



TRANSFORMAÇÕES GEOMÉTRICAS PARA OBTER MOSAICO DE IMAGENS DIGITAIS

Manolo Rodriguez Heredia (1); Luzivalda Araujo de Sousa(2); Ruan Lion (3).

- (1) Doutor em Matemática Aplicada, Professor do curso de Licenciatura em Matemática do IEA, manolorh@unifesspa.edu.br
(2) Graduanda no curso de Licenciatura em Matemática, luzy.araujo21@outlook.com, Universidade Federal do Sul e Sudeste do Pará - UNIFESSPA
(3) Graduando no curso de Licenciatura em Matemática, ruanlion.13@gmail.com, Universidade Federal do Sul e Sudeste do Pará - UNIFESSPA

RESUMO

O objetivo do trabalho é apresentar uma aplicação da matemática na visão computacional para obter um mosaico de imagens digitais. O mosaico se refere a conectar duas ou mais imagens e criar uma nova imagem de área ampla sem linhas de costura visíveis. Existem vários algoritmos para construir mosaicos a partir da sequência da imagem, onde o movimento da câmera é mais ou menos complexo. A maioria desses métodos baseiam-se na correspondência de pontos de interesse ou em modelos teóricos de canto. Este trabalho, por sua vez, descreve um método manual para determinar um mosaico de imagens com base na escolha de pontos que estão nos cantos de figuras planas, os quais servem como ponto de interesse para determinar a transformação geométrica que estabelece a correspondência de regiões. Serão estudados e implementados algoritmos que permitam estimar um alinhamento inicial das imagens junto com o enquadramento da janela de pesquisa de pontos de interesse e, portanto, reduzir a complexidade do algoritmo de correspondência de pontos de interesse.

Palavras-chave: Transformação geométrica, Visão computacional, Mosaico.

ABSTRACT

The objective of this work is to present an application of mathematics in computer vision to obtain a mosaic of digital images. Mosaic refers to connecting two or more images and creating a new wide area image with no visible sewing lines. There are several algorithms for building mosaics from the image sequence, where camera movement is more or less complex. Most of these methods are based on point-of-interest matching or theoretical corner models. This paper, in turn, describes a manual method for determining a mosaic of images based on the choice of points that are in the corners of flat figures, which serve as a point of interest for determining the geometric transformation that establishes region matching. Algorithms will be studied and implemented to estimate the initial alignment of the images along with the framing of the point of interest search window and thus reduce the complexity of the point of interest matching algorithm.

Keywords: Geometric transformation, Computational vision, Mosaic.

1. INTRODUÇÃO

Este trabalho buscar apresentar resultados preliminares da pesquisa realizada no projeto denominado "Estudo dos princípios de geometria projetiva com aplicações em visão computacional" e que integra o Programa Voluntário de Iniciação Científica – PIVIC da Universidade Federal do Sul e Sudeste do Pará - Unifesspa.

O trabalho irá discutir sobre as transformações projetivas, uma vez que estas modelam a distorção geométrica que surge quando um plano é fotografado por uma câmera em perspectiva. Sob uma perspectiva da imagem, certas propriedades geométricas são preservadas, como a colinearidade (uma linha reta é representada como uma linha reta), enquanto outras não são, por exemplo, linhas paralelas não são representadas como linhas paralelas em geral. A geometria projetiva modela essa imagem e fornece uma representação matemática para os cálculos, como abordado por Gomes e Velho (2004).

Um mosaico de imagens é a conexão de duas ou mais imagens e, portanto, a criação de uma nova imagem de área ampla sem linhas de costura visíveis, como apresentado no artigo de Zitova e Flusser (2003) e no artigo de Zagrouba, Barhoumi e Amri (2009). Neste trabalho utilizou-se a escolha manual de pontos em cada imagem para obter um mosaico. Embora existam vários algoritmos para construir mosaicos a partir da sequência da imagem, onde o movimento da câmera é mais ou menos complexo. A maioria desses métodos baseiam-se na correspondência de pontos de interesse ou em modelos teóricos de canto. Estudou-se e implementou-se algoritmos que permitiram estimar um alinhamento inicial das imagens junto com o enquadramento da janela de pesquisa de pontos de interesse.

