



UNIVERSIDADE FEDERAL DE SANTA CATARINA

Aluno: Aléxei Felipe Paim

Matrícula: 20250264

BLU3040-08754 (20222) - Visão Computacional em Robótica

Trabalho 1

Descrição do problema:

No primeiro trabalho da disciplina de Visão Computacional em Robótica, é solicitado que seja desenvolvido um sistema simplificado de reconhecimento óptico de caracteres. Para realizar esta tarefa, é imposto que seja feito o uso da técnica de *Template Matching*. O programa implementado deve ser capaz de realizar a leitura de imagens (como a da figura 1), a fim de reconhecer os caracteres e transcrevê-los para janela de comandos, assim como ilustrado na figura 1.



Figura 1 : (a)Exemplo da imagem de um etiqueta, cujo texto deve ser reconhecido pelo sistema projetado.(MATSUO, 2022). (b) Exemplo do resultado a ser obtido.

Descrição do algoritmo proposto:

Para a exemplificação de como será implementado o sistema simplificado de reconhecimento óptico de caracteres, todo o processo será realizado sobre a imagem da figura 2. O primeiro passo neste sistema é realizar uma transformação geométrica de homografia planar, a fim de reposicionar e redimensionar os caracteres da imagem, para realizar esta etapa de forma automática é preciso fazer com que o sistema encontre as coordenadas dos quatro pontos verdes da imagem.



Figura 2 : Imagem que o processo será demonstrado

A função $f_{pon_verde_ord}$, é a responsável por realizar o processo descrito acima, a função primeiramente realiza uma segmentação baseada em cor, para encontrar os pontos em verde, de tal forma que então seja possível extrair as características de cada região da imagem. A figura 3 demonstra este processo.

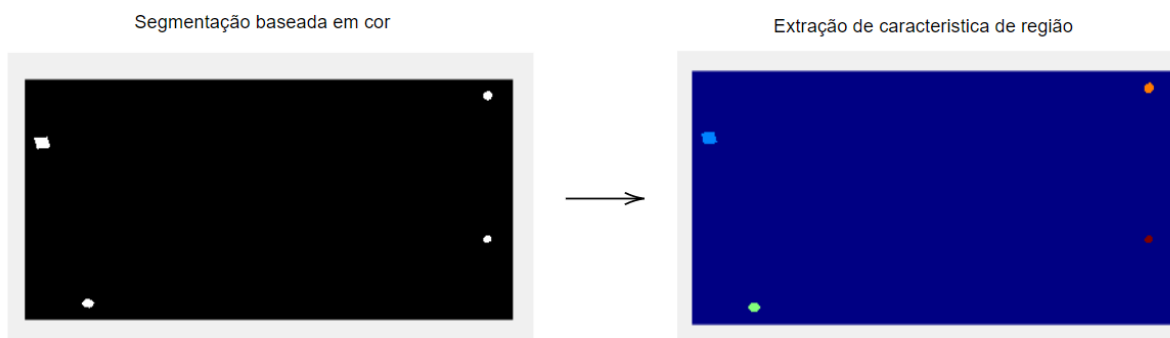


Figura 3 : processo de segmentação é baseada em cor e extração de característica

Como a função que realiza extração de características não ordena as regiões encontradas, a função $f_{pon_verde_ord}$, faz o chamado para a função f_{per_circ} que realiza o cálculo do perímetro e da circularidade para cada uma das regiões encontradas e então retorna somente o valor de circularidade, pois através deste parâmetro o sistema será capaz de identificar qual das regiões corresponde ao quadrado. Afinal o valor de circularidade para um quadrado é consideravelmente menor para que à de um círculo.

Após o sistema ter reconhecido o quadrado verde do canto superior esquerdo, realiza-se através da função f_{ctod} o cálculo da coordenada do centróide de cada uma das regiões encontradas, pois através destas coordenadas será feito comparações de posições, a fim de criar um vetor ordenado para o processo de transformação de homografia planar e assim encerrar a função $f_{pon_verde_ord}$.



Figura 4 : Imagem da transformação geométrica de Homografia Planar

Com as coordenadas obtidas na função descrita acima e a imagem original de entrada, torna-se possível realizar processo de transformação geométrica de Homografia Planar (figura 4), de tal forma que para a obtenção deste resultado é criado a função $f_{homografia}$ que será utilizada mais algumas vezes durante o desenvolvimento do sistema.

Na sequência do desenvolvimento é perceptível que a imagem de saída do processo de homografia, contém muita informação não relevante para o sistema, assim, novamente é utilizado a função $f_{pon_verde_ord}$, de tal forma que ela retorne as coordenadas dos pontos verdes da figura 4, para que então seja realizado um corte, uma limiarização e limpeza de borda, como ilustrado na figura 5.

