



CRICTE 2017

XXVIII Congresso Regional de Iniciação Científica e Tecnológica em Engenharia



UTILIZAÇÃO DOS ALGORITMOS DE K-MEANS E FUZZY C-MEANS PARA CONTROLE DE UM ROBÔ MÓVEL MEDIANTE O USO DE UM SENSOR DE PROFUNDIDADE

Evaristo J. do Nascimento

Acadêmico do curso de Engenharia de Computação, Universidade Federal de Santa Maria
evaristo.nascimento@ecomp.ufsm.br

Daniel F. T. Gamarra

Professor do Departamento de Processamento de Energia Elétrica, Universidade Federal de Santa Maria
daniel.gamarra@ufsm.br

Resumo. O trabalho apresenta o desenvolvimento de um sistema de reconhecimento de gestos utilizando o sensor de profundidade Kinect, aplicado em um robô móvel. O sistema é capaz de receber informações gestuais de um usuário, e enviar comandos para um robô para que realize movimentos específicos desejada por um usuário. A informação gestual é capturada medianteo sensor Kinect, o robô possui um sistema capaz de interpretar esse gesto utilizando algoritmos de agrupamento dados de Fuzzy C-Means e K-Means, após reconhecer o gesto lido, o robô executa a tarefa referente ao comando gestual recebido. O artigo descreve a arquitetura implementada com resultados experimentais.

Palavras-chave: Kinect, Reconhecimento Gestual, Agrupamento de Dados.

1. INTRODUÇÃO

O trabalho desenvolvido neste artigo visa desenvolver um sistema inteligente que permita o controle de um robô móvel a partir de comandos gerados por gestos de uma pessoa, os quais são obedecidos pelo robô fazendo com que ele execute movimentos específicos dependentes de qual comando foi recebido.

As técnicas a ser empregada para a classificação dos gestos de uma pessoa é o agrupamento de dados através do algoritmo de *K-means* e do algoritmo *Fuzzy C-means*. As referidas técnicas vão permitir classificar os gestos captados por um sensor de movimento, verificar qual comando foi recebido e qual é a tarefa, associada ao comando, que deverá ser realizada pelo robô. O trabalho também tem como objetivo comparar as técnicas que serão utilizadas na pesquisa e verificar a viabilidade do uso de cada uma.

Na literatura existem diversas aplicações utilizando as técnica de agrupamento de dados de *K-Means* e *Fuzzy C Means* para reconhecimento de gestos. Q. Zhang [1] utiliza o algoritmo de *Fuzzy C-means* para resolver problemas de reconhecimento gestual em local onde o fundo do ambiente, dificulta o contraste entre a pessoa e o ambiente, dificultando o analise do processo pelas técnicas de processamento de imagens; Pal [2], aplica a técnica para reconhecimento de gestos de pessoas com deficiência muscular, possibilitando, a partir destas informações, que um profissional de saúde possa realizar diagnósticos de uma possível evolução de uma doença que cause problemas físicos em seus pacientes; Panchal e Kandoriya [3] utilizam a técnica para reconhecimentos gestuais de uma mão

fixa, identificando as posições dos dedos e classificando-os em diferentes gestos e Gamarra [4], utiliza a técnica de agrupamento de dados, utilizando os algoritmos *K-means*, *Fuzzy C-means* e *K-medoids*, comparando o desempenho de diversos algoritmos de clusterização na aplicação da técnica de *Bag of Visual Words* para reconhecimento de objetos.

2.FUNDAMENTAÇÃO TEÓRICA

2.1 Algoritmo *K-means*

O Algoritmo de *K-means* é muito utilizado na área de *machine learning* utilizado para separação de dados em grupos, buscando classifica-los. Tomando como referência o trabalho de Gamarra em [4], o algoritmo funciona basicamente definindo os centroides que identificam cada grupo a ser formado recalculando as distancias entre cada dado a ser classificado e os centroides, onde o grupo em que o centroeide é o mais próximo do dado é o grupo ao qual este dado pertence.

2.2 Algoritmo *Fuzzy C-means*

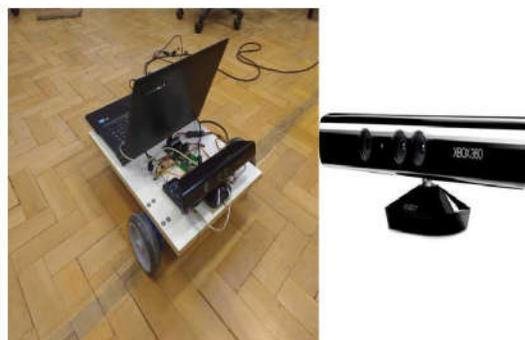
O algoritmo *Fuzzy C-means*, é uma variação do algoritmo de *K-means*, onde ele possibilita tratar incertezas, ou seja, dados que possuem características de dois diferentes grupos. Estes são classificados a partir de um “grau de pertinência” de cada grupo (Yonamine [5]). O grau de pertinência descreve o quanto de semelhança ou proximidade, determinado elemento tem de cada grupo, indicando que o grupo ao qual este determinado dado pertence é aquele que o que possui maior grau de pertinência.

