

**1.03.99 - Ciência da Computação.**  
**LUVA DE CAPTAÇÃO DE GESTOS PARA INTERAÇÃO HUMANO-COMPUTADOR.**

Felipe A. Quirino<sup>1</sup>, Alessandro Girardi<sup>2</sup>

1. Acadêmico de Ciência da Computação na Universidade Federal do Pampa (UNIPAMPA)
2. Professor na Universidade Federal do Pampa/Orientador

**Resumo**

O uso de dispositivos eletrônicos portáteis está difundido ao redor do mundo. No entanto, devido ao seu tamanho reduzido, a interação com o ser humano nem sempre é fácil. Além disso, é necessário desenvolver novas interfaces que permitam uma interação humano-computador mais eficiente. Este artigo descreve o desenvolvimento de uma luva de captura de gestos capaz de reconhecer padrões de movimento da mão e transformá-los em informações úteis. O sistema proposto consiste em seis acelerômetros e seis giroscópios localizados na ponta dos dedos e sobre a parte posterior da mão, permitindo a detecção precisa de movimentos no espaço tridimensional. Uma série de aplicativos foi desenvolvida, como o reconhecimento de padrões de gestos de linguagens de sinais. A luva pode substituir joysticks convencionais com a vantagem de inserir mais graus de liberdade no controle. Também foi desenvolvido um modelo de mão virtual em 3D controlado pela luva.

**Palavras-chave:** IHC; luva inteligente; LIBRAS

**Apoio financeiro:** CNPq

**Introdução**

O uso de dispositivos portáteis, como smartphones, smartwatches e tablets, está se tornando cada vez mais difundido entre os usuários comuns. Esses dispositivos facilitam o dia a dia das pessoas e são fundamentais para o desenvolvimento de novas aplicações (ROMANSSINI, 2018). No entanto, por causa do tamanho reduzido, a interação com o ser humano é difícil, especialmente a entrada de dados através de digitação em teclados virtuais na tela. Escrever um texto de tamanho razoável nesses dispositivos é desconfortável e pode causar danos físicos ao corpo - como LER (lesão por esforço repetitivo) e tendinite - se usado por um longo período. Assim, é necessário desenvolver novas interfaces para melhorar a interação humano-computador com menos esforço para o usuário (HAN, 2013).

Este projeto visa desenvolver uma luva de detecção de gestos capaz de reconhecer padrões de movimento da mão e transformá-los em informações úteis.

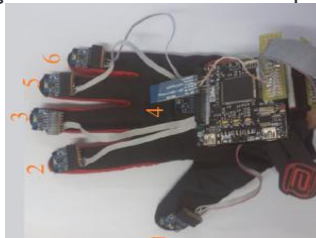
O sistema proposto consiste em seis acelerômetros e seis giroscópios (INVENSENSE, 2013) localizados na ponta dos dedos e sobre a parte posterior da mão, permitindo a detecção precisa de movimentos no espaço tridimensional. Os sensores são lidos por um microcontrolador, o qual processa os valores de aceleração linear e angular e envia as informações para o dispositivo final (computador, smartphone ou tablet).

A luva desenvolvida possibilita muitas aplicações, a exemplo do reconhecimento de padrões de movimento na língua de sinais para escrita em equipamentos digitais (CHOUHAN et al., 2014). Permite também que pessoas com necessidades especiais (cegas, surdas ou mudas, por exemplo) interajam mais eficientemente com um computador ou com outras pessoas. A luva pode substituir joysticks convencionais com a vantagem de inserir mais graus de liberdade na interação com o dispositivo. Outros usos também podem ser citados, como controle de jogos digitais, interface de realidade virtual, detecção de movimento para exercícios de fisioterapia, treinamento para acadêmicos de saúde e controle remoto de cirurgias. Além disso, a luva pode ser usada como uma plataforma para atividades de ensino através do desenvolvimento de aplicativos que podem explorar suas funcionalidades.

**Metodologia**

O gesto da mão é capturado por seis acelerômetros e giroscópios do tipo MPU 6050, fornecendo dados de aceleração linear e angular para cada dedo e para a área dorsal da mão. A figura 1 mostra a localização dos sensores na luva. O microcontrolador PAMPIUM, desenvolvido pelo nosso grupo de pesquisa (ENGROFF, 2014), é usado para controle, realizando configurações e capturando dados dos sensores. Ele está localizado sobre a mão, integrado em um FPGA. Um módulo Bluetooth transmite dados para um dispositivo final, como um computador ou smartphone, permitindo assim total liberdade de movimento para o usuário.

