

# V-Conga: Um Estudo de Caso Para Avaliar as Capacidades de Captura de Nuances Gestuais do Leap Motion® em um Sistema Híbrido

Eduardo S. Silva<sup>1</sup>, Giordano R. E. Cabral<sup>1</sup>, Geber L. Ramalho<sup>1</sup>

<sup>1</sup>Centro de Informática – Universidade Federal de Pernambuco (UFPE) Caixa Postal 50.740-560- Recife – PE – Brasil.

{ess, grec, glr}@cin.ufpe.br

**Abstract.** *Gesture interfaces have been expanding the possibilities of creating new Digital Musical Instruments (DMIs). The Leap Motion® Controller offers a fine-grained hand tracking capabilities. In this study, the gesture recognition precision and the system's latency are evaluated by simulating a conga drum. In addition, we present an experiment done with a professional musician to investigate the sensor's capabilities within the instrument's context of use.*

**Resumo.** *Interfaces gestuais têm expandido as possibilidades de criação de novos Instrumentos Musicais Digitais (DMIs). O Leap Motion® Controller oferece um rastreamento de mãos com alta granularidade. Neste estudo, avalia-se a precisão do reconhecimento de gestos e latência do sistema ao simular uma conga. Em seguida, apresentamos um experimento feito com um músico profissional para investigar as capacidades do sensor dentro do contexto de uso do instrumento.*

## 1. Introdução

Nos últimos anos, o surgimento de dispositivos como smartphones, tablets e controles de videogame influenciou o rompimento dos paradigmas de Interação Humano-Computador (HCI) consolidados ao longo do tempo, afastando-se das interfaces baseadas no paradigma da máquina de escrever. Essas interfaces gestuais despertaram o interesse da comunidade científica pelas possibilidades de aplicações em diversas áreas de pesquisa, como é possível encontrar na literatura.

Na área de computação musical, esses avanços possibilitam novas interações com a música e a criação dos chamados Instrumentos Digitais Musicais (DMIs). Diversos instrumentos têm sido desenvolvidos utilizando diferentes controles de movimento [Jordà et al., 2005; Silva, 2012; Rosa-Pujazón et al., 2013].

Apesar do crescente número de pesquisas e demonstrações (formais e informais<sup>1</sup>), poucos DMIs estão sendo utilizados por músicos profissionais. Limitações de algumas tecnologias dificultam o uso profissional [Rosa-Pujazón et al., 2013; Todoroff, Leroy and Picard-Limpens, 2011]. Pesquisas indicam problemas relativos a latência excessiva, ausência de *feedback* visual/háptico e carência de nuances gestuais (problemas relativos à expressividade) [Silva, 2012].

Novas tecnologias mostram grande potencial para construção de DMIs expressivos. Dispositivos como o Leap Motion Controller® oferecem um rastreamento de gestos com uma alta precisão.

---

<sup>1</sup> Demonstrações disponibilizadas em websites de compartilhamento de vídeo, e.g. Youtube e Vimeo.

Avaliações preliminares do dispositivo [Han and Gold, 2014; Silva et al., 2013] observaram diversas limitações. Porém, mostram o potencial de captura de nuances gestuais. Requer-se uma análise aprofundada para avaliar a captura dessas nuances em instrumentos que demandam maior precisão espacial, principalmente após as mudanças no algoritmo de rastreamento do Leap Motion© (SDK v2.0). Outras pesquisas também utilizam simulação de instrumentos de percussão para avaliar tecnologias [Collicutt, Casciato, and Wanderley, 2009].

Neste estudo, foi desenvolvido um DMI (V-Conga) que simula um instrumento de percussão. Utilizou-se os princípios de integração de múltiplos sensores, baixo custo de implementação e um processo de *design* centrado no usuário.

## 2. Escolha do Instrumento

As congas foram escolhidas por possuírem um vasto repertório de gestos e nuances gestuais, por terem características físicas e de geração sonora similares à diversos outros instrumentos (bongôs, atabaques, etc.) e por serem e utilizadas em diversos ritmos musicais.

