

Interação com o Ambiente Mediante Classificação por Distância Euclidiana

Werikson F. de O. Alves* Ketia S. Moreira**
Alexandre S. Brandão***

* Núcleo de Especialização em Robótica (NERO), Departamento de Engenharia Elétrica, Universidade Federal de Viçosa, MG, (e-mail: werikson.alves@ufv.br).

** Departamento de Engenharia Elétrica, Universidade Federal de Viçosa, MG, (e-mail: ketia@ufv.br)

*** Núcleo de Especialização em Robótica (NERO), Departamento de Engenharia Elétrica, Universidade Federal de Viçosa, MG, (e-mail: alexandre.brandao@ufv.br)

Abstract: Stock-based applications are increasingly common in everyday life. This is because it allows the user to activate equipment without the need to touch them and, moreover, it can be performed remotely. Thus, this work proposes a method of gesture recognition using a reduction of variables through Principal Component Analysis (PCA) followed by a classification by Euclidean distance. The purpose of this work is to interact with an environment through the actions of a user. For this, it was necessary to create a database using fourteen classes of shares. The validation of the classifier takes place through the online analysis of the operation and the results of the confusion matrix. In the end, it is concluded that the classifier had results above 80% for 6 classes. The main benefit of this work is to formalize a simplistic classifier, which makes use of a small database.

Resumo:

Aplicações baseadas em ações estão cada vez mais comuns no dia a dia. Isto porque possibilita ao usuário acionar equipamentos sem a necessidade de tocá-los e, além do mais, pode ser realizado remotamente. Desta forma, este trabalho propõe um método de reconhecimento de gestos utilizando uma redução de variáveis através da Análise das Componentes Principais (PCA) seguida de uma classificação por distância euclidiana. A proposta deste trabalho é interagir com um ambiente através de ações de um usuário. Para isso, foi necessária a criação de uma base de dados utilizando quatorze classes de ações. A validação do classificador se dá por meio da análise online de operação e pelos resultados da matriz de confusão. Ao final, conclui-se que o classificador teve resultados superior a 80% para 6 classes. O principal benefício deste trabalho é formalizar um classificador simplista, que faz uso de uma pequena base de dados.

Keywords: Action Recognition; Euclidean Distance; Interactive Robotics; Principal Component Analysis; OptiTrack;

Palavras-chaves: Reconhecimento de Ações; Distância Euclidiana; Robótica Interativa; Análise das Componentes Principais; OptiTrack;

1. INTRODUÇÃO

A robótica é uma das áreas de pesquisa sobre tecnologia muito importante e que está sempre em desenvolvimento, sendo muitas vezes usada para auxiliar o ser humano em tarefas que operam de forma automática e de forma repetida Campos (2019). Desde o seu surgimento, as aplicações da robótica sempre foram extensas começando na indústria e se espalhando para outras áreas, como por exemplo na saúde, segurança e na educação. Com o passar do tempo, a interação entre pessoas e tecnologias se desenvolveu e ainda continua de forma muito rápida, tanto para tarefas simples quanto para tarefas mais complexas. Com a busca por uma técnica simples e que possa atender diversas áreas,

várias técnicas de interação foram criadas e estudadas ao longo do tempo. Entre elas, aplicações baseadas em movimentos do corpo humano (reconhecimento de gesto, expressões e vozes) têm sido usadas cada vez mais em diversas áreas (jogos, acessibilidade, reabilitações, esportes, treinamentos, etc), ao permitirem uma interação mais humana com a tecnologia VIANA (2015) e uma maior liberdade ao não necessitar de cabos ou semelhantes ligados ao usuário.

Estas formas de controles às quais são baseadas em comandos realizados pelo corpo humano através da captação de gesto podem ser chamadas de Interface Natural do Usuário (INU) de Souza et al. (2016). Alguns exemplos destas aplicações são: sistema de INU desenvolvido para televisão interativa desenvolvido por Fan et al. (2013), controle de

* Este trabalho teve suporte financeiro da FAPEMIG.

diversos modelos de robôs ou até mesmo o controle de um ambiente de forma remota (iluminação, ar condicionado ou um portão eletrônico, por exemplo).