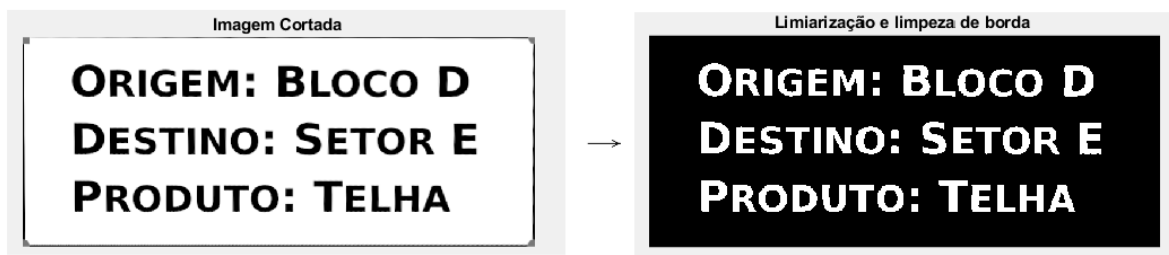
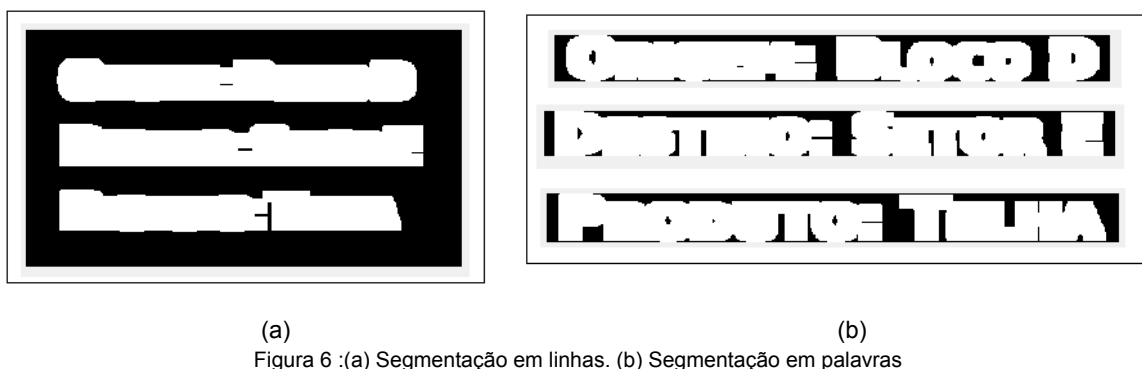


Figura 5 :Ilustração do processo de corte, limiarização e limpeza de borda.

Então, após o tratamento da imagem de entrada, se dá início a etapa de segmentação e rotulamento de regiões. Para isso, é criada a função $f_seggmen$. Esta função recebe três parâmetros, sendo eles, a imagem a ser segmentada e dois valores de parâmetro para a operação de dilatação, vertical e horizontal, respectivamente, assim obtendo os resultados para a segmentação das linhas e das palavras da figura 6. A função $f_seggmen$, então retorna uma struct contendo as imagens extraídas.



Assim como ilustrado na figura 6, o sistema repete a segmentação para encontrar cada linha de caracteres, cada palavra e para identificar cada caracter individualmente como mostra o exemplo da figura 7.

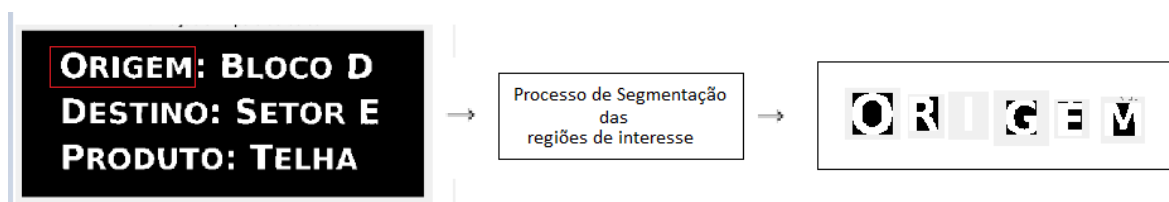


Figura 7 :Resultado do processo de segmentação das regiões de interesse

Descrito a etapa de segmentação e extração das regiões de interesse, pode-se dar início a explicação da função *f_trat_comp_print*. No entanto, como esta será a função responsável por realizar as comparações por similaridade, é preciso descrever como é obtido as imagens dos caracteres que servirão como template padrão para a comparação. Assim, para a realização deste trabalho o professor da disciplina, disponibilizou a imagem da figura 8, que contém as letras alfabéticas que deverão ser usadas como template padrão.

