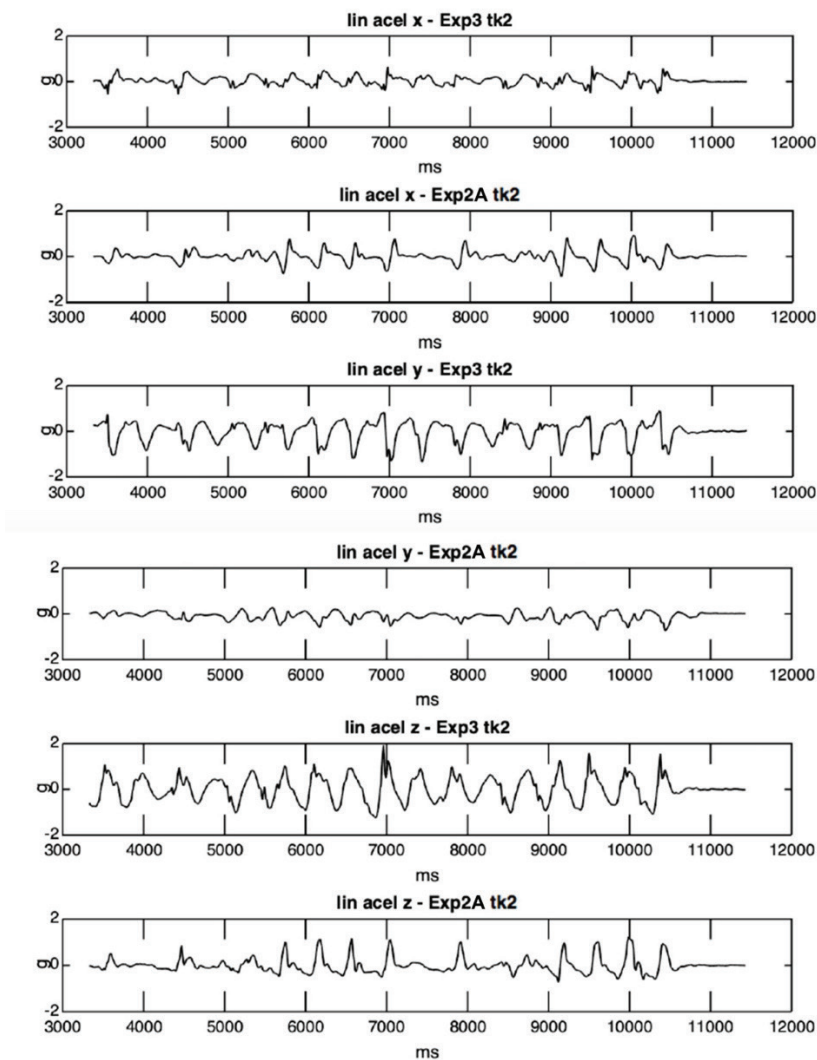
Primeiramente iremos comparar a relação dos acordes com os pulsos (subdivididos em semicolcheias), mostrado na Figura 44. Notamos que o músico Exp2 apresenta *spread* temporal maior no primeiro acorde que no segundo, quando deveria ser o oposto. Os ataques estão atrasados em relação ao pulso durante todo o trecho, e temos gestos que tocam apenas uma corda. A separação de planos do primeiro compasso é feita de maneira semelhante pelos dois músicos, fato que não ocorre no segundo compasso. A performance Exp2 é mais densa no grave e menos densa no agudo do que a Exp3.

Figura 44 – Acordes e pulso (em semicolcheias) dos músicos Exp2A e Exp3.



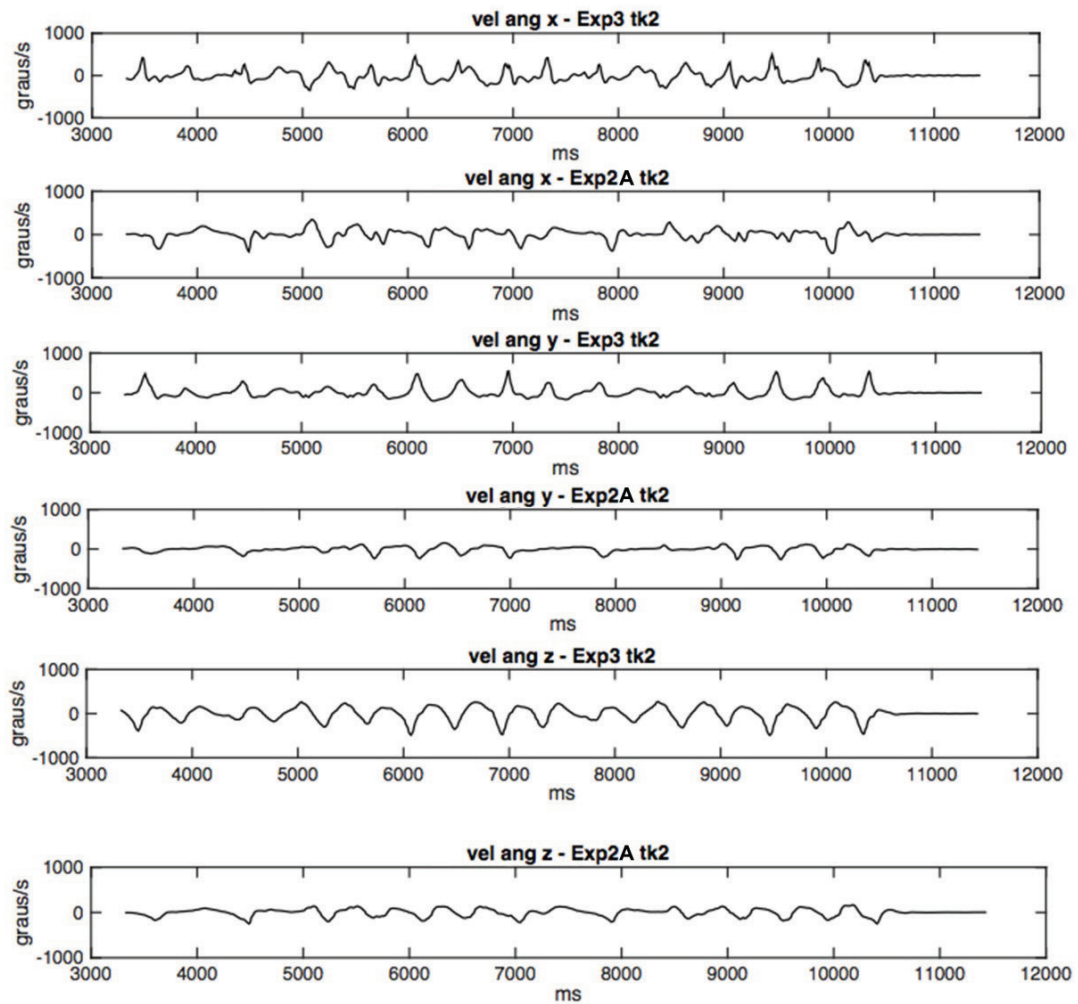
A Figura 45 mostra as acelerações lineares dos dois músicos na execução do ritmo 2. O primeiro ponto a ser destacado é o fato do músico Exp2A não explorar tanto os três eixos como Exp3, principalmente no eixo y, onde observamos a maior diferença. O eixo z em Exp3 apresenta valores maiores que Exp2A, assim como em x. Embora este fato possa ser relativizado pelas diferenças entre as medidas inerciais dos ângulos z durante as gravações, pode-se observar uma atividade global mais intensa no Exp3. Outro ponto a ser destacado é o fato do músico Exp2A não manter a movimentação constante da mão como Exp3, mostrado nas curvas entre 4000 ms a 5000ms e 8000 ms a 9000 ms, onde o primeiro não apresenta movimentação nestes intervalos.

