

# Computadores e Natural User Interface (NUI): Utilização do Microsoft Kinect em aplicações de NUI

Fernando Henrique Kreutz<sup>1</sup>, Gerson Battisti<sup>1</sup>

<sup>1</sup>Universidade Regional do Noroeste do Estado do Rio Grande do Sul (UNIJUI)  
Rodovia RS342, KM39 – 98900-000 – Santa Rosa – RS – Brasil

fernando.kreutz@me.com, battisti@unijui.edu.br

**Abstract.** *Microsoft Kinect, developed to be used alongside the Microsoft Xbox 360 videogame console, works as a natural motion sensor controller. With an depth sensor, this device is capable of tracking up to 2 active players with 20 joints each. It's launch in november 2010 caused a new revolution in the video-game market, with great ability to revolutionize the PC. This article describes the use and development of a NUI Microsoft Kinect enabled application on a Windows 7 PC.*

**Resumo.** *O Microsoft Kinect, desenvolvido para utilização no Microsoft Xbox 360, funciona como um controlador por movimentos naturais. Portando um sensor de profundidade, este dispositivo é capaz monitorar 2 pessoas com até 20 articulações cada. Seu lançamento em novembro de 2010 causou uma revolução no mercado de video-games, com grande capacidade de revolucionar o PC. Este artigo descreve a utilização e desenvolvimento de uma aplicação com o uso do Microsoft Kinect em aplicações NUI em um PC com Microsoft Windows 7.*

## 1. Introdução

A interação entre o homem e máquina é imprescindível para o funcionamento de um equipamento. Essa interação ocorre o tempo todo e estamos cercados dela. Nos últimos anos, a interação com os dispositivos eletrônicos tem ganhado evidência, mais especificamente os que tem como característica a interação natural, como dispositivos multi-touch. Michael destaca:

“Touch (e mais importante, multi-touch) como uma interface ganhou popularidade com os dispositivos móveis da Apple (iPhone, iPod Touch, iPad), que só entraram em cena a pouco mais de 4 anos. Até então, não haviam dispositivos direcionados ao consumidor dignos de uso baseados em toque. O movimento é similar, com o desenvolvimento inicial em tecnologia de movimento (interação natural) ligadas ao mercado de videogames. O Nintendo Wii era a primeira geração desta tecnologia, lançada para o natal de 2006. A Microsoft forneceu a próxima geração de interfaces de movimento com sua tecnologia Kinect no ano passado, que usa a metodologia de NUI.” [Fauscette 2011]

## 2. Natural User Interface - NUI

Daniel Wigdor e Dennis Wixon em seu livro NUI World destacam:

“Ironicamente, uma interface natural não ocorre naturalmente. Em nossa visão, criar uma NUI é uma meta de design. Alcançar esta meta requer um ponto de vista ilimitado, muito trabalho, design minucioso, testes rigorosos, e um pouco de sorte. O ponto de vista ilimitado começa com o entendimento e visão do que é uma NUI. Nossa visão é que uma NUI é a interface que fornece uma experiência ilimitada e agradável no decorrer de seu uso. Isso faz com que uma utilização qualificada pareça natural, tanto para aprender como para usar. Faz com que o aprendizado seja agradável e elimina as porcaria que distraem o usuário com experiência. [...] O significado é melhor entendido quando comparado à frase “aquela pessoa é tão natural”. Quando ouvimos alguém se referir desta maneira, temos a sensação de que sua atuação é ideal e que parece sem esforço e graciosa.” [Wigdor 2011]

Na interação natural com dispositivos, como Daniel [Wigdor 2011] descreve, natural refere-se ao comportamento e sentimentos do usuário durante uma experiência ao invés da interface ser o produto de algum processo orgânico.

## 3. Kinect

O Microsoft Kinect, periférico para Xbox 360, apresentado pela Microsoft na E3 Conference de Julho de 2009 sob o codinome “Project Natal”, foi desenvolvido pela Rare, uma subsidiária da Microsoft Game Studios.

