IFT-SLIC: A general framework for superpixel generation based on simple linear iterative clustering and image foresting transform

Eduardo Barreto Alexandre∗, Ananda Shankar Chowdhury†,
Alexandre Xavier Falcão‡, Paulo A. Vechiatto Miranda∗
∗ IME - University of São Paulo (USP)
† Jadavpur University, Kolkata, India
‡ IC - University of Campinas (UNICAMP)

28th Conference on Graphics, Patterns and Images

Introduction

A **superpixel** can be defined as a compact region of similar and connected pixels, which locally represent a same image structure. The desirable properties for superpixel generation methods are:

- **Adhesion to object boundaries in the image**,
- **Flexibility in the number of superpixels it generates**,
- **Efficiency** (to alleviate Computer Vision pipelines overhead, by replacing the rigid structure of the pixel grid).
- **Compactness**.

Related Work

Simple linear iterative clustering (SLIC) adapts a **k-means clustering** approach to efficiently generate superpixels.

- SLIC superpixels correspond to clusters in the **labxy feature space**.
- It has two parameters:
  - **k**: the desired number of approximately equal sized superpixels,
  - **m**: a parameter to offer control over their compactness.

Related Work

Simple linear iterative clustering (SLIC) adapts a **k-means clustering** approach to efficiently generate superpixels.

- SLIC superpixels correspond to clusters in the **labxy feature space**.
- It has two parameters:
  - **k**: the desired number of approximately equally sized superpixels,
  - **m**: a parameter to offer control over their compactness.

Disadvantage:

- It uses the direct distances between pixel and cluster centers,
- similar pixels may not group into one compact region,
- the problem has to be addressed by a post-processing in SLIC.

Technical Paper Contributions

We create a new Image Foresting Transform (IFT) operator that naturally defines **superpixels** as regions of strongly connected pixels:

- It extends a popular algorithm – Simple Linear Iterative Clustering (SLIC) – to consider minimum path costs between pixel and cluster centers rather than their direct distances.
- Non-smooth connectivity functions (NSCF) are also explored in our IFT-SLIC approach leading to improved performance.

Image Foresting Transform (IFT)

An image can be interpreted as a **graph** $G = (\mathcal{I}, \mathcal{A})$:

- The nodes are the image pixels in its image domain $\mathcal{I} \subset Z^n$.
- The arcs are the pixel pairs (s, t) in $\mathcal{A}$ (e.g., 4-neighborhood, or 8-neighborhood, in case of 2D images).

The adjacency relation $\mathcal{A}$ is a binary relation on $\mathcal{I}$. We use $t \in \mathcal{A}(s)$ and $(s, t) \in \mathcal{A}$ to indicate that t is adjacent to s.

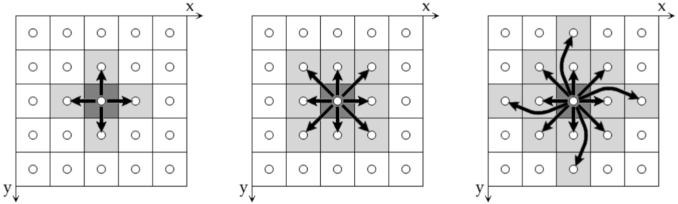


Image Foresting Transform (IFT)

For a given image graph $G = (\mathcal{I}, \mathcal{A})$:

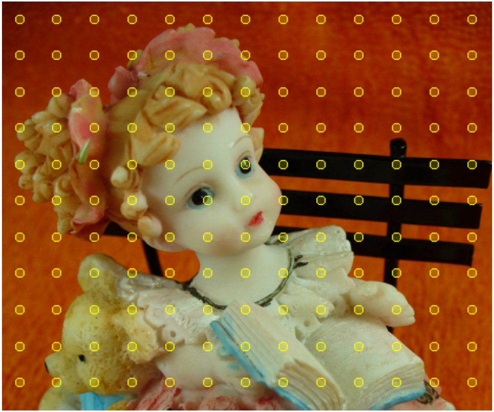
- A path $\pi_t = \langle t_1, t_2, \ldots, t_n = t \rangle$ is a sequence of adjacent pixels with terminus at a pixel t.
- A path is *trivial* when $\pi_t = \langle t \rangle$.
- A path $\pi_t = \pi_s \cdot \langle s, t \rangle$ indicates the extension of a path π_s by an arc (s, t).
- The notation $\pi_{s \rightsquigarrow t} = \langle t_1 = s, t_2, \ldots, t_n = t \rangle$ may also be used, where s stands for the origin and t for the destination node.

