

Universidade Federal de Juiz de Fora
Relatório Técnico do Departamento de Ciência da Computação

RelaTeDCC 002/2023

**Criação de imagem panorâmica utilizando algoritmos
SIFT e RANSAC**

Bruno Martins Bartolomeu

Prof. Dr. Luiz Maurílio da Silva Maciel

Agosto 2023

Trabalho Prático 01

1. Introdução
2. SIFT e RANSAC
3. Estratégias de implementação
4. Descrição dos experimentos
5. Conclusão
6. Referências

RESUMO

A visão computacional cresceu bastante nas últimas décadas, juntamente com o poder de processamento dos computadores atuais. Neste âmbito, vários sistemas de reconhecimento de objetos e algoritmos detectores de características foram propostos. Neste relatório, apresentamos um resumo sobre os resultados obtidos em um trabalho que tinha como objetivo a construção de uma imagem panorâmica.

O algoritmo SIFT foi utilizado para encontrar a correspondência de pontos-chave entre as imagens, enquanto o RANSAC foi utilizado para encontrar a matriz de homografia entre os pares correspondentes e reduzir a quantidade de falsas correspondências entre os pontos-chave das imagens. Uma descrição resumida de ambos os algoritmos, SIFT e RANSAC, é também fornecida neste artigo.

Palavras-Chave: Visão computacional. Reconhecimento de Objetos. Detecção de características. SIFT. RANSAC. Homografia. Imagem panorâmica.

ABSTRACT

Computer vision has grown a lot in recent decades, along with the power of current computer processing. In this context, several recognition systems of objects and feature detection algorithms were proposed. In this report, we present a summary of the results obtained in a work that had as The objective is to build a panoramic image.

The SIFT algorithm was used to find the correspondence of key points between the images, while RANSAC was used to find the homography matrix between the matching pairs and reduce the amount of false matches between the key points of the images. A brief description of both algorithms, SIFT and RANSAC, is also provided in this article.

Keywords: Computer vision. Object Recognition. Feature detection. SIFT. RANSAC. Homography. Panoramic image.

1 Introdução

A correspondência de imagens é fundamental em diversas situações da Visão Computacional como reconhecimento de objetos, reconhecimento de cenas, montagem automática de mosaicos, obtenção da estrutura 3D de múltiplas imagens, correspondência estéreo e perseguição de movimentos. Uma abordagem para se trabalhar com correspondência de imagens é se usar descritores locais para se representar uma imagem.

Descritores são vetores de características de uma imagem ou de determinadas regiões de uma imagem e podem ser usados para se comparar regiões em imagens diferentes. Este vetor de características é normalmente formado por descritores locais ou globais. Descritores locais computados em pontos de interesse provaram ser bem sucedidos em aplicações como correspondência e reconhecimento de imagens.

Descritores são distintos, robustos à oclusão e não requerem segmentação. Existem diversas técnicas para se descrever regiões locais em uma imagem. O mais simples descritor é um vetor com as intensidades dos pixels da imagem. A medida de correlação cruzada pode ser então usada para computar a similaridade entre duas regiões. Porém, a alta dimensionalidade de tal descritor aumenta a complexidade computacional da comparação. Então, esta técnica é principalmente usada para se encontrar correspondências ponto a ponto entre duas imagens. A vizinhança de um ponto também pode ser escalada de modo a reduzir sua dimensão. Outro descritor simples é a distribuição de intensidades de uma região representada por seu histograma. Trabalhos recentes na literatura têm se concentrado em fazer descritores invariáveis a transformações nas imagens.

2 SIFT e RANSAC

SIFT (Scale-Invariant Feature Transform) é uma técnica poderosa para correspondência de imagens que pode identificar e combinar recursos em imagens que são invariantes para dimensionamento, rotação e distorção afim. É amplamente utilizado em aplicações de visão computacional, incluindo correspondência de imagens, reconhecimento de objetos e reconstrução 3D. Descritores obtidos com a técnica **SIFT** são altamente distintos, ou seja, um determinado ponto pode ser corretamente encontrado com alta probabilidade em um banco de dados extenso com descritores para diversas imagens. Um aspecto importante da técnica **SIFT** é a geração de um número grande de descritores que conseguem cobrir densamente uma imagem quanto a escalas e localização.

A quantidade de descritores é particularmente importante para o reconhecimento de objeto, onde a capacidade de se encontrar pequenos objetos em ambientes desordenados requer ao menos 3 pontos encontrados em comum para uma identificação confiável.

