

Realidade Aumentada aplicada a uma biblioteca

Carlos Frias¹, Nuno Marques²

Mestrado Integrado em Engenharia Informática e Computação

Faculdade de Engenharia da Universidade do Porto

¹carlos.frias@fe.up.pt

²nuno.marques@fe.up.pt

Abstract— This report aims to document the implementation of an augmented reality application for the automatic detection of books in a shelf to be applied to a library. The application was developed in MATLAB using the tools of image processing and computer vision.

Keywords— Book detection; Augmented reality; MATLAB; Image Processing Toolbox.

I. INTRODUÇÃO

Este relatório pretende documentar a implementação de uma aplicação de realidade aumentada aplicada às bibliotecas para a deteção de livros numa estante, utilizando diversas técnicas de processamento e análise de imagens e de visão por computador. A aplicação foi desenvolvida em MATLAB, recorrendo à ferramenta Image Processing Toolbox. O programa permite acrescentar a uma imagem real, estática, de uma estante, uma marca virtual identificando um livro procurado por um utilizador, de forma semelhante à ilustrada na Figura 1.



Figura 1: Estante inicial e estante contendo uma marca no livro procurado.

II. ESPECIFICAÇÃO

O funcionamento do programa desenvolvido é o seguinte: É apresentada ao utilizador uma interface gráfica com se mostra na Figura 2:

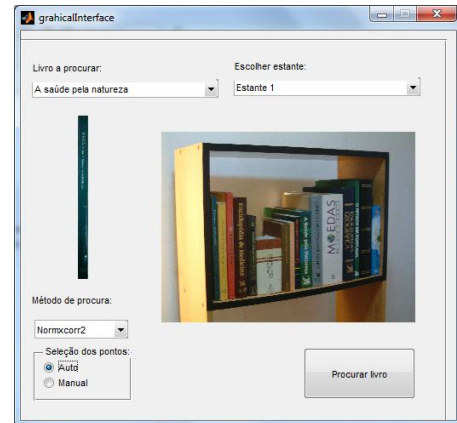


Figura 2: Interface da aplicação.

Nesta interface o utilizador pode seleccionar o nome do livro que pretende consultar no pop-up menu do canto superior esquerdo, pode seleccionar a imagem da estante no pop-up menu do canto superior direito, pode seleccionar o método de procura no pop-up menu do canto inferior esquerdo e ainda pode optar por uma seleção dos quatro pontos para realizar a homografia de um modo manual ou automático. O botão Procurar Livro inicia a procura do livro. Nesta interface o utilizador visualiza também a lombada do livro que está à procura, a estante com todos os livros, no final, visualiza a estante com uma marca sobre o livro procurado e com essa região clareada.

No desenvolvimento deste trabalho teve-se em conta os seguintes pressupostos simplificativos: todos os livros estão colocados em posição vertical e as suas lombadas estão alinhadas, isto é, são mais ou menos coplanares; para cada livro colocado na estante, existe um ficheiro contendo uma imagem da sua lombada em posição vertical, adquirida em pose frontal; as imagens da estante usadas para fazer a deteção do livro são adquiridas em posição estática mas não, necessariamente adquiridas em pose frontal; nas imagens adquiridas, são visíveis alguns pontos fiduciais da estante que permitam fazer a calibração do sistema.

III. DETALHES DA IMPLEMENTAÇÃO

A aplicação implementada percorre as seguintes etapas:

1. Escolha por parte do utilizador do livro a procurar e da estante onde essa procura irá ser realizada;
2. Realização da homografia que transforme a fotografia original da estante, tirada de uma posição

não frontal, numa imagem da estante em posição frontal;

3. Escalamento da imagem da lombada do livro para estar na mesma proporção da homografia realizada no passo anterior;
4. Detecção das coordenadas dos quatro cantos do livro procurado na imagem da homografia;
5. Realização da homografia inversa para esses quatro pontos, obtendo assim as coordenadas dos quatro cantos na imagem original;
6. Realce na imagem original do quadrilátero obtido, clareando a imagem, desenhando os lados desse quadrilátero e escrita da palavra DETECTED.

Na primeira etapa realizou-se a interface gráfica já apresentada e também uma interface em modo de texto. Os nomes dos livros e estantes, as suas dimensões reais, bem como os caminhos para os ficheiros das imagens das lombadas do livro e das imagens das estantes são carregados a partir dois ficheiros de texto, *dbbookcase.txt* e *dbbooks.txt*.