O Kinect é um periférico relativamente pequeno, mas extremamente complexo. Internamente [Marshall 2010], ele possui um projetor infravermelho, que utiliza-se da tecnologia proprietária LightCoding para rastreamento 3D, uma câmera infravermelha — que em conjunto ao projetor infra-vermelho compõe o sistema de sensoriamento de profundidade (depth sensing) — além de uma camera RGB, esta por sua vez opcional no sistema, e um arranjo de 4 microfones.

Mikkel Viager explica como funciona a tecnologia de sensoriamento de profundidade com Light Coding:

“Muitas tecnologias de busca de distância usam o tempo-de-vôo para determinar a distância de um objeto ao medir o tempo que leva para um raio de luz viajar e refletir devolta da superfície de um objeto. Light Coding usa uma abordagem totalmente diferente onde a fonte de luz está constantemente ligada, diminuindo bastante a necessidade de tempo de precisão para as medições. Uma fonte laser emite luz invisível (aproximadamente no comprimento de onda do infra-vermelho) que passa através de um filtro e é dispersa em um padrão semi-randômico, mas constante, de pequenos pontos que são projetados no ambiente à frete do sensor. O padrão refletido é então detectado por uma camera infra-vermelha e analisado.” [Viager 2011]

Para entender melhor como essa tecnologia funciona, basta observar a figura 1. A camera RGB não é necessária para a detecção de objetos e indivíduos no espaço 3D, sendo necessários apenas o projetor infra-vermelho, representado pelo número 1 e a camera que irá capturar o padrão de pontos projetados, representado pelo número 2.



**Figura 1. Tecnologia de processamento de imagem. [Primesense 2011]**

Os dados capturados são então processados pelo PS1080, resultando em uma imagem de profundidade, onde, na figura, diferentes indivíduos são representados por cores diferentes – de quentes para frias – de acordo com a distância do sensor – próximos e afastados.

#### **4. Kinect no PC**

Aplicações que utilizam o Kinect em computadores já são amplamente divulgadas em diversas comunidades e em websites especializados, como o <http://kinect.dashhacks.com> e <http://www.kinecthacks.com>, onde são divulgados softwares tanto desenvolvidos com o SDK oficial como softwares desenvolvidos com APIs de terceiros.

Os softwares divulgados variam desde softwares que adicionam objetos em uma cena real [Kinect Hacks 2012], capturada pela câmera, até softwares mais engenhosos que controlam robôs [Kinect Hacks 2012] e softwares de construção de modelos 3D a partir de objetos reais [Kinect Hacks 2012], capturados pela câmera do Kinect.

Aplicações como a de construção de modelos 3D a partir de objetos reais, seria extremamente cara, utilizando diversos equipamentos de alta complexidade, não fosse pelo Kinect, que possibilita que um objeto real seja escaneado para um ambiente virtual tridimensional utilizando apenas um sensor e uma mesa giratória, o que se traduz automaticamente em baixo custo.

Uma das maiores comunidades, o Open Kinect, é uma comunidade aberta para pessoas interessadas em utilizar o Kinect no PC, com bibliotecas open source disponíveis para Windows, Linux e Mac [Open Kinect 2012]. Outra grande comunidade, o CODING4FUN da Microsoft, é um HUB de projetos desenvolvidos com base no Kinect, com artigos, how to's e um fórum, permitindo um grande nível de interação entre os

interessados, onde os projetos são postados com código fonte aberto, permitindo o compartilhamento de conhecimento na comunidade [Coding4Fun 2012].