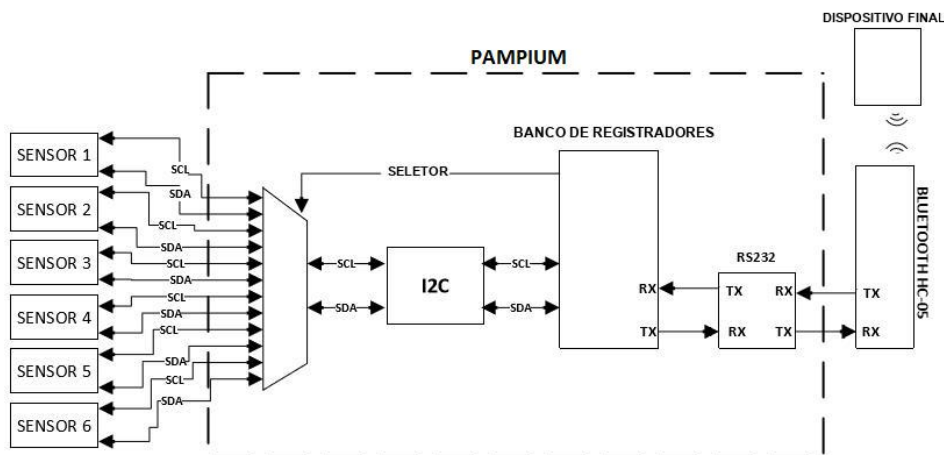
Figura 1: Implementação do hardware da luva para captação de gestos.



A figura 2 mostra o sistema implementado. A comunicação entre os sensores e o PAMPIUM é feita através do protocolo de comunicação I2C. A comunicação entre o PAMPIUM e a interface Bluetooth (HC05) é implementada pelo protocolo de comunicação RS-232. Finalmente, a comunicação com o dispositivo final acontece através do protocolo Bluetooth, que não requer nenhum fio. As interfaces de comunicação I2C e RS-232 são implementadas em FPGA em conjunto com o microcontrolador PAMPIUM. Para isso, usamos uma pequena placa de desenvolvimento FPGA 101 colocada no topo da mão.

Considerando que são utilizados seis sensores MPU 6050, é necessário adotar uma unidade multiplexador/demultiplexador para comutar a comunicação entre os sensores e o microcontrolador. Isso permite que apenas um único canal de comunicação I2C seja implementado no PAMPIUM. Considerando que todas operações do PAMPIUM ocorrem nos registradores, conectamos as interfaces de comunicação diretamente no banco de registradores, facilitando a implementação de protocolos de controle tanto para envio quanto para recebimento de dados.

Figura. 2: Diagrama de bloco da luva para captação de gestos.



Cada MPU 6050 possui um acelerômetro de três eixos, responsável por medir a aceleração linear nos eixos x, y e z, além de um giroscópio de três eixos que mede a aceleração angular sobre os eixos x, y e z, obtendo assim um boa noção de espaço tridimensional. O acelerômetro é capaz de medir acelerações nos três eixos separadamente até 16g (156,896 m/s<sup>2</sup>) e possui quatro faixas programáveis: 2g, 4g, 6g e 16g. A unidade g refere-se ao valor da aceleração gravitacional da Terra. O giroscópio é capaz de medir variações instantâneas nos três eixos separadamente com quatro faixas programáveis: 250°/s, 500°/s, 1000°/s e 2000°/s (INVENSENSE, 2013). A comunicação através do padrão I2C ocorre a uma taxa máxima de 400 kHz e opera em uma faixa de tensão de alimentação de 2,37 V a 3,46 V.

O MPU 6050 digitaliza os valores dos eixos x, y e z coletados dos giroscópios e acelerômetros por meio de seis conversores A/D (análogo/digital) - três A/D para o acelerômetro e três A/D para o giroscópio. Os valores de leitura são armazenados em registradores internos de 16 bits.

O HC05 é um módulo Bluetooth SPP (protocolo de porta serial), projetado para comunicação serial sem fio transparente, que facilita a interface com o dispositivo final.