A B C D E F G
H I J K L M N
O P Q R S T
U V X Z W

Figura 8 :Template das letras alfabéticas

Entretanto, para realizar as comparações de similaridade também é necessário que as letras de base estejam segmentadas. Então para isso é utilizado a função *f_templates_letras*, que para a imagem da figura 8, realiza a limiarização, segmentação das regiões de interesse, os redimensionamentos de 50x51 e o retorno de uma struct contendo, cada letra encontrada juntamente com seu respectivo caractere, como ilustra a figura 9.

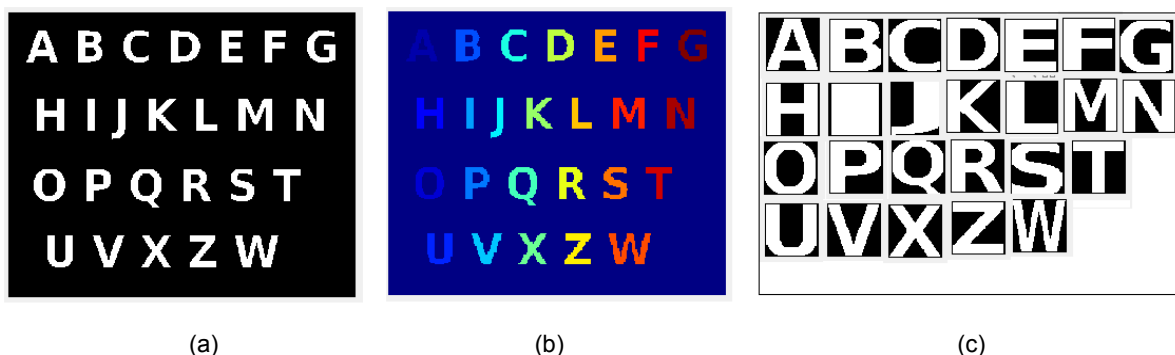


Figura 9 : (a) Limiarização da imagem template. (b) Extração das regiões de interesse. (c) Imagens extraídas e redimensionadas

Com a struct contendo as imagens de base para comparação e os respectivos caracteres, pode-se dar início à função *f_trat_comp_print* que nas imagens obtidas na etapa de segmentação e extração das regiões, realizará o redimensionamento das imagens. Assim, dessa forma, novamente é utilizado a função de transformação geométrica de Homografia Planar, em cada uma das imagens. Realizado este processo faz-se um ajuste para que todas as imagens fiquem exatamente com o tamanho de 50x51.como mostra a figura 10.

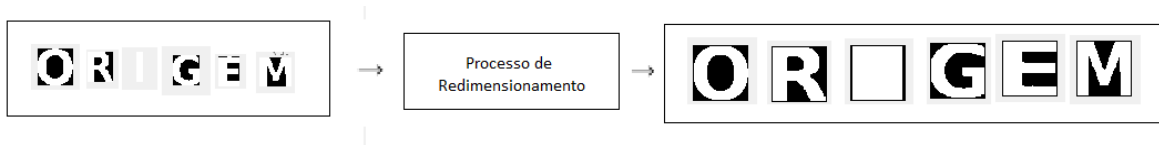


Figura 10 :Processo de redimensionamento dos caracteres extraídos da imagem de entrada

Ainda na função *f_trat_comp_print*, utiliza-se cada uma das imagens dos caracteres extraídos e redimensionados da imagem de entrada, para realizar a comparação das métricas de similaridade com as imagens dos caracteres extraídos do template base. A comparação das métricas de similaridade é realizado através da correlação cruzada normalizada (equação 1), Onde I_1 e I_2 são as imagens dos caracteres do template e dos caracteres da imagem de entrada, u e v são as linhas e colunas dos pixels das imagens respectivamente

$$(1) \quad \text{NCC} \quad s(I_1, I_2) = \frac{\sum_{(u,v) \in I_1} I_1(u, v) I_2(u, v)}{\sqrt{\sum_{(u,v) \in I_1} I_1^2(u, v) \sum_{(u,v) \in I_2} I_2^2(u, v)}}$$

A comparação das métricas de similaridade utilizando a correlação cruzada normalizada é a forma mais robusta de obter a similaridade entre as imagens, o processo de comparação é ilustrado na figura 11. A comparação é feita da seguinte forma, para cada imagem de carácter extraída da imagem de entrada é feito o cálculo da métrica de similaridade com cada uma das imagens de caracteres extraídas do template base. Assim, o maior valor entre as comparações corresponderá com o carácter encontrado na imagem.

