

Universidade de São Paulo
Instituto de Matemática e Estatística
Bachalerado em Ciência da Computação

Pedro Ferreira Alexandre

**Segmentação de imagens pela
Transformada Imagem-Floresta
em grafos de superpixels**

São Paulo
Fevereiro de 2017

Segmentação de imagens pela Transformada Imagem-Floresta em grafos de superpixels

Monografia final da disciplina
MAC0499 – Trabalho de Formatura Supervisionado.

Supervisor: Prof. Dr. Paulo A. V. de Miranda

São Paulo
Fevereiro de 2017

Resumo

A técnica de superpixels consiste em agrupar pixels similares de uma imagem em regiões compactas e conexas. As regiões idealmente devem preservar as bordas de transição entre os diferentes objetos presentes na imagem. Essa técnica pode diminuir a complexidade de tarefas de visão computacional e processamento de imagens, pois no nível de superpixels temos um número de elementos da imagem significativamente menor, consequentemente elevando a eficiência. A Transformada Imagem-Floresta (Image Foresting Transform - IFT) [Falcão *et al.* \(2004\)](#) é um algoritmo que calcula uma floresta de caminhos ótimos em um grafo derivado da imagem. A segmentação interativa de objetos via IFT é obtida através da seleção de pixels sementes dentro e fora do objeto, sendo o objeto segmentado definido pela união das árvores enraizadas em suas sementes internas. No artigo [Rauber *et al.* \(2013\)](#) foi proposto o uso da IFT sobre um grafo de regiões (superpixels) obtido via algoritmo de watershed. Segundo os autores, os resultados da segmentação via IFT no nível do grafo de regiões gerou melhores resultados quando comparado a IFT em um grafo no nível dos pixels. No entanto, existem algoritmos mais adequados para a geração de superpixels que o uso das bacias do watershed, tais como o SLIC [Achanta *et al.* \(2012\)](#). Esse trabalho visa avaliar o impacto do uso de um grafo de regiões obtido pelo método SLIC na segmentação via IFT. Os experimentos serão conduzidos usando sementes obtidas automaticamente de um gabarito de segmentação (ground truth), via técnica de erosão, gerando uma região de incerteza a ser segmentada entorno da borda.

Palavras-chave: superpixels, processamento de imagens.

Abstract

The superpixels technique is to group similar pixels of an image in compact and related regions. The regions should ideally preserve the transition edges between different objects in the image. This technique can reduce the complexity of tasks of computer vision and image processing, because superpixels have a significantly lower number of picture elements, thus increasing efficiency. The Transformed Image-Forest (Image foresting Transform - IFT) [Falcão *et al.* \(2004\)](#) is an algorithm that calculates a forest of optimal paths in a graph derived from the image. Interactive segmentation via IFT object is achieved through the selection of seed pixels inside and outside the object, and the segmented object defined by the union of trees rooted in its internal seeds. Article [Raubert *et al.* \(2013\)](#) has proposed the use of IFT regions on a graph (superpixels) obtained via watershed algorithm. According to the authors, the results of segmentation via IFT in regions graph level generated better results when compared to IFT in a graph at the level of pixels. However, there are more suitable algorithms for generating superpixels than the use of watershed basins, such as the SLIC [Achanta *et al.* \(2012\)](#). This study aims to evaluate the impact of using a graph of regions obtained by SLIC method when segmented via IFT. The experiments will be conducted using seeds obtained automatically from a target template (ground truth), via erosion technique, generating an uncertainty region to be targeted around the edge.

Keywords: superpixels, image processing.

Sumário

1	Introdução	1
1.1	Descrição do problema	1
1.1.1	Agrupamento de pixels	1
1.2	Objetivos	2
1.3	Organização do trabalho	3
2	Conceitos e ferramentas	5
2.1	Algoritmos	5
2.1.1	SLIC	5
2.1.2	IFT	7
2.2	Geração de sementes para a IFT	7
2.2.1	Erosão	7
2.2.2	Geração de sementes	11
2.3	Coeficiente de Dice	12
3	Experimentos	13
3.1	Descrição	13
3.1.1	Programas	13
3.1.2	IFT sobre pixels	13
3.1.3	IFT superpixels por imagem	14
3.1.4	IFT com 400 superpixels por imagem	14
3.1.5	IFT com 900 superpixels por imagem	14
A	Título do apêndice	17
	Referências Bibliográficas	19

Capítulo 1

Introdução

1.1 Descrição do problema

Quando trabalhamos com imagens nem sempre imaginamos quantos cálculos são necessários para realizar pequenas alterações. A maioria das fotos são compostas por mais de 1 milhão de pixels, e cada operação que envolva essa imagem, se não for devidamente otimizada, pode ter que ser repetida milhões de vezes.

