

CONIC-SEMESP 13º Congresso Nacional de Iniciação Científica

Anais do Conic-Semesp. Volume 1, 2013 - Faculdade Anhanguera de Campinas - Unidade 3. ISSN 2357-8904

TÍTULO: APLICAÇÃO PARA RECONHECIMENTO DE GESTOS, BASEADA EM OPENNI

CATEGORIA: CONCLUÍDO

ÁREA: ENGENHARIAS E TECNOLOGIAS

SUBÁREA: COMPUTAÇÃO E INFORMÁTICA

INSTITUIÇÃO: FACULDADE ANHANGUERA DE BELO HORIZONTE

AUTOR(ES): FILIPE SANTOS DE JESUS

ORIENTADOR(ES): SANDRO RENATO DIAS

COLABORADOR(ES): OTAVIANO MARTINS MONTEIRO

Realização:



Apoio:



Aplicação para reconhecimento de gestos, baseada em OpenNi

Resumo

No Brasil, a maioria da população é ouvinte e a cultura brasileira é marcada pelas representações dos ouvintes, aonde no dia-a-dia os surdos são obrigados a se contextualizar como se não possuíssem deficiência auditiva. Apesar de toda a sua expressividade, o população surda está imersa sob uma cultura dominante, ainda sendo notável uma dificuldade de inclusão. Visando contribuir com a sociedade, este projeto consiste no estudo do framework OpenNi, que é uma solução em termos de *software* que permite a captação de movimentos para que posteriormente possam ser comparados com LIBRAS (Língua Brasileira de Sinais), de modo que o estudo desta tecnologia possa servir de base para que futuramente o OpenNi juntamente com outros softwares e equipamentos de interação natural possam combinar em uma tecnologia assistiva que promova uma maior autonomia para as pessoas surdas. Até o momento foi iniciado o estudo sobre esse framework, analisando seus principais recursos e realizando testes práticos dele com o Kinect, que é um dispositivo capaz de captar movimentos. Com o sucesso obtido nos testes realizados até o momento é possível comprovar a viabilidade deste estudo.

Introdução

A cada novo avanço tecnológico percebe-se que a relação entre os seres humanos e a tecnologia se torna cada vez mais natural e necessária. Entende-se também, que a tendência dos dispositivos que atuam nesse intercâmbio homem-máquina não é guiada apenas pelo objetivo de enviar, receber, armazenar e computar informações, mas, sobretudo pela finalidade de interação com conteúdo da informação, com o espaço, e com os objetos físicos. Há uma ampla variedade de tecnologias relacionadas a esse tipo de interação melhorada, como por exemplo, telas sensíveis ao toque, dispositivos que interpretam comandos de voz, e mais recentemente, dispositivos que captam e interpretam movimentos.

Por meio dessas novas realidades e das possibilidades que surgem através dos avanços das empresas que atuam no campo tecnológico, especificamente nas áreas de entretenimento, jogos eletrônicos e dispositivos multimídia, é

possível estimar a aplicação desses dispositivos de interação, em especial captura de movimentos, no ensino e interpretação da língua de sinais LIBRAS.

As Línguas de Sinais (LS) são as línguas naturais das comunidades surdas, e LIBRAS é um acrônimo para Língua Brasileira de Sinais. Portanto, é uma língua nacional e possui o reconhecimento oficial do governo brasileiro pela Lei 10.436/2002. “As Línguas de Sinais não são apenas um conjunto de mímicas ou gestos soltos, possuem o status de língua porque também são compostas pelos níveis linguísticos: o fonológico, o morfológico, o sintático e o semântico.” (AJA, 2012). Assim, segundo FELIPE (2007) “seu uso [LIBRAS], como meio de comunicação e como língua de instrução, precisa ser implantado nas escolas brasileiras que registram a matrícula de alunos surdos”. Dessa forma é visível a grande importância do desenvolvimento e divulgação de tecnologias que facilitam o seu ensino. Além de, a importância de promover um meio de aumentar a interação e proximidade com os cidadãos surdos em uma sociedade que em grande parte é dominada pelo “ouvintismo”, que de acordo com Skliar (1998), “é um conjunto de representações dos ouvintes, a partir do qual o surdo está obrigado a olhar-se e narrar-se como se fosse ouvinte”. Assim, para integrar essas duas vertentes, o sistema linguístico da Linguagem de Sinais LIBRAS e a capacidade dos dispositivos de interação é necessária à existência de uma inteligência que trate as informações e produza a resposta desejada, que nesse caso, se concentra no software.

