

Interação em ambientes de Realidade Virtual: comparação de usabilidade entre Kinect e Leap Motion

Djenifer Renata Pereira
Grupo de Pesquisa IMAGO
Departamento de Informática
Universidade Federal do Paraná
Email: drp17@inf.ufpr.br

Olga Regina Pereira Bellon
Grupo de Pesquisa IMAGO
Departamento de Informática
Universidade Federal do Paraná
Email: olga@ufpr.br

Abstract—Interaction is a very important feature in a variety of applications, especially when Virtual Reality (VR) is used. To meet this need, alternative methods such as gestural interface from 3D sensors can be a good proposition. Therefore, this work aims to compare the Kinect and Leap Motion sensors related to their usability in natural interfaces to VR environments. The comparison will be made from assessments on the recognition of gestures of these devices and the feasibility of their use in short and long movements. The main goal is to guide the decision of what is the best option to be used in a 3D Virtual Museum application.

Resumo—A interação é uma característica muito importante em diversas aplicações, principalmente quando a Realidade Virtual (RV) é usada. Para satisfazer essa necessidade, métodos alternativos, como interface gestual a partir de sensores 3D pode ser uma boa proposta. Por isso este trabalho visa comparar os sensores Kinect e Leap Motion em frente à sua usabilidade em interfaces naturais para ambientes RV. A comparação será feita a partir de avaliações sobre o reconhecimento de gestos desses dispositivos e a viabilidade de seu uso em movimentos curtos e longos. O objetivo final é guiar a decisão de qual dispositivo utilizar em uma aplicação de Museu Virtual 3D.

I. INTRODUÇÃO

A interação de qualquer aplicação é muito importante para o seu funcionamento, pois é sua forma de comunicação com o usuário. A interação deve ser bem projetada para evitar os pontos negativos da experiência do usuário e melhorar os pontos positivos, desta forma as aplicações se tornam mais fáceis, efetivas e agradáveis de usar [1]. Isto é particularmente verdade com aplicações que utilizem Realidade Virtual (RV). A Interação Humano-Computador (IHC) é uma das maiores dificuldades em sistemas RV, no entanto a utilização de Interface Natural do Usuário (*Natural User Interface* – NUI) podem ajudar a superar esses problemas [2], [3].

Uma NUI é uma interface que possibilita a interação de forma intuitiva sem precisar de grandes esforços. Ela deve ser fácil de aprender para que exista uma comunicação espontânea entre o usuário e o sistema, ou seja, o usuário deve conseguir agir de forma **natural** perante a aplicação [3]–[5].

As interfaces gestuais são uma entre muitas categorias de NUIs. Para viabilizar seu uso é preciso rastrear e reconhecer gestos feitos pelo usuário para se comunicar com

a aplicação. Assim, o objetivo deste trabalho é apresentar os sensores Kinect e Leap Motion e compará-los em frente a sua usabilidade como interface natural para interação em ambientes de Realidade Virtual. Para isso, serão observados artigos que pesquisaram sobre reconhecimento de gestos com esses dispositivos.

A motivação para realizar esta pesquisa foi a procura de melhores formas de interação para o Museu Virtual 3D do IMAGO. O museu utiliza de realidade virtual para apresentar as obras do escultor brasileiro Antônio Francisco de Lisboa, conhecido como Aleijadinho. Apesar da imersão ser grande com a utilização do HMD (*head mounted display*), a sensação de presença poderia ser melhorada com a troca de interação por controle por uma interação natural. Portanto, esta pesquisa guiará a escolha de um sensor para a implementação de uma NUI para o Museu do Aleijadinho.

Este trabalho está organizado da seguinte forma: a seção dois apresenta os dispositivos e descreve suas características; a seção três elenca os trabalhos que avaliaram os sensores em reconhecimento de gestos; a seção quatro compara o Leap Motion e o Kinect; e, a seção cinco fecha este trabalho com a conclusão.

II. OS SENSORES

O foco neste trabalho é a aplicabilidade dos sensores na Realidade Virtual, porém os dispositivos também podem ser usados em diversas áreas como preservação cultural e histórica, engenharia e construção, robótica, saúde, etc [6]–[10]. Esses sensores utilizam infravermelho para capturar a profundidade da imagem, porém cada um usa um mecanismo diferente que serão apresentados a seguir.

