# LSC: Superpixel segmentation using linear spectral clustering

CVPR 2015

Paper by: Zhengqin Li, Jiansheng Chen

## REVIEW

## Introduction

Superpixel oversegmentation [2] is a popular technique to represent images using aggregate clusters of pixels which form perceptually consistent regions, such that all the image pixels contained within the superpixel are similar in terms of spatial location, colours, intensity, or any other measure of 'similarity'. A good quality segmentation algorithm should run with low computational complexity, conform to natural boundaries of objects, and use non-local information to group pixels with varying intensities [1]. However, for majority of superpixel segmentation algorithms, there exists a trade-off between inclusions of global (or non-local) information and the computational complexity induced due to such process. Here, I discuss this trade-off in context of two popular methods – NCuts [3] and SLIC [4], as these are relevant to stage the introduction of the paper [1] being reviewed. NCuts (Normalized Cuts) is a graph cut based method which computes superpixel segmentations using similarity among local as well as non-local pixels. However, this process makes NCuts computationally expensive. On the other hand, SLIC uses localized weighted K-means clustering (with the search space of cluster centroids restricted to a local region) which is computationally efficient. However, this approach may fail if there exist high variations among local pixels which need resolution in global context. Such methods may require excessive post-processing for merger of superpixels to preserve perceptual consistency.

In the paper [1] reviewed here, the authors address the trade-off between *capturing the non-local properties of image* and *time complexity of segmentation.* The **main contribution** of the paper is an algorithm which performs superpixel segmentation **using local as well as global relationships** between pixels, while maintaining **linear time complexity**. The method, called **Linear Spectral Clustering (LSC),** is based on the mathematical equivalence between the objective functions of **normalized cuts [3]** (optimal for global information inclusion) in the original pixel space and **weighted K-means** (optimal for low computational complexity) in a *high-dimensional feature space* for pixels. This equivalence between normalized cuts (NCuts) and weighted K-means exists under following conditions [1],

$$w(p)\phi(p) \, . \, w(q)\phi(q) = \boldsymbol{W(p,q)} \quad \forall\, p,q \tag{1}$$

$$w(p) = \textstyle\sum_{q\in V} W(p,q)\,, \ \forall\, p \tag{2}$$

where $V$ is set of all vertices/pixels, $w(p)$ is weight of pixel $p$, and $\phi(p)$ is a high-dimensional feature [1].

Equation (1) implies that the similarity between two pixels $W(p,q)$ should be equal to their weighted inner-product in new feature space $\phi$. Equation (2) implies that the weight $w(p)$ itself should be equal to total connectivity (local and non-local ) of a pixel over entire image.

To satisfy equations (1) and (2), the authors of LSC [1] derive a **10-dimensional sinusoidal feature space** $\phi$. An overview of this algorithm is shown in Figure 1.

## Comments on the Approach

The basis of LSC algorithm, which is the equivalence between objectives of Normalized Cut ($F_{NCUT}$) and Weighted K-means ($F_{km}$), is theoretically sound. The authors prove in the paper that if the conditions mentioned in equations (1) and (2) hold true, then, $\boldsymbol{F_{km} = C - K \times F_{NCUT}}$.

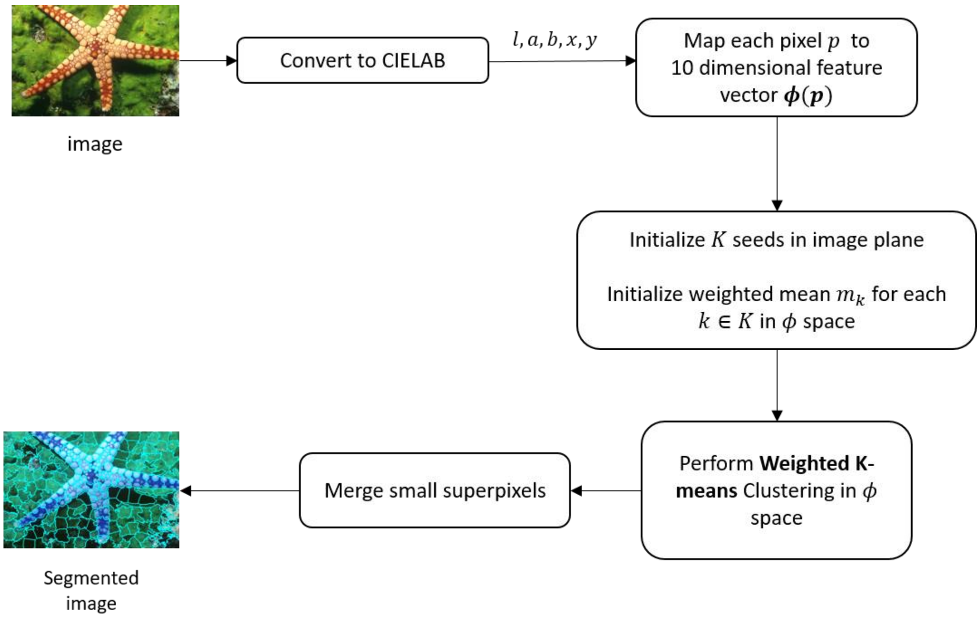


Figure 1: LSC algorithm sketch

Here, $C$ is the weighted sum of *magnitudes* of features vectors corresponding to all pixels and $K$ is the number of clusters. The proof has been omitted from this review due to space constraint, but can be found in [1]. Although authors of LSC are not first to recognize this equivalence [8], they are the first to develop a complete segmentation algorithm based on it. The feature space $\phi(p)$ is also derived in the context of satisfying equations (1) and (2).