Dado esta evolução tecnológica, novas formas de processar esses dados, que facilitassem e simplificassem, foram desenvolvidas ao longo do tempo como por exemplo a Análise de Componentes Principais (PCA) e a Análise Discriminante Linear (LDA) Martinez and Kak (2001). Com estas ferramentas, pode-se realizar a redução de dimensionalidade, compactação de dados, extração de recursos e visualização de dados.

O foco deste trabalho é a PCA, o qual é uma ferramenta cujo objetivo é encontrar a base de dados mais significativa para então retorná-la como um determinado conjunto de dados, desta forma revelando a estrutura oculta no conjunto de dados e filtrando o ruído Kurita (2019).

Neste contexto, o objetivo deste trabalho é a criação de um método de classificação de ações no qual se utilize poucos marcadores e que possua uma técnica de processamento simples no qual o usuário realize gestos naturais e intuitivos, predefinidos, que se relacionam com ações de alto nível. Como objetivos específicos tem-se a) a coleta de informações espaciais do usuário, b) o processamento destes dados e c) a aplicação da técnica desenvolvida.

A sequência deste trabalho está organizada como: A seção 2 apresenta trabalhos relacionados, compondo a base teórica deste trabalho. A seção 3 descreve as etapas do método desenvolvido. A seção 4 faz um estudo dos resultados obtidos. Por fim, a seção 5 apresenta as conclusões deste trabalho e sugestões para sua continuidade.

2. REFERENCIAL TEÓRICO

A medida em que os meios tecnológicos evoluem, também é possível observar uma crescente em relação a Interação Homem-Máquina (IHM) a qual pode ser aplicada em diversos setores, como, por exemplo, controle de robôs, plataformas ou de salas e/ou ambientes fechados de forma remota Costalonga et al. (2014). Dessa forma, pode-se dizer que a IHM se baseia em um sistema onde uma tarefa é executada pela cooperação do componente humano e de automação Ramaswamy et al. (2014).

Segundo Pereira et al. (2009), a busca por uma interface de comunicação mais natural se dá pela tentativa de deixar as pessoas cada vez mais à vontade ao interagir com uma máquina. Com isto em mente, diversas formas de interfaces foram criadas e são estudadas, como, por exemplo, reconhecimento de voz Hudson et al. (2017), periféricos Huttenrauch and Eklundh (2002) e reconhecimento de gestos Chen et al. (2019), dentre estas, no reconhecimento de gesto existem diversas de técnicas utilizáveis podendo ser baseadas na forma da mão, movimento do corpo ou de membros específicos do corpo.

De acordo com Basilio et al. (2019), devido a facilidade e a possibilidade de se trabalhar de forma remota e as mais variadas aplicações, o reconhecimento de gestos vem ganhando mais destaque no meio acadêmico. Dessa forma, a IHM por meio de gestos pode ser realizada através da linguagem corporal, viabilizando por exemplo, que uma

pessoa possa controlar um determinado ambiente fechado sem precisar estar presente no local.

Para realizar o reconhecimento de uma ação, inicialmente é necessário realizar o pré-processamento das informações coletadas para em seguida realizar o processo de reconhecimento. Sendo assim, existem diversos meios e formas de se fazer este processamento de imagens que possibilitam a captura de informações para a classificação dos gestos, como por exemplo o uso de sensores como o *OptiTrack*, o *Intel RealSense* e o *Kinect*. Este processamento digital de imagem se dá pelo rastreamento das informações relacionadas aos marcadores utilizados pelo usuário ao estabelecer uma correspondência da estrutura de uma imagem entre consecutivos quadros baseados nas suas características principais.

Levando em consideração que para se ter um reconhecimento de ação eficiente diversos dispositivos podem ser usados e que eles podem variar de situação para situação, o sensor o *OptiTrack* se torna uma excelente ferramenta para a coleta de dados, devido a sua alta precisão e pela quantidade de informação que se pode coletar. Em seu conjunto, o *OptiTrack* é composto por marcadores reflexivos e câmeras infravermelhas (que fazem a função de sensores de profundidade), capazes de inferir a posição tridimensional dos marcadores.

