

UNIVERSIDADE SAGRADO CORAÇÃO

FELIPE ANGELO DA COSTA PERIN

**USO DO KINECT NA INTERFACE NATURAL DE
USUÁRIO**

BAURU
2013

FELIPE ANGELO DA COSTA PERIN

**USO DO KINECT NA INTERFACE NATURAL DE
USUÁRIO**

Trabalho de Conclusão de Curso apresentado ao Centro de Ciências Exatas e Sociais Aplicadas como parte dos requisitos para obtenção do título de bacharel em Ciência da Computação, sob orientação da Prof.^a Dra. Patrícia Bellin Ribeiro.

BAURU
2013

FELIPE ANGELO DA COSTA PERIN

USO DO KINECT NA INTERFACE NATURAL DE USUÁRIO

Trabalho de Conclusão de Curso apresentado ao Centro de Ciências Exatas e Sociais Aplicadas como parte dos requisitos para obtenção do título de bacharel em Ciência da Computação, sob orientação da Prof.^a Dra. Patrícia Bellin Ribeiro.

Banca examinadora:

Prof.^a Dra. Patrícia Bellin Ribeiro
Universidade Sagrado Coração

Prof.^o M.Sc. Patrick Pedreira Silva
Universidade Sagrado Coração

Prof.^a Ms. Elaine Cecília Gatto
Universidade Sagrado Coração

Bauru, 12 de dezembro de 2013.

"Nunca deixe ninguém dizer que você não pode fazer alguma coisa. Se você tem um sonho, tem que correr atrás dele. As pessoas não conseguem vencer e, dizem que você também não vai vencer. Se quer alguma coisa, corre atrás". (À Procura da Felicidade, 2006)

RESUMO

O Kinect mudou a forma de jogar *videogame*, interagir com as pessoas, além de ajudar no tratamento de distúrbios mentais e até mesmo facilitou alguns procedimentos cirúrgicos. Essa poderosa ferramenta desenvolvida pela Microsoft revolucionou o modo não somente de interação, mas também todas as estruturas relacionadas a este procedimento. As pessoas atualmente estão acostumadas a utilizar outros dispositivos para se comunicar com o sistema operacional, mas pode-se verificar nesse estudo que com apenas um dispositivo é possível executar esse processo de forma natural. Em um futuro próximo, o meio de interação homem-máquina mais comum, será a Interface Natural de Usuário onde as pessoas se libertarão de sensores ou de *hardwares* intermediários que transmitem aos computadores o que o usuário deseja. Dessa forma, a tecnologia será capaz de “entender” os anseios das pessoas e executá-los com precisão e eficiência. Este projeto, portanto, traz uma maneira eficiente de manipular arquivos, conteúdos multimídia e programas de computador, utilizando a tecnologia do Kinect, ou seja, estes processos podem ser realizados através do movimento e gesto das mãos e comandos de voz. A aplicação possui uma interface simples onde o usuário cria uma independência no modo de manipular o sistema operacional e é possível ficar livre de “cliques”, teclas e toques, entrando assim, na nova era da interação em que apenas é necessária a linguagem corporal.

Palavras-chave: Interface Natural de Usuário. NUI. Kinect. Interação Natural.

ABSTRACT

Kinect has changed the way we play video games, interact with people, and help in the treatment of mental disorders and even facilitated some surgical procedures. This powerful tool developed by Microsoft revolutionized not only the way of interaction, but also all related to this procedure structures. People are currently accustomed to use other devices to communicate with the operating system, but can be seen in this study that with only one device can run this process naturally. In the near future, the means of man-machine interaction more common, is the Natural User Interface where liberate people from sensors or hardware intermediaries that transmit to computers what the user wants. Thus, the technology will be able to "understand" the aspirations of the people and execute them with accuracy and efficiency. This design, however, provides an efficient way to manipulate files, multimedia content and computer programs, using the technology of Kinect, i.e., these processes can be performed through movement and hand gesture and voice commands. The application has a simple interface where the user creates independence in order to manipulate the operating system and you can get free clicks, keys and a ring, thus entering in the new era of interaction that body language is necessary.

Keywords: Natural User Interface. NUI. Kinect. Natural Interaction.

LISTA DE ILUSTRAÇÕES

Figura 1 – Kinect	13
Figura 2 - Detecção 3D	14
Figura 3 - Calculo da diferença da profundidade de dois pixels	15
Figura 4 - Diagrama com cores	16
Figura 5 - Hospital brasileiro já utiliza Kinect em cirurgias	17
Figura 6 - Aplicação facilita interação de crianças com o jogo e gera relatórios	18
Figura 7 - Criança com autismo	19
Figura 8 - Painel interativo.....	20
Figura 9 - Tela Inicial	24
Figura 10 - Corpo do Usuário Detectado	25
Figura 11 - Área de Relaxamento	26
Figura 12 - Cursor em Movimento para Direita	27
Figura 13 - Mão Fechada	28
Figura 14 - Pegando Objeto	29
Figura 15 - Soltando Objeto Após Movimento	29
Figura 16 - Menu de Contexto	30
Figura 17 - Mãos Abertas e Próximas	31
Figura 18 - Mãos Fechadas e Afastadas	31
Figura 19 - Voz	32
Figura 20 - Atalho de Teclas “Windows + TAB”	34

SUMÁRIO

Capítulo 1 - Introdução.....	9
1.1 Objetivos.....	11
1.1.1 Objetivo Geral	11
1.1.2 Objetivos Específicos	11
1.2 Justificativa	11
Capítulo 2 – Interface Natural de Usuário (NUI).....	12
2.1 Kinect.....	12
2.1.1 O Sensor Kinect.....	13
2.1.2 Detecção 3D.....	14
2.1.3 Reconhecimento do Corpo.....	15
2.3 Kinect e suas Aplicações	16
2.3.1 Medicina.....	17
2.3.2 Educação	18
Capítulo 3 – Metodologia.....	21
3.1 Estudos das Tecnologias Associadas ao Kinect.....	21
3.2 Estudos dos Algoritmos Necessários.....	22
3.3 Implementação do Software	22
Capítulo 4 – Implementações e Funcionalidades.....	24
4.1 Tela Inicial.....	24
4.2 Detectando Usuário	25
4.3 Movimentações do Cursor	25
4.4 Funções de Clique.....	27
4.5 Funções do <i>Scroll</i> (Roda do <i>Mouse</i>).....	30
4.6 Comandos de Voz	31
Capítulo 5 – Resultados	35
5.1 Trabalhos Futuros.....	36
Considerações Finais	37
Referências	38

Capítulo 1 - Introdução

Durante o avanço da tecnologia nos anos 70, a noção de interface-usuário, identificada à de interface homem-máquina, começou a se tornar uma preocupação dos projetistas de software.

Moran (1981) definiu a interface-usuário como: “aqueles aspectos do sistema com os quais o usuário entra em contato”. O que, por sua vez, significa “uma linguagem de entrada para o usuário, uma linguagem de saída para o sistema, e um protocolo de interação” (CHI, 1985 apud PREECE et al., 1994, p. 7).

