

CAPTURA DE GESTOS ATRAVÉS DO KINECT™ E TRADUÇÃO PARA O WEKA: O ESTUDO DE CASO DO RECONHECIMENTO DE SINAIS DE LIBRAS

INESSA LUERCE¹; PABLO MARGREFF²; MARILTON AGUIAR³

¹Universidade Federal de Pelotas – idluerce@inf.ufpel.edu.br

²Universidade Federal de Pelotas – pmargreff@inf.ufpel.edu.br

³Universidade Federal de Pelotas – marilton@inf.ufpel.edu.br

1. INTRODUÇÃO

Os progressos recentes em sistemas de entretenimento e de jogos vêm trazendo novas possibilidades de interação do usuário com os dispositivos. Tecnologias como o Microsoft Xbox Kinect™, baseadas na captura de movimentos do corpo humano, possibilitam o reconhecimento de gestos, que podem ser interpretados como forma direta de comunicação. Este recente paradigma que adota sensores para o reconhecimento de gestos permite uma nova forma de comunicação com o computador, favorecendo a interação através do movimento corporal.

A integração de técnicas de mapeamento, reconhecimento de padrões e aprendizagem de máquina pode ser utilizada para interpretação de gestos específicos. É neste contexto que o projeto LIKI - *Libras e Kinect Aplicados às Tecnologias Assistivas*, em desenvolvimento pelo grupo de pesquisa, está inserido. O projeto busca propor modelos e ferramentas computacionais para auxiliar no processo de ensino-aprendizagem da Língua Brasileira de Sinais (LIBRAS) para crianças com algum tipo de deficiência auditiva.

Mais especificamente, uma das temáticas abordadas pelo projeto trata de capturar, especificar e armazenar de modo apropriado os movimentos do usuário utilizando o sensor Kinect™ (versão 2). Outras temáticas abordadas no projeto envolve o desenvolvimento e a aplicação de técnicas de reconhecimento de padrões para a identificação dos gestos e, então, integrá-las em um software que será utilizado em sala de aula.

Assim, este artigo pretende apresentar os resultados preliminares alcançados com a captura dos gestos (e seus dados) obtidos a partir do Microsoft Kinect™ e sua adaptação (tradução) para o formato de entrada do software de aprendizado de máquina WEKA (HALL et. al, 2009). A tradução dos dados capturados para o WEKA permitirá a agilização da etapa de definição da técnica de Inteligência Artificial mais apropriada para o reconhecimento dos sinais de LIBRAS.

2. METODOLOGIA

Como parte importante deste trabalho, foi investigado o uso de tecnologia Kinect™ e suas bibliotecas que permitem a captura de gestos. Posteriormente, o foco foi no tratamento e armazenamento dos dados capturados, onde foram definidos os meios para a aquisição, análise de elementos gestuais, processamento dos gestos e gerenciamento das informações.

O Kinect™ é formado por um conjunto de tecnologias capazes de permitir uma interação mais natural entre o homem e a máquina. Originalmente foi

lançado para a plataforma de jogos Xbox e, posteriormente, passou a ser compatível com computadores pessoais (BBC, 2012). Desde então, o uso não ficou restrito aos jogos, possuindo pesquisas e aplicações em diversas áreas, como vendas, terapia, saúde, educação e treinamento físico (MICROSOFT BUILD, 2014).

A segunda versão do dispositivo, utilizada neste projeto, possui uma câmera tridimensional (3D) possibilitando um maior alcance de profundidade e de cor; e, sua imagem tem três vezes mais precisão que a versão anterior do mesmo dispositivo. Essa câmera possui resolução de 1080p e a captura independe da luminosidade, uma vez que possui sensor infravermelho. Além de reconhecer 25 pontos do corpo, o Kinect™ ainda percebe três estados diferentes da mão: aberta, fechada e *lasso*, que constitui da mão parcialmente aberta, com apenas dois dedos visíveis (MICROSOFT, 2015).

Desta forma, a escolha por este dispositivo deu-se primordialmente em virtude da quantidade de sensores que ele oferece. Outros fatores decisivos para a escolha incluem a vasta documentação fornecida por seu desenvolvedor; e, o baixo preço de mercado, que facilita a aquisição do dispositivo, tanto para fins de pesquisa, quanto para o usuário final da aplicação.

A documentação do *software* é uma peça fundamental para um rápido aprimoramento das aplicações desenvolvidas. Empresas consolidadas, como a Microsoft, disponibilizam ferramentas e manuais de seus produtos a fim de incentivar programadores a colaborar em seus projetos. A comunidade desenvolvedora é, frequentemente, bastante participativa e disponibiliza códigos e informações para tornar o desenvolvimento mais ágil.