Image Foresting Transform (IFT)

- A connectivity function computes a value $f(\pi_t)$ for any path π_t, usually based on arc weights.
- A path π_t is optimum if $f(\pi_t) \leq f(\tau_t)$ for any other path τ_t in G.
- By taking to each pixel $t \in \mathcal{I}$ one optimum path with terminus t, we obtain the optimum-path value $V(t)$, which is uniquely defined by $V(t) = \min_{\forall \pi_t \text{ in } G}\{f(\pi_t)\}$.
- The Image Foresting Transform (IFT) takes an image graph $G = (\mathcal{I}, \mathcal{A})$, and a *smooth* path-value function f; and assigns one optimum path π_t to every pixel $t \in \mathcal{I}$ such that an optimum-path forest P is obtained.

Superpixel Generation by IFT

Similar to SLIC, we start with the same selection of k initial cluster centers $C_i = [l_i \; a_i \; b_i \; x_i \; y_i]^T$, which are sampled on a regular grid spaced $S = \sqrt{N/k}$ pixels apart.



- The main difference with SLIC lies in the assignment step. Instead of using an adaptive k-means clustering approach, we consider the computation of an IFT with the **non-smooth connectivity function** f_D:

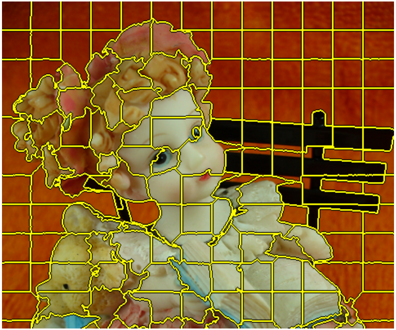
$$
\begin{aligned}
f_D(\pi_t = \langle t \rangle) &= \begin{cases} 0 & \text{if } t \in \mathcal{S} \\ +\infty & \text{otherwise} \end{cases} \\
f_D(\pi_{r \rightsquigarrow s} \cdot \langle s, t \rangle) &= f_D(\pi_s) + (\|I(t) - I_r\| \cdot \alpha)^{\beta} + \|s, t\|
\end{aligned}
$$

where $I(t)$ is the color vector at pixel t, i.e., $I(t) = [l_t \; a_t \; b_t]^T$, and I_r is the color vector of the cluster center of seed r (i.e., $I_r = [l_j \; a_j \; b_j]^T$ where $C_j = [l_j \; a_j \; b_j \; x_j \; y_j]^T$ and r is at the coordinate (x_j, y_j)).

Superpixel Generation by IFT

At the end of the assignment step, each **cluster/superpixel** will be represented by its respective **tree** in the spanning forest (i.e., the predecessor map P) computed by the IFT.

(a) Seeds

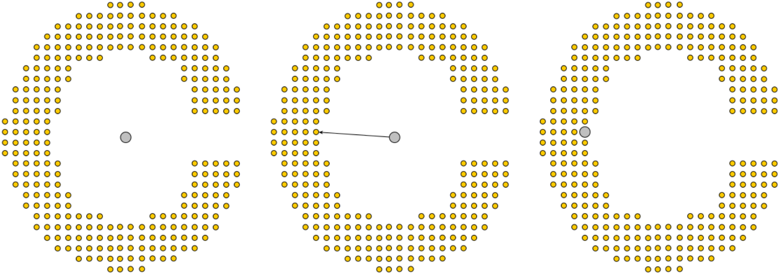


(b) Superpixels

Superpixel Generation by IFT

After that, an update step adjusts the **cluster centers**:

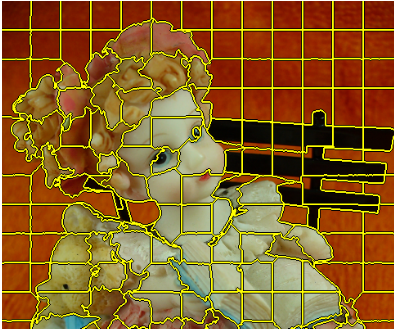
- Differently from SLIC, which considers the mean $[l\ a\ b\ x\ y]^T$ vector of all the pixels belonging to the cluster, we take for the (x, y) the coordinate of the cluster's pixel closest to the mean position **(to avoid an updated position outside its cluster)**.



The assignment and update steps are then repeated for a total of 10 iterations. IFT-SLIC does not require a **post-processing** step as the connectivity is already guaranteed by design.