A. Leap Motion

O Leap Motion (figura: 1) é um dispositivo de baixo custo, categorizado como sistema óptico de rastreamento baseado em Visão Estéreo, que consiste da utilização de duas ou mais imagens capturadas por câmeras diferentes de um mesmo ponto para determinar a profundidade [11]. O sensor é pequeno e possui duas câmeras e três luzes infravermelhas. O seu alcance

de visão é limitado a 60 cm devido a forma de captura de imagens que utiliza o infravermelho [6].

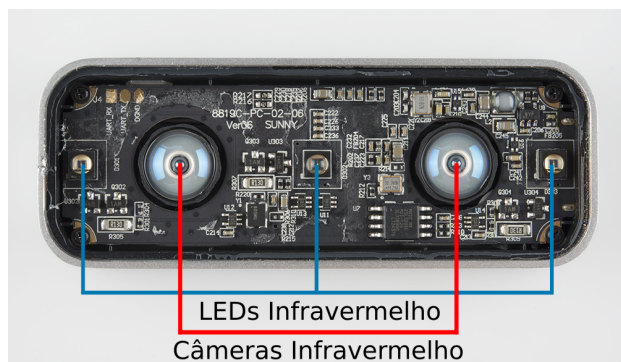


Figura 1. Leap Motion. (Adaptado de [12])

B. Kinect

O Kinect é um sensor que captura imagens de profundidade e cor. Desenvolvido pela Microsoft, este sensor possui duas versões, o Kinect v1 e o Kinect v2 (figuras 2a e 2b respectivamente). Ambas as versões possuem uma câmera RGB (Red-Green-Blue), um conjunto de microfones e o sensor de profundidade composto por um emissor e uma câmera infravermelhos. Para determinar a profundidade da imagem, o Kinect v1 projeta padrões de infravermelho ao que está a sua frente e a partir das distorções aplicadas a esses padrões determina a profundidade. Já a segunda versão utiliza o princípio de Tempo de Voo, que mede a profundidade a partir do tempo que a luz emitida leva para ir ao objeto e voltar [11], [13]–[15].

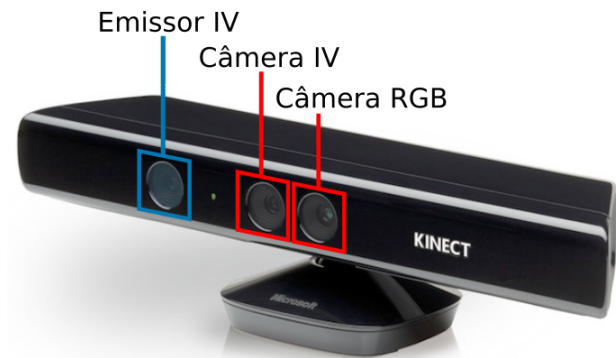
III. RECONHECIMENTO DE GESTOS

Alguns pesquisadores avaliaram o Kinect e o Leap Motion em relação ao reconhecimento de gestos. Essas pesquisas serão de grande importância para a comparação que será feita na Seção IV.

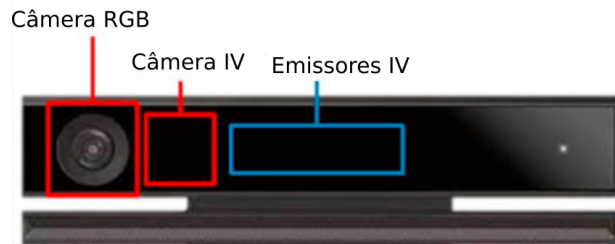
Potter, Araullo e Carter [18] pesquisam a viabilidade de utilizar o Leap Motion no reconhecimento da Língua de Sinais Australiana (Auslan). Para a avaliação foi usado o alfabeto Auslan para testar o funcionamento do dispositivo. Os resultados mostraram que, mesmo conseguindo reconhecer dedos sozinhos, o Leap Motion não seria viável para reconhecer o Auslan devido a interpretação incorreta de dedos juntos e rotações relativas ao centro da palma da mão.

Marin, Dominio e Zanuttigh [19] apresentam os dispositivos Kinect e Leap Motion e os avalia perante algoritmos diferentes de reconhecimento de gestos. O estudo apontou que ambos os dispositivos alcançaram uma boa acurácia, aproximadamente 80% e concluiu que juntar os dois dispositivos traria uma acurácia ainda melhor, mais de 90%.

O estudo de Penelle e Debeir [20] propõe a calibragem do Leap Motion e do Kinect para rastrear a mão. Foram marcados pontos de um alvo para serem rastreados nos dois aparelhos e eles retornarem as coordenadas para fazer a calibragem.