Para a expressão “interface-usuário” e para o campo da Interação Humano-Computador (IHC) surgiram: “um conjunto de processos, diálogos e ações por meio dos quais um ser humano utiliza e interage com um computador” (BAECKER e BUXTON, 1987).

Neste âmbito a interação acontece não apenas pelo mecanismo estrutural, mas também pelo aspecto visual. “A IHC é uma área preocupada com o design, a avaliação e a implementação de sistemas computacionais interativos para o uso humano e com o estudo dos fenômenos que o circundam.” (ACM SIGCHI, 1992).

Atualmente os usuários estão acostumados a se comunicar com os computadores por meio de *mouse*, teclado, *joysticks* e outros dispositivos. O próximo passo da IHC é a Interface Natural de Usuário ou NUI (*Natural User Interface*), como define o site oficial de Kinect para Windows: “Com NUI, os usuários podem interagir naturalmente com os seus computadores ao tocar suas telas ou através de gestos e comandos de voz para dizer ao computador o que eles querem fazer” (MICROSOFT, 2013).

Os computadores vão se adaptar aos anseios dos seres humanos, sem que haja necessidade de teclas, toques e cliques, apenas a linguagem corporal.

O Kinect para Windows remove a camada de abstração entre os seres humanos e a tecnologia, dando às pessoas a capacidade de usar a voz e os gestos para se comunicar naturalmente com seus computadores e não apenas os dedos. (MICROSOFT, 2013).

O Kinect, portanto, é um dispositivo capaz de suprir a necessidade de interatividade de um sistema operacional, de maneira que precisem mais se adaptar

à máquina, mas a máquina se adapta aos movimentos e gestos das pessoas, como conta o brasileiro por trás da criação do Kinect, Alex Kipman (MENEZES, 2010),

Acordei um dia e uma felicidade assim de não ter tecnologia em volta de mim. Então eu me perguntei: 'nós estamos ficando escravos da tecnologia'. E se a gente voltar a fazer um negócio onde a tecnologia é a nossa escrava.

O uso funcional do Kinect vai além de uma mera tecnologia para videogame, ele também pode ser inserido no mundo computacional e beneficiar o desenvolvimento de outras esferas como, por exemplo, a inserção deste em tratamentos de saúde e na reabilitação de pacientes através da interação com o dispositivo.

Na Espanha, é possível enxergar o Kinect sendo usado como uma ferramenta de referência durante as operações. Os cirurgiões podem referenciar raios-x, por exemplo, e evitar infecções bacterianas em tempo real. Este é um uso reconhecidamente primitivo do Kinect. O cirurgião Tedesys admite que ele esteja apenas fazendo o mínimo do que é capaz de fazer. Mas é um passo crucial em direção a um futuro incrível, onde os profissionais médicos podem fazer o seu trabalho mais rápido e fácil. Se ele faz tudo à risca muito melhor, vamos levá-lo adiante (DYER, 2011).

Diante dos diversos usos potenciais do Kinect, este projeto busca melhorar e aperfeiçoar a Interface Natural de Usuário, em que a pessoa se comunica de forma natural com o sistema operacional, neste caso o Windows, que obedece a linguagem corporal do usuário e executa suas solicitações. O *mouse* e o teclado podem ser substituídos, então, pelo Kinect e este é usado como recurso para manipular o computador, ou seja, os movimentos e cliques do *mouse* são captados através dos movimentos e gestos das mãos, e o teclado e seus atalhos por comandos de voz. Utilizando sensores, o Kinect é capaz de absorver essas informações e transmitir os comandos ao computador, segundo o site do Xbox (2013), "Kinect usa um sensor de movimento que rastreia seu corpo inteiro. Portanto, quando você joga, ele não somente reconhece suas mãos e pulsos, e sim você por completo".

1.1 Objetivos

1.1.1 Objetivo Geral

Verificar o uso potencial da ferramenta tecnológica Microsoft Kinect como dispositivo de Interface Natural de Usuário.

1.1.2 Objetivos Específicos

- Estudo da tecnologia Microsoft Kinect para o processo de desenvolvimento do projeto;
- Dar ao usuário a possibilidade de controlar o sistema operacional de forma natural, ou seja, com gestos e comandos de voz, por meio da implementação de um software;
- Aplicar testes com diferentes pessoas utilizando o Kinect como recurso de interação com o computador, e assim verificar sua eficiência;

1.2 Justificativa

O processo de interação vai muito além de uma mera substituição de ferramentas funcionais do computador, como o mouse, por exemplo, por movimentos do corpo. O objetivo é doutrinar essas tecnologias de acordo com os anseios de seus usuários, direcionando sua funcionalidade à situações corriqueiras. Além de ser uma forma de interação mais intuitiva por ser a partir de movimentos simples das mãos ou comandos de voz, é também uma forma mais completa de interação onde tecnologia e usuário se aproximam. Sendo assim, esse estudo pretende promover uma forma mais rápida e eficaz de interação natural homem-computador, sendo que essa interação seja no sentido unilateral do comando do usuário para a máquina e que esta se adapte às vontades da pessoa que a está utilizando.

Capítulo 2 – Interface Natural de Usuário (NUI)

A NUI é um dos últimos estágios dos processos de interação entre homem e computador, por meio de comportamentos naturais e cotidianos do usuário para interagir diretamente com o conteúdo digital. Em uma explicação mais teórica desse conceito, a NUI é "uma camada abstrata em que uma aplicação computacional que faz a mediação das interações do usuário com o computador de uma forma intuitiva e natural, provida através do corpo, gestos, voz e toque" (BUXTON, 2010).

O termo "natural" advindo da teoria de NUI é usado porque a maioria das plataformas de interface do computador são dispositivos de controle artificial por meio de ferramentas que devem ser dominadas de forma pragmática pelo usuário como *mouse*, teclado, entre outros, ou seja, requer uma adaptação e um tempo hábil para que se apreenda a função destes meios interativos. Já no conceito de NUI o usuário deve ser capaz de realizar movimentos, gestos, sons e toques de forma natural, intuitiva sem que se tenha um treinamento prévio ou um tempo de habilitação para se manusear ou interagir com os conteúdos digitais, sendo assim uma interface que é natural, atingindo o último estágio da interatividade (WIGDOR; WIXON, 2011).

2.1 Kinect

O Kinect (Figura 1), ideia que nasceu com o nome de "Projeto Natal", foi lançado, no Brasil e no mundo, em novembro de 2010. (MENEZES, 2010).

Segundo Crawford (2013), "o dispositivo em si é uma caixa preta com cerca de 30 cm de comprimento e três câmeras sensores, que é plugada na parte de trás do Xbox 360".

Essencialmente, este hardware é uma caixa com algumas câmeras que fazem uso de iluminação infravermelho (IR) para obter dados de profundidade de cores, imagens e som. (FAIRHEAD, 2013).



Figura 1 - Kinect
Fonte: Microsoft / Divulgação.

Crawford (2013), afirma que,

Ao que se revela, o Kinect não é tão mágico assim. É uma combinação altamente inovadora de câmeras, microfones e software que transforma seu corpo em um controle de videogame. O nome Kinect é inspirado pelas palavras em inglês "kinect" (cinético), que significa estar em movimento, e "connect" (conectar), que significa estar "conectado aos amigos e ao entretenimento que você ama".