Este sistema de rastreamento possibilita a calibração e a configurações de diversas funções que auxiliam a captura de informações, fornecimento de interfaces para captura e processamento de dados 3D, podendo ser gravado ou transmitido ao vivo para outros programas Motive (2020). Além destas funcionalidades, o *OptiTrack* pode ser utilizado em outras aplicações, como, por exemplo, produções visuais, estudo de movimentos de corpos, realidade virtual, animações e na robótica.

Para realizar a coleta de dados, é necessário fazer o rastreamento através de marcadores, podendo ela ser feita através de marcadores ativos ou passivos, sendo que os ativos emitem luz própria ao utilizar diodos emissores de luz colocados junto ao corpo do usuário, enquanto os marcadores passivos apenas refletem a luz proveniente das câmeras.

Ao receber esse sinal, emitido ou refletido, respectivamente, a câmera consegue identificar os marcadores mais brilhantes e diferencia-los dos objetos menos brilhantes. Outra diferença entre eles é que os marcadores passivos não necessitam de cabos ou baterias, contudo requerem mais pré-processamento Vital et al. (2015).

Portanto, de forma similar a de Carvalho et al. (2022), após a captura das imagens em profundidade pelo *OptiTrack*, o programa previsto neste trabalho identifica o movimento feito pelo usuário, faz o reconhecimento desse gesto e interage com o ambiente.

3. METODOLOGIA

3.1 Base de dados

Trabalhar com reconhecimento de gestos por meio de técnicas de aprendizado de máquinas, exige comumente uma grande quantidade de conjuntos de dados. Este trabalho,

entretanto, possui a proposta de interagir com um determinado ambiente através de gestos simples e de qualquer usuário. Assim, foi optado pela criação de uma base de dados própria, onde cada tarefa está associada a uma classe diferente, de forma a ganhar um pouco mais de liberdade e flexibilidade.

A fim de tornar mais dinâmico e simples o processo de reconhecimento do gesto, foram utilizadas duas fitas reflexivas no formato de pulseiras que são colocadas uma em cada pulso do usuário para serem os marcadores que serão rastreados pelo *OptiTrack*, conforme é apresentado na Figura 1. Em seguida, utilizando o programa foram feitas as gravações dos gestos, nas quais o início e o fim de cada ação foram considerados a posição de repouso (quando uma pessoa está com os braços para baixo). Ao todo, foram realizadas 70 gravações, sendo 14 classes de gestos diferentes e 5 repetições de cada classe.



Figura 1. Localização dos marcadores no usuário.

Para se detectar o início e o fim de um gesto, a abordagem escolhida foi a utilização de um gatilho duplo, no qual são avaliados a altura e a velocidade de ambas as pulseiras. Levando-se em conta que todos os movimentos necessitam que ao menos uma das pulseiras seja erguida em relação ao solo, foi escolhido a coordenada y de cada marcador, a qual está posicionada perpendicular ao solo, para se analisar. Portanto, ao se analisar uma janela fixa de f quadros, resultando em uma amostragem t milissegundos, é possível obter a altura e velocidade de cada pulseira naquele instante. Logo, se a velocidade ou a altura do mesmo alcançarem um determinado valor mínimo (V_{min} em metros por quadros e H_{min} em metros), serão armazenados os últimos d quadros anteriores ao acionamento de uma das condições e em seguida as janelas atuais e parando somente quando ambas as condições não forem mais cumpridas. Os valores utilizados para f , t , V_{min} , H_{min} e d foram 24, 100, 0,001, 0,95 e 23, respectivamente.

Cada classe de movimento pode ser usada para controlar uma determinada função dentro de uma sala. As classes utilizadas são apresentadas na Figura 2, sendo elas:

A) Levantar as duas mãos e depois abaixar a direita;

B) Girar os dois braços, terminando com eles na frente do peito, com a mão direita ligeiramente acima da esquerda;

C) Levantar as duas mãos e abaixar a esquerda;

D) Desenhar um círculo na frente do rosto com a mão direita, enquanto mantém o braço esquerdo ao lado do corpo;

E) Desenhar um círculo na frente do rosto com a mão direita, enquanto levanta o braço esquerdo até a altura do ombro;