**Preserving Global Relationships:** It can be easily shown that,

$F_{km} = C - \sum_{k=1}^{K} \frac{\sum_{p \in \pi_k} \sum_{q \in \pi_k} w(p)\phi(p) .w(q)\phi(q)}{\sum_{p \in \pi_k} w(p)}$ . Therefore, minimizing $F_{km}$ leads to maximizing the sum of $W(p,q)$ for points within a cluster $\pi_k$ as per equation (1). Hence, an optimal weighted k-means clustering in $\phi(p)$ space should provide similar global relations as obtained by $W(p,q)$ similarity measure in Normalized Cut. Thus, LSC theoretically preserves the global information without explicitly building a graph in $\phi$ space. However, authors mention in [1] that they restrict to localized K-means update similar to SLIC [4] in order to maintain "local compactness for superpixels" [1]. The size of this local search space might affect the global inclusion which, according to equation (2), requires search over entire image.

**Complexity**: Authors prove that LSC can be executed in linear time complexity of order $O(N)$ where $N$ is the number of pixels.

## Test and Comparison

In order to test the LSC algorithm, I use the Berkeley Segmentation Database [5] consisting of 300 RGB images with accompanying Ground Truth segments provided by human subjects. The **LSC** [1] algorithm is available via OpenCV XIMGPROC. It is compared with 4 popular superpixel segmentation methods: **SEEDS** [7] using implementation in OpenCV XIMGPROC, **Normalized Cuts (NCuts)** [3] using scikit-image implementation, **SLIC** [4] and **Felzenszwalb (FH)** [6] methods, both using scikit-image segmentation module. The results are shown below. All values are averaged over 300 images. The performance measures are discussed below,

**Undersegmentation Error** (Figure 3): This is the fraction of total number of pixels in segmented image which overlap with multiple regions in the Ground Truth [1], which is basically a measure of pixel leakage. In my implementation, first an overlapping superpixel in segmented image is identified and associated with largest overlapping ground truth segment. Then, pixels leaking out of this ground truth boundary are counted. It can be seen that LSC (blue plot) performs better than other methods by giving least error.

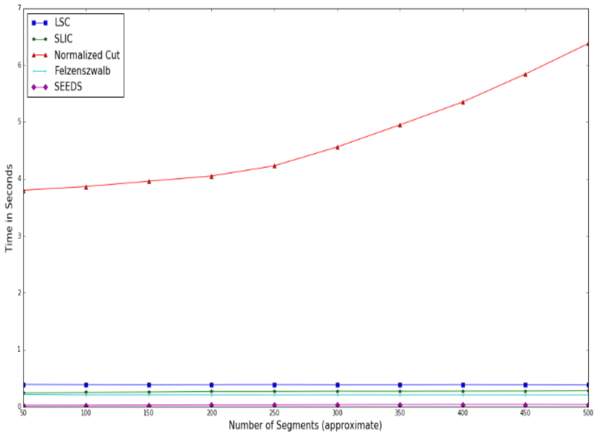


Figure 2: Runtime in seconds.

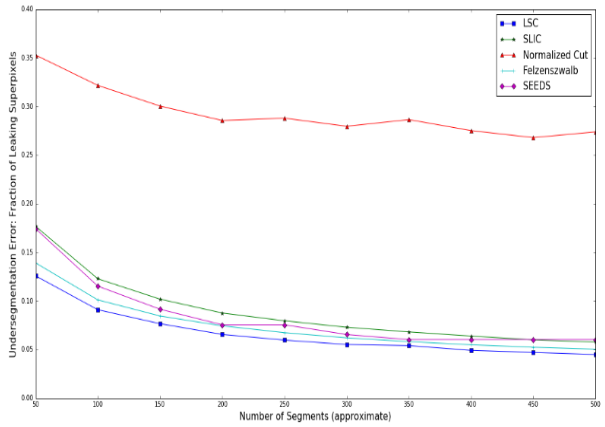


Figure 3: Undersegmentation Error

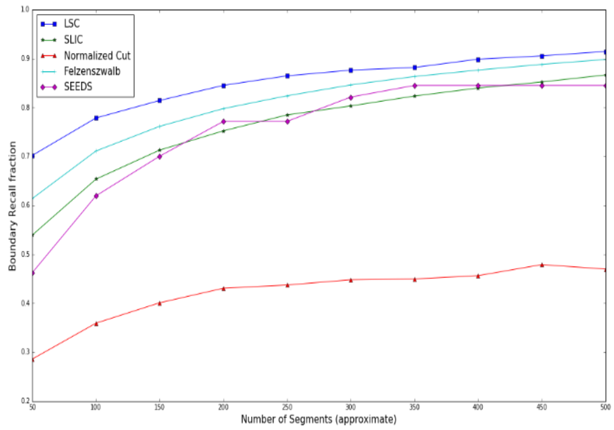


Figure 4: Boundary Recall

**Boundary Recall** (Figure 4)**:** This is the fraction of total number of ground truth boundary pixels which are within a 2-pixel distance from any superpixel boundary in segmented image [1]. Here, LSC provides best result (with highest recall) among all algorithms, implying that the method is able to recover (or conform to) the ground truth boundaries more efficiently.