A linguagem de programação técnica Octave é usada no processamento de dados e na análise de aplicativos relacionados à engenharia e à ciência. Neste trabalho usamos este software no processamento de imagens, em particular, foi utilizado esse software para marcar o ponto de interseção de duas imagens, as quais estão representadas como matrizes. Dados representados em uma matriz unidimensional costumam ser comparados plotando-se seus vetores e os de outra matriz num mesmo gráfico. O gráfico permite a visualização de dados e mostra os pontos relevantes, incluindo os máximos e mínimos e os pontos de interseção entre duas séries de dados. Depois de encontrar os pontos de interseção, também é possível rotulá-los. Isso é feito para renderizar o mosaico de imagens, assim, obter uma nova imagem de área ampla sem linhas de costura visíveis.

2. OBJETIVO

O objetivo do trabalho é apresentar uma aplicação da matemática na visão computacional para obter um mosaico de imagens digitais.

3. MÉTODO

Apresenta-se os conceitos de plano projetivo, transformações geométricas no plano projetivo para aplicar esses conceitos e obter um mosaico de imagens por meio da estimação de uma homografia.

1.1. Transformações geométricas no plano projetivo

Para entender o plano projetivo, denotado por \mathbb{RP}^2 , pode-se pensar em um conjunto de raios em \mathbb{R}^3 que passam pela origem. Na medida em que um valor real k varia, obtemos um conjunto de vetores $k(x_1, x_2, x_3)^T$ que passam pela origem de \mathbb{R}^3 . Logo, esse raio pode ser considerado como representante de um único ponto no \mathbb{RP}^2 , veja Figura 1.

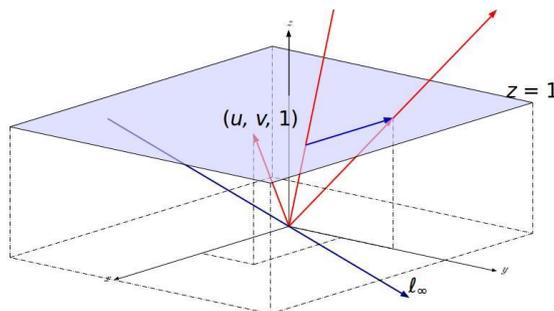


Figura 1 – Representação de \mathbb{RP}^2 , com $\mathbb{RP}^2 = (\mathbb{R}^2 \times \{1\}) \cup \ell_\infty$
(Fonte: Elaboração dos Autores.)

Além disso, nesse modelo, as linhas no \mathbb{RP}^2 são planos passando pela origem de \mathbb{R}^3 . Verifica-se que dois raios não idênticos estão exatamente em um plano e que quaisquer dois planos se cruzam em um raio. Dessa maneira, observa-se como o papel de pontos e linhas pode ser intercambiado em declarações sobre as propriedades de linhas e pontos. Conceitos primitivos de geometria projetiva como pontos, retas e planos utilizados em visão computacional estão descritas com mais detalhes em livros didáticos sobre geometria de múltiplas visualizações, veja Faugeras, Luong e Papadopoulos (2001) e Hartley e Zisserman (2004).

A geometria projetiva em duas dimensões pode-se considerar como o estudo das propriedades do plano projetivo \mathbb{RP}^2 que são invariantes por um grupo de transformações conhecidas como projetivas. Uma transformação projetiva é uma aplicação invertível que leva respectivamente pontos e linhas de \mathbb{RP}^2 a pontos e linhas em \mathbb{RP}^2 , isto é:

Definição 1. Uma aplicação $h : \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$ é uma transformação projetiva ou uma homografia se h é invertível e os pontos x_1, x_2, x_3 são colineares se, e somente se, $h(x_1), h(x_2), h(x_3)$ também são colineares.

As transformações mais simples que ocorrem no plano são ilustradas na Figura 2.