(a) Primeira versão. (Adaptado de [16])



(b) Segunda versão. (Adaptado de [17])

Figura 2. Sensor Kinect

Os resultados mostraram que, como cada sensor captura as imagens de forma diferente, um dispositivo pode compensar o outro em alguma falha.

O trabalho de Lan, Li e Ju [21] propõe um modelo de reconhecimento de gestos em tempo real que utiliza dados cruzados usando o Kinect v2. A partir das informações de profundidade obtidas e dados do esqueleto (estruturas 2D após o processo de afinamento), foram feitos testes com usuários diferentes para calcular o tempo que o método proposto consome. Os resultados apontaram que esse modelo é 18% mais rápido que os métodos do estado-de-arte.

Kim e Lee [22] projetam um método para reconhecer gestos a menos de 0.5 metros do Kinect v2, ou seja, fora do seu alcance de detecção. A detecção foi feita a partir de uma característica vinda da limitação do sensor, uma área branca. Já o reconhecimento dos gestos foram feitos desenhando um círculo preto no centro da mão e contando as áreas isoladas. Este método difere dos demais por conta da simplicidade em rastrear os dedos.

IV. COMPARAÇÃO DOS DISPOSITIVOS

Para fazer a comparação dos dispositivos é preciso considerar a particularidade de cada um, pois uma característica pode estar presente em um sensor e no outro não. Por isso, foram levantados dados sobre tais características [11], [15], [24]–[27] ou a falta delas para elaborar a tabela I que apresenta um comparativo dos sensores Kinect v1, v2 e do Leap Motion com a forma de captura das imagens, a resolução das imagens obtidas, o que rastreia e quanto.

A usabilidade dos sensores em ambientes virtuais pode variar conforme ao tipo de interface que pretende-se utilizar.



Figura 3. Movimentos para a interface. (Adaptado de [23])

Tabela I
COMPARATIVO DOS DISPOSITIVOS

	Kinect v1	Kinect v2	Leap Motion
Técnica de captura de imagem	Triangularização	Tempo de Voo	Visão Estéreo
Imagem colorida (px)	640 x 480	1920 x 1080	–
Imagem de profundidade (px)	640 x 480	512 x 424	–
Imagem em escala de cinza (px)	–	–	640 x 240
Alcance (cm)	40 a 4500	50 a 4500	2.5 a 60
Comunicação	USB 2.0	USB 3.0	USB 2.0
Rastreo de dedos	–	X	X
Rastreo de mãos	X	X	X
Rastreo de corpo	X	X	–
Esqueletos rastreados	2	6	–

Um exemplo é o trabalho de [23] que propôs uma interface baseada em gestos para uma visita virtual. Para escolher os gestos a serem utilizados foram consideradas as sugestões dos usuários, e os obtidos foram: movimentos pequenos, com variação em relação aos dedos (Figura 3a); movimentos médios, com variação em relação a um braço (Figura: 3b); e, movimentos grandes, com variação dos dois braços (Figura: 3c). Portanto os sensores serão avaliados pela capacidade de captura dos movimentos dos dedos, das mãos e dos braços para as interações em ambientes de realidade virtual.

O Leap Motion é um dispositivo de alta acurácia [11]. O reconhecimento dos dedos esparsos é bom, porém se os dedos estiverem juntos ou bem próximos, algum dedo pode sumir do reconhecimento. Na movimentação em si das mãos, não foi encontrado problema, com exceção de rotações que provoquem a obstrução do campo de visão do sensor. Devido ao campo de visão, o Leap Motion não suporta movimentos grandes, ou seja, com ambos os braços.

No Kinect v1 o reconhecimento de dedos não é muito viável devido aos dados do esqueleto rastreado não ter suporte para reconhecimento de dedos. No entanto, esta versão do kinect possui o maior alcance entre todos os demais sensores, o que possibilita variação da posição do usuário no seu espaço físico sem perder o rastreamento.

O kinect v2 consegue capturar todos os tipos de movi-

mentos (pequenos, médios e grandes). Alguns sinais possuem indicação própria durante o reconhecimento. São desenhados círculos em volta da mão do esqueleto e a cor do círculo indica que tipo de gesto o usuário está fazendo. O círculo verde indica mão aberta, o círculo vermelho indica mão fechada e o círculo azul indica dois dedos levantados (figuras: 4a, 4b e 4c respectivamente).



