Geometry-Based Superpixel Segmentation

Marie-Anne Bauda, Sylvie Chambon
Pierre Gurdjos, Vincent Charvillat

mbauda@enseeiht.fr

Problem

Obtaining a semantic segmentation from calibrated multi-view

Difficulties:

1. Low texture difference between two surfaces in natural and artificial scenes
2. Sparse 3D point clouds generally unstructured and unmeaning

Input Data:

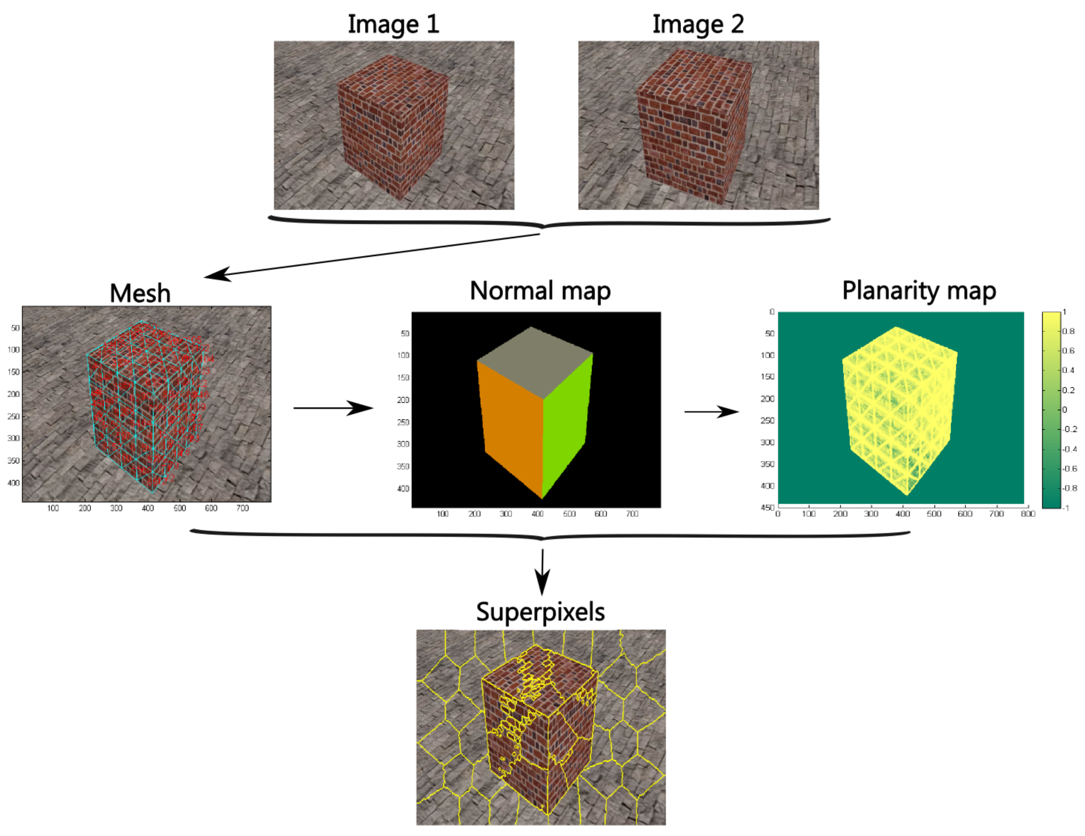
- Two calibrated images
- A sparse 3D point cloud

Overview

Proposition of a **new over-segmentation constructor** in the same way of the well known k-means superpixel approach [1], based on:

1. Geometry extraction
2. Superpixel construction

More precisely, we introduce a new distance that combines **photometric and geometric aspects** in the **aggregation step of the superpixel construction**.



1. Geometry Extraction

Two types of geometric information extracted (see overview)

- **Normal map** provides a surface normal $\vec{n}_i$ to each pixel p_i
- **Planarity map** gives similarity measure, from *Universal Quality Index* (UQI) [4] between a reference image and a warped adjacent image estimated through the homography induced by the plane of support [2, 3]

2. Superpixel Construction

Introduction of geometric information in the approach *Simple Linear Iterative Clustering*, (SLIC) [1]. Originally, it provides compact and uniform in size superpixels, based on a 5D k-means clustering (3D for colors and 2D for space), by following these three steps:

1. Initialise seeds on a regular grid
2. Compute iteratively superpixels until convergence
   (a) Aggregate pixels to a center by minimizing the **aggregation distance**
   (b) Update position center (mean on each cluster)
3. Enforce connectivity on small entities

Aggregation Distance Description

Aggregation distance between a seed p_i and $p_j \in \mathcal{V}(p_i)$ with $\mathcal{V}$ the considered neighbourhood