Figura 45 – Aceleração linear músicos Exp2A e Exp3 no take 2. Os músicos utilizam os eixos de maneira diferente na execução do ritmo 2.



Observamos também diferenças consideráveis nas velocidades angulares, indicadas na Figura 46. Sinais invertidos para x e y se devem ao posicionamento do sensor na mão. Principalmente em y e z, Exp2A utiliza menos os eixos que Exp3, indicando uma mão mais “dura”, que não gira na aplicação da técnica. Esse fato pode estar ligado à expertise dos performers, já que o primeiro é menos experiente no *strumming* nesse tipo de ritmo com ornamentos diferentes do que o segundo. Como no caso das acelerações, em Exp3 nota-se a manutenção dos movimentos mesmos nos momentos onde não há ataques de acordes.

Figura 46 – Velocidade angular músicos Exp2A e Exp3. Os músicos utilizam os eixos de maneira diferente na execução do ritmo 2.



## 2.4. Considerações

Neste capítulo apresentamos os experimentos realizados durante o mestrado. Apesar de terem ocorrido em momentos distintos da pesquisa, principalmente com programação diferente do *IMU*, discutiram de modo relevante, em seu escopo e na comparação entre si, as questões que envolvem o complexo gesto da técnica de *strumming*.

As ferramentas desenvolvidas para esse estudo contribuíram com a análise multimodal proposta da técnica e apresentam grande possibilidades para pesquisas futuras. Com o GuiaRT identificamos os *onsets* por corda, nos permitindo observar a expressão rítmica dos performers nos trechos sugeridos, além da extração de outros atributos sonoros que poderão ser utilizados futuramente. Com o vídeo em 240 fps fomos capazes de segmentar os gestos presentes na aplicação do *strumming* e correlacionar esses pontos de troca de sentido do movimento da mão com os dados do sensor de movimento (com destaque para a velocidade angular no eixo z), indicando uma possível segmentação através dos dados do *IMU*. O MetaMotionR nos permitiu analisar os aspectos dos movimentos conjugados de transação e rotação, nos dois modos de programação do sensor. A possibilidade de se colocar numa mesma representação gráfica dados de diferentes modalidades contribui para uma análise mais abrangente dessa técnica.

Apontamos aqui algumas discussões com os dados dos experimentos e observamos que é possível tanto identificar características comuns quanto identificar diferenças entre os performers. Esse fato nos mostra que na técnica de *strumming* o aspecto gestual é idiossincrático e, conseqüentemente, o resultado sonoro e musical obtido.

## Conclusão

Consideramos importantes para a discussão do *strumming* as questões levantadas aqui, contribuindo para um entendimento melhor da técnica na sua aplicação em ritmos populares, o que pode mudar o modo como ela é estudada e ensinada.

As pesquisas sobre o *strumming*, já citadas no início deste trabalho, abordam o tema em relação a um aspecto isolado da técnica, apresentando resultados importantes sobre os componentes sonoros, gestuais e neurológicos da aplicação, todos importantes para seu entendimento. Esse trabalho, a partir da análise multimodal, buscou uma aproximação desses elementos em um estudo conjunto, trazendo novas perspectivas para a compreensão sobre como o *strumming* é utilizado pelos performers.

No primeiro experimento, foi possível mostrar a efetividade da diferenciação entre três músicos realizando um ritmo simples apenas com a observação das velocidades angulares. O segundo experimento nos deu a dimensão da acurácia das medidas realizadas pelo sensor escolhido, além de alguns insights sobre características importantes desta técnica. O terceiro experimento, que se limitou à análise de apenas um músico, pelos motivos já relatados anteriormente, lidou com um ritmo mais complexo e em sua análise foi possível usar a combinação de dados variados: extraídos do áudio, do vídeo, das acelerações lineares, das velocidades angulares.

A possibilidade de diferenciar os performers, em consideração aos aspectos sonoros e gestuais, pode contribuir para a discussão sobre como a técnica é ensinada ou assimilada, visto que por muitas vezes é uma habilidade adquirida de forma autodidata, e podem indicar um norte para o desenvolvimento neste aspecto.

Os resultados aqui apresentados podem indicar um caminho para a busca de estratégias gestuais mais efetivas para o ensino e a performance da técnica, contribuindo para o modo como é estudada e transmitida. Além disso, nos ajudam a questionar o modo de representação gráfica, onde podem ser incluídos novos elementos, como a separação de planos e a indicação de acordes arpejados. A Figura 47 mostra um exemplo de representação para o ritmo 2 considerando esses pontos levantados.

Figura 47 – Exemplo de representação por setas do ritmo 2, em comparação com o modelo utilizado atualmente.

♩ = 70 Am G Am

Modelo atual

Modelo proposto

As ferramentas aqui desenvolvidas, tanto de extração como de análise dos dados, colaboram para um novo olhar sobre a técnica, abrindo um horizonte de possibilidades para pesquisas futuras. Os próximos passos incluem a construção de um banco de dados mais amplo, com gravações de músicos com diversos níveis de expertise no *strumming*, maior variedade de ritmos, e registro em situações diferentes de performance.



## REFERÊNCIAS

ABHAYASINGHE, N.; MURRAY, I.; SHARIF BIDABADI, S. **Validation of Thigh Angle Estimation Using Inertial Measurement Unit Data against Optical Motion Capture Systems**. *Sensors*, v. 19, n. 3, p. 596, 2019.