Dispositivos como o Kinect da Microsoft e o Xtion Pro da Asus são exemplos de sensores de movimento. Ambos possuem um fator em comum, a tecnologia de sensoriamento 3D da companhia israelita PrimeSense, que segundo a própria empresa é uma tecnologia que permite “que dispositivos percebam o mundo em 3D e traduzam essas percepções em uma imagem sincronizada, assim como as pessoas fazem.” (PRIMESENSE, 2012). Em dezembro de 2010, a PrimeSense lançou seus próprios drivers de código aberto junto com middleware de monitoramento de movimento chamado NITE. A PrimeSense depois anunciou que tinha uma parceria com Asus para desenvolver um aparelho compatível com PC semelhante ao Kinect, conhecido como Xtion. A companhia também lançou a OpenNI que é uma “organização liderada pela indústria, sem fins lucrativos, formada para certificar e promover a compatibilidade e a interoperabilidade dos dispositivos de interação naturais, aplicações e middleware.” (OpenNI, 2012).

Junto com o NITE a PrimeSense desenvolveu o OpenNI framework, que é um conjunto de APIs de código aberto, que segundo a documentação disponibilizada pela organização, promove suporte para reconhecimento de comandos por voz, captura de gestos de mão e de movimentos de corpo.

Objetivos

Esse projeto tem por objetivo geral estudar as possibilidades de utilização, aperfeiçoamento e adaptação de uma solução de software desenvolvida pela PrimeSense, empresa por trás da tecnologia usada no Kinect, denominada OpenNI. Com a finalidade de entender as características do framework OpenNI, suas bibliotecas e APIs. E por fim, verificar até que ponto esse framework pode contribuir para o desenvolvimento de um software que junto com um dispositivo de captura de movimentos será capaz de interpretar a língua brasileira de sinais LIBRAS.

Dentre os objetivos específicos estão: continuar o estudo sobre dispositivos e softwares de interação natural; aumentar o entendimento sobre o funcionamento do framework OpenNI; prosseguir com o estudo dos recursos do framework, incluindo a realização de outros testes práticos com o Kinect para o mapeamento do movimento das mãos e, posteriormente, expressões faciais; analisar os aperfeiçoamentos de software necessários para utilização do OpenNI como interface básica entre os dispositivos de captura de movimentos; caso necessário, identificar outras possíveis tecnologias de hardware para captura de movimentos.

Metodologia

A metodologia para o desenvolvimento do projeto consistiu inicialmente em uma pesquisa na literatura disponível, sobre dispositivos de interação natural, com o objetivo de adquirir um conhecimento básico sobre o funcionamento e arquitetura desses equipamentos.

Após esta análise, iniciou o estudo específico sobre o framework OpenNI, baseando principalmente na documentação oficial presente em seu site (<http://www.openni.org/reference-guide/#.Uc9rHju-ovw>). O estudo consistiu em primeiro obter uma visão geral do framework, em seguida, na compreensão do

seu funcionamento, dos seus conceitos e sobre a sua API. Além de conhecer as funcionalidades nativas e fazer um levantamento de suas bibliotecas de classes entendendo as funcionalidades disponíveis pelo framework.

Para obter um melhor entendimento do OpenNi, além do estudo citado acima, foram realizados testes práticos do framework juntamente com o dispositivo de interação natural Kinect, aonde além de realizar experimentos com o OpenNi, foram testados também com o Nite que é uma biblioteca deste framework. Os testes consistiram em obter um reconhecimento do dispositivo quanto aos softwares utilizados. Através das funções do OpenNi, foram realizadas gravações feitas com o Kinect e pequenas aplicações para avaliação das tecnologias e obtenção de aprendizado.