No entanto, uma imagem, geralmente, contém diversas áreas redundantes, sendo assim os mesmos cálculos podem estar sendo feitos repetidamente, e portanto são desperdício de tempo e computação.

1.1.1 Agrupamento de pixels

Para tentar solucionar isso poderíamos dividir a imagem, agrupando os pixels em regiões, e então economizaríamos operações aplicadas a pixels que estivessem na mesma região.

Abordagem trivial

Uma forma simples de representar as regiões seria através de linhas paralelas às bordas da imagem, como a seguir:

Pode-se notar com essa solução, que uma região específica pode conter um agrupamento de pixels que contenha dois pixels P_1 e P_2 tais que as características destes são muito distintas entre si, e portanto não seria correto tratá-los da mesma forma.

Abordagem elaborada

A técnica de superpixels consiste em agrupar pixels de uma imagem e classificá-los em regiões atômicas significantes, podendo substituir a estrutura rígida dos pixels, sempre



Figura 1.1: *Imagem com áreas redundantes.*

tentando preservar as características desta imagem, como bordas e cores.

Um superpixel é uma representação para uma região redundante da imagem. O uso dessa estrutura pode reduzir consideravelmente a complexidade de processamentos subsequentes que envolvam a imagem. A estrutura de super pixels se tornou parte importante de muitos algoritmos de computação visual, tais como: segmentação, estimativa de modelo corporal, e reconhecimento de objetos.

Pode-se notar a partir da imagem criada pelo SLIC que os contornos da imagem são preservados na maioria dos casos, e os pixels contidos em um superpixel podem receber o mesmo tratamento sem muita perda de informação.

Recentemente, superpixels tem atraído interesse por possibilitar rápida segmentação e uma descrição de características melhor do que janelas de tamanho fixo envolvendo os pixels.

1.2 Objetivos

A proposta deste trabalho é avaliar o desempenho em termos de acurácia para um algoritmo específico aplicado a uma imagem pré-processada e agrupada por regiões, superpixels, comparado com a aplicação do mesmo algoritmo na imagem a nível de pixels.

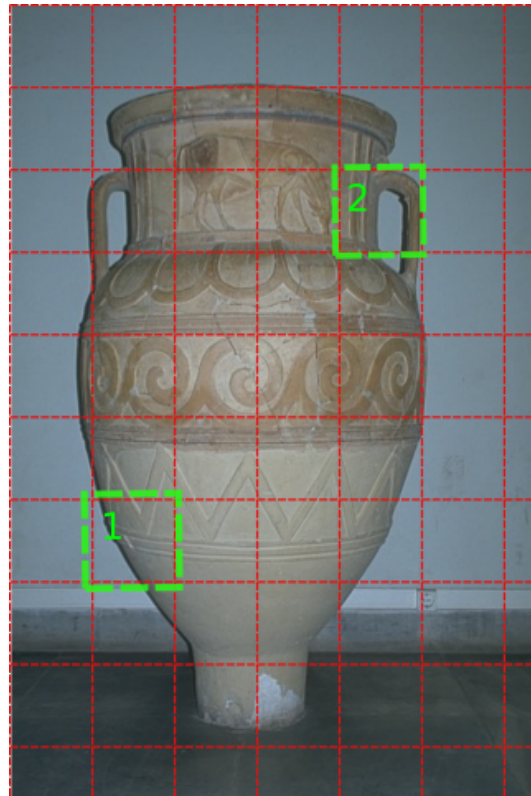


Figura 1.2: *Imagem superpixels em gride.*

O algoritmo aplicado para a comparação dos resultados será a Transformada Imagem-Floresta (IFT - Image Foresting Transform) [Falcão *et al.* \(2004\)](#) . O Algoritmo da IFT será usado para separar objetos do fundo em uma imagem. Para aplicar a IFT sobre os superpixels é preciso realizar um pré-processamento da imagem para a criação dessa estrutura. Existem alguns métodos conhecidos para a criação dos superpixels como:

No artigo [Rauber *et al.* \(2013\)](#) foi realizado um trabalho semelhante, da aplicação da IFT sobre imagens em pixels e superpixels. O estudo realizado por *Rauber* apresenta resultados satisfatórios embora o método de segmentação em superpixels utilizado não seja o mais atual.

Neste trabalho vou verificar o comportamento do algoritmo da IFT ao utilizar o SLIC [Rauber *et al.* \(2013\)](#) , um dos métodos mais atuais para a criação dos superpixels, e comparar os resultados ao desempenho da IFT a nível de pixels.

1.3 Organização do trabalho



Figura 1.3: *Imagem superpixels SLIC.*

Capítulo 2

Conceitos e ferramentas

2.1 Algoritmos

Essa seção explicará em detalhes as ferramentas e alguns conceitos utilizados nesse trabalho.

Primeiramente será explicado o que é o SLIC e algumas diferenças para seus principais concorrentes. Logo em seguida descreveremos em mais detalhes o algoritmo da IFT, que é o algoritmo que será utilizado nos experimentos.

Também será explicado o conceito de erosão de uma imagem e a geração de sementes para o algoritmo da IFT por erosão. Por fim, terá uma apresentação do Coeficiente de Dice, utilizado nesse trabalho para verificação da acurácia do método aplicado.

2.1.1 SLIC

Simple linear iterative clustering [Achanta *et al.* \(2012\)](#) é um método para geração de superpixels baseado em gradientes ascendentes. Algoritmos dessa classe utilizam-se de múltiplas iterações para refinar um conjunto de pixels, até que um critério seja satisfeito, para a formação de um superpixel.

Esse algoritmo é mais rápido, utiliza menos memória, realiza ótima aderência com bordas e fronteiras de objetos na imagem e melhora a performance de algoritmos de segmentação. Simple linear iterative clustering é uma adaptação do k-means para a geração de superpixels, com duas importantes distinções:

- O número de cálculos de distância é drasticamente reduzido ao limitar o espaço de busca para uma região proporcional ao tamanho do superpixel. Isso faz com que a complexidade seja reduzida a linear no número de pixels N e independente do número de superpixels k .

- Uma medida ponderada de distância que combina cor e proximidade espacial, promovendo controle sobre o tamanho e a compacidade dos superpixels.

2.1.2 IFT

A Transformada Imagem-Floresta (IFT - Image Foresting Transform) reduz problemas de processamento de imagem baseados em conexidade ao cálculo de uma floresta de caminhos ótimos no grafo derivado da imagem, seguido de um pós-processamento simples de atributos da floresta resultante, como visto em [Falcão *et al.* \(2004\)](#).

O Algoritmo da IFT pode ser usado para separar objetos do fundo em uma imagem. Nesse caso a IFT realiza uma busca através dos elementos da imagem, a partir de sementes iniciais que podem ser de objeto ou de fundo, por vizinhos similares. No decorrer da busca, elementos são conquistados virando objeto ou fundo e formando o conjunto final, que se trata de uma imagem binária.

2.2 Geração de sementes para a IFT

Nessa seção será discutido o uso de sementes na IFT, e a técnica de erosão de imagens para a geração dessas sementes. Para isso, será discutido o conceito de erosão de uma imagem, técnica utilizada nos experimentos da última seção para geração das sementes.

2.2.1 Erosão

A erosão (usualmente representada por \ominus) e a dilatação são consideradas as duas operações fundamentais do processamento morfológico de imagens, sendo assim, todas as outras operações são baseadas nelas.

Em morfologia binária, uma imagem é vista como um sub-conjunto de um espaço Euclidiano \mathbb{R}^d ou o conjunto de inteiros \mathbb{Z}^d , para alguma dimensão d .