2.1.1 O Sensor Kinect

O Kinect é composto por várias partes como a câmera RGB, o emissor de IR, o sensor de IR, o conjunto de microfones e, o gerenciamento desses dispositivos por um firmware interno é o que torna essa tecnologia inovadora (ROBERTO, 2011).

Ele opera com três componentes principais, assim define Crawford (2013):

Uma câmera de vídeo, chamada pela Microsoft de "câmera RGB", que reconhece três cores: vermelho, verde e azul. Esta auxilia no reconhecimento facial e de outras características;

O sensor de profundidade que utiliza um projetor IR e um sensor CMOS monocromático. Ambos funcionam para detectar o ambiente com uma visão tridimensional;

E o microfone multi-matriz que são quatro microfones com a capacidade de reconhecer as vozes dos usuários, independente do som ambiente.

2.1.2 Detecção 3D

Segundo Fairhead (2013), a Microsoft adquiriu a tecnologia de detecção em 3D, que é a chave para o hardware Kinect, da empresa israelense PrimeSense.

Muitos imaginam que para implementar essa tecnologia, utiliza-se um componente chamado câmera de tempo de voo. Assim explica Queiroz (2010), “Funciona baseado em eco: você emite um som (ou um pulso de rádio) e então se mede o tempo que demora até o som bater no objeto e voltar”.

Mas Fairhead (2013) afirma que, o Kinect utiliza um método muito mais simples, chamado “luz estruturada”. Se você tem um feixe de luz emitido para um receptor em uma pequena distância, então a luz projetada é deslocada de acordo com a distância que esta é refletida de volta (Figura 2).

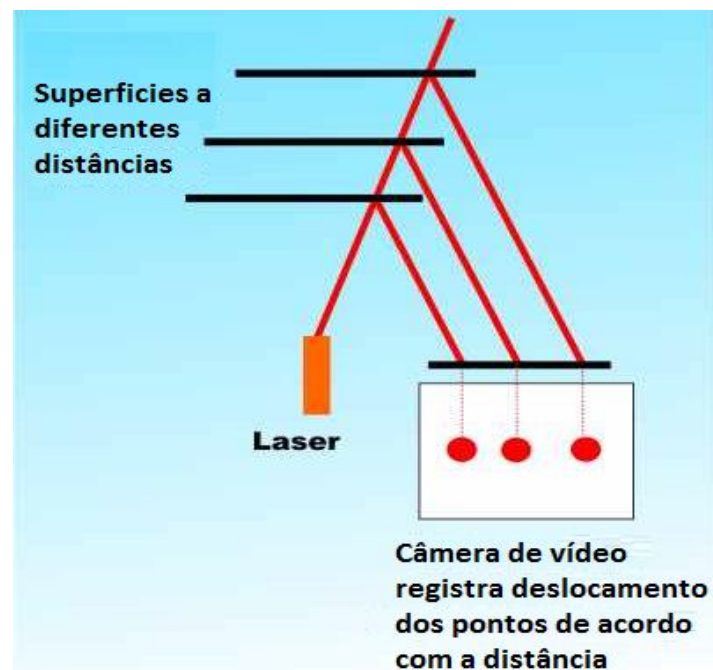


Figura 2 – Detecção 3D.
Fonte: Fairhead (2013).

Projetando uma grade fixa de pontos em uma cena, e medindo quanto cada um se deslocou, visto por uma câmera de vídeo, pode-se perceber a variação da distância em que cada ponto foi refletido de volta para o receptor.

2.1.3 Reconhecimento do Corpo

Fairhead (2013) explica que, os recursos que foram usados podem surpreender na medida em que eles são simples, e está longe de ser óbvio que eles contêm as informações necessárias para identificar as partes do corpo. As características são todas baseadas em uma fórmula simples:

$$f = d(x + u / d(x)) - d(x + v / d(x)) \quad (1)$$

Em que (u, v) são um par de vetores de deslocamento e $d(x)$ é a profundidade, ou seja, a distância do Kinect até o pixel em x . Este é um recurso muito simples, é a diferença na profundidade de dois pixels de deslocamento a partir do pixel alvo por u e v (Figura 3).

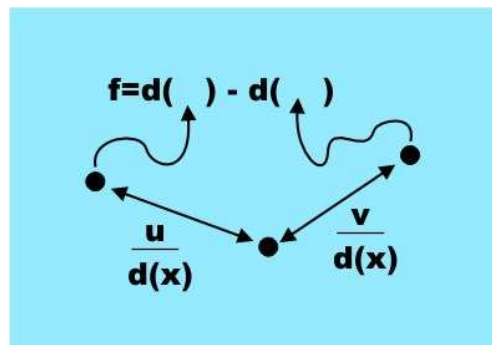


Figura 3 – Cálculo da diferença da profundidade de dois pixels.
Fonte: Fairhead (2013).

A única complicação é que o deslocamento é escalado com distância do pixel alvo, ou seja, dividido por $d(x)$.

Em seguida treinou-se um tipo de classificador chamado de floresta de decisão, ou seja, uma coleção de árvores de decisão. Cada árvore foi treinada num conjunto de características em imagens de profundidade, que foram pré-marcadas com as partes do corpo alvo. Essas árvores de decisão foram modificadas até que chegasse a classificação correta para uma parte específica do corpo em todo o conjunto de teste de imagens. O treinamento de apenas três árvores usando um milhão de imagens de teste levou cerca de um dia usando um *cluster* com mil núcleos.

Os classificadores treinados atribuem a probabilidade de um pixel estar em cada parte do corpo, e a próxima etapa do algoritmo é simplesmente escolher as áreas de probabilidade máxima para cada parte do corpo. Assim, uma área será classificada à categoria de "perna" se o classificador de perna der prioridade máxima naquela área. A etapa final é calcular as posições comuns sugeridas em relação às áreas identificadas como partes do corpo em particular.

A Figura 4 mostra as diferentes partes do corpo e a probabilidade máxima para cada área é indicada por cores.

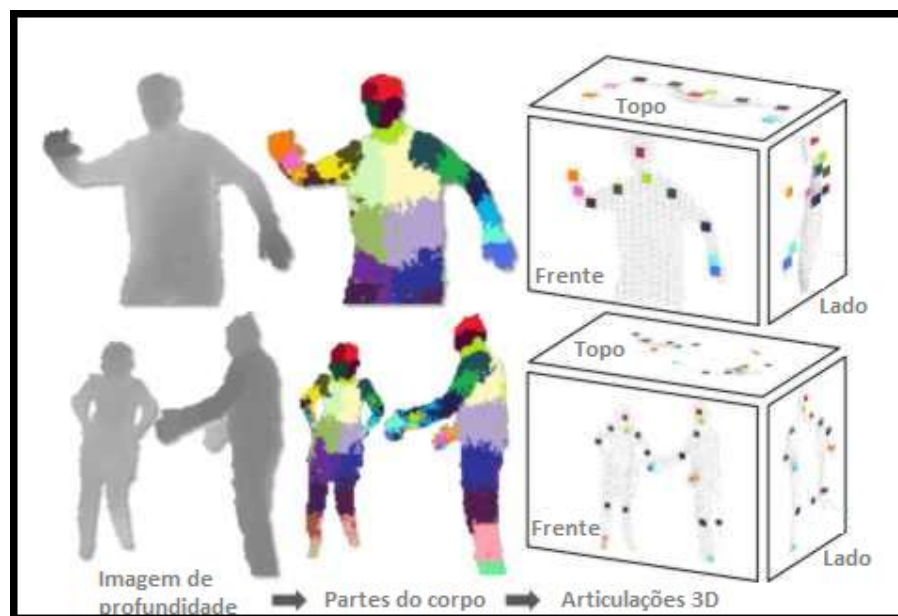


Figura 4 – Diagrama com cores.
Fonte: Fairhead (2013).