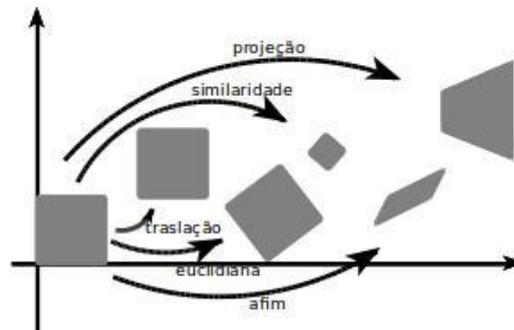


Figura 2 – Transformações geométricas no plano projetivo.
(Fonte: Elaboração dos Autores)

Uma homografia $h : \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$ pode ser representada matricialmente por $Hu = x$, isto é:

$$\begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

3.2. Estimação de uma homografia

Uma homografia tem 8 graus de liberdade, mas é um costume tratar todas as 9 entradas da matriz H como desconhecidas.

Dadas duas imagens diferentes para estimar a homografia que relaciona pontos dessas imagens, deve-se resolver um sistema linear, o qual determina as entradas da matriz H .

$$\begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\text{se, e somente se, } \begin{cases} uh_1 + vh_2 + h_3 = x \\ uh_4 + vh_5 + h_6 = y \\ uh_7 + vh_8 + h_9 = 1 \end{cases}$$

$$\text{se, e somente se, } \begin{bmatrix} 0 & 0 & 0 & -u & -v & -1 & yu & yv & y \\ u & v & 1 & 0 & 0 & 0 & -xu & -xv & -x \\ -yu & -yv & -y & xu & xv & x & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Observe que, a última linha é uma combinação linear de primeira e da segunda: $L_3 = -xL_1 - yL_2$. Assim, temos que resolver um sistema de 2 equações com 9 incógnitas. $Ah = 0$, onde:

$$A = \begin{bmatrix} 0 & 0 & 0 & -u & -v & -1 & yu & yv & y \\ u & v & 1 & 0 & 0 & 0 & -xu & -xv & -x \end{bmatrix} \text{ e } \mathbf{h}^T = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9)$$

Para determinar os valores das entradas da matriz H , escolhemos 4 pontos diferentes em cada uma das duas imagens tal que $H(u_i, v_i, 1) = (x_i, y_i, 1)$, com $i = 1, \dots, 4$. Além disso, pode-se procurar o valor de h_9 como 1, para tanto, considera-se o vetor $\mathbf{b}^T = (0, 0, 0, 0, 0, 0, 0, 0, 1)$. Assim, devemos resolver o sistema:

$$\begin{bmatrix} 0 & 0 & 0 & -u_1 & -v_1 & -1 & y_1 u_1 & y_1 v_1 & y_1 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -x_1 u_1 & -x_1 v_1 & -x_1 \\ \vdots & \vdots \\ 0 & 0 & 0 & -u_4 & -v_4 & -1 & y_4 u_4 & y_4 v_4 & y_4 \\ u_4 & v_4 & 1 & 0 & 0 & 0 & -x_4 u_4 & -x_4 v_4 & -x_4 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \\ h_6 \\ h_7 \\ h_8 \\ h_9 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

O sistema em (1), pode-se escrever como $M\mathbf{h} = \mathbf{b}$. Logo,

$$M\mathbf{h} = \mathbf{b} \quad \text{se, e somente se,} \quad \mathbf{h} = (M^T M)^{-1} M^T \mathbf{b}$$

Assim, para obter a solução deste sistema, pode-se usar a pseudo-inversa de Moore-Penrose, este conceito é estudado no livro de Watkins (2004).

4. RESULTADOS E DISCUSSÕES

Nesta seção, apresenta-se vários exemplos, nos quais se utiliza o código implementado. Nos dois primeiros exemplos se utilizam homografias conhecidas para testar o programa implementado. Nos dois últimos exemplos, procuram-se pontos em cada imagem para determinar a homografia que transforma a primeira imagem na segunda. Em particular, no último exemplo são procurados pontos tal que o resultado seja um mosaico de imagens.