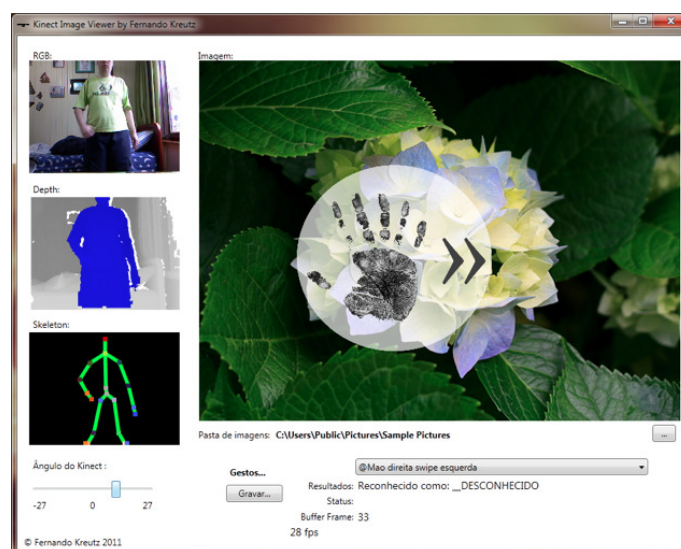
Recentemente a Microsoft anunciou o Kinect para PC, uma versão em que as lentes são ajustadas para focar em menor distância [PC Gamer 2012]. Também foi anunciado o portal Kinect Effect, em que a Microsoft mostra projetos que foram feitos utilizando o Kinect e que revolucionaram a maneira como o usuário cumpre tal tarefa. É dado grande destaque à aplicabilidade na área de saúde, onde um cirurgião pode interagir, por exemplo com exames do paciente, apenas com os gestos das mãos [Kinect Effect 2012].

#### 4.1. Prova de conceito

Durante o desenvolvimento, foram geradas várias versões do software, sempre aprimorando-se a funcionalidade do mesmo. Chamada de Kinect Image Viewer, a versão final é capaz de gravar e reconhecer gestos, estes utilizados para navegar pelas imagens de uma pasta, determinada pelo usuário.

A seção de software responsável pelo reconhecimento de gestos foi obtida através do software open source “Kinect SDK Dynamic Time Warping (DTW) Gesture Recognition” desenvolvido pelos usuários Rhemyst e Rymx do fórum MSDN. O software com código fonte em C# e WPF pode ser baixado no CodePlex [Codeplex 2011].

Navegar pelas imagens é muito fácil, basta mover uma das mãos da direita para a esquerda que a imagem é avançada na lista, e da esquerda para a direita que a imagem seja retrocedida. Durante a troca de imagens, é dado um feedback ao usuário com a imagem de uma mão e a direção do gesto reconhecida.



**Figura 2. Gesto à direita reconhecido. [Arquivo Pessoal]**

O sistema de reconhecimento de gestos, necessita de 32 quadros de dados para gravar ou reconhecer um gesto, os quais ficam armazenados em um buffer e são constantemente checados para verificar o reconhecimento do gesto.

No arquivo de gestos, são armazenados blocos, que representam cada quadro de dados capturados (são 32 por gesto). Cada grupo de blocos é iniciado pelo nome do

gesto, com a sintaxe “nome\_de\_gesto”. Ao final de cada um dos 32 blocos por gesto, o til representa o fim do mesmo. Ao final do grupo de blocos do gesto, “- - -” encerra o gesto, sendo seguido dos dados de um novo gesto. Cada linha do bloco representa uma junta/posição do esqueleto monitorada.

## 5. Conclusão

A evolução da interação entre o homem e o dispositivo computacional mudou substancialmente nos últimos anos, mas promete gradativamente revolucionar a maneira como estamos acostumados a interagir com tais dispositivos.

Sem sombra de dúvidas, as interfaces naturais são o futuro dessas interações, das quais hoje já podemos ter exemplos do que pode nos esperar em alguns anos, o que nos parece fantástico e inimaginável hoje pode ser a tecnologia que utilizaremos em nosso dia-a-dia no futuro.

Um excelente exemplo de NUI, que podemos carregar em nossos bolsos, é o iPhone 4S, que trouxe consigo uma nova tecnologia de reconhecimento de voz, em que conversamos com o aparelho ao invés de dar simples comandos. Essa tecnologia é conhecida como Siri.

Como o aparelho é capaz de entender o contexto, ao utilizar palavras relacionadas a familiares, como mãe, esposa ou filho, o aparelho irá entender quem são essas pessoas, assim como o aparelho entende que trabalho é o local onde o usuário trabalha, previamente indexado no dispositivo. Isso faz com que a interação com essa tecnologia natural pareça mais como uma conversa com uma secretária do que com um dispositivo eletrônico.