State of the art distance [1]

$$D_{5D}(p_j, p_i) = \sqrt{d_c + \frac{m}{S} d_s}$$

Colour distance

$$d_c(p_j, p_i) = (l_j - l_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2$$

Space distance

$$d_s(p_j, p_i) = (x_j - x_i)^2 + (y_j - y_i)^2$$

- (l, a, b) the 3 channels in Lab color space
- (x, y) the pixel coordinates
- m the weight of the relative importance between d_c and d_s
- $S = \sqrt{\frac{N}{K}}$
- N the number of pixel in the image
- K the number of desired superpixels

New energy introduced

$$D_{geom}(p_j, p_i) = \sqrt{d_{c_0} + \alpha . d_{s_0}^{\beta} + d_g}$$

Normalised colour distance

$$d_{c_0}(p_j, p_i) = \frac{d_c(p_j, p_i)}{\max_{p_k \in \mathcal{V}(p_i)}(d_c(p_k, p_i))}$$

Normalised space distance

$$d_{s_0}(p_j, p_i) = \frac{d_s(p_j, p_i)}{\max_{p_k \in \mathcal{V}(p_i)}(d_s(p_k, p_i))}$$

Geometric distance

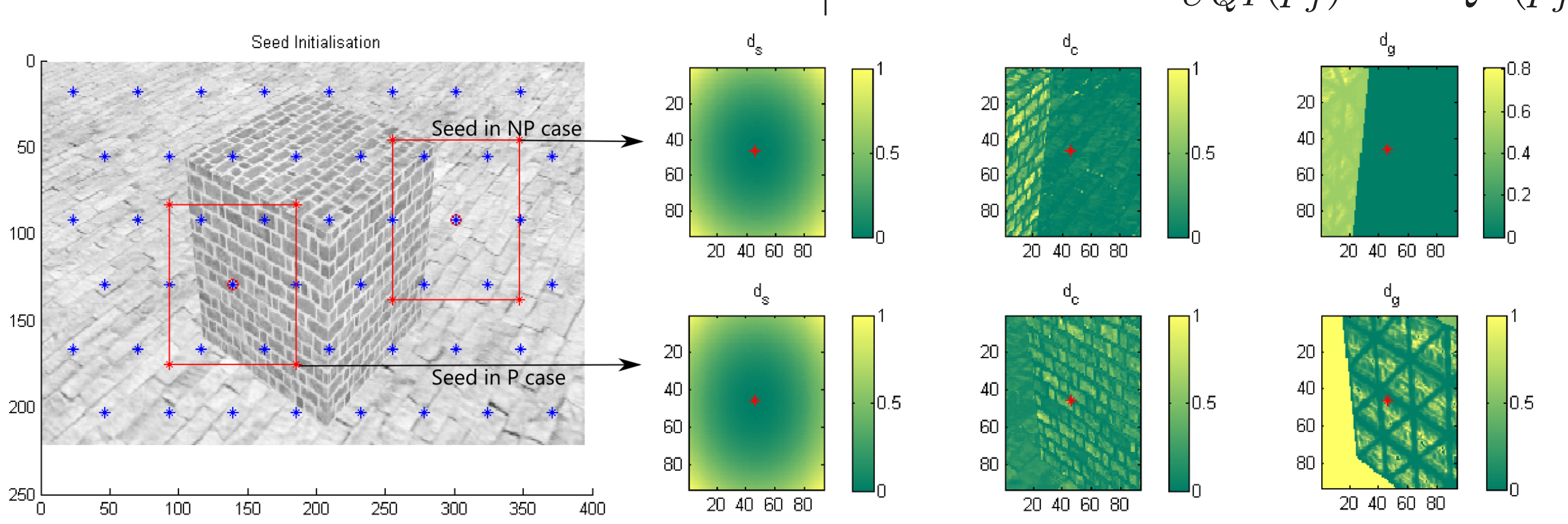
$$d_g(p_j, p_i) = 1 - d_{\vec{n}}(p_j, p_i) . d_{UQI}(p_j)$$