First iteration:

(a) Seeds

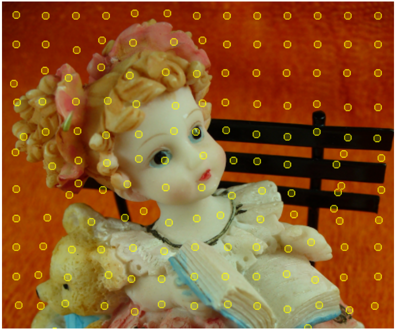


(b) Superpixels

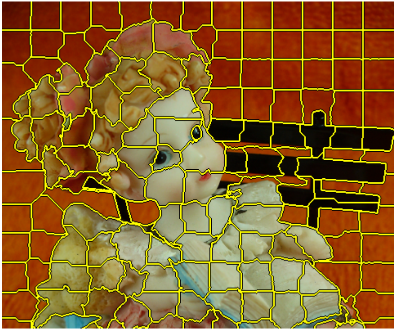
Superpixel Generation by IFT

The assignment and update steps are then repeated for a total of 10 iterations. IFT-SLIC does not require a **post-processing** step as the connectivity is already guaranteed by design.

Second iteration:



(a) Seeds

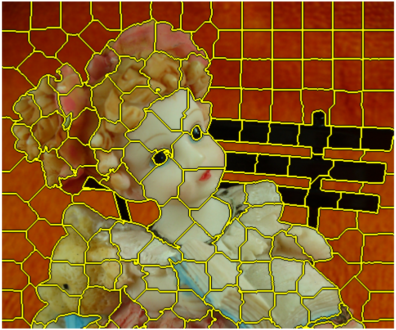


(b) Superpixels

The assignment and update steps are then repeated for a total of 10 iterations. IFT-SLIC does not require a **post-processing** step as the connectivity is already guaranteed by design.

Last iteration (10th):

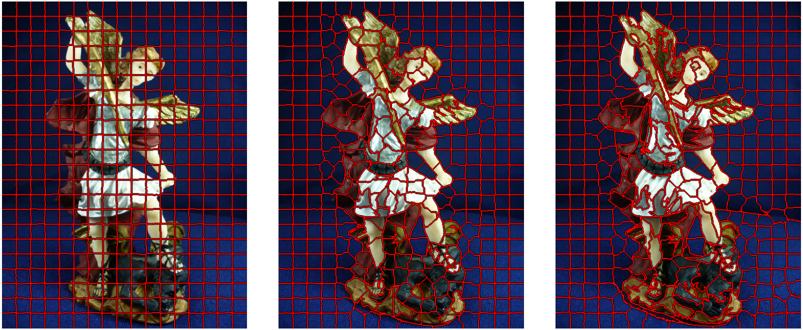
(a) Seeds



(b) Superpixels

Superpixel Generation by IFT

The effects of different values of α on the superpixels by IFT-SLIC. For higher values of α, we have a better adhesion to the image boundaries.



(a) $\alpha = 0.01$ (b) $\alpha = 0.04$ (c) $\alpha = 0.08$

Experiments and Results

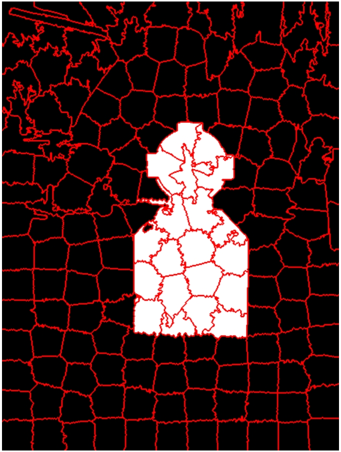
To measure the ability of the methods to adhere to image boundaries, we considered datasets with corresponding ground-truths.

- The superpixels by SLIC and IFT-SLIC are computed, and we assign to each superpixel the most frequent label of the ground truth occurring in its interior.
- The resulting segmentation is then compared to the ground-truth data using the **Dice coefficient**.