Implementamos a interface de comunicação I2C para que seu controle seja exercido diretamente pelos registradores do PAMPIUM. Para isso, alocamos alguns bits de um registrador específico e os conectamos à interface I2C. A comunicação começa quando um sinal de ativação é ativado. Então, a interface envia uma condição de partida (S) e entra em um loop esperando pelo sinal BUSY ir para o nível lógico 1. Analisando o bit R/W, que indica se a próxima ação será escrita ou leitura, o módulo de controle executa uma ação. Se R/W = 0, a interface executa uma rotina de escrita e verifica o bit de reconhecimento (ACK). Se ACK = 0, a condição de parada (P) é executada. No caso de ACK = 1, uma segunda verificação é feita, desta vez para o sinal S, que indica se uma nova condição inicial deve ser executada. Se R/W = 1, a interface executa uma rotina de leitura S. Se S = 1, a condição final da transmissão (NACK) é executada, seguida pela condição de parada. Se S = 0, a interface retorna ao loop e aguarda BUSY = 1 para executar uma nova ação. Esse ciclo se repete até que a rotina de escrita ou leitura de dados seja concluída.

Também implementamos a interface RS-232 de maneira semelhante à interface I2C, na qual todos os pinos de controle são conectados diretamente ao banco de registradores. Para isso, alocamos alguns bits da interface de controle para alguns registradores dedicados no banco de registradores. Para a comunicação entre a interface e o módulo Bluetooth, utilizamos dois pinos (RX e TX). A interface é iniciada quando um sinal de habilitação é definido para o nível lógico 1. A taxa de transmissão é detectada automaticamente por meio do procedimento de transferência automática. Para isso, a interface lê a entrada aguardando um byte específico do dispositivo conectado a ela. Este byte foi definido como 10101010 (caractere 'U' em código ASCII). A interface calcula o número de ciclos de clock necessários para receber este byte. Dividindo o número de ciclos por 8, obtém-se a velocidade de transmissão. Depois de ajustar a velocidade de comunicação, uma rotina de

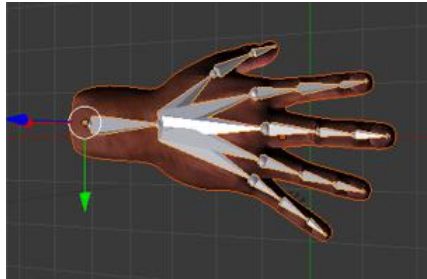
teste começa monitorando se a interface deve receber ou enviar os próximos dados. Primeiro, o pino RX é verificado. Se  $RX = 0$ , uma rotina de leitura é iniciada. Se  $RX = 1$ , o controle testa o bit GO. Se  $GO = 1$ , a rotina de escrita é executada. No final dessa rotina, o pino GO é redefinido e o controle volta para a etapa de verificação.

## Resultados

Implementamos uma interface para ler e processar dados da luva por um computador ou smartphone (dispositivo final). Esta interface recebe dados de acelerômetros e giroscópios da luva através do protocolo Bluetooth e processa as informações de acordo com a aplicação. Um pré-tratamento dos dados é necessário para a supressão de ruído e detecção de rotação. Os dados do sensor são normalizados em uma faixa de -10 a +10 para os acelerômetros e de  $-\pi$  a  $+\pi$  para os giroscópios.

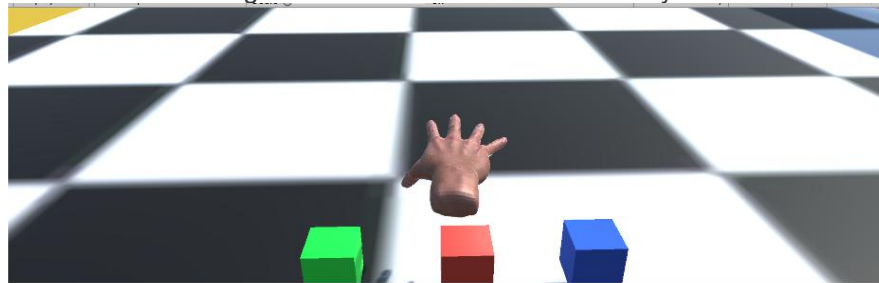
A capacidade total da luva pode ser demonstrada com um modelo de mão 3D. Este modelo virtual reproduz o movimento real da mão por meio dos valores lidos dos acelerômetros e giroscópios. Cada movimento dos dedos é independente dos outros, o que proporciona um controle realista do modelo 3D. O modelo 3D foi implementado no framework Unity. O modelo é composto de ossos para a montagem de cada parte do dedo. A figura 3 mostra a figura do modelo virtual da mão, onde os ossos são indicados em branco.