Para resolver o sistema linear (1) foi escrito na linguagem Octave um programa que determina uma homografia dados 4 pontos em cada imagem. Depois de obter uma transformação geométrica existe um segundo passo que corresponde a uma operação de interpolação. Visto que uma vez aplicada uma transformação geométrica, é necessário obter os valores de intensidade associados à transformação realizada, para que a imagem original apareça transformada geometricamente, mas com os valores de intensidade obtidos das correspondentes na imagem original. É nesta etapa que é gerada a modificação dos valores de intensidade da imagem original. Essa operação de interpolação também foi implementada no Octave. Para conhecer mais acerca de interpolação em imagens veja Manning e Dyer (1999).

Inicialmente, apresenta-se homografias quando é conhecida a matriz H , a qual representa uma homografia $h : \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$. Para tanto, considera-se as seguintes matrizes H_1 e H_2 , dadas

respectivamente por:

$$H_1 = \begin{bmatrix} \cos(\pi/6) & 0 & \text{sen}(\pi/6) \\ 0 & 1 & 0 \\ \text{sen}(\pi/6) & 0 & \cos(\pi/6) \end{bmatrix} \quad \text{e} \quad H_2 = \begin{bmatrix} 0,5 & 0 & 64 \\ 0 & 0,5 & 64 \\ 0 & 0 & 1 \end{bmatrix}$$



Figura 3 – Figura original.
(Fonte: <https://github.com/antimatter15/cameraman>)

Na Figura 3, apresenta-se a imagem original que será rotada no plano xz um ângulo de $\pi/6$, na Figura 4 está o resultado e a matriz H_1 é utilizada para isto.



Figura 4 – Figura 3 foi rotada $\pi/6$ radianos no plano xz por meio da matriz H_1 .
(Fonte: Elaboração dos Autores)

A matriz H_2 é utilizada para reduzir duas vezes a Figura 3 e o resultado é apresentado na Figura 5



Figura 5 – Figura 3 foi reduzida duas vezes desde o centro da imagem usando a matriz H_2 .
(Fonte: Elaboração dos Autores)

A seguir, apresenta-se o resultado de transformar imagens a partir da escolha de 4 pontos em cada imagem. Dada uma imagem, procura-se uma homografia que sirva para remover as distorções.

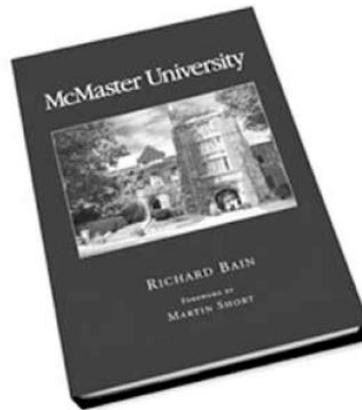


Figura 6 – Figura que será modificada.
(Fonte: https://www.mcmaster.ca/ua/alumni/services_coffeetable.html.)

Na imagem da Figura 6 foram escolhidos os seguintes pontos (59, 286), (16, 41), (189, 4) e (274, 231). Para obter o resultado na Figura 7 escolhemos os pontos (64, 312), (64, 70), (243, 70) e (243, 312). Assim, com esses 4 pontos em cada imagem é possível obter a representação matricial:

$$H_3 = \begin{bmatrix} 0,98883 & -0,1119 & 50,31 \\ 0,28347 & 1,2957 & 14,58 \\ -0,0001 & 0,0001 & 1 \end{bmatrix}$$

Observe que na Figura 6 as bordas do livro não determinam um retângulo. A matriz H_3 serve para encontrar esse retângulo formado pelas bordas do livro. O resultado é apresentado na Figura 7.

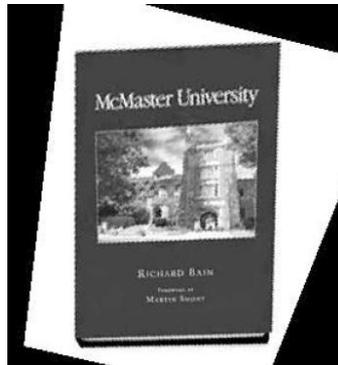


Figura 7 – Resultado de aplicar uma homografia na Figura 6
(Fonte: Elaboração dos Autores)

Finalmente, na Figura 10, apresenta-se um mosaico de imagens que é obtido utilizando as Figura 8 e Figura 9.