Os **Gestos Específicos** analisados produzem as notas: **Open Tone**, **Muffle Tone** (*Open Tone* abafado), **Slap Tone**, **Closed Slap Tone** (*Slap Tone* abafado), **Pressed Slap Tone** (*Slap Tone* abafado com a outra mão), **Bass Tone**, **Heel Tone**, **Tip Tone**, **Gliss** (ou *Glissando*) e algumas nuances como os **Flams** e uma técnica de mão fechada em punho (que foi denominado **Fist Note**). As diferenças nos gestos estão relacionadas com a posição da mão no instrumento, o formato da mão e o ângulo de ataque entre a mão e a pele do instrumento.

## 3. Sistema

O sistema (Figura 1) conta com uma superfície de duas camadas de madeira revestida de um emborrachado preto (para evitar reflexão no Leap Motion©), nela estão conectados dois sensores (situados entre as duas camadas). O Leap Motion© fica acima do sistema, em um pedestal comum e rastreando a cena (de cima para baixo). Mudanças no SDK permitiram uma maior a estabilidade em outras orientações de rastreamento.



Figura 1: Configuração do protótipo versão 1.0.

Os outros sensores integrados ao sistema foram em sensor piezoelétrico, ou piezo, para melhorar a latência do sistema (captura o ataque do gesto através da vibração na superfície) e um sensor de pressão FSR (*Force Sensing Resistor*), que funciona como um resistor variável, para captar nuances de abafamento das notas do instrumento (pela variação da resistência). Ambos os sensores foram integrados por meio de uma placa Arduino.

## 4. Resultados

Os aspectos focados foram a latência e o reconhecimento dos gestos específicos da conga. Além disso, foi feito um experimento com um músico profissional utilizando uma abordagem qualitativa [Barbosa et al., 2011]. Os valores apresentados representam a média de 30 medidas (para a latência) e 60 medidas (para cada gesto da conga).

Na **avaliação da latência** (Tabela 1), utilizou-se um microfone comum para capturar os sons do ataque na superfície e do *samples* disparados. A latência foi aprimorada ao longo da construção do artefato, escolhendo novas abordagens e implementando novos sensores.

**Tabela 1: Latência do artefato desenvolvido.**

Versão	Crystal Piano (Java)	0.1	0.2 <sup>2</sup>	1.0 <sup>3</sup>	V-Conga
Latência (ms)	150	71	17	7	7

A **avaliação dos gestos** (Tabela 2) ocorreu em duas etapas: com “**uma mão**” e depois com “**duas mãos**” dentro do campo de visão do sensor. Na segunda etapa, cada gesto foi analisado com cada uma das mãos. Erros são notas falso-negativas e disparo de *samples* incorretos.

**Tabela 2: Reconhecimento dos gestos da conga com apenas uma mão.**

Notas	Primeira Etapa (Percentagem Erro)		Segunda Etapa (Percentagem Erro)	
	Mão 1	Mão 2	Mão 1	Mão 2
Bass	6,67 %		100 %	100 %
Open	1,67 %		3,33 %	95 %
Muff	0 %		0 %	91,67 %
Slap	1,67 %		3,33 %	93,33 %
Closed Slap	0 %		8,33 %	86,67 %
Pressed Slap	-		0 %	91,67 %
Heel	0 %		0 %	95 %
Tip	3,33 %		0 %	100 %
Fist	0 %		0 %	98,33 %

Com uma mão, todos os gestos são capturados com poucos erros. Flutuações podem afetar o reconhecimento dos gestos (e.g. as notas *Tip* e *Bass*). Maior tolerância no rastreamento pode resolver o problema. Na segunda etapa, quase todos os gestos foram reconhecidos com estabilidade na “mão 1”, enquanto que a “mão 2” apresentou diversas falhas, por oclusão e perdas de rastreamento da segunda mão. A baixa robustez apresentada pelo Leap Motion®, na orientação “de cima para baixo” e rastreando as duas mãos, pode alterar os parâmetros de descrição da cena, causando erros. Apesar dos erros, o sistema foi capaz de detectar a maioria dos gestos e nuances do instrumento. Em contrapartida, a presença de erros, mesmo que poucos, pode tornar o instrumento imprevisível e inviável no contexto profissional.