A ideia básica em uma morfologia binária é examinar uma imagem utilizando um objeto simples com formato pré-definido, e decidir se esse objeto se enquadra nos objetos presentes na imagem. Esse objeto de examinação é chamado de elemento estrutural, e é uma imagem também.

Seja E um espaço Euclidiano ou um gride de inteiros, e A uma imagem binária em E . A **erosão** de uma imagem binária A produzida pelo elemento estruturante B é definida por:

$$A \ominus B = \{z \in E | B_z \subseteq A\} \quad (2.1)$$

Onde B_z é a translação de B dada pelo vetor z , i.e.,

$$B_z = \{b + z | b \in B\}, \forall z \in E \quad (2.2)$$

Exemplo:

Dada a imagem binária A (2.1) a seguir e o elemento estruturante B (2.1)

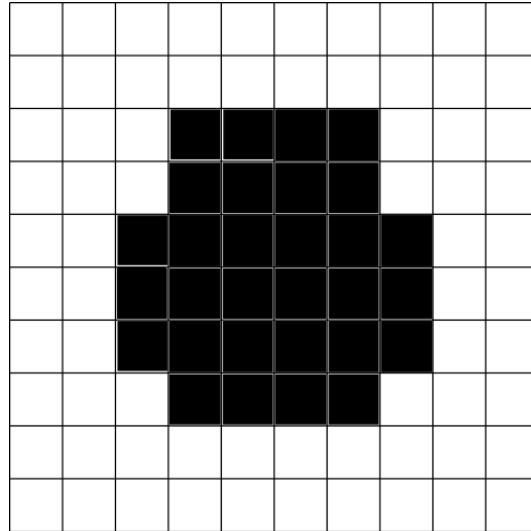


Figura 2.1: A

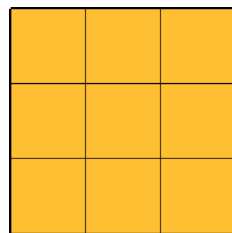


Figura 2.2: B

Aplicando-se o elemento estruturante B em cada elemento da imagem A , encontraremos elementos que não pertencem a $A \ominus B$, por exemplo o elemento centralizado no meio do elemento estruturante da imagem (2.3)

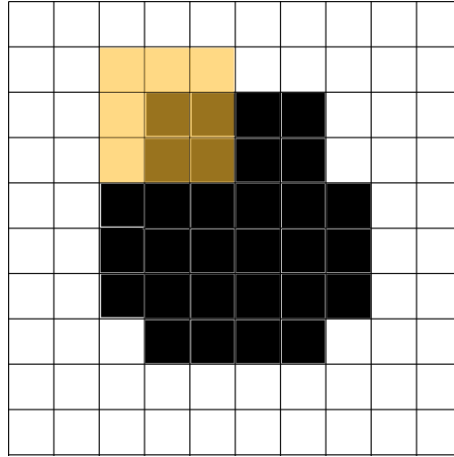


Figura 2.3

E também podem existir elementos que pertencem a $A \ominus B$, por exemplo o elemento centralizado no meio do elemento estruturante, como mostrado na imagem (2.4).

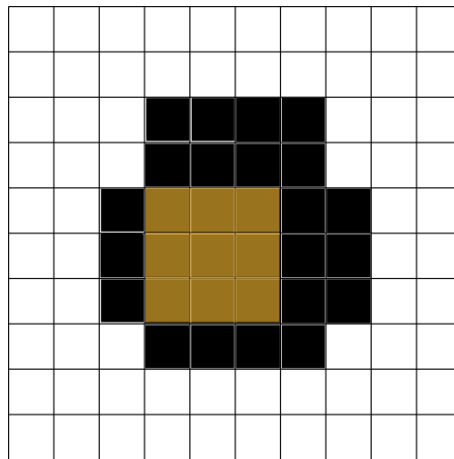


Figura 2.4

O resultado da operação de erosão, $A \ominus B$, será dado pela imagem (2.5).