Figura 8 – Imagem que serve para formar um mosaico
(Fonte: <http://bit.ly/2QILVyD>)

A matriz H_4 é obtida a partir da escolha de 4 pontos tanto na imagem da Figura 8 como na imagem da Figura 9 e resolver o sistema linear 1.



Figura 9 – Imagem para transformar e formar um mosaico.
(Fonte: <http://bit.ly/2QILVyD>)

Na Figura 8 foram escolhidos os pontos (164, 172), (164, 71), (245, 79) e (245, 173) e na Figura 9 os seguintes pontos (5, 175), (5, 63), (91, 75) e (91, 172). Com esses 4 pontos em cada imagem, obtém-se a matriz:

$$H_4 = \begin{bmatrix} 6.8586e - 01 & 3.4216e - 12 & 1.5973e + 12 \\ -1.1363e - 01 & 8.9716e - 01 & 1.4683e + 01 \\ -1.0252e - 03 & 3.1589e - 14 & 1.0000e + 00 \end{bmatrix}$$



Figura 10 – Mosaico de Imagens
(Fonte: Elaboração dos Autores)

5. CONSIDERAÇÕES FINAIS

Uma vez conhecida a representação da matriz H nos diferentes exemplos deste trabalho, é possível constatar que a aplicação das diferentes transformações geométricas no plano projetivo RP^2 em figuras resultou em outras figuras com a característica desejada, isto mostra que o código implementado no Octave, explicado na pg.4, para determinar a homografia de duas imagens funcionou. Portanto, determinar o mosaico apresentado na Figura 10 foi uma consequência de uma escolha adequada dos 4 pontos, tanto na imagem da Figura 8 como na imagem da Figura 9. Assim, o objetivo proposto para este trabalho foi alcançado.

Deve-se ressaltar que esta é uma aplicação, havendo ainda outras aplicações da geometria projetiva no contexto da visão computacional, por exemplo: a identificação de objetos e inclusive a implementação automática que defina um mosaico de imagens. Estas aplicações e algumas outras serão estudadas e apresentadas em futuros trabalhos desenvolvidos pela equipe do Projeto de Pesquisa que desenvolveu este trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- FAUGERAS, O.; LUONG, Q.; PAPADOPOULOU, T. **The Geometry of Multiple Images**: the laws that govern the formation of multiple images of a scene and some of their applications. London: MIT Press, 2001. (Mit Press).
- GOMES, J.; VELHO, L. **Fundamentos da computação gráfica**. Rio de Janeiro: IMPA, 2004. (Série de computação de matemática).
- HARTLEY, R.; ZISSERMAN, A. **Multiple view geometry in computer vision**. United Kingdom: Cambridge University Press, 2004.
- MANNING, R. A.; DYER, C. R. Interpolating view and scene motion by dynamic view morphing. In: IEEE. **Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)**. [S.l.], 1999. v. 1, p. 388–394.
- WATKINS, D. S. **Fundamentals of Matrix Computations**. New York: Wiley, 2004. (Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts).
- ZAGROUBA, E.; BARHOUMI, W.; AMRI, S. An efficient image-mosaicing method based on multifeature matching. **Machine Vision and Applications**, Springer, v. 20, n. 3, p. 139–162, 2009.
- ZITOVA, B.; FLUSSER, J. Image registration methods: a survey. **Image and vision computing**, Elsevier, v. 21, n. 11, p. 977–1000, 2003.

AGRADECIMENTOS

Ao Instituto de Engenharia do Araguaia - IEA por propor este evento e incentivar a produção intelectual na região araguiense. À Pró-Reitoria de Pós-Graduação, Pesquisa e Inovação Tecnológica - PROPIT da UNIFESSPA por incentivarem a pesquisa na nossa instituição. E a professora Dra. Cecília Orellana Castro por ser prestativa e que contribuiu grandemente na caminhada da pesquisa em Matemática Aplicada no IEA.