ADORNO, T. W.; PADDISON, M. **On the Problem of Musical Analysis**. *Music Analysis*, v. 1, n. 2, p. 169–187, 1982.

AHMAD, N. et al. **Reviews on Various Inertial Measurement Unit (IMU) Sensor Applications**. *International Journal of Signal Processing Systems*, v. 1, p. 256–262, 2013.

**ANÁLISE DE COMPONENTES PRINCIPAIS**. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2019. Disponível em: [https://pt.wikipedia.org/w/index.php?title=An%C3%A1lise de componentes principais&oldid=55520598](https://pt.wikipedia.org/w/index.php?title=An%C3%A1lise_de_componentes_principais&oldid=55520598). Acesso em: 18 jun. 2019.

ARMONDES, A. **GuiaRT, um sistema em tempo real de extração e re-criação de texturas musicais utilizando um violão de nylon**. In: 3º NAS NUUVENS... CONGRESSO DE MÚSICA: ANAIS, 2017.

ARMONDES, A. et al. **Uma primeira aproximação multimodal da técnica de strumming**. In: XXX CONGRESSO DA ANPPOM - MANAUS/AM, 2020.

ARMONDES, A.; FREIRE, S. **Ferramentas para uma abordagem multimodal da técnica de strumming no violão**. In: 5º NAS NUUVENS... CONGRESSO DE MÚSICA: ANAIS, 2019.

BARKOSKIE, A. T. **Rasgueado for flute, clarinet in Bb, violin, cello, percussion and piano**. Thesis - University of Texas, 2011.

BEANGE, K. **Validation of Wearable Sensor Performance and Placement for the Evaluation of Spine Movement Quality**. Thesis - University of Ottawa, 2019.

BELLO, K.; MAYOL, P. **Classification of Acoustic Guitar Strum using Convolutional Neural Networks and Long-Short-Term-Memory**. *Philippine e-Journal for Applied Research and Development*, 9, 49-57, 2019.

BESSONE, V. et al. **Validation of a new inertial measurement unit system based on different dynamic movements for future in-field applications**. *Sports Biomechanics*, p. 1–16, 2019.

BISHOP, L.; JENSENIUS, A. R. **Reliability of two infrared motion capture systems in a music performance setting**. *Proceedings of the 17th Sound and Music Computing Conference*. Torino, Italy, 2020.

BREGMAN, A. **Auditory scene analysis: the perception and organization of sound**. MIT Press, 1990.

CADOZ, C.; WANDERLEY, M. M. **Trends in gestural control of music**. Paris Ircam: Centre Georges Pompidou, 2000.

CAMURRI, A. et al. **EyesWeb: Toward Gesture and Affect Recognition in Interactive Dance and Music Systems**. *Computer Music Journal*, v. 24, n. 1, p. 57–69, 2000.

- CARAMIAUX, B.; BEVILACQUA, F.; SCHNELL, N. **Towards a Gesture-Sound Cross-Modal Analysis.** (S. Kopp, I. Wachsmuth, Eds.) *Gesture in Embodied Communication and Human-Computer Interaction. Anais...: Lecture Notes in Computer Science.* Berlin, Heidelberg, Springer, 2010
- CHEN, Y.-Y. et al. **Comparison of Dynamic Balance under Dual and Simple Task Conditions in Children with Attention Deficit Hyperactivity Disorder.** *Journal of Physical Therapy Science*, v. 24, n. 8, p. 633-637, 2012.
- CLARK, C. **Introduction: Sorcerers and Sorceresses.** *University of Toronto Quarterly*, 2006.
- CLARKE, E. **Empirical Methods in the Study of Performance.** Clarke, E., & Cook, N. (Eds.) *Empirical musicology : Aims, methods, prospects.* Oxford University Press, 2004.
- Oxford University Press, 2004.
- COOK, N. **Beyond the score: Music as performance.** Oxford University Press, 2013.
- DAHL, K. D. et al. **Wearable sensor validation of sports-related movements for the lower extremity and trunk.** *Medical Engineering & Physics*, v. 84, p. 144–150, 2020.
- FREIRE, S. et al. **Evaluation of Inertial Sensor Data by a Comparison with Optical Motion Capture Data of Guitar Strumming Gestures.** *Sensors*, 20, 5722, 2020.
- FREIRE, S.; ARMONDES, A.; SILVA, R. **GuiaRT, a Setup for Real-Time Transcription and Transformation of Music Played on a Hexaphonic Nylon Guitar.** Under evaluation, 2020.
- FREIRE, S.; ARMONDES, A.; VIANA, J. **Ritmo, microtempos e acentos na técnica violonística de strumming: ferramentas de análise e resultados preliminares.** In: XXVII CONGRESSO DA ANPPOM - CAMPINAS/SP, 2017
- FREIRE, S.; NÉZIO, L. **Study of the tremolo technique on the acoustic guitar: Experimental setup and preliminary results on regularity.** In: PROC. INT. CONF. SOUND AND MUSIC COMPUTING, STOCKHOLM, 2013
- GLOWINSKI, D. et al. **Analysis of leadership in a string quartet.** In: THIRD INTERNATIONAL WORKSHOP ON SOCIAL BEHAVIOUR IN MUSIC AT ACM ICMI, 2012.
- HSU, W. **Managing Gesture and Timbre for Analysis and Instrument Control in an Interactive Environment.** In: NIME. Citeseer, 2006
- ISHII, R. **Development of Extended Piano Techniques in Twentieth-Century American Music.** Thesis - Florida State University, 2005.
- JENSENIUS, A. R.; WANDERLEY, M. M. **Musical gestures: Concepts and methods in research.** In: *Musical Gestures.* Routledge, p. 24–47, 2010.
- KITTLEBERGER, B. **Acoustic Guitar Strumming Patterns.** Guitar Chalk, 2018. Disponível em: <<https://www.guitarchalk.com/acoustic-guitar-strumming-patterns/>>. Acesso em: 5 out. 2020
- KRENZ, S. **Gibson's Learn & Master Guitar.** Box Pck Pa Edição ed. Nashville, Tenn.: Legacy Learning, 2010.
- LARSEN, J. V.; OVERHOLT, D.; MOESLUND, T. B. **The Actuated Guitar: A platform enabling alternative interaction methods.** . In: SMC PROCEEDINGS OF THE SOUND AND MUSIC COMPUTING CONFERENCE. Logos Verlag Berlin Stockholm, Sweden, 2013