F) Desenhar um círculo na frente do rosto com a mão direita, enquanto levanta o braço esquerdo até a altura da cabeça;

G) Desenhar um círculo na frente do rosto com a mão esquerda, enquanto mantém o braço direito ao lado do corpo;

H) Desenhar um círculo na frente do rosto com a mão esquerda, enquanto levanta o braço direito até a altura do ombro;

I) Desenhar um círculo na frente do rosto com a mão esquerda, enquanto levanta o braço direito até a altura da cabeça;

J) Cruzar os dois braços formando um X na frente do peito;

K) Mover o braço direito para frente e depois deslizar para a esquerda;

L) Levantar os dois braços até a altura do estômago e depois levantar a mão direita;

M) Levantar os dois braços até a altura do estômago e depois levantar a mão esquerda;

N) Mover o braço esquerdo para frente e depois deslizar para a direita.

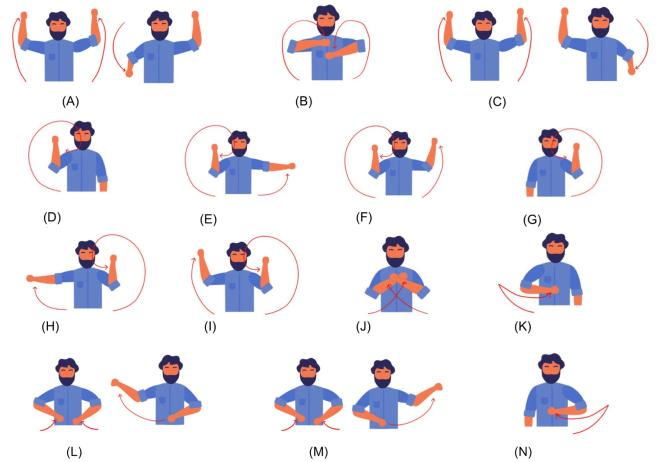


Figura 2. Classes de gestos no conjunto de dados de Carvalho et al. (2022)

Utilizando o *OptiTrack* para extrair as informações dos marcadores, cada amostra foi capturada com uma taxa de 240 FPS (quadros por segundo). Cada quadro consiste em um conjunto de características que juntos compõem a matriz de características, sendo ela composta por um marcador

que foi posicionado na região da cabeça (O), por uma pulseira no braço esquerdo (P) e por uma pulseira no braço direito (R),

$$F_Q = [O_x \ O_y \ O_z \ P_x \ P_y \ P_z \ R_x \ R_y \ R_z]. \quad (1)$$

Em outras palavras, esta matriz é composta por três sinais ao todo, onde Q informa o quadro atual e as colunas representam as coordenadas x , y e z de cada marcador, onde y é perpendicular ao solo. Em seguida, a amostra completa é armazenada como a concatenação subsequente da matriz de características,

$$A_{Q,U} = [F_1 \ F_2 \ \dots \ F_Q]^T, \quad (2)$$

onde U é o número da coluna.

Para realizar o reconhecimento da ação, foram coletadas as informações das duas pulseiras, ao longo do gesto. Já o marcador que está posicionado na região da cabeça, foi utilizado como referência, afinal ao realizar um mesmo gesto em posições diferentes, o movimento relativo será idêntico. Este marcador também serviu para centralizar todas as características (em seus respectivos eixos) em relação às coordenadas do marcador posicionada na cabeça, corrigindo uma eventual movimentação da cabeça em decorrência do movimento de outra parte do corpo. Para isto, adotou-se

$$B_{Q,U} = A - \bar{O}, \quad (3)$$

onde \bar{O} é a média das coordenadas das informações relacionadas ao marcador O, A é a matriz contendo todas as informações dos três marcadores referentes aquela amostra e B é o novo conjunto de dados gerado.

As amostras foram normalizadas segundo,

$$C_{Q,U} = \frac{B_{Q,U} - \bar{B}_U}{S_B}, \quad (4)$$

onde S_B é desvio padrão da amostra analisada, \bar{B}_U é a média da coluna do elemento em análise, centralizando a variável e resultando em uma variável com média zero e variância igual a unidade, dessa forma com desvio padrão igual a um. Isto visa evitar ter um conjunto de informações com excessiva desproporcionalidade, evitando redundâncias e tornando-o mais organizado. Assim, as variáveis padronizadas podem ser diretamente comparáveis, devido ao fato de não possuírem dimensão e serem expressas em unidades de desvio padrão.