Figura 4. Gestos pré-definidos para reconhecimento do Kinect v2. (Adaptado de [28] e [29])

Logo, para interfaces de movimentos longos ou grandes

o Kinect é a opção a ser escolhida, seja a primeira ou a segunda geração. Já para movimentos pequenos e próximos, o Leap Motion é o mais adequado devido a sua acurácia. Porém, o uso do Kinect v2 também é viável para captura de movimentos pequenos, mas não para movimentos próximos devido o seu alcance. Para movimentos médios, ou seja, variação de movimentos feitos com a mão, qualquer um dos sensores podem ser escolhidos.

Os movimentos que serão usados para a interação no Museu do Aleijadinho serão variados (dedos, mãos e braços) devido a presença das obras e de seus tamanhos distintos. Por exemplo, as Figuras de Presépio são obras de mesa que não chegam a metade da altura do Anjo Tocheiro que tem aproximadamente 170 cm [30]. Portanto, o Kinect v2 é o mais viável para o museu, já que o rastreo engloba todos os movimentos necessários.

V. CONCLUSÃO

Neste trabalho foram apresentados os sensores Leap Motion e Kinect como uma opção para interface gestual. Para poder avaliá-los foram levantados algumas informações sobre como os dispositivos capturam imagens e como são utilizados no reconhecimento de gestos. Os resultados mostraram que ambos possuem suas vantagens e desvantagens, porém o Kinect v2 se sobressai devido a grande gama de possibilidades que abrange, já que o dispositivo consegue rastrear movimentos de dedos e mãos com um alcance maior que o do Leap Motion.

Um possível trabalho futuro é implementar a calibragem do Kinect v2 com o Leap Motion, já que este tipo de fusão de dados só foi realizado com a primeira versão do Kinect e o Leap Motion. A combinação dos dois poderão resultar em dados mais precisos para a interação, proporcionando mais naturalidade.