- LEMAN, M. **Embodied music cognition and mediation technology**. MIT press, 2008.
- MAIDHOF, C.; KÄSTNER, T.; MAKKONEN, T. **Combining EEG, MIDI, and motion capture techniques for investigating musical performance**. Behavior Research Methods, v. 46, n. 1, p. 185–195, 2014.
- MARIN, O. S. M.; PERRY, D. W. 17 - **Neurological Aspects of Music Perception and Performance**. In: DEUTSCH, D. (Ed.). The Psychology of Music (Second Edition). San Diego: Academic Press, p. 653–724, 1999.
- MATSUSHITA, S.; IWASE, D. **Detecting Strumming Action While Playing Guitar**. Proceedings of the 2013 International Symposium on Wearable Computers. Anais...: ISWC '13. New York, NY, USA, 2013.
- MENESES, E. A.; FREIRE, S.; WANDERLEY, M. M. **GuitarAMI and GuiaRT: two independent yet complementary augmented nylon guitar projects**. In: PROCEEDINGS OF NIME. 2018
- MORLEY, I. **Evolution of the Physiological and Neurological Capacities for Music**. Cambridge Archaeological Journal, v. 12, n. 2, p. 195–216, 2002.
- MOURA, H. **Bringing multimodality and affordances to design of Technology Enhanced Learning Environment**. PhD Dissertation, Illinois Institute of Technology, Institute of Design, 2008.
- P. HUANG et al. **Singing-voice separation from monaural recordings using robust principal component analysis**. IEEE International Conference on Acoustics, Speech and Signal, 2012.
- PANAGAKIS, Y.; BENETOS, E.; KOTROPOULOS, C. **Music Genre Classification: A Multilinear Approach**. 9th International Conference on Music Information Retrieval, Drexel University, Philadelphia, PA, USA, 2008.
- PANAGAKIS, Y.; KOTROPOULOS, C.; ARCE, G. **Non-Negative Multilinear Principal Component Analysis of Auditory Temporal Modulations for Music Genre Classification**. Audio, Speech, and Language Processing, IEEE Transactions on, v. 18, p. 576–588, 2010.
- PASSARO, V. et al. **Gyroscope technology and applications: A review in the industrial perspective**. Sensors, v. 17, n. 10, p. 2284, 2017.
- PERSINGER, S. **Why Is Rhythm Guitar So Hard?**. Premiere Guitar, 2019. Disponível em: <<https://www.premierguitar.com/articles/29746-why-is-rhythm-guitar-so-hard>>. Acesso em: 5 out. 2020.
- PORNPRASIT, K. **Playing Techniques of New Pieces of Koto Music in Japanese Urban Culture**. Journal of Urban Culture Research, v. 1, p. 178–188, 2010.
- QUALISYS. **Qualisys Track Manager User Manual**. Gothenburg, Sweden, 2011.
- RAE, G.; MCCAMBRIDGE, K. **Correlates of performance anxiety in practical music exams**. Psychology of Music, v. 32, n. 4, p. 432–439, 2004.
- RAHMAN, M. M. et al. **Analysis of dexterous finger movement for piano education using motion capture system**. International Journal of Science and Technology Education Research, v. 2, n. 2, p. 22–31, 2011.
- RASAMIMANANA, N.; KAISER, F.; BEVILACQUA, F. **Perspectives on gesture-sound relationships informed from acoustic instrument studies**. Organised Sound, v. 14, n. 2, p. 208–216, 2009.

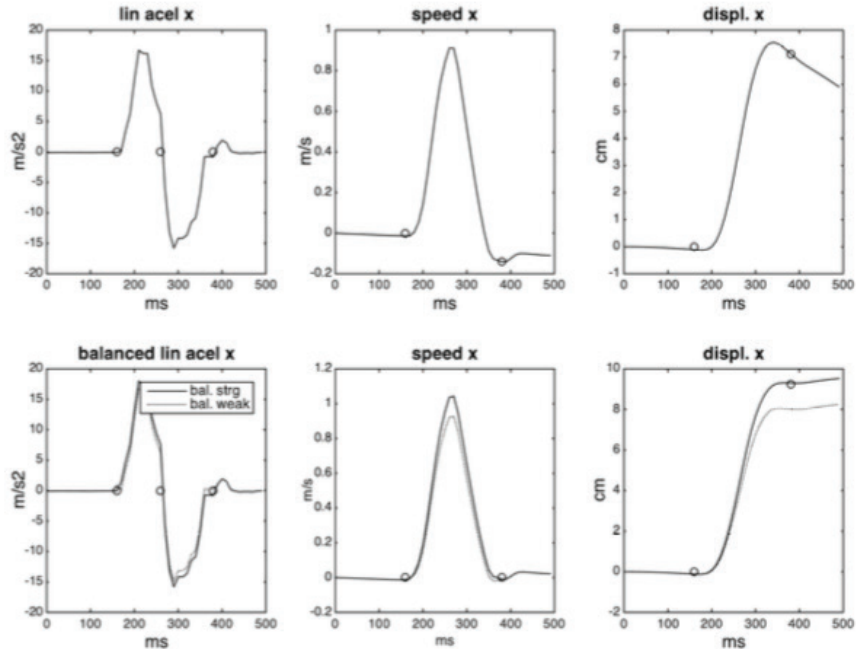
- RIBEIRO, T. S. et al. **Effects of a training program based on the Proprioceptive Neuromuscular Facilitation method on post-stroke motor recovery: A preliminary study.** *Journal of Bodywork and Movement Therapies*, v. 18, n. 4, p. 526–532, 2014.
- RINK, J. **Musical performance: A guide to understanding.** Cambridge University Press, 2002.
- ROBERT STRIZICH; JAMES TYLER. **Rasgueado.** Grove Music Online, 2001.
- ROLFHAMRE, R. **French baroque lute music from 1650-1700.** Saarbrücken, Germany: VDM Verlag, 2010.
- ROSSING, T. **Springer handbook of acoustics.** Springer Science & Business Media, 2007.
- ROTH, D. **SwiftOSC.** Framework OSC para linguagem Swift, 2006.
- SAWICKI, A.; ZIELIŃSKI, S. K. **An Algorithm for Detecting the Expressive Musical Gestures of Violinists Based on IMU Signals.** (K. Saeed, R. Chaki, V. Janev, Eds.) *Computer Information Systems and Industrial Management. Anais...: Lecture Notes in Computer Science.* Cham: Springer International Publishing, 2019
- SCHAEFFER, P. **Traité des objets musicaux.** Le Seuil, 2016.
- SCHIESSER, S.; SCHACHER, J. **SABRe: The Augmented Bass Clarinet.** NIME, Ann Arbor, Michigan, USA, 2012
- SÉRGIO FREIRE et al. **Strumming on an Acoustic Nylon Guitar: Microtiming, Beat Control and Rhythmic Expression in Three Different Accompaniment Patterns.** SMC, 2018.
- SILVA, R.; FREIRE, S. **Nibanna.** 4º Nas Nuvens... Congresso de Música: Anais, 2018
- THOMPSON, M. R.; LUCK, G. **Exploring relationships between pianists' body movements, their expressive intentions, and structural elements of the music.** *Musicae Scientiae*, v. 16, n. 1, p. 19–40, 2012.
- VISI, F.; SCHRAMM, R.; MIRANDA, E. **Gesture in Performance with Traditional Musical Instruments and Electronics: Use of Embodied Music Cognition and Multimodal Motion Capture to Design Gestural Mapping Strategies.** *Proceedings of the 2014 International Workshop on Movement and Computing. Anais...: MOCO '14.* New York, NY, USA, 2014
- ZHOU, L.; BAI, S.; LI, Y. Zhou. **Energy Optimal Trajectories in Human Arm Motion Aiming for Assistive Robots.** *Modeling, Identification and Control: A Norwegian Research Bulletin*, v. 38, n. 1, p. 11–19, 2017.
- ZÜGNER, R. et al. **Validation of inertial measurement units with optical tracking system in patients operated with Total hip arthroplasty.** *BMC Musculoskeletal Disorders*, v. 20, n. 1, p. 52, 2019.