RANSAC (RANdom SAmple Consensus) é um algoritmo iterativo para a estimativa robusta de parâmetros a partir de um subconjunto de inliers do conjunto completo de dados. Ele é geralmente usado em visão computacional, como por exemplo, para resolver simultaneamente o problema de correspondência entre pontos de duas imagens e estimar a matriz fundamental relacionada ao par de imagens estéreo.

Uma vantagem do **RANSAC** é a sua habilidade de realizar a estimativa de parâmetros de um modelo de forma robusta, ou seja, ele pode estimar parâmetros com um alto grau de acerto mesmo quando um número significativo de “outliers” esteja presente nos dados analisados. Uma desvantagem do algoritmo é que não há um limite superior de tempo para que ele possa computar tais parâmetros. Quando um limite superior é usado (número máximo de iterações) a solução obtida pode não ser a melhor existente.

Os métodos clássicos procuram utilizar o maior número de pontos para obter uma solução inicial e, então, eliminar os pontos inválidos. O **RANSAC**, ao contrário desses métodos, utiliza apenas o número mínimo e suficiente de pontos necessários para uma primeira estimativa, aumentando o conjunto com novos pontos consistentes sempre que possível.

3 Estratégias de implementação

- Detecção de extremos: Nesta primeira etapa é feita procura para todas as escalas e localizações de uma imagem. Isto é feito utilizando-se a diferença de filtros gaussianos de modo a se identificar pontos de interesse invariáveis à escala e rotação.

- Localização de pontos chave: Para cada localização em que foi detectado um extremo, um modelo detalhado é ajustado de modo a se determinar localização e escala. Pontos chaves, ou pontos de interesse, são então selecionados baseando-se em medidas de estabilidade.

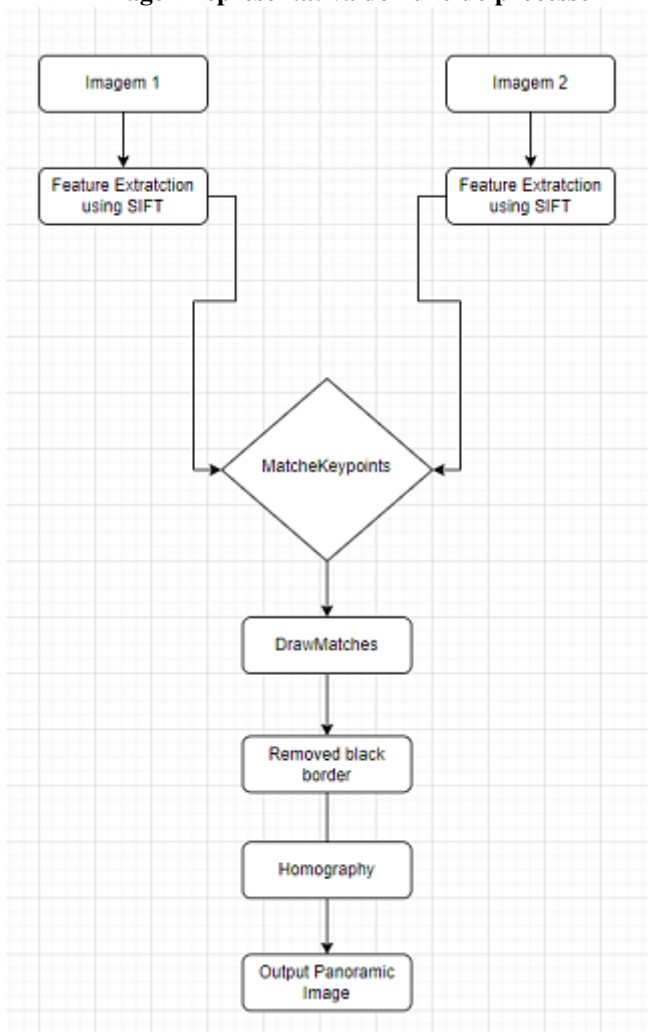
- Definição de orientação: É definida a orientação de cada ponto chave através dos gradientes locais da imagem. Toda operação a partir de então será feita com relação a dados da imagem transformados em relação à orientação, escala e localização de cada ponto chave. Desta maneira se obtém invariância a estas transformações.

- Descritor dos pontos chaves: Nesta etapa é feita a construção dos descritores ao se medir Gradientes locais em uma região vizinha a cada ponto de interesse.