Ainda será necessário analisar através de testes práticos até em qual ponto os recursos nativos do OpenNi são favoráveis para a coleta de movimentos quando principalmente feitos pelas mãos, levando em consideração os movimentos da Linguagem de Sinais Libras. Assim, será verificada a necessidade de desenvolver complementos para aprimorar as soluções fornecidas pelo framework.

Desenvolvimento

O primeiro passo da pesquisa consistiu em uma análise na literatura disponível sobre sensores de captura de movimentos e sobre a biblioteca OpenNi. A partir do levantamento realizado sobre os sensores, foi identificado que o dispositivo mais adequado para se iniciar a pesquisa seria o dispositivo Kinect da Microsoft. Devido a algumas características favoráveis, como por exemplo: preço, acessibilidade, configuração de hardware, documentação disponível e compatibilidade com o framework OpenNi. Em seguida, iniciou-se um estudo focado nesse dispositivo e em sua relação com o framework, onde buscou-se identificar quais seriam as ações necessárias para integrá-lo com a API fornecida pela biblioteca estudada. Após análise, foi possível constatar que devido a biblioteca OpenNi oferecer suporte a uma vasta gama de sensores 3D de captura de movimentos, para utilizá-la com o Kinect, seria necessário um driver específico do dispositivo, que é um requisito para se implementar a biblioteca com qualquer outro sensor compatível. Um driver é um programa que atua como um tradutor entre determinado dispositivo e as aplicações, dessa forma ele quem

possibilita a utilização de uma abstração de mais alto nível, que neste caso é o framework OpenNi, sobre o hardware, compreendido pelo sensor Kinect. Para se obter esse driver optou-se por fazer uso do SDK Kinect for Windows da Microsoft, que é um kit gratuito de desenvolvimento proprietário de aplicações não comerciais para o Kinect. Além da API em .NET disponível para o Kinect, o kit inclui os drivers necessários para se utilizar o dispositivo no Windows. Após se instalar o SDK, são visíveis os seguintes componentes conforme a Figura 1.

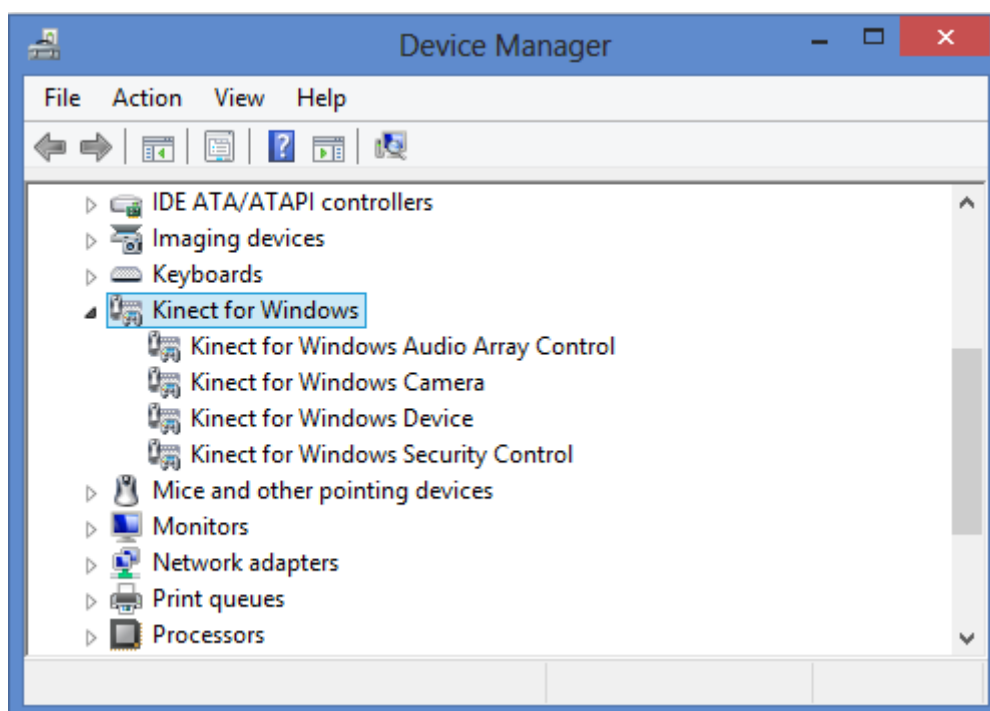


Figura 1 – Drivers para Kinect instalados a partir do SKD Kinect for Windows.