## APÊNDICE A – Integração dos dados de aceleração linear gerados pelo IMU

Podemos estimar o deslocamento linear em cada eixo através da integração da aceleração, com foco nos gestos periódicos (ou quase) do *strumming*. Para utilização nas comparações que fundamentam os experimentos, desenvolvemos um método aproximativo de integração de curvas de aceleração para estimativa de velocidades lineares (e por consequência, também de integração de curvas de velocidade para estimativa de deslocamentos). Mostraremos um exemplo simples, que apresenta o procedimento utilizado para equilibrar os ciclos de aceleração variando no tempo.

Uma trajetória linear simples tem uma curva de um lóbulo para a velocidade - que começa em zero, sobe ou desce até a velocidade máxima e retorna a zero - e uma curva simétrica (ou quase) de dois lóbulos para a aceleração: uma aceleração seguida por uma desaceleração na direção escolhida, com a velocidade máxima no cruzamento da curva com o eixo horizontal (valor zero). Deslizar o *IMU* na direção positiva de seu eixo x produz uma curva como a primeira ilustrada na Figura 1. Os valores são multiplicados pela constante de gravidade de 9,81. A integração desta curva usando o método trapezoidal fornece uma estimativa da velocidade e uma segunda integração fornece o deslocamento. A primeira linha da figura ilustra esse processo e revela um problema comum a esse procedimento: é um desafio para os humanos gerar uma curva de aceleração totalmente simétrica, porque a força aplicada, devido a diversos fatores, geralmente não é simétrica. Além disso, os lóbulos em cada ciclo podem ser representados por um número diferente de amostras. Portanto, optamos por aplicar um fator de compensação ao lóbulo mais fraco (ou mais forte), com base no número de pontos e valores médios de cada lóbulo. Ambas as opções são mostradas na figura. Antes desta compensação, é necessário definir os limites da curva através da detecção de três cruzamentos por zero: no início, meio e final do gesto. Os resultados da aplicação deste procedimento podem ser vistos na segunda linha da mesma figura. O procedimento de balanceamento nos permite zerar a velocidade no final do ciclo de aceleração.

Figura 1 – Dupla integração da curva de aceleração de uma trajetória simples. Na segunda linha, a linha sólida indica o balanceamento do lóbulo forte, e a linha pontilhada do lóbulo fraco. Os pequenos círculos indicam o zero-crossing extraídos da curva de aceleração.



### a) Deslocamento com rotação

A maioria dos gestos com as mãos combina translação e rotação de muitas maneiras diferentes, incluindo mudanças de forma, coordenação com outras partes do corpo e controle de forças dinâmicas (variáveis no tempo). Antes de investigar este gesto na aplicação do *strumming*, analisamos os dados gerados pelo sensor em um movimento oscilatório relacionado com essa técnica.

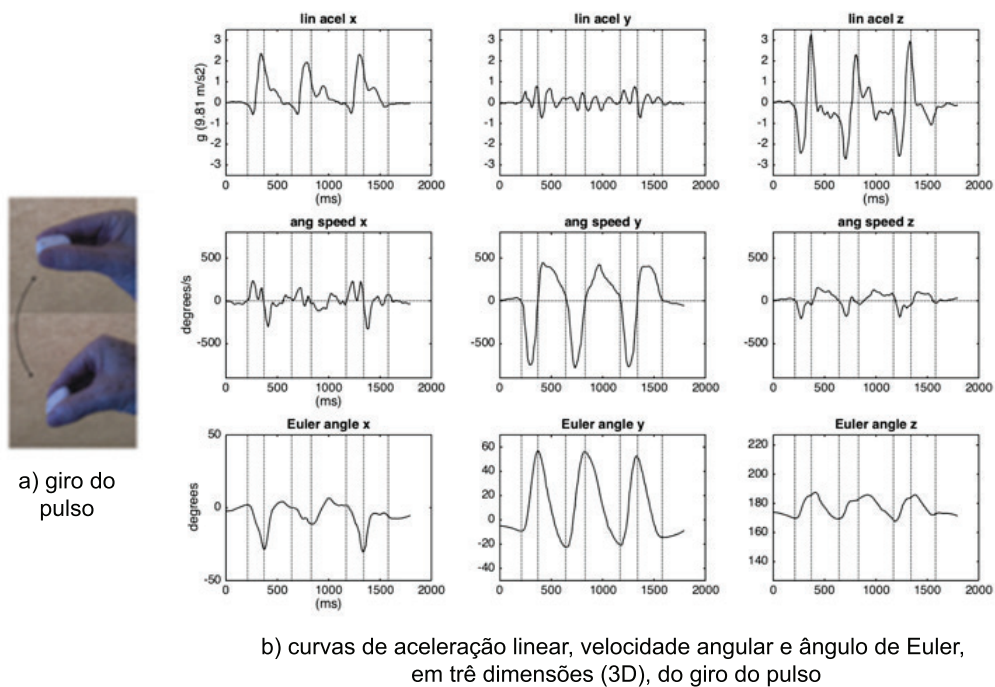
Registramos no *Max/MSP* uma série de seis rotações do punho (três para baixo e três para cima) com o *IMU* entre o polegar e o indicador, seu plano xy paralelo ao chão e o eixo z apontando para cima, conforme ilustrado na Figura 2a. De forma muito esquemática, esse gesto consiste em uma rotação em torno do eixo y, e uma translação ocorrendo principalmente no eixo z.