REFERÊNCIAS

- [1] Y. Rogers, H. Sharp, and J. Preece, *Interaction design: beyond human-computer interaction*, 4th ed. John Wiley & Sons, 2015, ch. 1, p. 2, (Seção I).
- [2] A. Jaimes and N. Sebe, "Multimodal human-computer interaction: A survey," *Computer vision and image understanding*, vol. 108, no. 1-2, pp. 116–134, 2007, (Seção I).
- [3] M. Roupé, P. Bosch-Sijtsema, and M. Johansson, "Interactive navigation interface for virtual reality using the human body," *Computers, Environment and Urban Systems*, vol. 43, pp. 42–50, 2014, (Seção I).
- [4] S. Stannus, D. Rolf, A. Lucieer, and W. Chinthammit, "Gestural navigation in google earth," in *Proceedings of the 23rd Australian Computer-Human Interaction Conference*. ACM, 2011, pp. 269–272, (Seção I).
- [5] D. Wigdor and D. Wixon, *Brave NUI world: designing natural user interfaces for touch and gesture*. Elsevier, 2011, ch. 1–2, (Seção I).
- [6] S. G. Barsanti, G. Caruso, L. Micoli, M. C. Rodriguez, and G. Guidi, "3d visualization of cultural heritage artefacts with virtual reality devices," *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 5, p. 165, 2015, (Seção II).
- [7] A. Boyali, N. Hashimoto, and O. Matsumoto, "Hand posture control of a robotic wheelchair using a leap motion sensor and block sparse representation based classification," in *SMART 2014, The Third International Conference on Smart Systems, Devices and Technologies*, 2014, pp. 20–25, (Seção II).
- [8] B. J. Fernandez-Palacios, D. Morabito, and F. Remondino, "Access to complex reality-based 3d models using virtual reality solutions," *Journal of cultural heritage*, vol. 23, pp. 40–48, 2017, (Seção II).
- [9] T. Hilfert and M. König, "Low-cost virtual reality environment for engineering and construction," *Visualization in Engineering*, vol. 4, no. 1, p. 2, 2016, (Seção II).
- [10] M. Iosa, G. Morone, A. Fusco, M. Castagnoli, F. R. Fusco, L. Pratesi, and S. Paolucci, "Leap motion controlled videogame-based therapy for rehabilitation of elderly patients with subacute stroke: a feasibility pilot study," *Topics in stroke rehabilitation*, vol. 22, no. 4, pp. 306–316, 2015, (Seção II).
- [11] F. Weichert, D. Bachmann, B. Rudak, and D. Fisseler, "Analysis of the accuracy and robustness of the leap motion controller," *Sensors*, vol. 13, no. 5, pp. 6380–6393, 2013, (Seção II e Seção IV).
- [12] SFUptownMaker, "Leap motion teardown," <https://learn.sparkfun.com/tutorials/leap-motion-teardown>, Data de acesso: 13 de junho de 2018. (Seção II).
- [13] A. Kolb, E. Barth, R. Koch, and R. Larsen, "Time-of-flight sensors in computer graphics," in *Eurographics (STARs)*, 2009, pp. 119–134, (Seção II).
- [14] N. Silberman and R. Fergus, "Indoor scene segmentation using a structured light sensor," in *Computer Vision Workshops (ICCV Workshops)*, 2011 *IEEE International Conference on*. IEEE, 2011, pp. 601–608, (Seção II).
- [15] O. Wasenmüller and D. Stricker, "Comparison of kinect v1 and v2 depth images in terms of accuracy and precision," in *Asian Conference on Computer Vision*. Springer, 2016, pp. 34–45, (Seção II) e (Seção IV).
- [16] S. Dent, "Wave goodbye to microsoft's original kinect for windows," <https://www.engadget.com/2014/12/31/original-kinect-discontinued/>, Data de acesso: 14 de junho de 2018. (Seção II).
- [17] L. Caruso, R. Russo, and S. Savino, "Microsoft kinect v2 vision system in a manufacturing application," *Robotics and Computer-Integrated Manufacturing*, pp. 174–181, 2017, (Seção II).
- [18] L. E. Potter, J. Araullo, and L. Carter, "The leap motion controller: a view on sign language," in *Proceedings of the 25th Australian computer-human interaction conference: augmentation, application, innovation, collaboration*. ACM, 2013, pp. 175–178, (Seção III).
- [19] G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with leap motion and kinect devices," in *Image Processing (ICIP)*, 2014 *IEEE International Conference on*. IEEE, 2014, pp. 1565–1569, (Seção III).
- [20] B. Penelle and O. Debeir, "Multi-sensor data fusion for hand tracking using kinect and leap motion," in *Proceedings of the 2014 Virtual Reality International Conference*. ACM, 2014, p. 22, (Seção III).
- [21] Y. Lan, J. Li, and Z. Ju, "Data fusion-based real-time hand gesture recognition with kinect v2," in *Human System Interactions (HSI)*, 2016 *9th International Conference on*. IEEE, 2016, pp. 307–310, (Seção III).
- [22] M.-S. Kim and C. H. Lee, "Hand gesture recognition for kinect v2 sensor in the near distance where depth data are not provided," *International Journal of Software Engineering and Its Applications*, vol. 10, no. 12, pp. 407–418, 2016, (Seção III).
- [23] V. M. Manghisi, A. E. Uva, M. Fiorentino, M. Gattullo, A. Boccaccio, and G. Monno, "Enhancing user engagement through the user centric design of a mid-air gesture-based interface for the navigation of virtual-tours in cultural heritage expositions," *Journal of Cultural Heritage*, 2018, (Seção IV).
- [24] M. Szymczyk, "How does the kinect 2 compare to the kinect 1?" <https://zugara.com/how-does-the-kinect-2-compare-to-the-kinect-1>, Data de acesso: 09 de julho de 2018. (Seção IV).
- [25] Skarredghost, "The difference between kinect v2 and v1," <https://skarredghost.com/2016/12/02/the-difference-between-kinect-v2-and-v1/>, Data de acesso: 09 de julho de 2018. (Seção IV).
- [26] L. Motion, "API Overview – Leap Motion C# SDK v3.2 Beta documentation," https://developer-archive.leapmotion.com/documentation/csharp/devguide/Leap_Overview.html, Data de acesso: 16 de julho de 2018. (Seção IV).
- [27] Raymond Ma, "Leap motion as a rudimentary depth sensor," <http://blog.rymnd.com/leap-motion-depth/>, Data de acesso: 31 de julho de 2018. (Seção IV).
- [28] L. B. Life, "Learn ASL alphabet video," https://www.youtube.com/watch?v=6_gXiBe9y9A, Data de acesso: 18 de agosto de 2018. (Seção IV).
- [29] A. MEGDICHE, "Programming kinect v2 for windows tuto1," <https://www.youtube.com/watch?v=GpJS0SBtHwY>, Data de acesso: 18 de agosto de 2018. (Seção IV).
- [30] IMAGO, "Esculturas de Aleijadinho," <https://www.imago.ufpr.br/MuseuApp/>, Data de acesso: 22 de agosto de 2018. (Seção IV).