2.3 Kinect e suas Aplicações

O Kinect revolucionou o mundo dos videogames e superou suas expectativas. Mas não é só nesta área que ele está se destacando, além dos jogos suas funcionalidades foram ampliadas em várias áreas do conhecimento, é o que o site do próprio Xbox chama de “*Kinect-Effect*” (Efeito-Kinect).

Construímos o Kinect para revolucionar o modo que jogamos e sua experiência de entretenimento. Mas as pessoas começaram a usar Kinect de maneira que nunca imaginamos. Desde ajudar crianças

com autismo, até ajudar doutores em salas de operações, levando o Kinect além dos jogos. É isso que chamamos de Efeito Kinect. (XBOX, 2013).

2.3.1 Medicina

Em um hospital de Londrina, Paraná, o Kinect foi testado com o intuito de diminuir o tempo das cirurgias. No decorrer deste processo, os médicos têm a necessidade de visualizar as imagens do paciente, e caso ele toque algum objeto é preciso fazer uma higienização para evitar o risco de infecção no paciente. Mas agora com essa nova tecnologia, é possível passar os exames e dar zoom nas imagens (Figura 5), sem tocar nada, utilizando apenas gestos (TERRA, 2013).



Figura 5 – Hospital brasileiro já utiliza Kinect em cirurgias.
Fonte: Hospital Evangélico de Londrina / Divulgação.

Pesquisadores da Universidade de Minnesota utilizam o Kinect para detectar possíveis distúrbios mentais em crianças. Segundo eles, o dispositivo consegue obter dados mais objetivos que auxiliam no diagnóstico final. Com isso, muitos gastos são poupados com sistemas de diagnóstico (FERREIRA, 2011).

2.3.2 Educação

A empresa brasileira Agência Detalhes, de Belo Horizonte, desenvolveu estudos nesta plataforma e começou a aplicar jogos educativos (Figura 6). Segundo o diretor de tecnologia da agência, os sensores de movimento e voz do Kinect possibilitaram um *game* mais atrativo (TERRA, 2013).

As instituições acharam a ideia fenomenal, pois enquanto a criança joga é possível pegar informações comportamentais de cada um, de acordo com seus erros e acertos. Além disso, dá para fazer relatórios comparativos entre alunos e turmas, o professor está lidando com turmas do mesmo colégio, com crianças de mesma característica econômica, mas elas têm desempenhos diferentes. Com a análise dos dados, ele pode saber onde errou no processo.

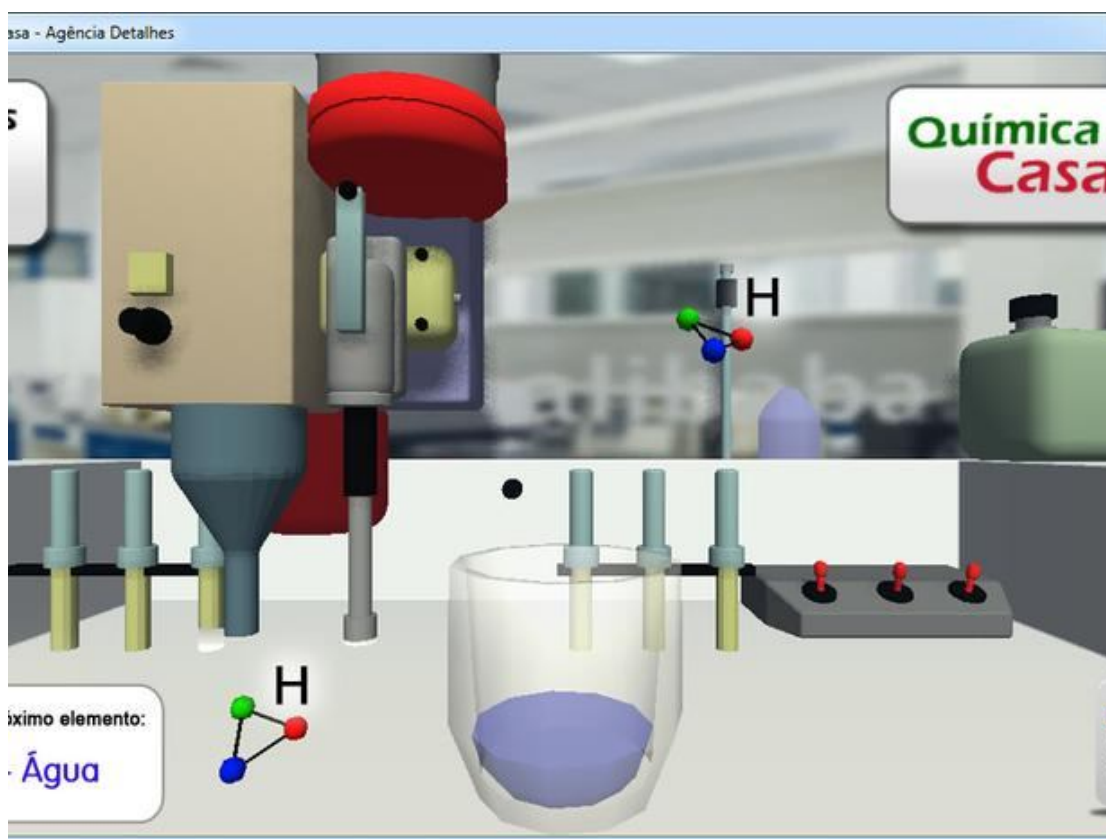


Figura 6 – Aplicação facilita interação de crianças com o jogo e gera relatórios.
Fonte: Agência Detalhes / Divulgação.

Segundo Dyer (2011), no Centro de Lakeside para o autismo, a educação e a terapia facilitada em Seattle, WA, professores estão usando o Kinect para quebrar as barreiras provocadas pelo desvio comportamental. A equipe está usando o Xbox

360 juntamente com o Kinect para ensinar a resolução de problemas e fortalecer habilidades sociais. Dinâmicas básicas de jogo e a criação de *Avatar*, por exemplo, deixam as crianças tranquilas para se comunicar. Em uma história especialmente tocante, um dos educadores cita o Kinect para equilibrar a coordenação do corpo de uma criança com o movimento do braço de forma assíncrona (Figura 7) “essas pessoas estão fazendo um trabalho comovente, e é incrível ver esses resultados”.