A Figura 2b mostra as curvas nos três eixos para aceleração linear, velocidade angular e atitude (ângulos de Euler). Pode-se notar que as curvas com as maiores amplitudes são as que acabamos de mencionar (rotação y e aceleração z). As linhas verticais pontilhadas são alinhadas com a mudança de direção da rotação do eixo y.



Embora esses seis gestos tenham sido realizados com a intenção de regularidade, os dados apresentam uma assimetria temporal, sendo os gestos para baixo mais rápidos que os para cima. Outro ponto que chama a atenção imediatamente é o sinal reverso entre velocidade angular e atitude, que provavelmente vem do firmware do *IMU*. Em cada eixo, um deslocamento angular positivo corresponde a uma velocidade negativa e vice-versa.

Figura 2 – Dados do IMU de seis rotações de pulso.

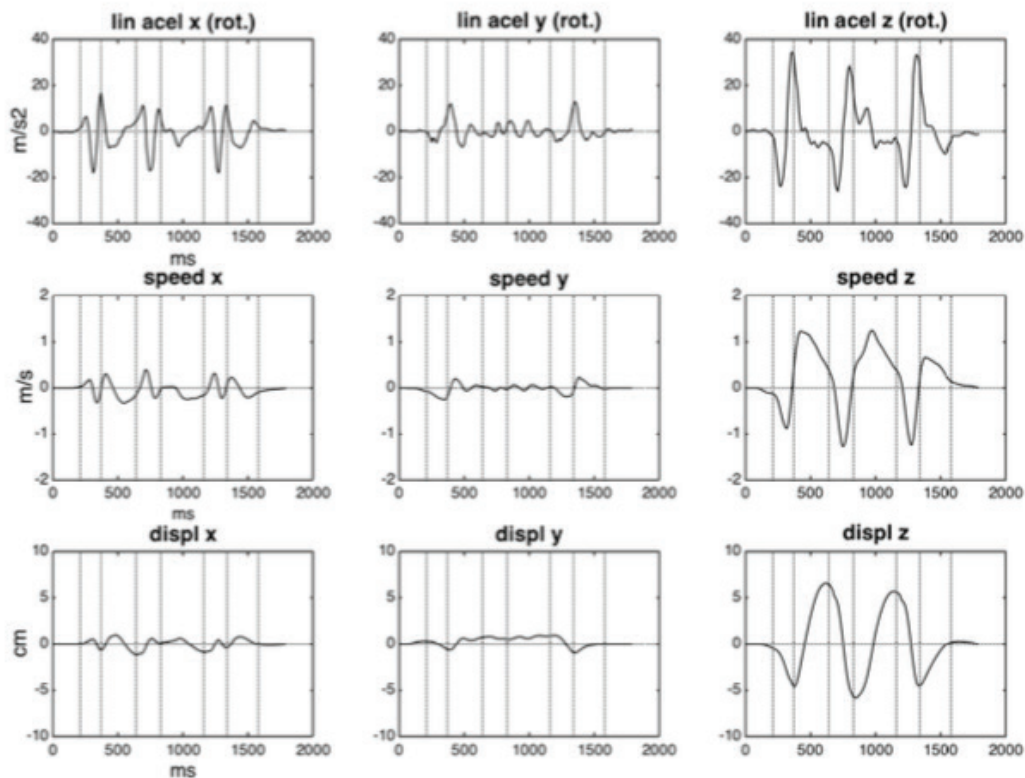


Existe uma sobreposição das curvas de aceleração dos gestos em movimentos cíclicos, com a desaceleração do gesto anterior já contribuindo para a aceleração do próximo. Portanto, apenas o primeiro e o último gesto teriam parte de sua aceleração não compartilhada com os outros. Esta sobreposição é bastante evidente em gestos periódicos mais rápidos e regulares, mas não tão evidente em oscilações mais lentas ou irregulares. Por exemplo, na Figura 2, é possível ver pequenos picos positivos e negativos entre dois cruzamentos de zero na curva de aceleração do eixo z durante a rotação que ocorre em torno de 1000 ms. Essa observação ajuda a decidir qual método de

balanceamento usar: se houver uma sobreposição substancial, podemos balancear para o lóbulo mais forte; caso contrário, o equilíbrio deve ser baseado no lobo mais fraco, a fim de evitar uma injeção virtual e artificial de energia.

A aceleração no eixo x mostra um pequeno aumento positivo e nenhuma compensação no lado negativo. Se relacionado a um movimento linear, isso indicaria um movimento sem descanso, o que não é o caso. Uma imagem mais confiável do que acontece com a aceleração linear é girar seus valores de acordo com um referencial fixo. Aqui, o *IMU* fornece essa estrutura fixa em sua posição de repouso no início do movimento; as matrizes de rotação são calculadas a partir dos ângulos de Euler para cada ponto de tempo. O resultado dessa rotação pode ser visto na primeira linha da Figura 3. Depois disso, lobos positivos e negativos na aceleração podem ser detectados em cada eixo, embora ainda sejam consideravelmente assimétricos.

Figura 3 - Na primeira linha, as curvas de aceleração 3D giraram para o referencial do IMU. As linhas a seguir representam as curvas de velocidade e deslocamento obtidas por integração.





Depois de detectar o *zero-crossing* na curva de aceleração de cada eixo (na verdade, cruzamentos de uma faixa de  $\pm$  limiar, com um valor mínimo de piso), dividimos a curva em dois grupos de ciclos parcialmente sobrepostos, um com seu primeiro lóbulo com sinal negativo e o outro com o primeiro lóbulo com sinal positivo. Em seguida, aplicamos o mesmo procedimento de compensação para cada ciclo, conforme descrito acima, escolhendo o equilíbrio de acordo com o lobo mais fraco. A estimativa da velocidade é feita pela média desses dois grupos de curvas. O mesmo procedimento foi aplicado para estimar o deslocamento em cada eixo, representado na terceira linha.