Relativamente à segunda etapa a realização da homografia pode ser feita de dois modos: modo manual e modo automático: No modo manual é pedido ao utilizador a seleção de quatro pontos na imagem. Esses pontos são os quatro cantos exteriores da estante, escolhidos consecutivamente no sentido dos ponteiros do relógio; No modo automático, esses quatro cantos seriam detetados pela aplicação utilizando a ferramenta de *cornerdetection* ou a função *normxcorr2* do MATLAB para comparar a imagem da estante com imagens templates dos quatro cantos da estante. Após a seleção dos quatro pontos, é realizada a homografia, fazendo uma chamada à função *cp2tform*, obtendo assim a imagem na pose frontal e a matriz de transformação. A Figura 3 mostra uma estante na sua posição original e a homografia obtida:



Figura 3: (a) Antes da homografia. (b) Depois da homografia.

Depois de realizada a homografia há que redimensionar o template da lombada do livro para que tenha as mesmas dimensões da imagem do livro presente na imagem em pose frontal. Efetua-se esse redimensionamento da seguinte forma: as dimensões reais da estante, comprimento e largura, são carregadas do ficheiro de texto, bem como a largura da lombada do livro. É calculada a razão de semelhança entre as dimensões reais, em centímetros, da estante e o número de pixéis, na horizontal e na vertical, da imagem da estante na posição frontal, (para isso efetua-se uma chamada à função *imcrop*) e com simples operações aritméticas determina-se quais terão de ser as dimensões da imagem template da

lombada do livro. Essa redimensão é realizada por uma chamada à função *imresize*.

Na quarta etapa, deteção da posição do livro na imagem da estante na sua pose frontal, foram utilizadas três técnicas distintas:

1. Utilização da matriz de correlação, (cross correlation matrix): efetuando uma chamada à função *normxcorr2*, passando como argumentos a imagem da estante na pose frontal e o template da lombada do livro na dimensão correta, ambos em escala de cinzentos. Da matriz obtida é selecionado o valor máximo, correspondendo os índices da posição desse máximo às coordenadas do ponto da imagem com maior semelhança com o template da lombada do livro. Se esse valor for superior a um valor pré-definido, é considerado que a deteção ocorreu com sucesso. Esse ponto é canto inferior direito do livro na imagem e, tendo as dimensões da lombada do livro, são calculadas, facilmente, as coordenadas dos restantes três cantos do livro;
2. BackProjection: efetuando a conversão da representação da imagem da estante e da lombada do livro de RGB para HVS. Efetua-se a contagem da cor (Hue) e da saturação do template da lombada do livro para uma matriz (256x256). Realiza-se uma normalização dessa matriz para os valores variarem de entre 0 a 255. De seguida realiza-se a comparação com a imagem original e obtém-se um histograma, em que, para cada posição se a cor e saturação da imagem nesse pixel for muito próxima da cor e saturação apresentada na matriz calculada, então é colocado a branco, senão colocado a preto.

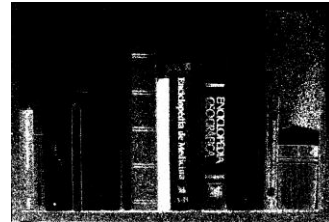


Figura 4: O livro procurado está todo branco.

3. Deteção das linhas de Hough: Neste processo é realizada a deteção de arestas (edges), efetuando a chamada à função *edge*, utilizando o método de Canny para a sua deteção, pois foi este o método que melhores resultados apresentou. É calculada a transformada de Hough, chamando a função *hough*, passado como argumento a imagem obtida no passo anterior e retornando a matriz da transformada de Hough, um vetor com os valores dos ângulos teta e outo com os valores de rho, de seguida determina-se os picos de Hough, chamando a função *houghpeaks* e com esses valores determinar-se as linhas de Hough com a função *houghlines*. Essas linhas são retornadas num structarray. O passo seguinte é selecionar dessas linhas apenas a que são verticais, uma vez que as fronteiras das lombadas dos livros são