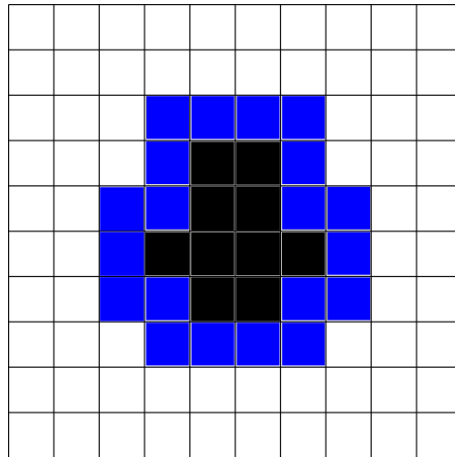


Figura 2.5

2.2.2 Geração de sementes

Sementes na IFT

O algoritmo da IFT é inicializado a partir de sementes do fundo e do objeto. Através dessas sementes o algoritmo busca por regiões da imagem que pertençam ao mesmo objeto que a semente especificada. Por exemplo, se a IFT receber uma semente de fundo em uma iteração, ela vai percorrer os arredores daquela semente se guiando através de regiões semelhantes para classificá-las como também sendo regiões pertencentes ao fundo da imagem.

Sementes por erosão

Uma forma de gerar as sementes é através da erosão da imagem do gabarito (imagem binária que separa o fundo do objeto). Aplicando a erosão no gabarito de uma imagem e utilizando raios diferentes para o elemento estruturante gera-se um conjunto de imagens para serem utilizadas como sementes da IFT. Se for realizado um processo de erosão, partindo do raio 0 até o raio máximo de erosão, serão geradas imagens distorcidas, gradualmente, sendo a imagem de raio zero uma semente extremamente significativa (gabarito sem erosão), até se tornar uma semente bem pouco significativa quando o raio é máximo.

Dada uma imagem, por exemplo



Figura 2.6: *scissors, base de dados pública do grabcut*

Se for aplicada a erosão no respectivo gabarito, será obtida uma semente diferente para os diversos raios de erosão. Na imagem seguinte será apresentada uma imagem variando seu raio de erosão de 0 (imagem sem erosão) até 45.

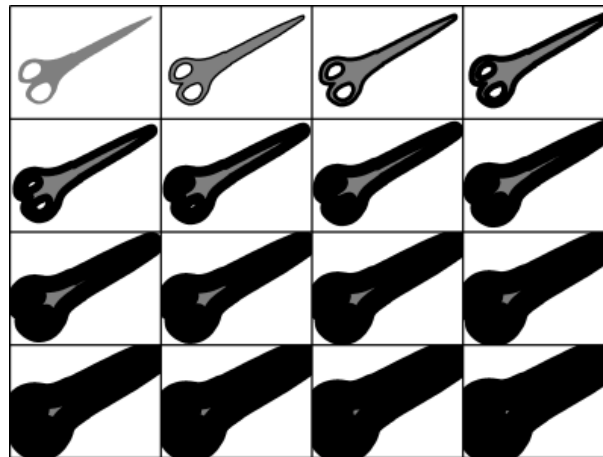


Figura 2.7: *raio 0 até 45, de 5 em 5.*

2.3 Coeficiente de Dice

Esse coeficiente também é conhecido por outros nomes, tais como: Índice de Sorensen ou Coeficiente de Sorensen-Dice. Essa medida é uma estatística usada para comparar a semelhança entre duas amostras. Também é conhecido como F1 score.

A aplicação da fórmula original era destinada para a verificação de presença ou ausência de uma informação específica.

$$CSD = \frac{2|X \cap Y|}{|X| + |Y|}$$

Capítulo 3

Experimentos

3.1 Descrição

Os experimentos consistiram em medir a acurácia da IFT ao separar objeto e fundo de uma imagem. Os experimentos foram rodados para o banco de imagens públicas do grabcut que contém 50 imagens. Foram feitas duas classes de rotinas principais:

- IFT sobre uma imagem bruta, sem pré-processamento em regioes (nível de pixels).
- IFT sobre regioes geradas pelo método SLIC, com 100, 400, 900 e 1600 superpixels.

3.1.1 Programas

Para rodar os experimentos foram utilizadas implementações do SLIC, da IFT e um programa para a produção de sementes com erosão. As bibliotecas utilizadas foram escritas em C++.