Fonte: A autoria própria

## APÊNDICE B – Resultados da comparação de dados gerados pelo MetaMotionR e pelo Qualisys

A comparação de diferentes séries temporais foi baseada em *RMSE* (sigla em inglês para a raiz quadrada do erro-médio) e covariâncias cruzadas normalizadas, ambas calculadas com defasagens. Decidimos usar os dois métodos, pois baixos valores de *RMSE* também podem aparecer em séries fracamente correlacionadas e vice-versa. Encontramos valores de referência na literatura apenas para dados rotacionais, pois comparações baseadas em dupla integração ou derivação são menos comuns. Observamos também as diferenças entre os valores máximos e mínimos presentes em cada série temporal, indicados por  $\Delta_{max}$ , que são úteis para interpretar alguns resultados. A escolha pela covariância, e não pela correlação, justifica-se por sua maior sensibilidade às diferenças de escala. Além disso, valores normalizados podem ser úteis para comparações entre curvas geradas por diferentes quantidades físicas, como acelerações e deslocamentos nas translações. O valor máximo para os atrasos usados em todas as comparações foi 12. Nas subseções a seguir, apresentaremos os resultados dos *takes* selecionados dos músicos 1 e 2.

### a) Dados de rotação

A comparação entre os dados rotacionais de ambos os sistemas é realizada após o alinhamento temporal, e um *yaw shift* realizado para cada *take*. É importante lembrar que os ângulos a serem comparados são expressos em relação a um referencial móvel, ou seja, o próprio *IMU*.

A Tabela 1 apresenta os resultados selecionados obtidos para as rotações em cada eixo e o valor médio e o desvio padrão para cada série temporal. A comparação completa mostra 73% do *RMSE* abaixo de  $5^\circ$  e 88% abaixo de  $10^\circ$ , bem como dois terços dos valores de covariância acima de 0,95. O *RMSE* mais alto e os valores de covariância mais baixos são encontrados nas rotações do eixo z. Este eixo também mostra valores mais altos para defasagens. Como o sensor foi posicionado para cada músico com uma rotação de  $180^\circ$ , os valores médios do eixo z têm valores semelhantes com sinais opostos. Os desvios-padrão podem ser associados a diferentes técnicas utilizadas para o *strumming* (ritmos 1 e 2). O músico 1 possui valores mais elevados para o eixo x,

indicando movimentos rotacionais que ocorrem principalmente no plano sagital. Por outro lado, o músico 2 apresenta valores maiores não só para este eixo, mas também para o eixo z, indicando rotações mais proeminentes também relacionadas ao plano vertical.

Tabela 1 – RMSE e covariâncias (por eixo) entre os dados rotacionais gerados por cada sistema dos takes dos músicos 1 e 2.

Take	Eixo	RMSE(°)	Lag	X-Covar	Média Mocap (°)	Stdev	Média IMU (°)	Stdev
<b>m1r1t1</b>	x	4,42	0	0,997	80,7	17,4	76,5	17,4
	y	0,68	0	0,993	29,1	5,8	29,1	5,7
	z	3,48	-8	0,877	-293,8	6,6	-293,7	5,5
<b>m2r1t2</b>	x	4,22	2	0,996	-78,1	19,7	-82,2	19,7
	y	1,14	2	0,996	-38,4	8,5	-39,1	8,9
	z	4,37	0	0,974	217	13,4	218,7	10,5
<b>m1r2t1</b>	x	3,24	0	0,993	83,3	17,3	80,9	17,9
	y	1,33	0	0,991	38,6	10,1	38,5	10,1
	z	6,25	-4	0,820	74,4	10,1	73,4	10
<b>m2r2t2</b>	x	6,69	-1	0,996	-79,7	14,7	-86,2	14,2
	y	0,79	-1	0,996	-42,3	9,3	-42,4	9,6
	z	5,13	-3	0,890	212,8	10,8	212,3	11
<b>m1r3t1a</b>	x	4,43	-1	0,972	-68,3	3,9	-72,5	4,4
	y	2,36	0/-1	0,994	-23,9	7,7	-26	7,7
	z	14,47	-1	0,839	-119,7	9,6	-106,9	3,7
<b>m2r3t2a</b>	x	5,68	0/1	0,948	-67,8	5,8	-73,1	6,1
	y	2,54	1	0,958	-30,6	7,9	-31,4	8,5
	z	11,20	2/1	0,730	-150,1	7,6	-140,2	6,6
<b>m1r3t1b</b>	x	7,87	0/-1	0,989	-73,6	9,4	-81,2	8,2

	y	4,64	-1	0,866	-4	1,6	0,6	0,9
	z	2,74	-1	0,987	-96,8	9,6	-94,8	9,8
<b>m2r3t2b</b>	x	3,29	0	0,985	-61,4	7,7	-63,6	9,6
	y	7,60	0	0,994	-8,7	9,6	-16,3	9,8
	z	13,39	2/1	0,928	-131,8	11,5	-121,6	3,1
<b>m1r3t1c</b>	x	11,09	-2	0,990	-76,8	8,0	-87,8	7,7
	y	1,56	-2	0,990	-1,1	4,1	-2,5	4
	z	6,83	-2	0,505	-93,5	4,7	-88	2,7
<b>m2r3t2c</b>	x	4,40	1	0,987	-64,4	11	-68,4	11,4
	y	7,64	1	0,993	-5,3	7,1	-12,9	7,4
	z	8,62	0	0,580	-121,6	6,9	-115,2	5,4

### a) Comparação das translações

Não há uma maneira direta de comparar os dados posicionais 3D gerados pelo *Qualisys* com as acelerações lineares 3D fornecidas pelo *IMU*. Portanto, tivemos que convertê-los para uma mesma grandeza física usando dois métodos diferentes: um que parte dos dados posicionais levando à aceleração (por derivação dupla) e outro dos dados de aceleração para o deslocamento (usando o método de integração compensativa proposto).

Para o método de derivação, as tabelas 2 e 3 mostram os resultados da comparação estatística entre as séries temporais resultantes de cada eixo, após os valores de aceleração do *IMU* terem sido rotacionados de acordo com os ângulos de atitude fornecidos pelo *Qualisys* para o corpo rígido. Nos trechos 1 e 2, ambos os músicos alcançaram altos valores de covariância para os eixos y e z. Os valores obtidos para o eixo x mostram uma correlação mais fraca, principalmente para o Músico 2. Este eixo é menos explorado na técnica de *strumming*, que é basicamente um movimento 2D ocorrendo no plano vertical. Valores mais precisos para a aceleração do *IMU* nestes eixos podem ser prejudicados pelo fato de dependerem da rotação do eixo z, que apresentou

a pior correlação na comparação das rotações. Outro fator que pode interferir nesses resultados é o fato de fixarmos a posição da cadeira para a performance, mas não a orientação espacial do violão, que é algo altamente pessoal.

