



# CRICTE 2017

XXVIII Congresso Regional de Iniciação Científica e Tecnológica em Engenharia



## IMPLEMENTAÇÃO DE SOFTWARE DE LEITURA GESTUAL PARA O APRENDIZADO DE LIBRAS

**Lori R. F. Machado Filho, Gerson Battisti, Edson L. Padoin**

Professor Dr. Edson L. Padoin do curso de Ciência da Computação da Universidade Regional do Noroeste do Estado do Rio Grande do Sul - Unijui  
lori.jar@hotmail.com, battisti@unijui.edu.br

**Resumo.** Nesse trabalho foi implementado um software para o equipamento Kinect de leitura gestual que facilita o uso da língua brasileira de sinais (LIBRAS). Nesse documento são abordados as técnicas utilizadas para interpretação das imagens captadas pelo dispositivo Kinect bem como a tentativa de ensinar LIBRAS a população em geral.

**Palavras-chave:** LIBRAS, Kinect, Interpretação de Gestos.

### 1. INTRODUÇÃO

A área da tecnologia voltada para a interação humano-computador tem sido abordada de diversos aspectos. Um deles é a interação física com a máquina, isto é, a interação que a pessoa utilize do seu próprio corpo para manipular comandos na máquina, seja ela um computador, um console ou simplesmente um comando para que as luzes da casa sejam apagadas.

Foi com esse objetivo que a Microsoft desenvolveu o Kinect, um hardware capaz de reconhecer as vozes de seus usuários e captar, em profundidade, a presença de até seis pessoas em sua frente. Esse dispositivo desenvolvido primeiramente para o console XBOX 360, teve uma versão criada para seu sucessor, o XBOX ONE, e uma versão para computadores, o XBOX For Windows. Esse último citado foi desenvolvido com o foco para o desenvolvimento de aplicações

utilizando-se do Kinect. Seguindo a ideia de desenvolvimento, a Microsoft lançou em sequência ferramentas para auxiliar nessa implementação, e posteriormente disponibilizou guias e manuais de desenvolvimento para aplicações voltadas para a educação em seu site.

Diante disso, este trabalho propõe o desenvolvimento de uma ferramenta que faz uso do hardware Kinect da Microsoft para criar uma alternativa amigável para apresentações de projetos profissionais ou particulares, como a exposição de um trabalho em uma feira por exemplo, de maneira mais interativa, fazendo com que os gestos realizados pelo usuário disparem a passagem entre as transparências através de um algoritmo de interpretação de partículas

### 2. LIBRAS - LÍNGUA BRASILEIRA DE SINAIS

A comunicação entre a sociedade surda e muda é realizada através do espaço visual, por meio de gestos, movimentos e, não obrigatoriamente, feições. Tais gestos são assemelhados aos que usamos no cotidiano para auxiliar a fala, conforme comenta Alex T. S. Carneiro *et al.* [1]. A língua brasileira de sinais (LIBRAS) surgiu para obter um padrão na comunicação dessa comunidade e possibilitar a inclusão da mesma na sociedade. Porém, ainda que a LIBRAS tenha sido oficializada como nossa segunda língua em 2002, Godoy *et al.* [2] afirma que



seu uso ainda é muito restrito em meios de comunicação e ferramentas, inclusive nas informatizadas.

Com base nos conhecimentos citados, aliados à tecnologia, o ensino de LIBRAS com o uso de uma ferramenta interativa seria de suma importância para a elevação dos níveis de inclusão social da sociedade surda e muda na população.

### 3. MODELO PROPOSTO

O trabalho tem como base a utilização de um hardware com sensores aptos a captar gestos de uma pessoa em conjunto com bibliotecas de código fonte livre para a implementação do software. Foi utilizada uma Interface de Programação de Aplicativos (API) que faz uso de duas outras APIs internamente, funcionando como uma espécie de driver para possibilitar a extração dos dados de dispositivo com sensores de leitura gestual e processá-los. Também foi utilizado uma biblioteca para reconhecer os dados binários para que fosse possível utilizar um algoritmo de interpretação de partículas.

#### 3.1 OpenKinect for Processing

Para a implementação do sistema, uma das interfaces de programação utilizadas foi a desenvolvida por Shiffman *et al.* [3], cuja funcionalidade é integrar a API OpenKinect, que tem a função de extrair os dados do dispositivo de hardware Kinect da Microsoft, e a API Processing, responsável pelo processamento desses dados.

Partindo da ideia de possuir dados processados provenientes do Kinect, foi empregado uma biblioteca chamada BlobScanner, capaz de auxiliar na criação do algoritmo de interpretação de partículas porposta para o software de leitura gestual.

#### 3.2 Microsoft Kinect

O dispositivo de interação humano-computador utilizado é conhecido como Kinect, da Microsoft. Tal dispositivo é utilizado, principalmente, para fins de entretenimento sem a utilização de um Joystick (Controle geralmente utilizado para ativar comandos de um Console), comenta da Silva *et al.* [4].

