

Segmentação de Objetos com Aplicação em Robótica

Francisco de Assis Zampiroli¹

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil
fzampiroli@ufabc.edu.br

Vinícius da Fonseca Simioni

Centro de Matemática, Computação e Cognição
Universidade Federal do ABC
Santo André, Brasil
vinicius.simioni@aluno.ufabc.edu.br

Resumo — Este artigo teve como objetivo capturar uma bola utilizando um braço mecânico criado pelo *LEGO MindStorms*. Tal procedimento foi realizado utilizando-se processamento de imagens digitais para localizar o objeto. Foi construído um programa, com código em Matlab para o robô localizar a bola, retirar os ruídos da imagem obtida e por fim deslocar-se à bola e acionar o motor para capturar a bola usando comunicação *bluetooth*. Para localizar a bola utilizou-se também um código em Matlab para criar um retângulo em torno do objeto a ser capturado¹.

Palavras-chave - visão computacional, segmentação, robótica.

I. INTRODUÇÃO

O Processamento de Imagens Digitais consiste em aplicar transformações às imagens. As imagens podem ser formadas por elementos pictóricos, elementos de imagem, *pels* e *pixels*, sendo o termo pixel o mais utilizado atualmente [1].

O sentido mais avançado, e consequentemente o de maior importância para nós, seres humanos, é a visão. Portanto, o processamento de imagens digitais torna-se um ramo importante de estudo e compreensão. Porém, os aparelhos que utilizam o processamento de imagens digitais cobrem toda a banda visual do espectro eletromagnético, variando de ondas gamas a ondas de rádio, enquanto a visão humana é limitada apenas a uma pequena parte do espectro. Um grande objetivo das pesquisas científicas atualmente, que utilizando o processamento de imagens digitais e a visão computacional, é emular a visão humana, incluindo a capacidade de aprendizado, inferências e agir com base nas informações visuais obtidas.

O processamento de imagens digitais tem um amplo campo da aplicação, necessitando uma forma de organização para melhor aprendizado. Uma forma básica de organização do processamento de imagens digitais para facilitar sua compreensão nos diferentes ramos de aplicações é categorizar as imagens de acordo com sua fonte, por exemplo, visual, raios X, etc. Os raios tipo gama são utilizados na medicina com objetivo de coletar imagens do esqueleto humano para identificar possíveis doenças ósseas, infecções ou tumores. Uma das fontes de imagens do espectro eletromagnético mais antigo são os raios X. São utilizados em diagnósticos médicos, mas também é utilizada em várias áreas da indústria e em outras áreas, como a astronomia. As imagens da banda ultravioleta têm várias

aplicações, como a litografia, inspeção industrial, microscopia, lasers, imagens biológicas e observações, astronômicas. E por fim, as imagens dentro da banda visível do espectro eletromagnético são as com maiores aplicações, pois a convivência com esse tipo de imagens é frequente. As aplicações podem ser em microscopia ótica, astronomia, sensoriamento remoto, indústria e policiamento [1].

II. PRELIMINARES

O termo robótica surgiu pela primeira vez em uma obra ‘Runaround’ de Isaac Asimov, em 1941, e designa a área da ciência e tecnologia em que estuda, desenvolve e aplica robôs em diversas instâncias. O primeiro robô manipulador criado por George Devol teve como objetivo a manipulação de objetos e substância perigosas, integrando aos sistemas mecânicos de teleoperação tipo mestre-escravo. Os robôs manipuladores rapidamente entraram na indústria pela sua capacidade de alta precisão e repetição. Utilizados em tarefas como soldagem, pintura, montagem de dispositivos em placas de circuito impresso, por exemplo. Portanto a robótica trata-se da interação de máquinas com o mundo, e com isso rapidamente percebeu-se que essa interação apresentava inúmeros problemas e seria uma abordagem complexa. Os problemas foram divididos em duas grandes subáreas, percepção e atuação [2].