também verticais. Essa seleção é feita na função, criada para o efeito, *selectVerticalLines* que apenas retorna as linhas em que o valor de teta é 0° (ou muito próximo), isto é, as linhas verticais. Depois, estas linhas são ordenadas por ordem crescente de abscissa, para facilitar o seu acesso. São efetuados dois refinamentos: um para unir linhas verticais com a mesma abscissa, ou abscissas muito próximas, e um outro para remover linhas demasiado curtas. Efetuados estes refinamentos, são calculados os retângulos que contém uma lombada de um livro, isto percorrendo as linhas verticais e unindo-as como num retângulo. Finalmente, é executado a função *normxcorr2* e determinado o ponto com maior correlação com o template da lombada do livro e, de seguida, identificado qual dos retângulos contém esse ponto. Na Figura 5 estão apresentados os passos de execução neste processo:

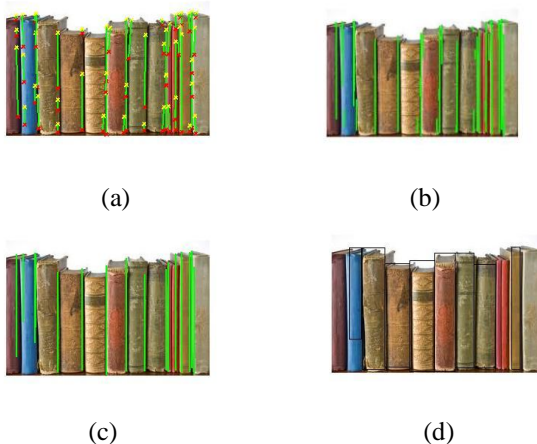


Figura 5: (a) detecção das linhas verticais
(b) união das linhas verticais com a mesma abscissa
(c) eliminação das linhas curtas
(d) detecção dos retângulos contendo livros

Após a detecção das coordenadas dos cantos do livro na imagem da estante em pose frontal, passamos à quinta etapa do desenvolvimento que é o cálculo das coordenadas desses pontos na imagem original da estante. Para isso, efetua-se a homografia inversa da homografia já realizada, com a chamada à função *tforminv*, passando como argumentos a matriz de transformação e as coordenadas do ponto na imagem em pose frontal, é retornada as coordenadas desse ponto na imagem original. Realiza-se essa operação para os quatro pontos.

Finalmente, resta a última etapa, de realce na imagem original do quadrilátero contendo o livro procurado. Para tal, é chamada a função *selection*, criada para esse efeito, recebendo como argumentos, a imagem original e as coordenadas dos vértices do quadrilátero, e retornando a imagem com o realce desse quadrilátero e com a palavra DETECTED, como mostra na Figura 6:

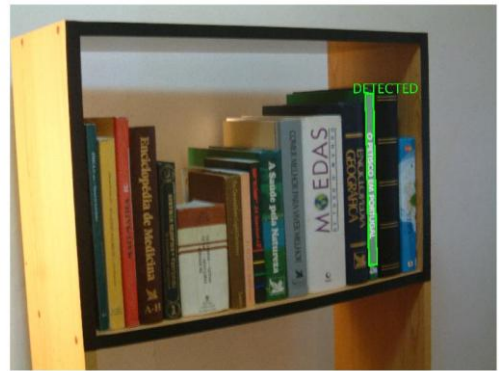


Figura 6: Estante com a marca da detecção do livro procurado.

IV. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Relativamente aos três processos de detecção utilizados, aquele que nos pareceu mais robusto, menos sensível a alterações de iluminação e que impunha um menor número de restrições foi o método que utiliza a matriz de correlação (cross correlation matrix), com este método, mesmo em condições diferentes de iluminação, com maior ou menor quantidade de reflexo e com maior ou menor existência de sombras o método permite identificar a posição de qualquer livro na maioria dos casos.

Na detecção pelas linhas de Hough as condições necessárias para a boa detecção dessas linhas são restritas, a saber: as lombadas deverão ter a menor quantidade de texto possível, ou mesmo a ausência de texto, pois não sendo esse o caso são detetadas linhas correspondentes ao texto da lombada, como se pode observar na Figura 7:

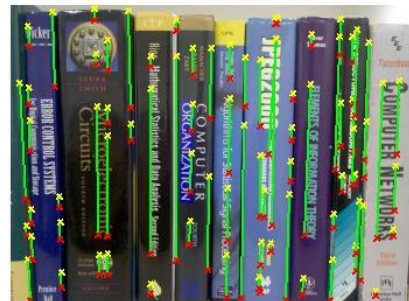


Figura 7: Má detecção das linhas de Hough.