Outro sistema natural de interação, a touchscreen, vem se popularizando de forma massiva no mercado, devido à maneira diferente e mais natural com que o usuário interage com o dispositivo. Ao invés de usar um dispositivo apontador, como o mouse, basta que o usuário toque nos ícones exibidos na tela. Sensores como o giroscópio e acelerômetro aumentam ainda mais a qualidade da interação natural, bastando girar o aparelho para que uma foto seja girada automaticamente, por exemplo.

Essas tecnologias naturais, antes limitadas pelo hardware com pouca capacidade de processamento, hoje podem ser aplicadas inclusive em dispositivos móveis, que carregamos no bolso, dispositivo móvel que facilmente possui mais poder computacional do que o melhor computador pessoal em 2000.

Interação natural com gestos é outra tecnologia que pode vir a revolucionar o mercado. Recentemente a Samsung, fabricante de dispositivos eletrônicos, colocou no mercado a Samsung Smart TV [Samsung 2012], que permite que o usuário interaja com o dispositivo utilizando gestos espaciais ou comandos de voz. O dispositivo também é capaz de reconhecer o usuário através do reconhecimento facial.

Uma possibilidade para o uso de NUI no âmbito comercial é o uso do Kinect para promover um produto. Por exemplo, supomos que uma grande loja de roupas utilize-se desta tecnologia para atrair consumidores. Bastaria então que o usuário se posicionasse à frente do dispositivo que o sistema poderia reconhecer e estimar suas medidas, oferecendo-lhe sugestões de compra. Ali mesmo o usuário poderia prová-las, através de um provador virtual.

A lista de possibilidades com tecnologias de interação natural é infinita, e com o nível de desenvolvimento atual e mais incentivo ao desenvolvimento de novas tecnologias do que nunca, acreditamos que o futuro das NUI's é brilhante.

## **Referências**

- Codeplex. Kinect SDK Dynamic Time Warping (DTW) Gesture Recognizer. Disponível em: (<http://goo.gl/oUazE>). Acessado em 20/09/2011.
- Coding4Fun. MSDN. Disponível em: (<http://goo.gl/RETHI>). Acessado em 12/07/2012.
- Fauscette, Michael. Evolution of the User Interface. Disponível em: (<http://goo.gl/07H3V>). Acessado em 22/11/2011.
- IHCUC - Interactive Human-Robot Collaboration Using Kinect. Kinect Hacks. Disponível em: (<http://goo.gl/0hyU5>). Acessado em 10/04/2012.
- Kinect Effect. Microsoft. Disponível em: (<http://goo.gl/xWorh>). Acessado em 12/07/2012.
- Marshall, Davie. Inside the Kinect hardware - 11/05/2010. Disponível em: (<http://goo.gl/C7OUA>). Acessado em 11/22/2011.
- Open Kinect. Disponível em: (<http://goo.gl/whB4e>). Acessado em 12/07/2012.
- Kinect on PC: more expensive, less useful, still exciting. PC Gamer. Disponível em: (<http://goo.gl/dQBzQ>). Acessado em 12/07/2012.
- Primesense. PrimeSensor Reference Design. Disponível em: (<http://goo.gl/qoo6V>). Acessado em 12/08/2011.
- Viager, Mikkel. Analysis of Kinect for mobile robots - Technical University of Denmark. Disponível em: (<http://goo.gl/7UmXo>). Acessado em 22/11/2011.
- Wigdor, Daniel e Wixon, Dennis. Brave NUI World: Designing natural user interfaces for touch and gesture IX. Elsevier Science (2011).
- Kinect Ninja Attack. Kinect Hacks. Disponível em: (<http://goo.gl/zuGI1>). Acessado em 10/04/2012
- Real-Time 3D Reconstruction with Kinect Turntable. Kinect Hacks. Disponível em: (<http://goo.gl/wRVOv>). Acessado em 10/04/2012.
- Samsung, Smart TV. Samsung do Brasil. Disponível em: (<http://goo.gl/sqUCY>) Acessado em: 11/07/2012.