3.2 Algoritmo de reconhecimento de ações

Agora, tendo esses dados em mãos, a PCA seguido pelo método da distância euclidiana para classificar a base de dados foram aplicadas, sendo este processo detalhado a seguir.

Para simplificar a classificação de cada amostra, em vez de se analisar todas as informações capturadas pelo sensor, optou-se por aplicar a técnica de PCA na qual é possível realizar a redução de dimensão da matriz de características, sendo ela originalmente $A \in R^{Q \times 9}$ (Q é o número de quadros da amostra em análise).

Para isto, após a centralização dos dados, são desconsideradas as informações relacionadas ao marcador da cabeça, devido suas informações permanecerem praticamente constante ao longo do tempo em relação as outras informações,

gerando $B \in R^{Q \times 6}$, para em seguida, aplicar a normalização obtendo $C \in R^{Q \times 6}$.

Em seguida, aplicando PCA, obtém-se os autovalores referentes as variáveis em ordem decrescente, conforme

$$\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4 > \lambda_5 > \lambda_6, \quad (5)$$

onde λ_i representa a importância de cada variável para aquela amostra. Dessa forma, optou-se por escolher os três maiores autovalores, aos quais representam mais de 90% das informações, para cada amostra.

Outro resultado que a PCA entrega é o conjunto de dados $D \in R^{6 \times 6}$ que contém as componentes principais. Como foram escolhidas as três componentes mais importantes para cada amostra (λ_1 , λ_2 e λ_3), ou seja, as variáveis que possuem maior significado para aquele gesto, foi obtido $E \in R^{6 \times 3}$, que contém um conjunto de dados reduzido de $D \in R^{6 \times 6}$.

O processo de classificação é baseado na comparação de uma ação qualquer com uma outra ação de referência. Para realizar isto, são utilizados os dados contidos na matriz $E_{6,3}$ obtido da amostra de referência e da amostra em análise.

O reconhecimento do gesto pode ser obtido por meio da comparação da série de informações do conjunto de dados reduzidos da ação a ser classificada com a da ação da referência. Ao alterar a ação de referência para diferentes classes, aquela em que se obter a menor diferença representa a classe da ação a ser classificada. Em outras palavras, a classificação se dá entre a classe corrente e as classes de referência, conforme,

$$Dist = \sum_{j=1}^3 ||D_{jReferencia} - D_{jCorrente}|| \quad (6)$$

A Figura 3 ilustra um fluxograma que resume as etapas do subsistema de classificação realizadas neste trabalho e seus respectivos resultados.

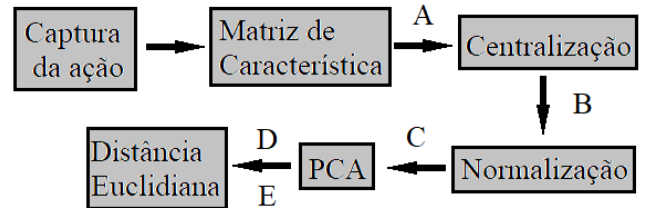


Figura 3. Fluxograma com as etapas e respostas do classificador proposto.

4. DISCUSSÃO E RESULTADOS

Para a validação do método proposto, foi realizado um teste online, no qual foram realizadas diversas simulações, onde uma amostra aleatória seria comparada com o banco de dados construído, verificando se o conjunto de dados possui alguma amostra similar. Este teste online foi dividido em duas etapas, sendo a primeira referente as comparações entre as amostras e a segunda referente a classificação da ação.