Os robôs são classificados de acordo com sua função. A classificação é, portanto, os robôs manipuladores, robôs móveis, percepção robótica e a visão robótica. Os robôs manipuladores têm como fundamentos a cinemática e a dinâmica ligados ao problema de controle. Na maioria das aplicações desse tipo de robô consiste na execução de um movimento específico planejado pelo manipulador robótico. O movimento é classificado como sendo livre ou restrito. Movimento livre não existe interação física entre o manipulador e o ambiente, e o movimento restrito ocorre quando há forças de contato entre o operador e o ambiente. Os robôs móveis são aqueles que não possuem uma base fixa, possibilitando a locomoção. Utilizam-se dispositivos para possibilitar esses movimentos, eles são as rodas, esteiras e pernas. Os robôs móveis apresentam os robôs humanoides que são os robôs que tentam replicar a anatomia humana. A percepção robótica utiliza-se de sensores para permitir ao robô que obtenha as informações do ambiente em que ele se encontra ou sobre si próprio. Os sensores são classificados como sensores internos ou proprioceptivos, que fornecem a informação sobre o estado robô [2].

¹ O autor agradece o apoio da FAPESP, processo 2009/14430-1.

A visão robótica é um meio sensorial mais complexo de ser realizado, pois emular a visão humana é uma tarefa extremamente complicada. Apesar de parecer que a visão humana seja trivial, é o sentido que necessita a metade ou mais do cérebro para realizar essa tarefa. A visão computacional refere-se ao processamento de dados captados por um sensor que use o espectro eletromagnético para produção de imagens. As imagens são representadas por meio de um conjunto de pixels (contração de *Picture element*) e são denominadas imagens digitais [1].

III. RESULTADOS

A. Estrutura

Este artigo teve o objetivo a criação de um robô LEGO [3] com a capacidade, através do processamento de imagens digitais, de selecionar e capturar uma bola, utilizando o software Matlab [4]. Utilizou-se neste software a *toolbox* chamada *Morphological Image Processing*, com o intuito de auxiliar na construção do algoritmo. O código criado em Matlab consiste em identificar um objeto no campo de visão da câmera instalado no robô, ativar os motores para ir em direção ao objeto e por fim, com o terceiro motor realizar a captura do objeto pelo braço mecânico. O vídeo processado em um desktop usando o Matlab foi capturado por uma webcam e os controles de movimento foram realizados por bluetooth [5], veja Figura 1.

B. Processamento de inicializações e controle – Módulo 1

Primeiramente foi realizada a conexão do LEGO com o computador, feito via Bluetooth, conforme [5]. A conexão via *bluetooth* foi a mais viável, pois a conexão via USB gera uma limitação da movimentação do robô via cabo. Após a conexão LEGO foi realizado a verificação das portas de saída do robô. Após a verificação dos sensores, é inicializada a conexão da câmera de vídeo com o Matlab. Ou seja, é realizada uma conexão da câmera de vídeo com um computador com o Sistema Operacional Windows (para usar Mac OS ou outro sistema, basta alterar uma variável *winvideo* para *macvideo* no código Matlab), na segunda porta (isso significa que foi usada a segunda câmera do computador). Em seguida, é realizado um ‘preview’, ou seja, uma visualização da câmera para ajustes finos antes de começar a processar o algoritmo. Esta visualização pode ser vista na Figura 2, caso contrário, o robô iria girar aleatoriamente até achar uma bola.

Após todos os requisitos citados estiverem corretos, o código inicia a captura do vídeo a fim de realizar processamento de imagens digitais. A primeira ação, após a captura de um frame do vídeo, é achar um objeto que seja capaz de ser capturado pelo braço mecânico. Ou seja, o algoritmo captura (imagens) e realiza um teste com esta imagem, para determinar se possui algum objeto com nível de cinza maior que um valor entre 0 e 255 (*threshold*), conforme a iluminação do ambiente e o contraste entre a bola e uma mesa, onde se encontra.

Em seguida é chamada um segundo módulo, para o reconhecimento da bola. Após este reconhecimento o

controle volta para o módulo 1 e acionar o braço mecânico para capturar a bola.