3. SETUP EXPERIMENTAL

O sistema está constituído por um robô móvel com três rodas, duas rodas são tracionadas por um motor em cada roda, e uma terceira roda independente auxilia o

equilíbrio do robô; Um sensor de profundidade *Kinect* permite que possamos capturar movimentos de uma pessoa obtendo as coordenadas X, Y, Z das juntas de uma pessoa, assim como também sua imagem. Ambos o robô e o *Kinect* são mostrados na figura 1.

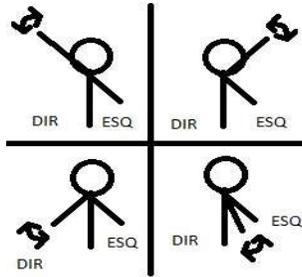
Figura 1. Robô móvel e Sensor *Kinect* utilizados



Para obtenção dos dados do *Kinect* é utilizado o ambiente de programação *Processing*, criado pelo MIT (*Massachusetts Institute of Technology*). O *Processing* é um software que possibilita o desenvolvimento de projetos de processamento de imagens utilizando *Java*, *Python* ou *JavaScript* (Reas e Fry [7]). Também foi utilizado o software *MATLAB*, da empresa *Mathworks* para o desenvolvimento do sistema inteligente para a classificação dos gestos baseado nas técnicas de *clustering* utilizando os algoritmos *K-means* e *Fuzzy C-means*, e também é feita a associação de cada gesto com cada comando para ser enviado ao robô. No *Processing* foi desenvolvida a obtenção de dados do *Kinect*, e a preparação dos dados para serem enviados para o *MATLAB*. O processamento dos comandos é realizado mediante o microcontrolador *Arduino* (*ATMEGA 2560*) fixado no robô.

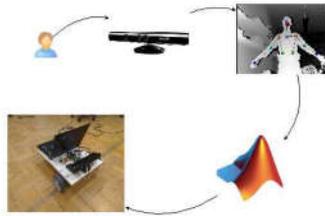
A figura 2 apresenta os gestos que foram utilizados para a realização da pesquisa.

Figura 2: Gestos utilizados para realização do estudo.



A figura 3 mostra o funcionamento do sistema como um todo. Primeiramente o *Kinect* reconhece o indivíduo e monta o mapa das principais juntas do corpo da pessoa. A partir de então ele obtém as informações de posição das juntas (no nosso experimento a mão esquerda) e envia as informações para o MATLAB. O MATLAB realiza o processo de classificação do gesto recebido e então envia o respectivo comando ao robô para que realize a atividade desejada.

Figura 3. Esquema de Funcionamento do Sistema com o robô móvel



4.RESULTADOS

O funcionamento do sistema consiste em reconhecer a região da imagem lida pelo *Kinect*. A região da imagem é identificada e associada a um determinado movimento que será executado pelo robô. Os comandos referentes para cada região de leitura do *Kinect* foram definidos na implementação do sistema. Os algoritmos trabalham com matrizes bidimensionais de dados, onde o número de colunas define o número de coordenadas de cada posição lida de uma junta.

O algoritmo de *clustering* recebe as coordenadas X, Y e Z da posição onde se encontra a mão esquerda do usuário (direita na visão do *kinect*), e essas informações são enviadas ao sistema construído no

MATLAB para serem classificados. As figuras 4 e 5 apresentam os resultados de classificação utilizando os algoritmos de *K-means* e o *Fuzzy C-means* respectivamente.

As imagens mostram um processo de *clustering* dos dados das imagens do *kinect* em cinco grupos. Cada região de dados mais estrema significa um comando gestual diferente e o grupo central representa um gesto nulo. As marcações em “X” representam os valores centroeide de cada grupo.