Ainda se livros consecutivos tiverem cores similares, a detecção da linha entre eles pode não se efetuar.

Na detecção pelo método de backprojection as condições de utilização também teriam de ser algo restritas, a saber: a necessidade de as tonalidades das lombadas dos livros terem de ser diferentes e a alta sensibilidade a alterações das condições de iluminação, brilho e sombras. Com base no numero de ocorrências das cores de um template, esta abordagem apresenta outra fragilidade, a impossibilidade de distinguir livros diferentes mas com quantidade de cores iguais.

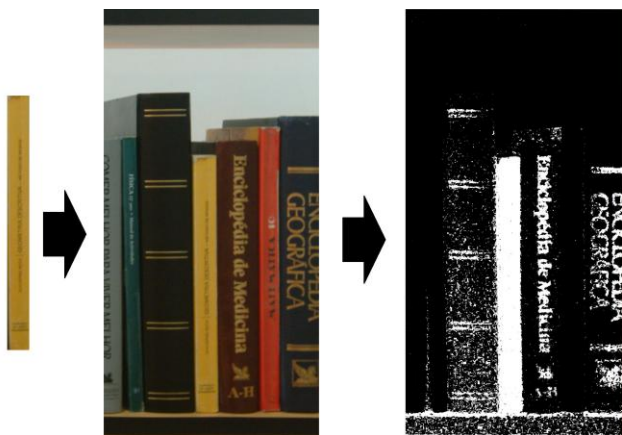


Figura 9: O backprojection parte das duas imagens a direita e devolve a da esquerda.

Relativamente à deteção dos quatro cantos da estante automaticamente, tentou-se, com resultados insatisfatórios, utilizar a ferramenta de *cornerdetection*. No entanto, esse problema pode ser resolvido com a deteção dos cantos com matriz de correlação (cross correlation matrix), usando a função *normxcorr2* e um template para cada um dos cantos da estante.

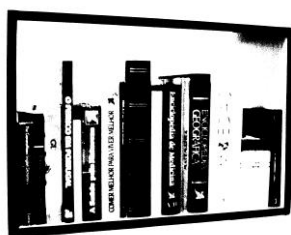


Figura 8: Prateleira com contorno preto para auxílio da sua deteção.

Na interface gráfica é apresentada a possibilidade de escolha do método de deteção, no entanto, apenas o primeiro método está implementado para essa interface, em virtude dos dois restantes exigirem condições mais restritas na escolha da imagem da estante. A verificação desses dois métodos terá de ser realizada manualmente.

V. CONCLUSÕES E MELHORAMENTOS

Com a realização deste trabalho, além do desenvolvimento de aptidões na programação em MATLAB, permitiu elaborar um programa estável, para deteção de livros numa estante, e ainda o aprofundamento de conhecimentos de processamento e análise de imagens e visão por computador. Algumas melhorias para este projeto passariam pelo refinamento da deteção dos livros pelos dois métodos que apresentaram piores resultados e exigiam maior número de condicionantes. E ainda, a deteção automática dos pontos críticos (cantos da estante) para o processo de execução ser mais fluído.

Em suma, pensamos que o trabalho realizado apresenta qualidade e que atingiu claramente os objetivos mínimos que se propunha.

BIBLIOGRAFIA E REFERÊNCIAS

- [1] Professores António Augusto de Sousa e Jorge Alves da Silva, Página da disciplina de Realidade Virtual e Aumentada, <https://moodle.fe.up.pt/1213/course/view.php?id=570>
- [2] Documentação do Matlab Image Processing Toolbox, <http://www.mathworks.com/products/image/>
- [3] Template matching video tutorial using the *normxcorr2* function, <http://www.youtube.com/watch?v=Q-OzmDen4HU>
- [4] Back Projection tutorial, http://docs.opencv.org/doc/tutorials/imgproc/histograms/back_projection/back_projection.html
- [5] Xilin Chen, Member, IEEE, Jie Yang, Member, IEEE, Jing Zhang, and Alex Waibel, Member, IEEE, Automatic Detection and Recognition of Signs From Natural Scenes, <http://www.jdl.ac.cn/user/xlchen/Paper/TIP04.pdf>