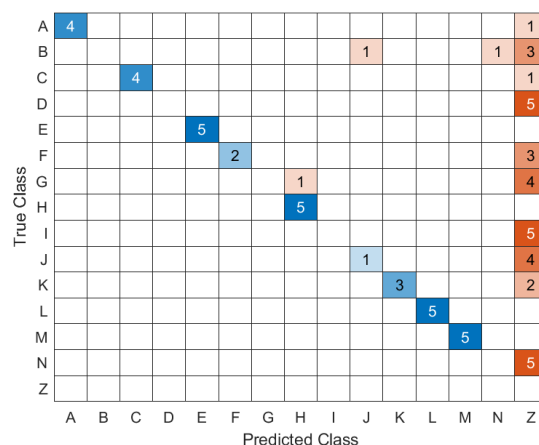
Na primeira etapa, cada amostra foi comparada com todas as outras amostras da base de dados. Feito isto, para cada

True Class	A	B	C	D	E	F	G	H	I	J	K	L	M	N
A	17									1	1			1
B		7				3				3	3	1		3
C	2		17	1										
D		2		1	2	7	2	1	5					
E					19								1	
F		2				10	1		7					
G	1			3	1	3	2	3	5			2		
H								20						
I		2		2	3	6			7					
J		3			4	2				9			1	1
K		1						1			12		4	2
L												20		
M													20	
N	4	2	2		2	1	2				1			6

Por meio da Figura 4, observa-se que houveram 3 classes que tiveram a classificação das amostras excelente, 100% de acerto (H, L e M), 4 classes com a classificação das amostras boa, acima de 60% de acerto, (A, C, E e K), 2 classes tiveram desempenho razoável (F e J), entre 40 e 60%, enquanto 5 tiveram o desempenho abaixo de 40% (B, D, G, I e N). A classificação entre as amostras, portanto, obteve desempenho médio 59.6%.

A segunda etapa do teste, trata-se da classificação final do gesto. Para se realizar a classificação de uma ação, correta ou errada, é necessário que a maioria (metade mais um) das respostas sejam iguais. Logo, sabendo que para cada amostra simulada foram obtidas 4 classificações, é preciso que no mínimo 3 respostas sejam classificadas como a mesma ação, para que ela receba uma classificação final e caso essa condição não seja atendida o gesto será classificado como inconclusivo. O resultado para esta etapa é apresentado na Figura 5.

Por meio da Figura 5, pode-se perceber que houveram 4 classes tiveram um desempenho excelente (E, H, L e M), 3 classes tiveram um desempenho bom (A, C e K), 1 classe teve um desempenho razoável (F), 3 classes tiveram o desempenho abaixo de 40% (B, G e J) e 3 classes tiveram resultados inconclusivos (D, I e N).



Um ponto importante a se perceber é que 33 ações das 70 foram classificadas como inconclusivos enquanto as outras 37 obtiveram uma classificação dentre as classes como resposta, ou seja, 52,86% das ações tiveram uma classificação final. Destas, apenas 3 foram classificados erroneamente, logo, pode-se dizer que o método teve um desempenho de 48,57% em relação ao total de amostras e um desempenho de 91,9% entre as amostras que receberam uma classificação dentre as classes predefinidas.

Um outro ponto que vale ressaltar é que ao comparar as informações contidas em ambas as figuras pode-se notar que apenas a ação J, que apesar das amostras terem 45% de acerto entre elas, teve uma classificação final distante do esperado tendo apenas 20% de acerto.

5. CONSIDERAÇÕES FINAIS

Ao observar as informações relacionadas as posições dos marcadores, foi possível concluir que o *OptiTrack* se mostra ser um bom dispositivo para coleta de dados relacionadas com rastreamento de corpos, devido a precisão e a grande quantidade de informações que se é possível extrair.

O classificador apresentou uma taxa de acerto média de 91,9% o qual pode ser considerado como tendo um excelente desempenho e juntamente a isto, a classe inconclusiva se provou ser um recurso muito útil, para reduzir a quantidade de erros, reforçando a viabilidade e a eficiência.

do mesmo, devido a simplicidade e do pequeno banco de dados construído.

Desta forma, ao associar as classes com tarefas de alto nível, em algum ambiente fechado, espera-se que a maioria das ações sejam executadas corretamente. Logo, para trabalhos futuros, espera-se que ao se aumentar a base de dados, espera-se um aumento no desempenho médio e que menos classes sejam classificadas como inconclusivas e associar as classes com tarefas ou missões que deverão ser realizadas por robôs móveis em uma arena estruturada.