Experiments and Results

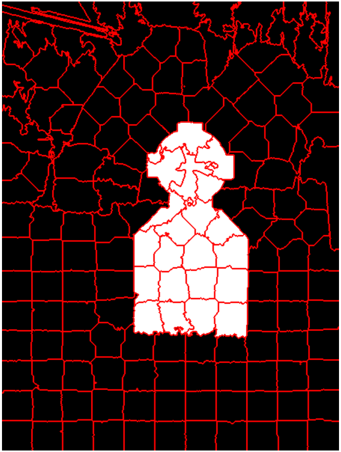
(a) Input image (b) Ground truth

(c) SLIC superpixels



(d) SLIC segmentation

(e) IFT-SLIC superpixels



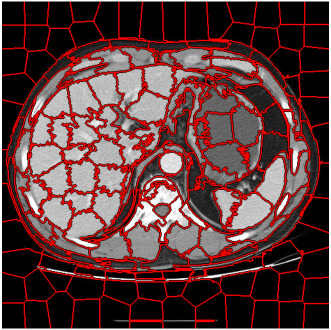
(f) IFT-SLIC segmentation

Experiments and Results

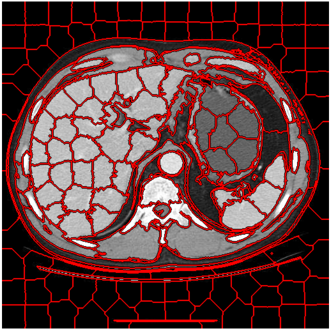
- In the first experiment, we used the test set of 50 natural images of the public **GrabCut dataset**.
- For the second dataset, we conducted quantitative experiments, using a total of 40 image slices of 10 thoracic CT studies to segment the **liver**.
- In the third experiment, we performed the segmentation of the **talus bone**, using 40 slices from MR images of the foot.

Experiments and Results

A liver from a CT abdominal study. Superpixel results by:

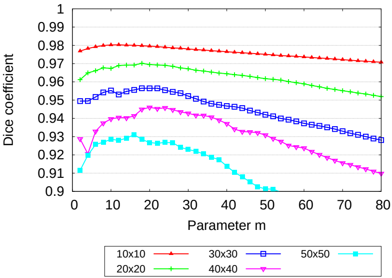


(a) SLIC ($m = 18$)

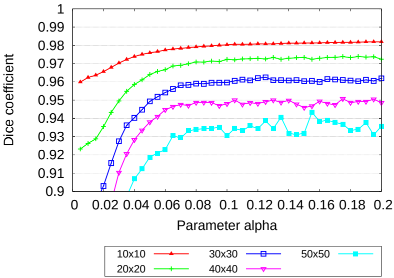


(b) IFT-SLIC ($\alpha = 0.08$)

The mean accuracy curves for segmenting the **GrabCut dataset** for different superpixel sizes.



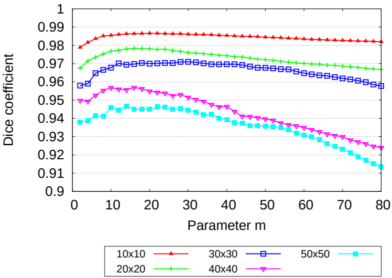
(a) SLIC



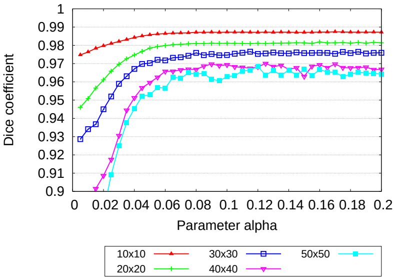
(b) IFT-SLIC

Experiments and Results

The mean accuracy curves for segmenting the **liver dataset** for different superpixel sizes.



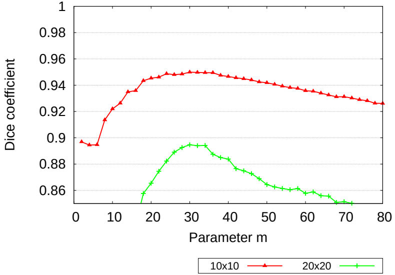
(a) SLIC



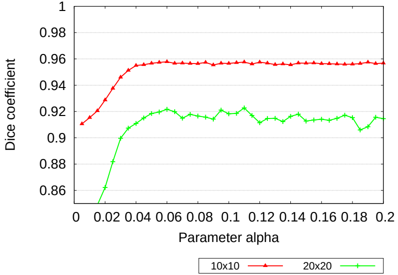
(b) IFT-SLIC

Experiments and Results

The mean accuracy curves for segmenting the **talus dataset** for different superpixel sizes.



(a) SLIC



(b) IFT-SLIC

Conclusion

- Clearly, the accuracy decreases as we increase the superpixel size for both methods, but IFT-SLIC presents a better performance compared to SLIC.
- As future work, we intend to test IFT-SLIC with other path-cost functions and seed selection procedures, to cope with the particularities of a given application.

Acknowledgment

The authors thank CNPq (305381/2012-1, 486083/2013-6, 302970/2014-2, 479070/2013-0, FINEP 1266/13), FAPESP grant # 2011/50761-2, CNPq, CAPES, NAP eScience - PRP - USP, and Dr. J. K. Udupa (MIPG-UPENN) for the images.