**Runtime** (Figure 2)**:** While LSC requires more processing time than SEEDS, SLIC, and FH, the true comparison can be seen between LSC and NCuts. As I have discussed under core contribution of the paper [1], LSC tries to optimize a similar objective as NCuts but with higher computational efficiency, which can be observed here as a significantly less runtime of LSC compared to NCuts.

The results are not exact reproduction of the paper [1]. The differences might be due to different implementations used in the original paper and in this review. The code for my implementation can be found at https://github.com/ozpicium/Linear-Spectral-Cluster-Superpixel. The methods are imported from various open-source libraries as mentioned previously. The code to run all methods to together, to calculate various performance measures, and to obtain results is my own.

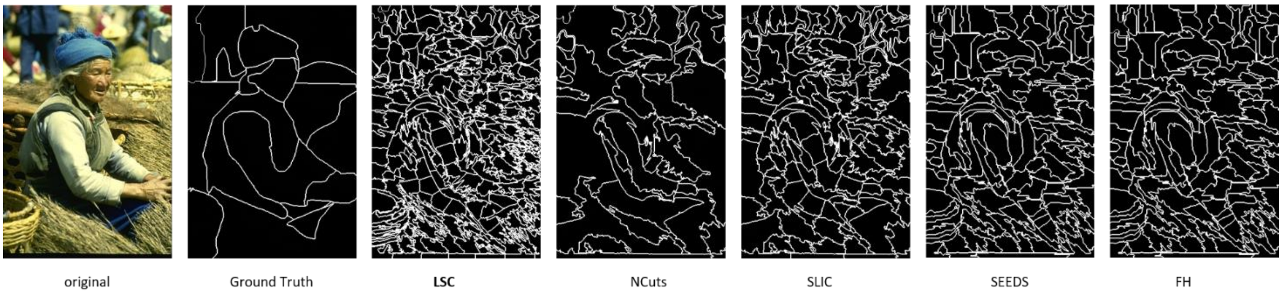


Figure 5: Superpixel results of various algorithms, for segmentation size = 200

## Conclusive Remarks

LSC is a theoretically sound algorithm which aims to solve the problem of globally consistent superpixel segmentation while maintaining low computational complexity. It shows good improvement in terms of *runtime* over standard Normalized Cuts method which segments images using local as well as non-local similarity, but is computationally costly. Moreover, LSC maintains the quality of segmentation in terms of conformity to natural boundaries (ground truth) as seen in Boundary Recall performance (Figure 4).

While the core method itself is mathematically supported, the actual *approximate* implementation of the algorithm in [1] has a *disadvantage* that it performs K-means clustering by restricting the search space of cluster centroids to local neighbourhood of pixels. This might contribute to some runtime improvement and local compactness, but such approximation might provide bad results in special cases where the global context needs to be expanded. In such scenario, normalized cuts with global connectivity graphs might be better. However, this is a minor technical issue, which can be easily addressed by adjusting the search space for K-means. **The authors do provide a parameter $\tau$ to balance local compactness and global optimality.**

## References

[1] Li, Zhengqin, and Jiansheng Chen. "Superpixel segmentation using linear spectral clustering." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.

[2] X. Ren and J. Malik, "Learning a Classification Model for Segmentation," Proc. IEEE Ninth Int'l Conf. Computer Vision, vol. 1, pp. 10-17, 2003.

[3] J. Shi and J. Malik, "Normalized cuts and image segmentation", IEEE PAMI, vol. 22, no. 8, Aug. 2000.

[4] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and Sabine Süsstrunk, SLIC Superpixels Compared to State-of-the-art Superpixel Methods, IEEE T-PAMI, vol. 34, num. 11, p. 2274 - 2282, May 2012.

[5] Martin, David, et al. "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics." *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*. Vol. 2. IEEE, 2001.

[6] P. Felzenszwalb and D. Huttenlocher, "Efficient Graph-Based Image Segmentation," Int'l J. Computer Vision, vol. 59, pp. 167-181, 2004.

[7] Van den Bergh, Michael, et al. "Seeds: Superpixels extracted via energy-driven sampling." *European conference on computer vision*. Springer, Berlin, Heidelberg, 2012.

[8] I. Dhillon, Y. Guan, and B. Kulis. Weighted graph cuts without eigenvectors: a multilevel approach. IEEE Trans. on PAMI, 29(11):1944–1957, 2007.