AGRADECIMENTOS

Os autores agradecem a FAPEMIG - Fundação de Amparo à Pesquisa de Minas Gerais pelo suporte financeiro. Sr. Alves agradece em especial a FAPEMIG pela bolsa de Iniciação Científica concedida, a qual lhe propiciou a participação neste trabalho.

REFERÊNCIAS

- Basilio, V.T., Carvalho, K.B., and Brandao, A. (2019). Reconhecimento de ações por rna em aplicações de robótica social. *XIV Simpósio Brasileiro de Automação Inteligente, Ouro Preto/MG*.
- Campos, F.R. (2019). *A robótica para uso educacional*. Editora Senac São Paulo.
- Chen, F., Lv, H., Pang, Z., Zhang, J., Hou, Y., Gu, Y., Yang, H., and Yang, G. (2019). Wristcam: A wearable sensor for hand trajectory gesture recognition and intelligent human-robot interaction. *IEEE Sensors Journal*, 19(19), 8441–8451. doi:10.1109/JSEN.2018.2877978.
- Costalonga, T., Ávila, L., Muniz, L., and Brandao, A. (2014). Interface homem-máquina utilizando sensor kinect para control de um quadrotor ar. drone parrot. In *XX Congres so Brasileiro de Automática*.
- de Carvalho, K.B., Villa, D.K.D., Sarcinelli-Filho, M., and Brandao, A.S. (2022). Gestures-teleoperation of a heterogeneous multi-robot system. *The International Journal of Advanced Manufacturing Technology*, 118(5), 1999–2015.
- de Souza, P.H.R., de Jesus Melo, W.A.L., de Santana, I.M.F., Doria, N.S.F., and de Matos, L.B.C. (2016). Controle de modelos robóticos por meio de gestos naturais. In *Anais do III Encontro Nacional de Computação dos Institutos Federais*, 732–735. SBC.
- Fan, Y.C., Chen, C.L., and Huang, S.S. (2013). Natural user interface for interactive television. In *2013 IEEE International Symposium on Consumer Electronics (ISCE)*, 189–190. doi:10.1109/ISCE.2013.6570177.
- Hudson, C., Bethel, C.L., Carruth, D.W., Pleva, M., Juhar, J., and Ondas, S. (2017). A training tool for speech driven human-robot interaction applications. In *2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA)*, 1–6. doi:10.1109/ICETA.2017.8102488.
- Huttenrauch, H. and Eklundh, K. (2002). Fetch-and-carry with cero: observations from a long-term user study with a service robot. In *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*, 158–163. doi:10.1109/ROMAN.2002.1045615.
- Kurita, T. (2019). *Principal Component Analysis (PCA)*, 1–4. Springer International Publishing, Cham. doi:10.1007/978-3-030-03243-2_649-1. URL https://doi.org/10.1007/978-3-030-03243-2_649-1.
- Martinez, A. and Kak, A. (2001). Pca versus lda. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2), 228–233. doi:10.1109/34.908974.
- Motive (2020). Motive Documentation - naturalpoint products. URL https://v22.wiki.optitrack.com/index.php?title=Motive_Documentation. Accessed: 2022-04-18.
- Pereira, F.G., Schmitz, N., Vassallo, R.F., and Berns, K. (2009). Cooperação entre homens e robôs baseada em reconhecimento de gestos. *IX Simpósio Brasileiro de Automação Inteligente, Brasília/DF*.
- Ramaswamy, A., Monsuez, B., and Tapus, A. (2014). Model driven software development for human-machine interaction systems. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, 270–271.
- VIANA, D.L. (2015). *Uma linguagem de domínio específico para descrição e reconhecimento de gestos usando sensores de profundidade*. Master’s thesis, Universidade Federal de Pernambuco.
- Vital, J., Couceiro, M.S., Dias, G., and Ferreira, N. (2015). Tecnologias para a análise do movimento humano. In *R. Ruben, M. Vieira, C. Campos, H. Almeida, J. Siopa, P. Bártolo, & J. Folgado (Edits.), 6º Congresso Nacional de Biomecânica*, 1–6.