- Implantação de homografia: são estimadas entre imagens para detectar características correspondentes, nesse projeto, foi utilizado o algoritmo RANSAC porém existem outros como por exemplo o DLT.

- Imagem blending/Output: Fase final do projeto e resultado com a imagem em formato panorâmico.

Imagem representativa do fluxo do processo



4 Descrição dos experimentos

Segue abaixo as principais descrições relacionados ao processo.

- A quantidade de correspondências utilizadas:

- Características da imagem que teve sucesso no processo
- Left img size (397 * 529)
- Right img size (397 * 529)
- Right img levemente com borda escura

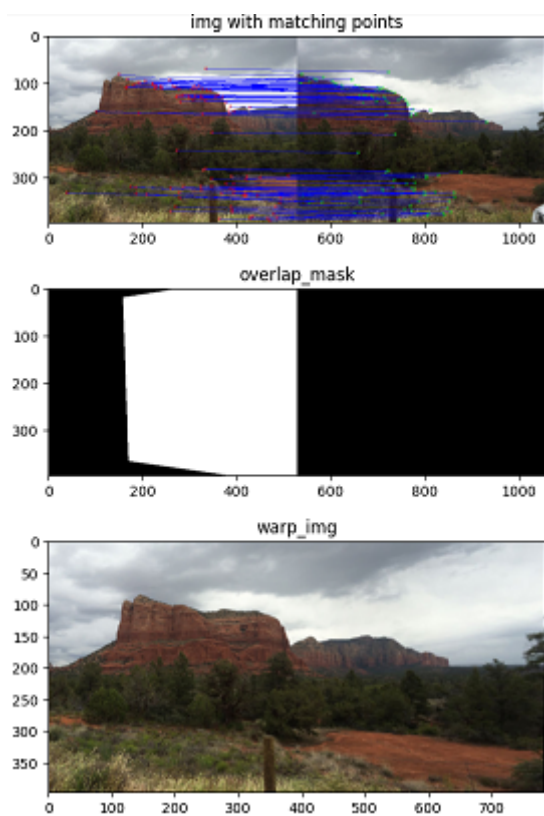
- Quantidade de imagens testadas e de onde foram obtidas:

- Foram testados 3 pares de imagens e obtidas diretamente do Google
- 1 imagem com total sucesso e outras com erros devido a classe Blender

Os resultados obtidos para as métricas avaliadas

- The number of matching points: 106
- O percentual de inliers/outliers ao final do RANSAC
- The Number of Maximum Inlier: 97

Imagem representativa do resultado



5 Conclusão

O projeto teve como objetivo criar uma imagem panorâmica através de análise de um conjunto de keypoints entre duas imagens utilizando algoritmos já conhecidos da área. Segue abaixo 2 pontos essenciais para o desenvolvimento:

- Detecção de características – Utilização do algoritmo SIFT para a extração dos descritores de características da imagem.

- Correspondência entre essas características – Utilização do RANSAC para descarte dos “outliers”, ou seja, dos falsos “matches”.

Pode-se observar também que quanto maior for a imagem mais tempo o algoritmo leva para trazer as informações e montar a panorâmica.

De acordo com os resultados obtidos, verifica-se a possibilidade destas ferramentas funcionarem de maneira satisfatória, com geração de resultados confiáveis e de maneira automática, auxiliando com rapidez no propósito.

A ideia também pode ser modificada para atender outras necessidades, necessitando para isso somente adaptações para a nova aplicação em questão, baseada na análise das características que se deseja analisar.

Link para repositório: https://github.com/Bmartins25/Projeto01_UFJF_SIFT_RANSAC

6 Referências

[1] - **Feature Matching and RANSAC** - https://people.cs.umass.edu/~elm/Teaching/ppt/370/370_10_RANSAC.pptx.pdf

[2] - **Comparação de Algoritmos para Detecção de Pontos de Interesse** - <https://www.lume.ufrgs.br/bitstream/handle/10183/236205/001137526.pdf?sequence=1>

[3] - **Computer Vision: Algorithms and Applications, 2nd ed** - <https://szeliski.org/Book/?authuser=0>

[4] - **Mosaico de imagens com OpenCV e python** - <https://dev.to/tassi/mosaico-de-im>

[5] - **Homography Estimation** - https://cseweb.ucsd.edu/classes/wi07/cse252a/homography_estimation/homography_estimation.pdf