Fundamentado nas características do dispositivo citadas Zhang *et al.* [5] e Correia *et al.* [6] como, o sensor de profundidade, a câmera de cor, e uma matriz de quatro microfones que fornecem captura de movimento 3D de corpo inteiro, reconhecimento facial, e as capacidades de reconhecimento de voz, que foi escolhido o equipamento Kinect para a implementação desse sistema.

### 4. IMPLEMENTAÇÃO

O software proposto baseou-se na premissa de que o usuário não possuiria conhecimento inicial algum em LIBRAS, obtendo todo seu saber a partir da ferramenta desenvolvida, ou seja, assimilaria a língua como L2, isto é, seria sua segunda língua. A partir disso, o sistema disponibilizará de três funcionalidades principais, sendo elas:

- O ensinamento das letras, em sequência, de L-I-B-R-A-S;
- Um exercício composto de cinco letras ou números aleatórios para que o usuário realize os seus respectivos gestos na Língua Brasileira de Sinais;
- Uma tarefa para expor o aprendizado obtido, possibilitando o usuário se comunicar em LIBRAS com o Kinect e o mesmo responder o que o usuário está tentando gesticular na linguagem em questão.

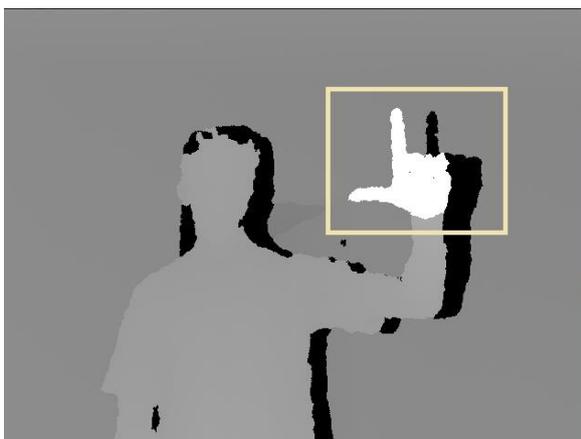
Além disso, o sistema desenvolvido é disposto de uma configuração que viabiliza o usuário as seguintes alternativas:



- Visualizar o percentual de comparação entre o gesto do usuário com a respectiva letra ou número proposto para execução;
- Visualizar um contorno gráfico ao redor da mão do usuário indicando os pontos de colisão com a letra ou número proposto para execução;
- Visualizar a imagem da letra ou número em forma de contorno no momento de execução da tarefa para auxiliar como o usuário deve efetuar o respectivo gesto;
- Visualizar a imagem da letra ou número que é utilizada como base para realização do algoritmo de colisão de pontos com o fim de identificar o respectivo gesto em LIBRAS;
- Selecionar a leitura gestual para destros ou canhotos;

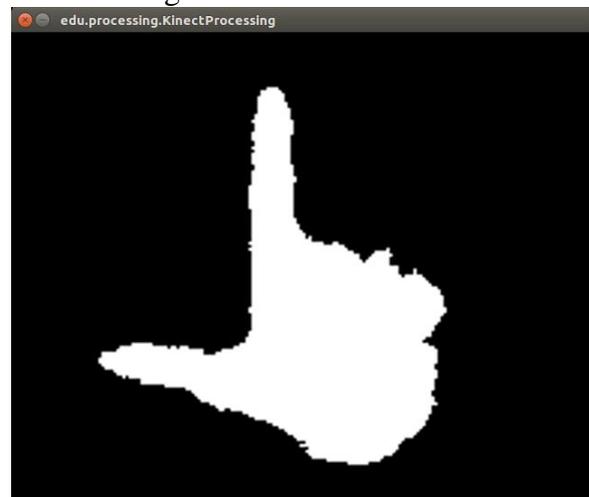
Para desempenhar os processos mencionados, foi realizada uma função na ferramenta para que a mesma execute uma aproximação diretamente no quadro de maior quantidade de pontos, no caso, a mão do usuário. Para tal, deve-se posicionar a mão rente ao Kinect em uma distância aproximada de 90 centímetros. Feito isso, o sistema realizará um escalonamento da imagem captada pelo sensor e mostrará na Interface Gráfica do Utilizador (GUI - *Graphical User Interface*) apenas uma região de interesse (ROI - *Region of Interest*) de 194 x 158 pixels. O tamanho do ROI foi escolhido de forma aleatória como ilustração seguinte na Fig. 1, onde o quadro em amarelo é a região de interesse do sistema

Figura 1. ROI – Região de Interesse



A figura 1 está retornando cores com tons de cinza para o fundo e branco para a região onde é realizado o algoritmo de colisão de pontos. Para facilitar o confronto de pontos, a região distinta do ROI é escurecida totalmente, fornecendo maior contraste para o branco como é muito bem evidenciado na fig.2.