Figura 3 - Modelo de mão 3D. Os ossos estão indicados em branco.



Os ossos são conectados uns aos outros para que as articulações reproduzam um movimento realista da mão. A referência fixa está localizada no pulso. Este modelo é um protótipo para qualquer projeto usando um espaço 3D, como jogos de realidade virtual, simulações ou controle remoto. A Figura 4 ilustra o modelo da mão 3D virtual renderizada controlada pela luva no Unity.

Figura 4 - 3D Modelo da mão no Unity.



O Unity é uma ferramenta popular para criação de jogos. Isso facilita o trabalho porque parte do cenário do jogo é feito arrastando e soltando objetos. Além disso, fornece métodos nativos para o controle dos ossos da mão, o que permite que os dedos se movam em uma determinada distância, mantendo a integridade da mão.

O movimento de cada dedo é calculado pela diferença entre os valores atual e anterior, lidos pelo giroscópio relacionado. Essa diferença é adicionada à posição atual do osso na ponta do dedo. Cada dedo tem três falanges, exceto o polegar, que tem duas. A posição dos ossos é estimada para dar a impressão de fechar ou abrir a mão. Esta estimativa é feita pela multiplicação da posição do osso por um coeficiente da diferença calculada do giroscópio. Os coeficientes foram definidos como 1,3 para o distal, 1 para o médio e 0,7 para a falange proximal. O movimento da mão inteira é calculado de acordo com o valor lido no acelerômetro localizado no sensor 4 na parte posterior da mão.

Para pegar um objeto, usamos um método de colisão como um gatilho que é ativado quando a mão está aberta e se fecha perto de um objeto. Os eixos locais do objeto tornam-se filhos da mão, seguindo assim o movimento dela. Quando a mão é aberta, os eixos do objeto deixam de ser nó filho, fazendo com que ela libere o objeto.

Uma aplicação interessante para a luva desenvolvida é o reconhecimento de gestos como letras, em que o usuário faz um certo movimento e a aplicação entende como uma tecla pressionada no teclado. Adotamos o alfabeto da língua brasileira de sinais (LIBRAS) (FELIPE, 2001), como padrão para o reconhecimento de gestos. A LIBRAS é a língua brasileira oficial de sinais e seu uso é difundido em todo o país. Com a combinação de vários gestos, é possível soletrar palavras e frases completas.

O reconhecimento automático dos gestos com a luva e a tradução para uma letra no teclado permite a digitação em um dispositivo portátil (um smartphone, por exemplo) de uma maneira intuitiva para uma pessoa cega, surda ou muda.

O reconhecimento de gestos é feito pelo uso de algoritmos de classificação. Existem vários algoritmos para essa finalidade na literatura, por isso é importante fazer uma análise de algoritmos de classificação para determinar o mais adequado para esta aplicação. O treinamento desses algoritmos requer a divisão de dados em classes. O espaço de projeto é dividido em 28 classes correspondentes a todas as letras do alfabeto, além dos símbolos para "espaço" e "inativo". Os dados de entrada possuem 18 variáveis, compostas pelas medições de seis giroscópios de 3 eixos previamente descritos, capturados da luva em um dado instante.

Selecionamos seis algoritmos de aprendizado de máquina para avaliar qual se adapta melhor a nossa aplicação e produz o melhor desempenho em termos de precisão e tempo de previsão. Eles são os seguintes: Máquina de Vetor de Suporte (SVM), Vizinhos mais Próximos (KNN), Análise de Discriminantes Lineares (LDA), Análise de Discriminantes Quadráticos (QDA), Regressão Logística (LR) e Floresta Aleatória (RF) (JAMES, 2013).

As métricas para avaliar o desempenho do algoritmo de classificação são precisão, recall, F1-Score e o tempo de predição. Uma classe prevista pode resultar em quatro casos: verdadeiro positivo (TP), falso positivo (FP), verdadeiro negativo (TN) e falso negativo (FN). Essas métricas definem precisão (PPV), recall (TPR) e F1-Score (SOKOLOVA, 2006). Para o tempo de predição, é necessário que o algoritmo leve menos de 10 ms, porque este é o tempo de atualização dos sensores.