3. RESULTADOS E DISCUSSÃO

Como resultado deste trabalho, foi desenvolvida uma aplicação que utiliza o sensor do Kinect™ para coletar os dados dos gestos capturados e traduzi-los para o formato de entrada do *software* WEKA. Essa ferramenta tomou como base uma *feature* do projeto LightBuzz, que converte os dados adquiridos pelo Kinect™ para o formato CSV (LIGHTBUZZ, 2014). Neste caso, a ferramenta gera arquivos no formato ARFF, que podem ser lidos pelo *software* WEKA de forma mais eficaz, comparada a outros formatos de arquivos, como o CSV. Um trecho do formato ARFF pode ser observado na Figura 1, onde NeckX indica a localização no eixo x do pescoço, seguida do tipo de dado (numérico) e os valores para aquele ponto.

```

@RELATION gesto

@ATTRIBUTE SpineBaseX NUMERIC, -0.490301, 1.613542, 0.2432985
@ATTRIBUTE SpineBaseY NUMERIC, -0.4903264, 1.611722, 0.2435638
@ATTRIBUTE SpineBaseZ NUMERIC, -0.4903753, 1.611909, 0.2435709
@ATTRIBUTE SpineMidX NUMERIC, -0.4904682, 1.612097, 0.2435234
@ATTRIBUTE SpineMidY NUMERIC, -0.4906055, 1.612812, 0.2434609
@ATTRIBUTE SpineMidZ NUMERIC, -0.4906742, 1.613139, 0.243412
@ATTRIBUTE NeckX NUMERIC, -0.4908054, 1.613542, 0.2432985
@ATTRIBUTE NeckY NUMERIC, -0.4908366, 1.613801, 0.2431799
@ATTRIBUTE NeckZ NUMERIC, -0.4908977, 1.614082, 0.243052
@ATTRIBUTE HeadX NUMERIC, -0.4909391, 1.614676, 0.2427911
@ATTRIBUTE HeadY NUMERIC, -0.4911005, 1.615549, 0.2425586
@ATTRIBUTE HeadZ NUMERIC, -0.4911621, 1.616527, 0.241918
@ATTRIBUTE ShoulderLeftX NUMERIC, -0.4912354, 1.616923, 0.2413603
@ATTRIBUTE ShoulderLeftY NUMERIC, -0.4913764, 1.617304, 0.2411804
@ATTRIBUTE ShoulderLeftZ NUMERIC, -0.4915635, 1.617863, 0.2409935

```

Figura 1 – Exemplo de arquivo no formato .arff, com dados extraídos do Kinect™

Quando a aplicação é iniciada, automaticamente inicia-se a captura dos gestos através do sensor Kinect™; e, os dados dos eixos x, y e z são gravados. Todos os frames coletados através desta gravação são salvos em um arquivo, o qual é utilizado para seleção dos dados que serão utilizados propriamente. A captura dos dados utiliza-se da abscissa, ordenada e cota (eixos x, y e z, respectivamente), das 25 juntas do corpo e possuem seus valores medidos em metros. O Kinect™ fornece diversas informações para realizar a análise de gestos como: a imagem do esqueleto com 25 junções, a profundidade dos elementos no seu campo de visão, uma imagem em alta definição e uma imagem em infravermelho. Um exemplo do que pode ser capturado com o sensor pode ser observado na Figura 2.



Figura 2 - Mapeamento do corpo com as visões de Color, Infrared, Body, e BodyIndex no Microsoft Kinect Studio™

Entretanto, após a realização destes testes preliminares, pode-se perceber que a Língua Brasileira de Sinais é bastante complexa (ACESSIBILIDADE BRASIL, 2008), de modo que as informações obtidas apenas pelo mapeamento destes 25 pontos não são suficientes para classificação e reconhecimento de gestos de forma precisa.

Assim, torna-se necessária a avaliação dos outros tipos de dados capturados pelo Kinect™, que possam ser relevantes para a classificação do movimento e o reconhecimento como um gesto do alfabeto LIBRAS.

4. CONCLUSÕES

O presente trabalho possibilitou a captura de dados extraídos através do Microsoft Kinect™ e a geração de um arquivo possível de ser lido pelo software de aprendizado de máquina WEKA. Neste estágio de desenvolvimento, a contribuição deste trabalho foi de possibilitar a manipulação dos dados obtidos pelo sensor, para que possam ser utilizados para treinamento e classificação, a fim de gerar ou definir um modelo que possa reconhecer os movimentos obtidos e classificá-los dentro do conjunto de gestos pertencentes à Lingua Brasileira de Sinais.

Em um contexto mais amplo, este trabalho pode ser utilizado para qualquer projeto que necessite extrair informações possíveis de serem obtidas pelo Kinect™, como mapeamento ou posicionamento do corpo e movimentos; e que tenha a finalidade de utilizar esses dados para treinamento, classificação ou avaliação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

Acessibilidade Brasil. **Dicionário da Língua Brasileira de Sinais.** Dicionário Libras Web, 2008. Acessado em 9 jun. 2015. Online. Disponível em: <http://www.acessibilidadebrasil.org.br/libras/>

BBC. **Kinect for Windows gesture sensor launched by Microsoft.** BBC News Technology, 8 mar. 2012. Acessado em 16 jul. 2015. Online. Disponível em: <http://www.bbc.com/news/technology-16836031>

LightBuzz. **Kinect2CSV.** 2014. Acessado em 16 jul. 2015. Online. Disponivel em: <https://github.com/LightBuzz/Kinect-2-CSV.git>

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten. **The WEKA Data Mining Software: An Update;** SIGKDD Explorations, 2009. Volume 11, Issue 1.

Microsoft. **Kinect for Windows.** Acessado em 9 jun. 2015. Online. Disponível em: <https://www.microsoft.com/en-us/kinectforwindows/>

Microsoft Build. **Kinect 101: Introduction to Kinect for Windows.** Youtube, 21 abr. 2014. Acessado em 16 jul. 2015. Online. Disponivel em: <http://www.microsoft.com/en-us/kinectforwindows/develop/how-to-videos.aspx>