Figura 7 – Criança com autismo.
Fonte: Dyer (2011).

Engenheiros e tecnólogos do instituto de tecnologia FIT, desenvolveram um “painel interativo” (Figura 8) com o Kinect, que está sendo testado por pessoas com deficiências. Foi criado um mapa do mundo no qual os usuários navegam pelos continentes utilizando somente as mãos ou a fala. A uma pequena distância, as pessoas podem conhecer as diferentes culturas e até responder a um jogo geográfico sobre os países, suas bandeiras e capitais (OLHAR DIGITAL, 2012).



Figura 8 – Painel interativo.
Fonte: OLHAR DIGITAL (2012).

Capítulo 3 – Metodologia

O referencial teórico deste trabalho foi construído a partir de citações, estudos e pesquisas de outros profissionais que desenvolvem ferramentas para o Kinect e aprimoraram a usabilidade da ferramenta.

A fim de se estudar a tecnologia Microsoft do Kinect, este foi testado por pessoas aleatórias para comprovarem a eficiência do processo de interação com o uso desse hardware, em detrimento do uso de outros dispositivos, como o *mouse* por exemplo. A metodologia utilizada foi estabelecida por meio de estudos comparativos e coleta de dados a partir da observação do uso do Kinect por usuários comuns, ou seja, indivíduos que não trabalham na área de tecnologia da informação. Assim foi possível melhorar e aperfeiçoar suas características e corrigir erros.

No caso dessa pesquisa, o objetivo foi criar um modo eficiente e funcional de interação natural com o computador, onde a tecnologia consegue captar a linguagem corporal das pessoas e transmiti-la á maquina.

Portanto, a teoria estudada juntamente com novas pesquisas neste âmbito, é fundamental para o aperfeiçoamento deste projeto que será aplicado concretamente em outras áreas de interesse e necessidade do usuário.

3.1 Estudos das Tecnologias Associadas ao Kinect

O sistema foi desenvolvido a partir de um notebook Toshiba Satellite A665 com as seguintes configurações: processador Intel Core i5, 4 GB de memória RAM, placa de vídeo Nvidia GeForce GTS 350M com 1GB de memória dedicada e HD de 500 GB.

Este notebook utiliza o sistema operacional Windows 7 Professional 64 bits pois ele é adequado para a instalação dos softwares (mencionados posteriormente neste tópico) que serão utilizados no desenvolvimento do projeto.

Para desenvolver o projeto, foram utilizadas pesquisas e ferramentas como o SDK (versão 1.8) do Kinect, que são bibliotecas específicas do dispositivo que facilitam e auxiliam na programação, disponibilizado pela Microsoft, juntamente com

a ferramenta Visual Studio 2012 utilizando a linguagem C# (CSharp) que serão essenciais neste procedimento.

A escolha desses recursos foi estabelecido a partir da compatibilidade de procedimento destes com o Kinect pelo fato de serem criados e disponibilizados pela mesma empresa, a Microsoft. Assim, facilitou o processo de desenvolvimento do projeto.

3.2 Estudos dos Algoritmos Necessários

O objetivo inicial foi eliminar o mouse do computador, ou seja, a sua função de movimento foi substituído pelo movimento da mão direita ou esquerda. Para isso utilizou-se um algoritmo do próprio SDK do Kinect, que rastreia 48 pontos do corpo e também suas coordenadas no espaço. Como ele atingiu os objetivos desejados, não foi necessária a busca ou criação de outro algoritmo.

Para programar a função de clique, a aplicação deveria reconhecer os gestos de abrir e fechar das mãos do usuário, sendo o estado inicial de mão aberta e o de selecionado de mão fechada. O algoritmo do SDK que reconhece estes gestos, também foi muito satisfatório.

3.3 Implementação do Software

Inicialmente foi instalado e configurado a ferramenta Visual Studio 2012 e a posteriori foi feita a instalação do SDK do Kinect para que o desenvolvimento do projeto pudesse ser executado.

Após a implementação desses programas, a aplicação iniciou seu desenvolvimento utilizando a linguagem C# por meio do Visual Studio 2012, sendo que foi importado à biblioteca disponibilizada pelo SDK.

Na primeira fase foi programada a função onde a aplicação iria reconhecer o corpo do usuário, sendo que o movimento da mão direita e esquerda foi isolado. E de acordo com sua posição em relação aos eixos X e Y, dentro do espaço, foi calculado e transmitido ao cursor do computador, e este passa a se movimentar, então, de acordo com a posição das mãos.

Na segunda fase foi implementada a função de clique em que o sistema iria detectar se a mão do usuário está fechada ou aberta. Caso estivesse fechada, o sistema transmitiria ao cursor o estado de clicado (selecionado). Porém, se estivesse aberta, o cursor voltaria ao seu estado inicial.

Na terceira fase, a primeira e segunda fases foram executadas simultaneamente, e assim, o *mouse* pode ser eliminado, sendo possível controlar o cursor do sistema operacional com as mãos.

Na fase final, o comando de voz foi programado para substituir alguns atalhos úteis do teclado.

Capítulo 4 – Implementações e Funcionalidades

Como a intenção do projeto foi proporcionar ao usuário uma forma de interação natural e simples, as funções de movimento e gestos foram implementadas e melhoradas de acordo com as necessidades que surgiram durante o desenvolvimento e os testes. Os atalhos por comandos de voz também foram criados para facilitar a interação e na maioria deles utilizou-se o idioma português e com isso foi possível aproximar ainda mais o usuário da tecnologia.

4.1 Tela Inicial

A tela inicial (Figura 9) é a única da aplicação, como o programa trabalha diretamente com os recursos do sistema operacional, não foi necessária a criação de outras telas.

Nela tem-se o botão “Ligar” / “Desligar”, que inicia e interrompe, respectivamente, as funcionalidades de aplicação e os indicadores de sensibilidade X e Y do cursor que podem ser modificados de acordo com a necessidade do usuário. Do lado direito a câmera de profundidade - a parte superior central - descreve o estado em que se encontra a mão do usuário durante a interação, ou seja, se está aberta ou fechada, e a parte inferior central descreve qual das mãos (direita ou esquerda) está sendo utilizada naquele momento para movimentar o cursor.



Figura 9 – Tela Inicial.
Fonte: Autoria própria.

4.2 Detectando Usuário

Ao clicar no botão “Ligar”, o Kinect inicia seu funcionamento. Ao detectar um ou mais corpos, a aplicação atribui uma cor aleatória a ele (Figura 10). A partir disso, o programa fica aguardando algum usuário posicionar as duas mãos a frente do corpo para ativar o controle do sistema operacional.

Assim que alguém posiciona as duas mãos à frente do corpo, a aplicação é ativada e permite o controle do sistema a este usuário. No mesmo instante, o ícone do cursor é substituído por uma imagem de mão aberta na cor azul, dando a ele um sentido mais real do processo.



Figura 10 – Corpo do Usuário Detectado.
Fonte: Autoria própria.