C. Processamento de reconhecimento - Módulo 2

Este módulo começa processando a imagem capturada após o *threshold* realizada pelo módulo 1, com o comando do Matlab ‘*imfill*’ que preenche buracos na imagem. Em seguida, ‘*imclearborder*’ elimina os objetos que tocam a borda da imagem. Então, o comando ‘*bwareaopen*’ é usado para eliminado objetos com área inferior a 600 pixels, eliminado assim pequenos ruídos e então é usado o comando ‘*bwlabel*’ para rotular os objetos (é desejado neste momento ter apenas um objeto, senão o controle retorna para o módulo 1 e obter um novo *threshold*). Em seguida é realizada uma lógica para criar um retângulo contendo a bola: primeiramente com o comando ‘*regionprop*’, defini-se um vetor com características geométricas dos objetos da imagem, tais como o centro geométrico. Assim, é possível guarda a informação do centro da imagem, e a partir deste centro é possível criar um retângulo que contenha a imagem com a bola. Em seguida é calculada a área do objeto a ser capturado e retornado este valor para o módulo 1.

Portanto, o objetivo deste módulo 2 é, com a imagem fornecida pelo módulo 1, criar o menor retângulo que contenha o objeto a ser capturado, calcular a área deste objeto e retornar este valor ao módulo 1. A Figura 3 indica o retângulo criado.

Depois de identificado o objeto a ser capturado, o módulo 1 indica ao robô LEGO comandos, utilizando os motores para chegar ao objeto: 1) A lógica utilizada para esta ação é a seguinte, há um ‘ponto’ pré-definido no módulo 2, sendo a metade na horizontal e $\frac{3}{4}$ na vertical. Quando o centroide do retângulo do objeto coincidir (com um erro pré-definido de 15 pixels) com este ponto é acionado o braço mecânico para a captura da bola. Na Figura 3 está indicada a congruência destes pontos. Para que haja esta congruência entre os pontos, são acionados os motores do LEGO por comandos simples de ‘se/senão’, no módulo 1, em referência as diferenças das coordenadas entre esses pontos. Ou seja, esta parte do código, aproxima o ponto pré-definido com o centroide do objeto. Quando os dois pontos tiverem coordenadas com distâncias inferior a 15 pixels é acionado o terceiro motor, a garra, com a finalidade de captura do objeto. A Figura 4 indica o momento após a captura da bola² e a Figura 5 resume o fluxograma do módulo principal.

IV. TRABALHOS CORRELATOS E CONCLUSÕES

A. Trabalhos correlatos

Existem inúmeros trabalhos correlatos de visão computacional com aplicações em robótica. Por exemplo, em [6] apresenta uma UAV para reconhecimento de pista de pouso para um avião de brinquedo. Em [7] é desenvolvido um veículo para locomoção em terrenos irregulares e a sua localização ocorre através de GPS, infravermelho e bússula. Em [8] desenvolve um UAV para ambientes com obstáculos.

² O link a seguir mostra um vídeo com a execução dos módulos apresentados: <http://youtu.be/G-NC3tKF1kU>.

Porém este robô se destaca por ser de baixo custo e didático, onde é possível usar em laboratórios de processamento de imagens digitais e inteligência artificial, além de laboratórios de programação básica, onde os módulos complexos são fornecidos aos alunos, aplicando de forma lúdica o ensinamento de estruturas de repetição e de controle.

B. Conclusões

Este artigo teve como objeto a criação de um robô 'LEGO' com a função de encontrar uma bola e locomover-se a ela para capturá-la. Foi utilizado dois módulos em Matlab, um de controle e outro de segmentação para que uma bola fosse capturada, utilizando técnicas de processamento de imagens digitais. O módulo principal, teve como função verificar todos os motores do robô, iniciar a câmera e capturar imagens a ser processada. O módulo 2 processa a imagem capturada pelo módulo 1, realizando técnicas de processamento de imagens digitais, como por exemplo excluindo os objetos que tocam a borda da imagem e também os objetos com área inferior a 600 pixels, deixando a imagem sem os ruídos. Após o reconhecimento da bola, o seu centroide é retornada ao módulo 1 para acionar os motores de locomoção. A garra é acionada quando um ponto pré-estabelecido (sendo a metade na horizontal e $\frac{3}{4}$ na vertical) coincide com o centroide do objeto, calculado no código 2, com uma distância inferior a 15 pixels.