Os dois algoritmos se comportam de uma forma muito semelhante apresentando variações somente em regiões de fronteira entre os grupos. Neste espaço de amostra os algoritmos apresentam algumas dificuldades de interpretação dos dados nas fronteiras entre o grupo central e o inferior direito. Nos grupos 1, 2 e 4 os dois algoritmos sem comportam de forma muito semelhante, apresentando os mesmos resultados, compondo esses grupos com 56, 58 e 56 elementos respectivamente. Já com os grupos 3 e 5, devido a algumas incertezas com dados de fronteira, há uma variação na composição dos grupos onde que com *K-means* os grupos possuem 56 e 51 dados respectivamente e com o *Fuzzy C-means* os grupos possuem 58 e 49 elementos respectivamente. Esse fenômeno pode ocasionar erros no momento de execução do sistema. Isso se deve ao fato de que o sistema está levando em conta variações em profundidade (Eixo “Z”) e esses dados apresentam características de profundidade semelhantes aos dados do seu grupo vizinho. A figura 6 apresenta a comparação do resultado apresentado por cada algoritmo.

Figura 4. Resultados do *clustering* obtidos utilizando *K-means*

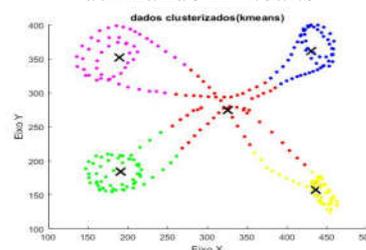


Figura 5. Resultados do *clustering* obtidos utilizando *Fuzzy C-means*

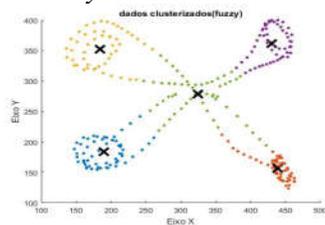
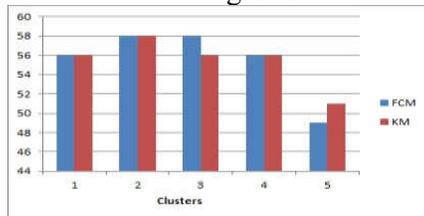


Figura 6: Comparação entre os resultados dos dois algoritmos.



5. CONCLUSÃO

O trabalho desenvolvido neste artigo tem como principais resultados, a criação de um sistema que integra o comando dos movimentos do robô mediante o uso de um sensor *Kinect* para reconhecimento de gestos de um usuário. O referido sistema utilizou diversas tecnologias as quais foram integradas para o controle de movimentos do robô. Um segundo aporte do artigo é a implementação de um sistema que mediante o uso de técnicas de *clustering* como o *K-means* e *Fuzzy C-means*. O sistema consegue processar e discriminar gestos diferentes de uma pessoa baseado na extração das informações das juntas do usuário. Foi possível também obter dados sobre as diferentes técnicas utilizadas na pesquisa possibilitando uma análise de desempenho de cada algoritmo e verificando a viabilidade do uso dos mesmos. As técnicas utilizadas para identificação dos gestos tiveram resultados muito semelhantes o que não nos possibilita dizer se um algoritmo é melhor que o outro. Para isso é necessário realizar novos experimentos com dados que possuem diferentes características e comportamento o que está a ser considerado numa futura evolução do estudo. Trabalhos futuros poderiam incluir o uso de redes neurais para as tarefas de classificação.

6. REFERÊNCIAS

- [1] Zhang, Q., Lu, J., Wei, H., Zhang, M., Duan, H., “Dynamic hand gesture segmentation method based on unequal-probabilities background difference and improved FCM Algorithm”, In: International Journal of Innovative Computing, Information and Control, 2015.
- [2] Pal, M., Saha, S., Amit, K., “A Fuzzy C Means Clustering Approach for Gesture Recognition in Healthcare”, In: International Journal of Enhanced Research in Science Technology & Engineering, 2014.
- [3] Panchal, B. J., Kandoriya, K. P., “Hand Gesture Recognition Using Clustering Based Technique”, In: International Journal of Science and Research, 2013.
- [4] Gamarra, D. F. T., Cuadros, M. A. de S. L., “Comparisson of Fuzzy C Means, K-Means and K-MEdoids for Clustering in the Bag of Visual Words Algorithm”, In: Biblioteca Digital Brasileira de Computação, 2015.
- [5] Yonamine, F. S., Specia, L., Carvalho, V. O., Nicoletti, M. C. (2002) “Aprendizado não supervisionado em domínios fuzzy – algoritmo fuzzy c-means”, Universidade Federal de São Carlos.
- [6] Cruz, L., Lucio, D., Velho, L. (2012) “Kinect and RGBD Images: Challenges and Applications”, In: SIBGRAPI-T, Conference on Graphics, Patterns and Images Tutorials. p 36-49.
- [7] Reas, C., Fry, B., “Processing: A programming Handbook for Visual Designers and Artists”, Massachusetts Institute of Technology, 2007.