Dessa forma, finalizou-se a instalação dos componentes necessários para a utilização da biblioteca OpenNi com o sensor Kinect no sistema operacional Windows. A escolha da plataforma Windows como base para o desenvolvimento e testes ocorreu devido ao maior suporte disponibilizado pela Microsoft em relação à distribuição de drives, e a opção por se utilizar o Visual Studio como IDE para o desenvolvimento com o OpenNi.

Após as etapas de instalação e configuração dos drivers e softwares necessários, iniciou-se o estudo e aprendizagem da biblioteca. Por meio da documentação disponível no site da organização identificou-se as convenções da API em C++, descrição e detalhamento das funções das classes e dos métodos disponíveis, além de exemplos de códigos e aplicações. Depois dessa etapa de aprendizagem, focou-se em desenvolver pequenas aplicações para teste, com o

objetivo de analisar a estrutura do código, os dados de entrada e saída, e de estimar quais as possibilidades e funcionalidades que o OpenNi proporciona. Um exemplo de aplicação de teste que foi desenvolvida, foi um código no qual se fez uso de uma classe responsável por efetuar a gravação de um stream de vídeo. Antes de utilizar a classe foi necessário efetuar a chamada de um método que carrega o framework, e também foi preciso instanciar classes que representavam o dispositivo utilizado e qual tipo de sensor do dispositivo que seria utilizado na gravação. Abaixo, na Figura 2, é possível visualizar um frame capturado a partir de um stream de vídeo utilizando o sensor de profundidade.



Figura 2 – Drivers para Kinect instalados a partir do SKD Kinect for Windows.

Foi desenvolvida também uma simples aplicação que por meio de chamadas a métodos nativos, os quais retornam dados contendo propriedades dos dispositivos suportadas, exibe em uma tela de console apresentando as configurações disponíveis e ativas do dispositivo utilizado, que no ocasião foi o Kinect. A Figura 3 exibe um trecho da tela contendo as configurações detectadas.

```
C:\Windows\system32\cmd.exe
Inicializacao completa - OpenNI Versao 2.2.0.21
Dispositivos encontrados:
Name: [Kinect]
Uri: USB\VID_0409&PID_005A\6&295F6A4E&0&3
Fabricante: Microsoft
Usb Product Id: <null>
Vendor Id: <null>

Sincronizacao dos frames <Color/Depth>: False
Registration <Depth to Color>: OK

Dispositivo [Kinect] inicializado com sucesso.

Sensores identificados:

[Cor]
Modos de video suportados:
FPS: 12
Formato do pixel: 200
Resolucao X: 1280
Resolucao Y: 960

FPS: 30
Formato do pixel: 200
Resolucao X: 640
Resolucao Y: 480

FPS: 15
Formato do pixel: 201
Resolucao X: 640
Resolucao Y: 480

[Profundade]
Modos de video suportados:
FPS: 30
Formato do pixel: 100
Resolucao X: 640
Resolucao Y: 480

FPS: 30
Formato do pixel: 100
Resolucao X: 320
Resolucao Y: 240

FPS: 30
Formato do pixel: 100
Resolucao X: 80
Resolucao Y: 60

[Infravermelho]
Modos de video suportados:
FPS: 30
Formato do pixel: 200
Resolucao X: 640
```

Figura 3 – Tela do Console Application contendo as configurações do Kinect.