4.3 Movimentações do Cursor

Para movimentar o cursor na tela, o usuário deve decidir qual das mãos irá utilizar para tal processo e posicioná-la um pouco a frente do seu corpo.

Posteriormente, deve-se verificar o “centro” do espaço onde irá realizar o movimento. Ao movimentar a mão é possível encontrar uma área onde o cursor não se move, sendo esta a área de “relaxamento” (Figura 11).

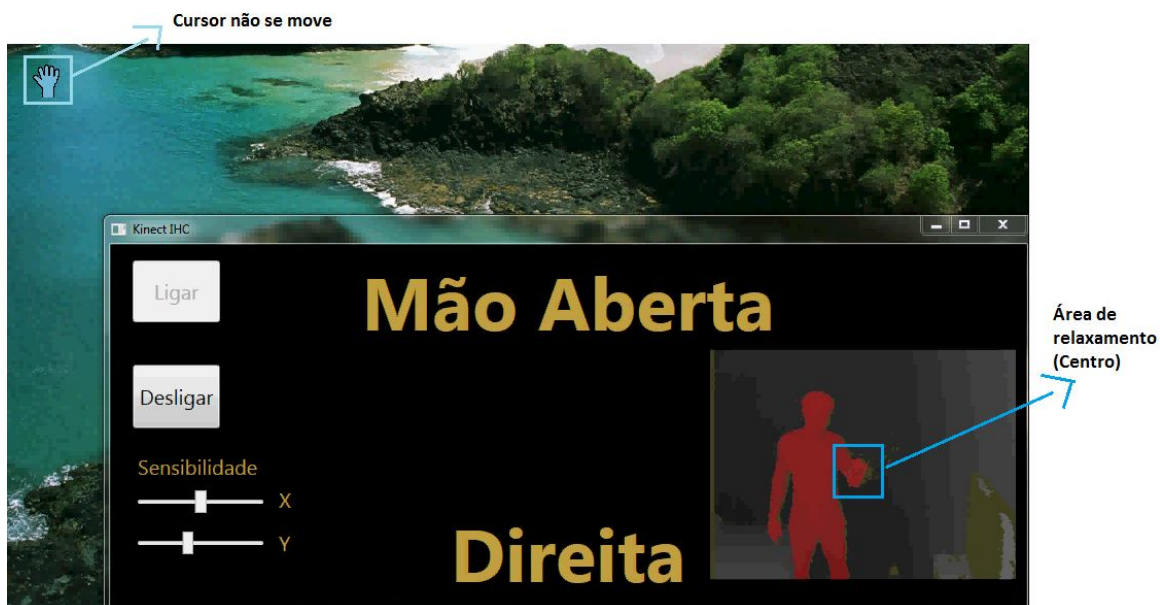


Figura 11 – Área de Relaxamento.
Fonte: Autoria própria.

Diferentemente do *mouse*, que move o cursor utilizando a posição absoluta, nesta aplicação este processo foi adaptado. Movendo a mão levemente para a direita (Figura 12) o cursor se deslocará na tela para a direita e continuará até que a mão volte ao seu estado inicial (centro). Caso o movimento da mão seja mais rápido, o cursor também se deslocará mais rapidamente, ou seja, quanto mais longe a mão estiver de sua posição inicial, mais rápido será o deslocamento do cursor. Este processo também se aplica às demais coordenadas: esquerda, para cima e para baixo.

Desta forma, quando houver a necessidade de acessar um conteúdo que esteja nas extremidades da tela, o usuário não permanecerá com o braço esticado, ele apenas deslocará o cursor até o local desejado e voltará para a “área de relaxamento”, evitando assim, o cansaço do membro durante a interação.



Figura 12 – Cursor em Movimento para Direita.
Fonte: Autoria própria.

4.4 Funções de Clique

Com intuito de facilitar a interação e deixá-la o mais natural possível, as funções de clique foram substituídas pelo gesto da mão, ou seja, ao fechá-la, por exemplo, a aplicação transmite ao cursor o estado de selecionado. Quando esta função está sendo executada, o programa troca o ícone do cursor por uma imagem de mão fechada, na cor azul e num tom mais escuro; transmitindo assim, ao usuário, a sensação de estar “pegando” um objeto do mundo virtual (Figura 13).



Figura 13 – Mão Fechada.
Fonte: Autoria própria.

Ao abrir a mão novamente, o cursor volta ao seu estado inicial (dê selecionado), e o ícone também é trocado pelo anterior.

Diante disso, para abrir uma pasta ou executar qualquer aplicativo do sistema operacional, basta dar o duplo clique: abrindo e fechando a mão duas vezes, em um curto intervalo de tempo, pois a ação será executada.

Para arrastar qualquer ícone ou objeto na tela, deve-se ir até o conteúdo desejado, fechar a mão (Figura 14), movê-lo até o local de destino e por fim soltá-lo (Figura 15).



Figura 14 – Pegando Objeto.
Fonte: Autoria própria.



Figura 15 – Soltando Objeto Após Movimento.
Fonte: Autoria própria.

Também houve a necessidade de implementar a função do menu de contexto, que por padrão no sistema operacional deste projeto (Windows 7) é acionado pelo botão direito do *mouse*.

Como nessa aplicação tem-se a possibilidade de utilizar tanto a mão direita quanto a mão esquerda, como recurso de interação, para exibir o menu de contexto (Figura 16), é necessário colocar as duas mãos um pouco a frente do corpo e depois fechar e abrir uma das mãos. Neste momento a função de movimento é interrompida para aumentar a precisão do ato.

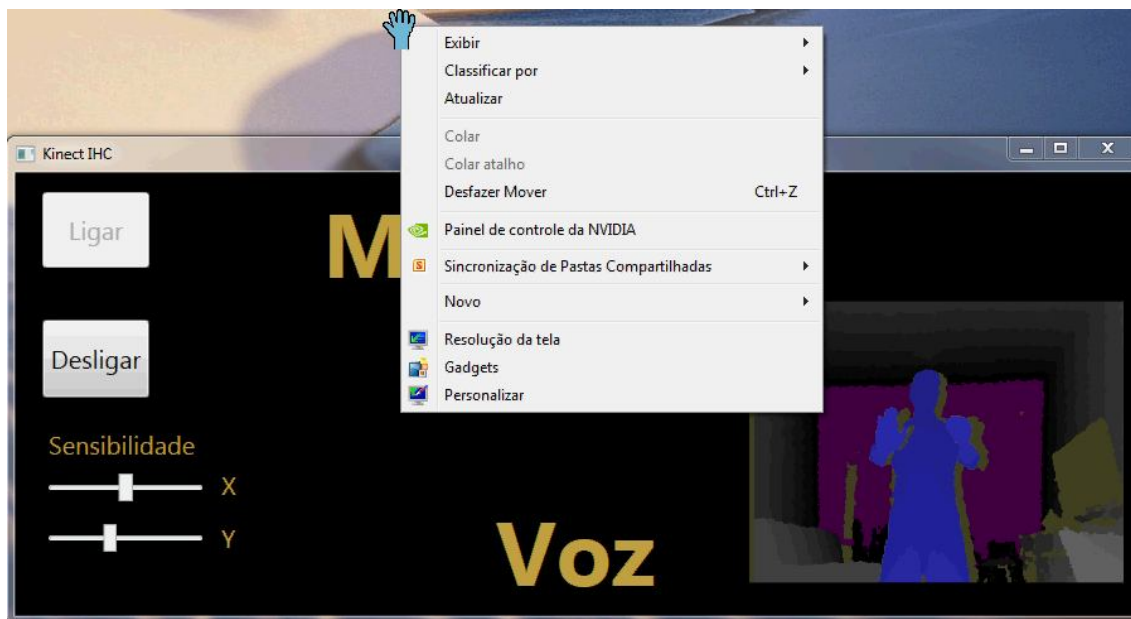


Figura 16 – Menu de Contexto.
Fonte: Autoria própria.