O Leap Motion® não foi capaz de capturar algumas nuances de expressividade, como o abafamento das notas, necessitando que outro sensor fosse integrado. Entretanto, outros parâmetros podem ser utilizados para isso (mão aberta/fechada).

<sup>2</sup> Piezo integrado ao sistema desenvolvido em puredata.

<sup>3</sup> Sistema portado para C/C++. Driver ASIO integrado ao sistema.

Outras nuances não foram capturadas devido à arquitetura do sistema e às limitações do campo de visão do sensor (*Glissando* e abafamento com o cotovelo, respectivamente).

**O experimento com o músico profissional**, de avaliação qualitativa, envolve 3 etapas: exploração livre, realização de tarefas musicais (ritmos simples) e entrevista semiestruturada.

A instabilidade do sistema dificultou a execução das tarefas pelo usuário, durante a segunda etapa, salvo naquela realizada usando apenas uma mão. Segundo o músico o DMI possui pouca tolerância a erros de execução, o que não ocorre com o instrumento acústico. Há falhas se os gestos são executados com pouca precisão. Isso afeta seu raciocínio e gera a necessidade de ele reaprender o instrumento que já domina.

## 5. Conclusões e Trabalhos Futuros

Neste artigo, avaliou-se a capacidade do Leap Motion© de capturar nuances gestuais para construção de DMIs. O Leap Motion© foi essencial para o reconhecimento dos gestos e nuances das congas. O sistema conseguiu capturar a maioria das nuances, com baixa percentagem de erro, mostrando o potencial de captura desse dispositivo.

A instabilidade observada indica que novos avanços devem ser realizados para a construção de um DMI para músicos profissionais. Outras técnicas (e.g. visão computacional, aprendizagem de máquina) devem ser investigadas para solucionar o problema da segunda mão. Entretanto, a implementação de algoritmos mais complexos pode afetar a latência do sistema. Outra possibilidade é modificar a superfície de contato, para eliminar possíveis problemas de rastreamento. Ademais, pode-se usar dois sensores Leap Motion©. Outros experimentos podem ainda ser realizados com músicos profissionais.

## 6. Referências

- Barbosa, J., Calegario, F., Magalhães, F., Teichrieb, V., Ramalho, G. & Cabral, G. (2011). Towards an evaluation methodology for digital music instruments considering performer's view: a case study. In *Proceedings of 13th Brazilian Symposium on Computer Music*.
- Collicutt, M., Casciato, C., & Wanderley, M. M. (2009, June). From real to virtual: A comparison of input devices for percussion tasks. In *Proceedings of NIME* (pp. 4-6).
- Jordà, S., Kaltenbrunner, M., Geiger, G. and Bencina R. (2005). The reactable. *Proceedings of the international computer music conference (ICMC 2005)*, Barcelona, Spain, pages 579-582. CiteSeer.
- Rosa-Pujazón, A., Barbancho, I., Tardón, L. J., & Barbancho, A. M. (2013). Drum-hitting gesture recognition and prediction system using kinect. In *I simposio espanol de entrenamiento digital seed* (pp. 108-118).
- Silva<sup>1</sup>, E. S., de Abreu<sup>1</sup>, J. A. O., de Almeida<sup>1</sup>, J. H. P., Teichrieb, V., & Ramalho, G. L. (2013). A preliminary evaluation of the leap motion sensor as controller of new digital musical instruments.
- Silva, J. V. S., (2012). *Avaliando Interfaces Gestuais Para Prática de Instrumentos Virtuais de Percussão*. Dissertação de Mestrado. Universidade Federal de Pernambuco, Recife.