- `ift_sp.cpp` implementação da IFT sobre superpixels utilizando uma implementação do SLIC.
- `ift.cpp` implementação da IFT sobre pixels.
- `eroeval.cpp` realiza chamada para `iftsp` e `ift` para varios raios de erosão da imagem.
- Também foi utilizada a biblioteca externa `gft`, onde foi utilizada implementações de erosão de imagem, leitura e escrita de imagem.

3.1.2 IFT sobre pixels

Foi criado uma rotina para rodar o executável do `eroeval` para todas as imagens do grabcut, inicialmente passando como parâmetro a opção da `ift` a nível de pixels. Representado na imagem dos resultados pela curva roxa tracejada, a IFT a nível de pixels obteve bons resultados quando as sementes eram mais significativas, raios menores do que 10 pixels, sendo

a primeira colocada entre as curvas geradas no intervalo de raio 0 a 5 pixels. No entanto, ela piora sua acurácia ao longo da diminuição do raio da erosão.

3.1.3 IFT superpixels por imagem

E em seguida, também executada a mesma rotina, mas trabalhando com a IFT a nível de superpixels, recebendo o número de superpixels a serem gerados pelo SLIC como parâmetro.

100 superpixels

A curva obteve um dos melhores comportamentos, e no início, quando todas as outras curvas de super pixels ficaram mais afastadas da curva da IFT sobre pixels, a curva verde se manteve bem próxima a curva da IFT sobre pixels. No final, assim como todas as curvas de super pixels, se manteve estável e bem melhor do que a curva tradicional, sobre pixels.

400 superpixels

Um pouco pior no início, quando comparado a IFT sobre a imagem com 100 superpixels, no entanto, logo se mantém junto as outras.

900 e 1600 superpixels

Ambas obtiveram desempenho bem similar, e mostram que quando maior o número de superpixels, de acordo com esse experimento, pior foi a acurácia para raios pequenos de erosão, sementes que dão uma boa noção do objeto verdadeiro.

3.1.4 IFT com 400 superpixels por imagem

3.1.5 IFT com 900 superpixels por imagem

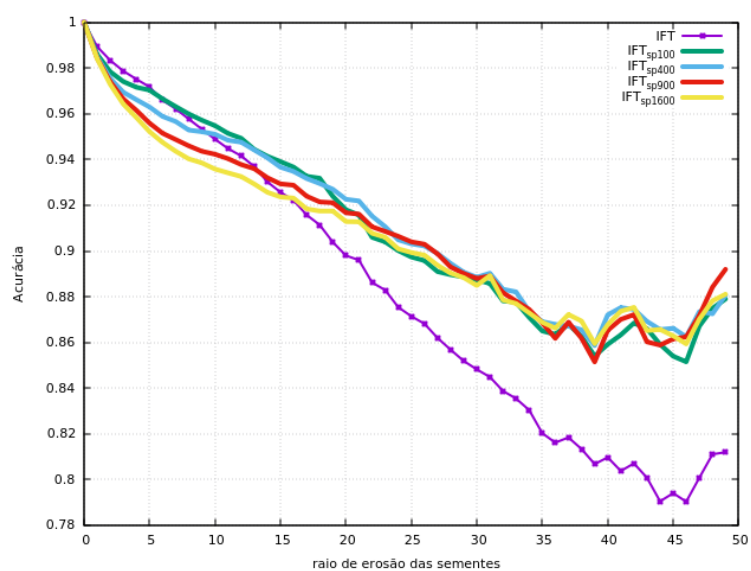


Figura 3.1: *resultados*

Apêndice A

Título do apêndice

Referências Bibliográficas

- Achanta et al.(2012)** Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, Fellow e Sabine Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. 34(11). Citado na pág. [i](#), [iii](#), [5](#)
- Falcão et al.(2004)** Alexandre X. Falcão, Jorge Stolfi e Roberto de Alencar Lotufo. The image foresting transform: Theory, algorithms, and applications. 26(1):19–28. Citado na pág. [i](#), [iii](#), [3](#), [7](#)
- Rauber et al.(2013)** Paulo E. Rauber, Alexandre X. Falcão, Thiago V. Spina e Pedro J. de Rezende. Interactive segmentation by image foresting transform on superpixel graphs. Em *XXVI Conference on Graphics, Patterns and Images*, páginas 131–138. Citado na pág. [i](#), [iii](#), [3](#)