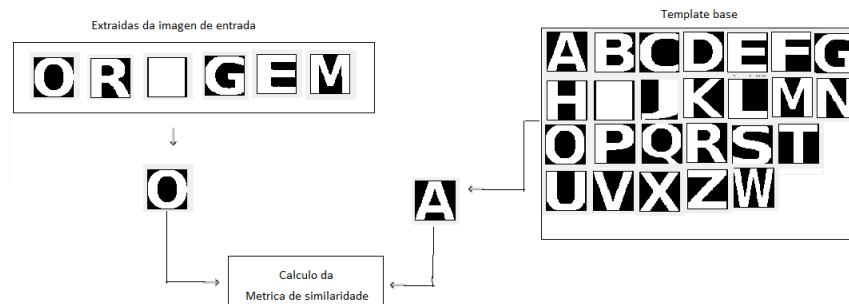


Figura 11 :Ilustração do processo de comparação de similaridade

Por tanto, identificado cada caracter da imagem de entrada, realiza-se a impressão do caracter correspondente na janela de comandos, assim como mostra a figura 12. O mesmo processo é aplicado para todas as imagens do banco de imagens disponibilizado.

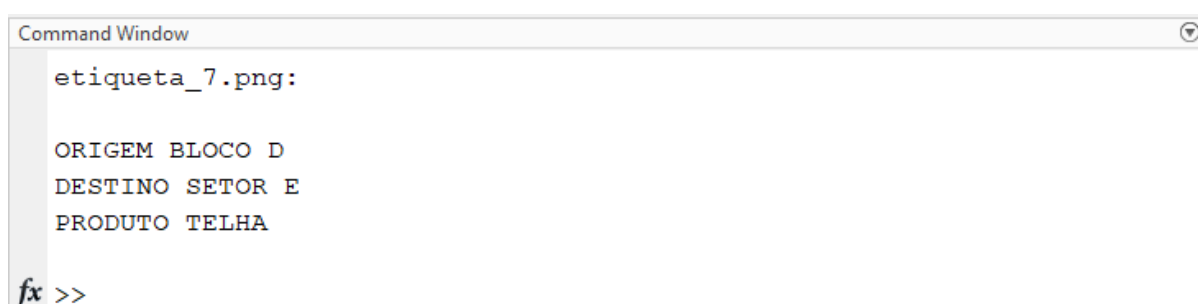


Figura 12 :Resultado final

Por fim, na figura 13, é apresentado um fluxograma simplificado de todo o processo realizado pelo sistema simplificado de reconhecimento óptico de caracteres implementado.

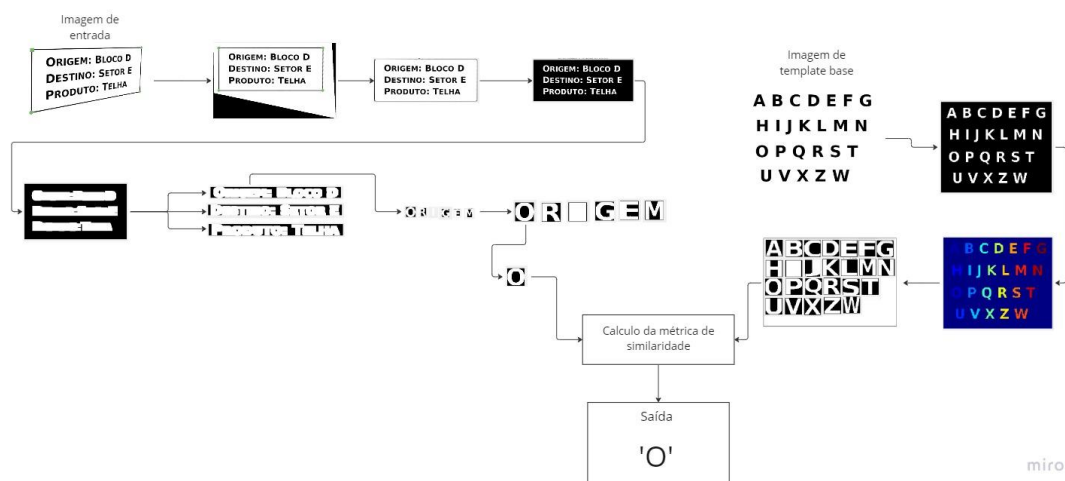


Figura 13: Fluxograma simplificado do sistema simplificado de reconhecimento óptico de caracteres

Resultados:

Com o sistema simplificado de reconhecimento óptico de caracteres implementado, é obtido os resultados que estão ilustrados na figura 14, o que indica que o sistema corresponde com o resultado que foi solicitado na descrição do problema deste trabalho.



Figura 13: Resultados do sistema implementado

Conclusão:

Por fim, conclui-se que este trabalho de implementação de um sistema simplificado de reconhecimento óptico de caracteres, é de suma importância para a fixação dos conteúdos ministrados na disciplina de Visão computacional em Robótica, de tal forma que este trabalho também apresenta uma aplicabilidade do grande campo de Visão computacional.

Referências:

MATSUO, Marcos. **Trabalho** 1. Disponível em: https://moodle.ufsc.br/pluginfile.php/5835233/mod_assign/introattachment/0/BLU3040%20-%20Trabalho%201.pdf?forcedownload=1. Acesso em: 26 nov. 2022.