Orientation term

$$d_{\vec{n}}(p_j, p_i) = \frac{1 + \cos(\vec{n}_j, \vec{n}_i)}{2}$$

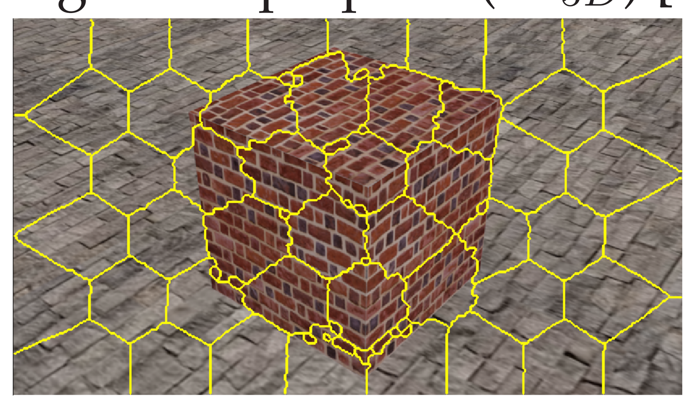
Planarity term

$$d_{UQI}(p_j) = UQI(p_j) . \mathbb{1}_{UQI>0}$$

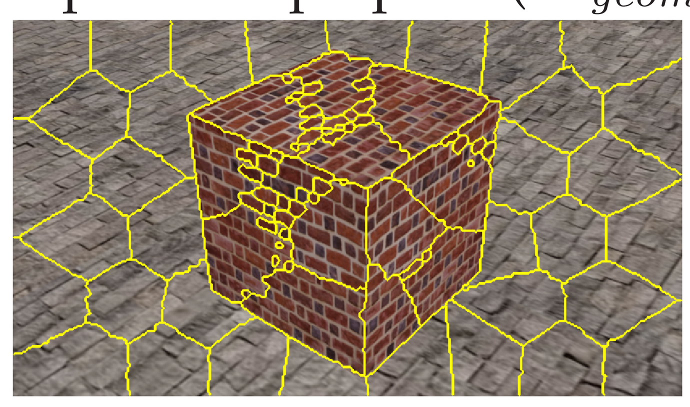


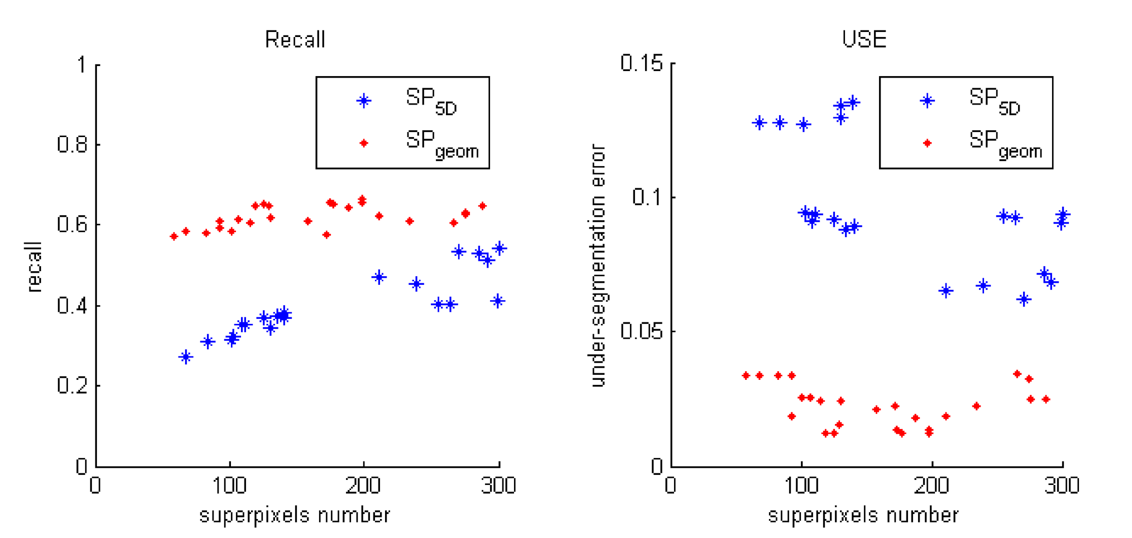
Results

Original Superpixel (SP_{5D}) [1]



Proposed Superpixel (SP_{geom})





Conclusion and Perspectives

This approach provides superpixels consistent with the scene geometry when the mesh matches with the real surfaces. In perspective, we have to include a cutting process on non-planar faces and to generalise the approach on real images.

References

[1] R. Achanta and A. Shaji and K. Smith and A. Lucchi and P. Fua and S. Susstrunk SLIC Superpixels Compared to State-of-the-art Superpixel Methods, In IEEE PAMI, 2012

[2] M.-A. Bauda and S. Chambon and P. Gurdgos and V. Charvillat Image Quality Assessment for Photo-consistency Evaluation on Planar Classification in Urban Scenes, In ICPRAM, 2015

[3] R. I. Hartley and A. Zisserman Multiple View Geometry in Computer Vision, 2004

[4] Z. Wang and A. Bovik A universal image quality index, In IEEE Signal Processing Letters, 2002