Os próximos passos do desenvolvimento da pesquisa consistiram em mais estudos e análises dos resultados obtidos através de testes com framework OpenNi, e também com alguns midlewares disponíveis, que implementam sobre a biblioteca funções mais avançadas e específicas de captura de movimentos com as mãos. Como por exemplo, o middleware Nite, que segundo a Primesense (2013) é o mais avançado e robusto middleware de visão computacional 3D disponível atualmente. E também o 3D Hand Tracking Library, “que possui um software que controla a posição em 3D, a orientação e a articulação de uma mão humana, a partir de observações visuais sem marcadores”. (OpenNi, 2013)

Resultados

Os resultados obtidos até então são animadores, pois além da realização do estudo teórico do OpenNi e de dispositivos de interação natural, conseguiu-se realizar testes práticos com ambos. Sendo o Kinect o equipamento escolhido, foi possível observar o reconhecimento deste sensor quanto ao framework e ainda desenvolver pequenas aplicações utilizando as funções nativas do OpenNi.

Com o objetivo de adquirir um domínio e maior conhecimento do OpenNi com o aparelho, desenvolveu-se uma pequena aplicação que exibe na tela as informações referentes às configurações do dispositivo utilizado. Esta codificação foi feita com a utilização de métodos nativos do OpenNi.

Outro exemplo do que foi construído, é um código-fonte que capta frames dos movimentos realizados através da gravação com o Kinect, permitindo escolher qual tipo de sensor do equipamento será utilizado.

Foi possível realizar de uma forma eficiente a gravação dos movimentos, levando em consideração os recursos fornecidos pelo OpenNi e pelo Kinect, como o sensor de profundidade. Pode ser um passo importante para a captação de sinais de Libras, pois alguns sinais necessitam da utilização das duas mãos com uma sobrepondo a outra, assim tornando necessário uma noção de profundidade.

Considerações Finais

Os resultados obtidos estão dentro do esperado. Foi realizado o estudo sobre o OpenNi e dispositivos de interação natural, incluindo a realização de testes práticos. Estes testes ocorreram entre o framework OpenNi com o dispositivo Kinect e houve o reconhecimento do aparelho quanto ao software e ainda foi viável desenvolver pequenas aplicações utilizando ambos, dessa forma, é possível observar a viabilidade deste estudo.

É importante que este trabalho tenha uma continuidade, pois embora o Kinect sozinho possua um bom poder de captação de movimentos, ele deixa a desejar quanto ao mapeamento de movimentos feitos pelos dedos e estes são primordiais ao se tratar de Libras. O framework OpenNi pode ser a saída para esta deficiência do dispositivo, possibilitando a captação destes movimentos.

Ao dar prosseguimento a este estudo, será possível observar em até qual ponto o OpenNi poderá ser útil para traduzir os sinais de Libras para português, e

assim documentar os resultados obtidos para que futuramente possa ser desenvolvida uma aplicação que possa contribuir para o povo surdo, que segundo Strobel (2006) é um grupo composto por “indivíduos que estão ligados por uma origem, tais como a cultura surda, costumes e interesses semelhantes, histórias e tradições comuns e qualquer outro laço”, que apesar de alguns esforços visíveis, ainda assim, enfrenta certas dificuldades de inclusão.

Fontes Consultadas

AJA - Associação do Jovem Aprendiz. Disponível em <<http://www.libras.org.br/libras.php>>. Acesso em 18 nov. 2012.

FELIPE, Tanya A. **LIBRAS em contexto: Curso básico: Livro do estudante.** 8ª ed. Rio de Janeiro: WalPrint, 2007.

OpenNI. Disponível em: <<http://openni.org/>>. Acesso em: 18 nov. 2012

PRIMESENSE. **Get a PrimeSense Sensor.** Nov 2012. Disponível em: <<http://www.primesense.com/en/getsensorsol>>. Acesso em: 19 nov. 2012.

SKLIAR, Carlos. **A Surdez: um olhar sobre as diferenças.** Porto Alegre: Editora Mediação, 1998.

STROBEL, **Surdos: vestígios culturais não registrados na história.** Dissertação de mestrado em fase de elaboração, na área de educação GES/UFSC, 2006.