4.5 Funções do *Scroll* (Roda do *Mouse*)

Juntamente com o comando de voz “zoom”, que será explicado no próximo tópico, e a interação das mãos, este processo funciona da seguinte maneira: para simular a rolagem vertical para cima é necessário que o usuário posicione as duas mãos abertas a frente do corpo e perto uma da outra (Figura 17). Depois fechando e afastando as mãos em sentidos opostos (Figura 18), o processo se inicia. Já para simular a rolagem vertical para baixo, a ação é inversa, ou seja, posicionando as mãos abertas e longe uma da outra, ao fechar e aproximá-las, inicia-se o processo. Em ambos, quanto mais rápido for o movimento das mãos de aproximar ou afastar, maior será a quantidade de pontos de giro transmitida ao sistema.



Figura 17 – Mãos Abertas e Próximas.
Fonte: Autoria própria.



Figura 18 – Mãos Fechadas e Afastadas.
Fonte: Autoria própria.

4.6 Comandos de Voz

Como um dos objetivos deste projeto é dinamizar e facilitar a interação com o usuário, houve a necessidade de implementar os atalhos úteis do teclado, como por exemplo, o de copiar e colar, e adaptá-los para comando de voz.

A intenção foi colocar todos os comandos em português, mas no reconhecimento de voz utilizado (SDK do Kinect), não há compatibilidade, até o

momento, para esta língua. Neste caso, os comandos foram adaptados para que o programa pudesse reconhecer algumas palavras neste idioma.

A solução encontrada foi “remanejar” o reconhecimento de voz na língua inglesa, fazendo com que a pronúncia das palavras em inglês chegasse o mais próximo possível da pronúncia em português; por exemplo, para implementar o reconhecimento da pronúncia “fechar”, o programa tenta reconhecer em inglês a palavra “feshar”. Desse modo, foi possível adaptar todos os comandos de voz deste projeto, no idioma desejado.

Posicionando as duas mãos abertas e a frente do corpo (Figura 19) a aplicação ativa o comando de voz e neste mesmo momento o movimento do cursor é interrompido, fazendo com que o mesmo pare em sua última posição.



Figura 19 – Voz.
Fonte: Autoria própria.

Os comandos implementados foram:

Apresentar: ao pronunciar esta palavra a aplicação simula a tecla “F5”, que por padrão inicia uma apresentação do Microsoft PowerPoint. Este atalho foi implementado visando facilitar a exibição de conteúdo, em uma palestra ou aula. Enquanto estiver neste modo, basta fechar e abrir uma das mãos para transcorrer entre os *slides*;

Colar: este comando simula as teclas “CTRL + V”, que cola da área de transferência para o local desejado;

Copiar: simula as teclas “CTRL + C” do teclado. Sua função é copiar um conteúdo para área de transferência;

Desktop: mostra a área de trabalho do sistema;

Enter: simula a tecla “ENTER”;

Explorer: abre a pasta de navegação “Windows Explorer”;

Fechar: simula as teclas “CTRL + F4”. Este atalho tem a função de fechar a janela ativa do sistema operacional;

Finish: tem a função de desligar a aplicação e desativar o Kinect;

Facebook: este comando simula o texto “www.facebook.com.br” mais a tecla “ENTER”. Com isso, o usuário pode acessar a rede social acima, em qualquer navegador, com apenas um comando;

Iniciar: simula a tecla “Windows”, exibindo o “Menu Iniciar”;

Mapas: simula o texto “maps.google.com.br” mais a tecla “ENTER”. Com isso, o usuário pode acessar o link acima, em qualquer navegador, com apenas um comando;

Max: este comando maximiza a janela ativa do sistema;

Mini: minimiza a janela ativa do sistema;

Música: ao pronunciar este comando a aplicação perguntará ao usuário qual música ele deseja ouvir. As opções serão aquelas que estiverem na pasta pessoal, do sistema operacional: “Minhas músicas”. Para executar o som desejado, basta dizer o primeiro (e se houver), o segundo nome do arquivo;

Selecionar: simula as teclas “CTRL + A”. Este atalho seleciona todo o conteúdo da tela;

Web: ao pronunciá-lo o navegador padrão do sistema é executado;

Windows: este comando simula as teclas “Windows + TAB”. Através deste atalho pode-se navegar, de forma mais intuitiva, entre as janelas abertas do sistema operacional (Figura 20);



Figura 20 – Atalho de Teclas “Windows + TAB”.
Fonte: Autoria própria.

Zoom: ativa o *scroll* (roda do *mouse*).

Os comandos acima foram programados diante da sugestão de usuários que testaram a aplicação e durante este processo tiveram a necessidade de usá-los para facilitar e aperfeiçoar sua interação, permitindo ao computador ser capaz de “entender” os desejos das pessoas e executá-los.

Capítulo 5 – Resultados

Durante o desenvolvimento e os testes, observou-se que a aplicação permitia apenas que um usuário estivesse na área de “visão” do Kinect, para interagir com o computador. Ao detectar um novo usuário o programa se perdia, pois o mesmo executava as solicitações de ambos os usuários. Para resolver esta dificuldade, encontrou-se uma solução tecnicamente fácil em termos lógicos. A biblioteca do hardware define um “id”, ou seja, um código para cada usuário detectado, desta maneira foi possível distingui-los e atribuir o controle do sistema operacional apenas à uma pessoa; neste caso, aquela que posicionar as duas mãos a frente do corpo primeiro.

A resposta do cursor à movimentação da mão foi bem precisa e funcional, tanto para a mão direita quanto para a mão esquerda. Houve algumas dificuldades para implementar a área de “relaxamento”, onde a mão pode fazer movimentos leves sem que o cursor seja afetado, aumentando assim, a precisão do clique. Durante os testes o processo foi melhorado e aperfeiçoado, permitindo que o resultado final fosse muito satisfatório.

Nas funções de clique também houve êxito e o resultado foi positivo. O algoritmo do SDK do Kinect funciona de maneira sincronizada para detectar os gestos de mão aberta e mão fechada. Durante os testes, não houve erros de detecção nem a necessidade de pesquisar outras soluções para executar tal processo, pois este atendeu a todos os objetivos propostos.