A Tabela 1 mostra o comparativo para precisão, recall, F1-Score e o tempo de predição para os algoritmos implementados. O SVM e o LDA alcançaram melhor precisão, com 94% de previsões corretas. O pior resultado foi obtido pela KNN, com 88%. Considerando o F1-Score e o recall, o SVM e o LDA apresentaram novamente os melhores valores. SVM alcançou 0,93 para TPR e F1-score, e o LDA obteve 0,91 para ambas as métricas. O alto F1-Score demonstra que poucas classes possuem alta taxa de erro e que em geral todas apresentam baixa taxa de erro. Estes resultados demonstram que todas as heurísticas podem ser usadas para previsão em tempo real na luva, porém o SVM demonstrou o melhor desempenho geral.

Tabela 1: Métricas de performance estimadas para os algoritmos de classificação.

Algoritmo	PPV	TPR	F1-score	Tempo
KNN	0.88	0.87	0.87	0.65ms
RF	0.91	0.90	0.90	6.55ms
QDA	0.92	0.90	0.90	0.84ms
LR	0.93	0.91	0.91	0.05ms
SVM	0.94	0.93	0.93	0.14ms
LDA	0.94	0.91	0.91	0.10ms

## Conclusão

Desenvolvemos uma luva para captura de gestos baseada em acelerômetros e giroscópios. A comunicação entre a luva e um computador é sem fio, realizada através do protocolo Bluetooth, proporcionando uma boa mobilidade ao usuário. É um sistema de baixo custo e pode ser usado por qualquer pessoa com tempo de treinamento curto. O projeto é totalmente personalizável e inclui um microcontrolador dedicado desenvolvido pelo nosso grupo, bem como as interfaces de comunicação para os protocolos I2C e RS-232. A parte de controle digital embutida na luva é sintetizada em FPGA.

O desenvolvimento de aplicativos, como movimento do mouse, controle do joystick, reconhecimento de alfabeto LIBRAS e modelo 3D, mostra o potencial da luva como um dispositivo de entrada. Os seis graus de liberdade fornecidos pela luva permitem uma detecção natural do movimento. A luva é especialmente adequada para pessoas cegas, surdas ou mudas interagirem com um computador ou smartphone de uma maneira mais eficiente. O impacto social é imediato, uma vez que o sistema proposto pode mitigar a dificuldade das pessoas com deficiência física na comunicação.

## Referências

- CHOUHAN, Tushar et al. Smart glove with gesture recognition ability for the hearing and speech impaired. In: 2014 IEEE Global Humanitarian Technology Conference-South Asia Satellite (GHTC-SAS). IEEE, 2014. p. 105-110.
- ENGROFF, A. M. PAMPÍUM I - Pampíum I-Projeto e implementação de um microcontrolador de 16 bits em arquitetura RISC. Dissertação (Mestrado) — Universidade Federal do Pampa – Unipampa, 2014.
- FELIPE, Tanya A. Libras em contexto: curso básico, livro do estudante cursista. Brasília: Programa nacional de apoio à educação dos surdos, MEC, 2001.
- HAN, Jungong et al. Enhanced computer vision with microsoft kinect sensor: A review. IEEE transactions on cybernetics, v. 43, n. 5, p. 1318-1334, 2013
- INVENSENSE, I. MPU-6000 and MPU 6050 Product Specification Revision 3.4. 2013.
- JAMES, Gareth et al. An introduction to statistical learning. New York: Springer, 2013.
- SOKOLOVA, Marina; JAPKOWICZ, Nathalie; SZPAKOWICZ, Stan. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: Australasian joint conference on artificial intelligence. Springer, Berlin, Heidelberg, 2006. p. 1015-1021.
- ROMANSSINI, Marcelo; QUIRINO, Felipe Antunes; GIRARDI, Alessandro Gonçalves. System Implementation of a Gesture Detection Glove for Human-Computer Interaction. In: WORKSHOP ON CIRCUITS AND SYSTEMS DESIGN, 8., 2018, Bento Gonçalves. Anais... . Bento Gonçalves: Cass, 2018.