Figura 2. ROI escalonado



Por fim, a próxima ilustração será da execução da primeira tarefa com todas as configurações acionadas, isto é, será delineado elipses no perímetro onde há pontos de colisão, desenhando um contorno de auxílio para o usuário e adicionado um percentual para verificação de comparação com a letra ou número proposto, a apresentação de uma imagem de auxílio que é utilizada como base para a realização do algoritmo de colisão de pontos, ao qual é demonstrado na fig. 3. Contudo, a opção para o tipo de usuário será para pessoas destros. Portanto, o usuário terá total auxílio na execução das tarefas propostas. A sequência referida é a execução das letras L-I-B-R-A-S em sequência, então será demonstrado a tentativa de execução apenas da letra L, conforme ilustração na fig. 4:



Figura 3. Letra L utilizada como base na identificação do gesto

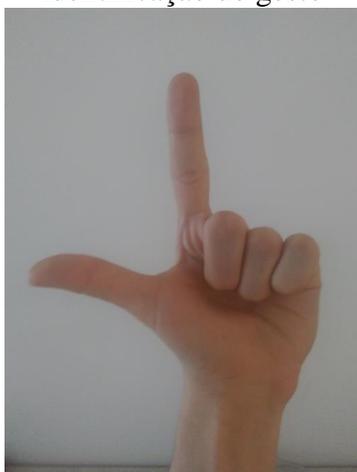
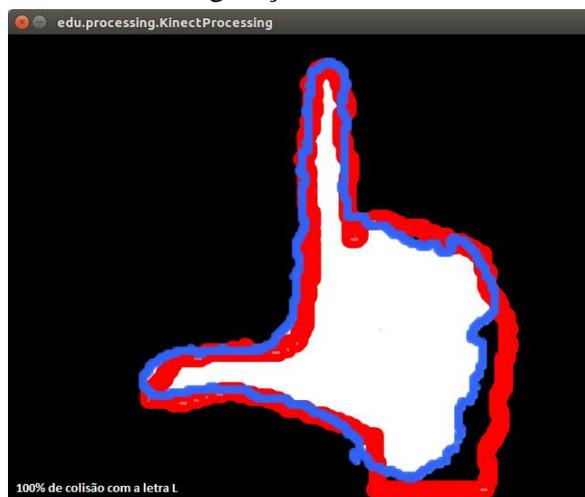


Figura 4. Demonstração com todas as configurações auxiliares



## 5. CONCLUSÃO

Efetuuou-se a análise da viabilidade do desenvolvimento de uma ferramenta com a capacidade de integrar sua aplicação com um dispositivo de hardware responsável pela leitura gestual de uma pessoa, fazendo dessa união um software de ensino da língua brasileira de sinais. Essa análise obteve um resultado positivo, com base no sucesso do desenvolvimento da ferramenta proposta.

A ferramenta desenvolvida para o ensino de LIBRAS foi baseada em um dicionário limitado de letras ou números obtidos através do próprio software. Ademais, o tempo de resposta do software é fundamentado nesse mesmo dicionário. A partir dessa ideia, é possível concluir que o tempo de resposta do algoritmo utilizado aumenta linearmente conforme o tamanho do dicionário.

Contudo, limitando-se as disponibilidades de processamento do hardware utilizado, o tempo de resposta foi considerado hábil para uma boa interação entre o usuário e o software, evitando o desinteresse do usuário pela ferramenta.

## 6. REFERÊNCIAS

- [1] A. Carneiro and P. Cortez and R. Costa “Reconhecimento de Gestos da LIBRAS com Classificadores Neurais a partir dos Momentos Invariantes de Hu”. Interaction, pp. 190-195, 2009.
- [2] V. de M Godoy. “Método para Classificação de um Conjunto de Gestos usando Kinect”. Tese de Doutorado. Pontifícia Universidade Católica do Paraná. 2013
- [3] D. Shiffman. “Getting Started with Kinect and Processing”. <http://shiffman.net/p5/kinect/>.
- [4] R. K. da Silva. “Interfaces Naturais e o Reconhecimento das Línguas de Sinais”. Tese (Mestrado em Tecnologias da Inteligência e Design Digital) - Pontifícia Universidade Católica de São Paulo, São Paulo. 2013
- [5] Z. Zhang. “Microsoft kinect sensor and its effect”. MultiMedia, IEEE, v. 19, n. 2, 2012, pp. 4-10.



# CRICTE 2017

XXVIII Congresso Regional de Iniciação Científica e Tecnológica em Engenharia



- [6] M. Correia. “Reconhecimento de elementos da língua gestual portuguesa com Kinect”. Tese (Mestrado Integrado em Engenharia Eletrotécnica e de Computadores) - Faculdade de Engenharia da Universidade do Porto,