O *scroll* respondeu bem aos testes, mas pode ser melhorado e aperfeiçoado, pois algumas funções não se aplicam de maneira prática e interativa, como por exemplo, para rolar um texto de várias páginas ou um site com barra de rolagem. Pois o intuito deste processo foi substituir a ferramenta de *zoom*, ou seja, para ampliar ou reduzir algo que esteja na tela.

Já o comando de voz variou muito durante os testes, visto que, cada pessoa possui um tom de voz e sotaques diferentes para pronunciar certas palavras. Outro ponto a ressaltar, é que o reconhecimento de voz foi adaptado ao português, limitando ainda mais este recurso. Apesar dos principais comandos (copiar, colar, fechar, música, selecionar e etc.) apresentarem uma resposta satisfatória, este processo como um todo é algo que precisa ser melhorado e aperfeiçoado.

De modo geral, a aplicação foi satisfatória em quase todas as suas funcionalidades, proporcionando ao usuário uma forma natural e simples de interação com a máquina.

5.1 Trabalhos Futuros

Para trabalhos futuros, deseja-se que este projeto seja testado e aplicado em ambientes em que o computador é manipulado a uma distância considerável, ou seja, onde não é possível ter contato com outro hardware, o mouse, por exemplo. E com isso, analisar a sua eficiência como produto.

O comando de voz poderia ser “treinado” para reconhecer a voz de cada usuário, fazendo com que a chance de acerto seja quase total.

Além disso, novas funcionalidades e atalhos poderão ser adicionados e outros melhorados conforme surgir as necessidades das pessoas e o anseio pelo aperfeiçoamento da Interface Natural de Usuário.

Considerações Finais

Em um futuro próximo, o meio de interação homem-máquina mais comum, será a Interface Natural de Usuário onde as pessoas se libertarão de sensores ou de *hardwares* intermediários que transmitem aos computadores o que o usuário deseja; ou seja, a tecnologia será capaz de “entender” o desejo das pessoas e executá-lo com precisão e eficiência.

Neste trabalho, foi possível verificar o quão a tecnologia está avançada e o quanto ela ainda pode avançar. Nesse âmbito, o Kinect é uma ferramenta incrível de interação natural e é capaz de beneficiar muitas áreas que vão além da tecnologia e das ciências exatas.

Pode-se constatar que os benefícios de se trabalhar com bibliotecas próprias do produto, ferramentas e sistemas que são disponibilizados pela mesma empresa que o fabrica, no caso a Microsoft, são muito satisfatórios, tanto na produção quanto na qualidade final, o resultado é muito positivo.

Por outro lado, a companhia citada acima, tem o comando de voz do produto limitado. Até o momento, as grandes potências econômicas, ou seja, o Brasil e a língua portuguesa são desfavorecidos para este recurso, o que limitou o reconhecimento de voz do produto e obrigou a adaptação da aplicação para realizar tal tarefa.

Por fim, em relação à aplicação desenvolvida, verificou-se que a interação natural já é uma realidade e que funciona muito bem diante da diversidade de características físicas das pessoas, sendo que qualquer um pode usar essa interação sem restrição. Além disso, essa ferramenta pode evoluir muito e ajudar pessoas de várias outras formas, basta ter criatividade, imaginação e um pouco de programação.

Referências

ACM SIGCHI (1992). **Curricula for human-computer interaction**. Technical report, ACM, NY, 1992. Disponível em: <<http://www.acm.org/sigchi/>>. Acesso em: 08 abr. 2013.

BAECKER, R. M. & BUXTON, W. A. S. (Eds.) (1987). **Readings in human-computer interaction**: A multidisciplinary approach. San Mateo, CA.

BUXTON, Bill. **CES 2010: NUI with Bill Buxton**, 2010. Disponível em: <<http://channel9.msdn.com/blogs/larrylarsen/ces-2010-nui-with-bill-buxton>> Acesso em: 21 de setembro de 2013.

CRAWFORD, Stephanie. **Como funciona o Microsoft Kinect**. Disponível em: <<http://eletronicos.hsw.uol.com.br/microsoft-kinect.htm/printable>>. Acesso em: 04 abr. 2013.

DYER, Mitch. **Kinect: Year One**, 31 outubro 2011. Disponível em: <<http://www.ign.com/articles/2011/10/31/kinect-year-one>>. Acesso em: 10 abr. 2013.

FAIRHEAD, Harry. **All About Kinect**. Disponível em: <<http://www.i-programmer.info/babbages-bag/2003-kinect-the-technology-.html?start=0>>. Acesso em: 28 maio 2013.

FERREIRA, C. E. **Kinect é utilizado para diagnosticar distúrbios mentais em crianças**, 16 março 2011. Disponível em: <<http://www.baixakijogos.com.br/noticias/13535-kinect-e-utilizado-para-diagnosticar-disturbios-mentais-em-criancas>>. Acesso em: 03 jun. 2013.

MENEZES, César. Sistema que usa o corpo como joystick chega ao Brasil por R\$ 600. **Jornal da Globo**, 05 novembro 2010. Disponível em: <<http://g1.globo.com/jornal-da-globo/noticia/2010/11/sistema-que-usa-o-corpo-como-joystick-chega-ao-brasil-por-r-600.html>>. Acesso em: 29 abr. 2013.

MICROSOFT. **About Us**. Disponível em: <<http://www.microsoft.com/en-us/kinectforwindows/news/about.aspx>>. Acesso em: 18 maio 2013.

MORAN, T. (1981). **The Command Language Grammars**: a representation for the user interface of interactive computer systems. *International Journal of Man-Machine Studies*, 15, 3-50.

OLHAR DIGITAL. **Kinect é utilizado para ajudar na educação de pessoas com deficiências**: Engenheiros usam acessórios para promover interatividade, 12 setembro 2012. Disponível em: <<http://olhardigital.uol.com.br/noticia/kinect-e-utilizado-para-ajudar-na-educacao-de-pessoas-com-deficiencias/29006>>. Acesso em: 27 maio 2013.

QUEIROZ, Murilo. **Um cientista explica o Microsoft Kinect**, 29 outubro 2010. Disponível em: <<http://blog.vettalabs.com/2010/10/29/um-cientista-explica-o-microsoft-kinetic/>>. Acesso em: 03 jun. 2013.

ROBERTO, Carlos. **Microsoft publica artigo sobre as entranhas do Kinect**, 06 abril 2011. Disponível em: <<http://kinect-i.blogspot.com.br/2011/04/microsoft-publica-artigo-sobre-as.html>>. Acesso em: 25 maio 2013.

TERRA. **Tecnologia do Kinect revoluciona atividades cotidianas**, 20 agosto 2012. Disponível em: <<http://tecnologia.terra.com.br/tecnologia-do-kinect-revoluciona-atividades-cotidianas,910829b92696b310VgnCLD200000bbcceb0aRCRD.html>>. Acesso em: 22 ago. 2013.

WIGDOR, Daniel; WIXON, Dennis. **Brave NUI world: designing natural user interfaces for touch and gesture**, 2011.

XBOX. **Efeito Kinect**: como o mundo usa o Kinect. Disponível em: <<http://www.xbox.com/pt-BR/Kinect/Kinect-Effect>>. Acesso em: 01 maio 2013.