Figure 1. Robô LEGO e a bola a ser capturada.

C. Trabalhos Futuros

Como trabalhos futuros, espera-se aprimorar o módulo de reconhecimento de objetos para identificar outros tipos. Por exemplo, para cada tipo de objeto o robô deve ter uma ação distinta. Espera-se também adaptar o módulo 1 para outros tipos de veículos autônomos (*Unmanned Aerial Vehicle* - UAV), como um quadróptero.

REFERÊNCIAS

- [1] Gonzales, R.C. e Woods, R.E. "Digital Image Processing". Addison-Wesley Pub. Company, 3rd, pg. 976, 2007.
- [2] Aguirre, A.L. "Enciclopédia de Automática Controle & Automação", Vol. 3. 1ª Edição, 2007. Editora Blucher.
- [3] LEGO, NXT Vision Ball Picker. Disponível no site http://www.roborealm.com/tutorial/ball_picker/slide100.php. Acessado em maio de 2012.
- [4] LEGO MINDSTORMS NXT. Disponível no site: <http://mindstorms.lego.com>. Acessado em maio de 2012.
- [5] Matlab. Disponível no site: <http://www.mathworks.com>. Acessado em maio de 2012.
- [6] Frew, E., Mcgee, T., Kim, Z., Xiao, X., Jackson, S., Morimoto, M., Rathiam, S., Padiyal, J. and Sengupta, R. Visionbased road-following using a small autonomous aircraft, Aerospace Conference, 2004. Proceedings. 2004 IEEE 5: 3006-3015.
- [7] Kendoul, F., Nonami, K., Fantoni, I., Lozano, R. An adaptive vision-based autopilot for mini flying machines guidance, navigation and control. Autonomous Robots v.27-n.3:165-188, 2009.
- [8] Neto, A.A., Macharet, D.G. and Campos, M.F.M. On the Generation of Trajectories for Multiple UAVs in Environments with Obstacles. Intell Robot Syst, 57:123-141, 2010.



Figure 2. Visualização da câmera fixada no robô LEGO.

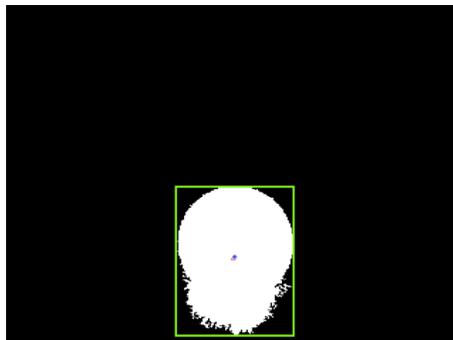


Figure 3. Segmentação da bola e contornando com um retângulo.



Figure 4. Captura da bola.

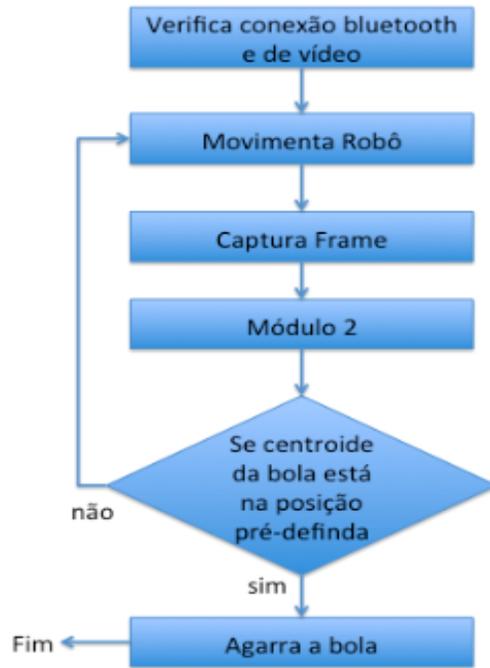


Figure 5. Fluxograma do módulo principal.