Tabela 2 – Covariâncias (por eixo) entre acelerações lineares do IMU e derivadas duplas de dados posicionais do mocap - takes selecionados dos músicos 1 e 2.

<b>Eixos</b>	<b>x</b>		<b>y</b>		<b>z</b>	
	<b>x-Covar</b>	<b>Lag</b>	<b>x-Covar</b>	<b>Lag</b>	<b>x-Covar</b>	<b>Lag</b>
<b>m1r1t1</b>	0,808	1	0,948	0	0,978	-1
<b>m2r1t2</b>	0,549	-3	0,876	-2	0,970	-2
<b>m1r2t1</b>	0,854	0	0,948	-1	0,968	-1
<b>m2r2t2</b>	0,447	0	0,956	1	0,989	1
<b>m1r3t1a</b>	0,952	0	0,899	1	0,973	0
<b>m2r3t1a</b>	0,787	-2	0,930	-2	0,955	-2
<b>m1r3t1b</b>	0,809	0	0,970	0	0,936	1
<b>m2r3t2b</b>	0,679	-1	0,658	-1	0,848	-1
<b>m1r3t1c</b>	0,719	-1	0,788	0	0,627	1
<b>m2r3t2c</b>	0,753	-1	0,738	-1	0,837	-1

Tabela 3 - RMSE e faixas máximas (por eixo) na comparação dos dados de aceleração de ambos os sistemas - takes selecionados dos músicos 1 e 2.

Eixos	x			y			z		
	RMSE	$\Delta$ Max		RMSE	$\Delta$ Max		RMSE	$\Delta$ Max	
	(m/s <sup>2</sup> )	Mocap	IMU	(m/s <sup>2</sup> )	Mocap	IMU	(m/s <sup>2</sup> )	Mocap	IMU
<b>m1r1t1</b>	0,98	5,32	7,40	0,66	10,83	10,73	0,82	16,19	15,40
<b>m2r1t2</b>	0,63	2,41	3,64	1,43	10,33	11,87	0,91	13,53	14,07
<b>m1r2t1</b>	0,61	6,54	6,77	0,77	12,85	12,88	0,72	15,57	15,83
<b>m2r2t2</b>	1	4,73	5,04	1,03	12,71	13,61	0,83	17,74	17,82
<b>m1r3t1a</b>	0,49	8,23	8,15	0,89	14,05	10,02	0,33	8,65	9,45
<b>m2r3t1a</b>	0,99	7,45	7,13	0,98	14,79	9,06	0,72	11,09	11,11
<b>m1r3t1b</b>	3,46	22,72	13,05	1,34	13,75	13,07	1,59	10,85	13,08
<b>m2r3t2b</b>	2,27	11,69	10,76	2,33	9,87	12,15	1,98	15,98	13,22
<b>m1r3t1c</b>	1,73	14,30	6,96	1,17	10,09	11,9	1,05	7,37	6,11
<b>m2r3t2c</b>	1,55	11,28	11,69	2,19	12,14	17,26	1,35	13,43	11,65

Para o método de integração, aplicado apenas aos ritmos 1 e 2, as Tabelas 4 e 5 mostram os resultados. As acelerações foram novamente rotacionadas de acordo com os ângulos fornecidos pelo IMU antes de serem integradas. Os deslocamentos medidos pelo Qualisys foram subtraídos de sua posição de repouso inicial. Podemos observar uma aproximação razoável (covariância acima de 0,8) para os eixos y e z no ritmo 1. Aqui, o pequeno *RMSE* no eixo x pode enganar a interpretação. Neste caso, as observações sobre a característica bastante plana do gesto do *strumming* também se

aplicam, e as diferenças no eixo x entre os dois músicos podem estar relacionadas a idiossincrasias técnicas.

Tabela 4 – Covariâncias (por eixo) entre dados posicionais de mocap e dados de IMU duplamente integrados - takes selecionados dos músicos 1 e 2.

Eixos	x		y		z	
	x-Covar	Lag	X-Covar	Lag	x-Covar	Lag
<b>(lóbulo mais forte)</b>						
<b>m1r1t2</b>	0,605	-2	0,820	-3	0,865	-3
<b>m2r1t1</b>	0,169	-12	0,841	-1	0,844	0
<b>m1r2t2</b>	0,116	-1	0,603	0	0,673	0
<b>m2r2t1</b>	0,107	-5	0,843	-1	0,840	-1
<b>(lóbulo mais fraco)</b>						
<b>m1r1t2</b>	0,313	-4	0,798	-3	0,875	-3
<b>m2r1t1</b>	0,258	-12	0,807	-1	0,872	0
<b>m1r2t2</b>	0,451	0	0,526	-1	0,672	0
<b>m2r2t1</b>	0,138	-6	0,777	-1	0,856	-1

Tabela 5 - RMSE e faixas máximas (por eixo e método) na comparação dos dados de deslocamento de ambos os sistemas - takes selecionados dos músicos 1 e 2.

Eixo	x			y			z		
Take	RMSE (cm)	$\Delta$ Max		RMSE (cm)	$\Delta$ Max		RMSE (cm)	$\Delta$ Max	
		Mocap	IMU		Mocap	IMU		Mocap	IMU
<b>(lóbulo mais forte)</b>									
<b>m1r1t2</b>	0,93	1,842	2,64	1,62	3,48	6,55	2,56	6,96	9,75
<b>m2r1t1</b>	0,67	0,882	3,43	1,94	4,75	7,49	1,29	7,94	8,86
<b>m1r2t2</b>	3,23	2,926	16,56	2,89	5,86	15,9	2,62	8,28	16,76
<b>m2r2t1</b>	1,96	2,210	9,11	1,31	5,13	8,91	3,08	9,17	10,67
<b>(lóbulo mais fraco)</b>									
<b>m1r1t2</b>	0,89	1,842	1,87	1,52	3,48	2,81	2,48	6,96	5,81
<b>m2r1t1</b>	0,26	0,882	0,98	1,44	4,75	3,97	1,11	7,94	5,58
<b>m1r2t2</b>	1,21	2,926	2,55	2,37	5,86	3,18	1,96	8,28	5,90
<b>m2r2t1</b>	1,20	2,210	1,68	0,95	5,13	3,78	3,13	9,17	6,12

Fonte: Autoria própria