IMPROVING DEPTH ESTIMATION USING SUPERPIXELS

Cambra A.B., Muñoz A., Murillo A.C., Guerrero J.J., Gutierrez D.
{acambra, adolfo, acm, jguerrer, diegog}@unizar.es

Escuela de Ingeniería y Arquitectura
Universidad Zaragoza
1542

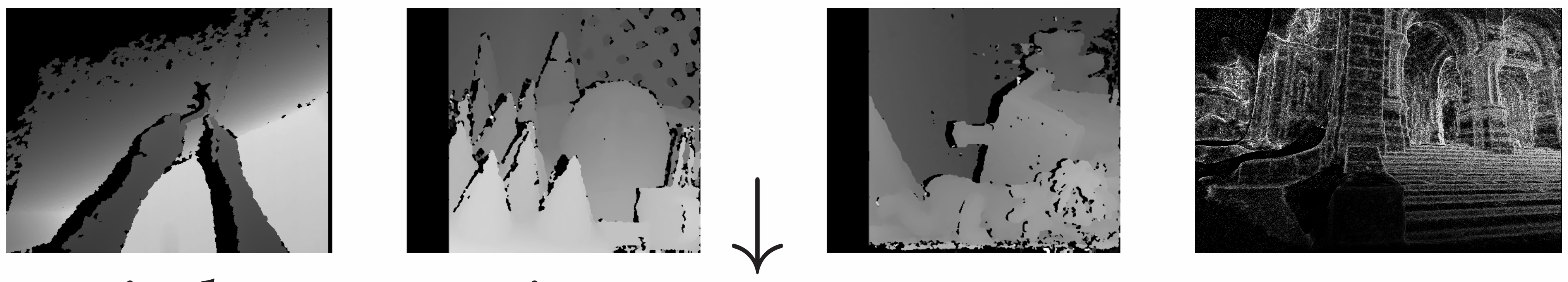
Abstract

The depth information obtained from multiple view algorithms or RGB-d sensors is frequently incomplete [3]. We study how to improve the scene depth estimation combining any kind of rough initial estimation with a pipeline for pixel-wise labeling optimization [5]. Our preliminary pipeline makes use of superpixel image segmentation and Markov-Random-Field solvers, both of them very powerful tools frequently used to obtain a robust and consistent labeling in an image. We propose and analyze how to modify the MRF cost functions and superpixel description to improve the performance.

Pipeline and Sample Results

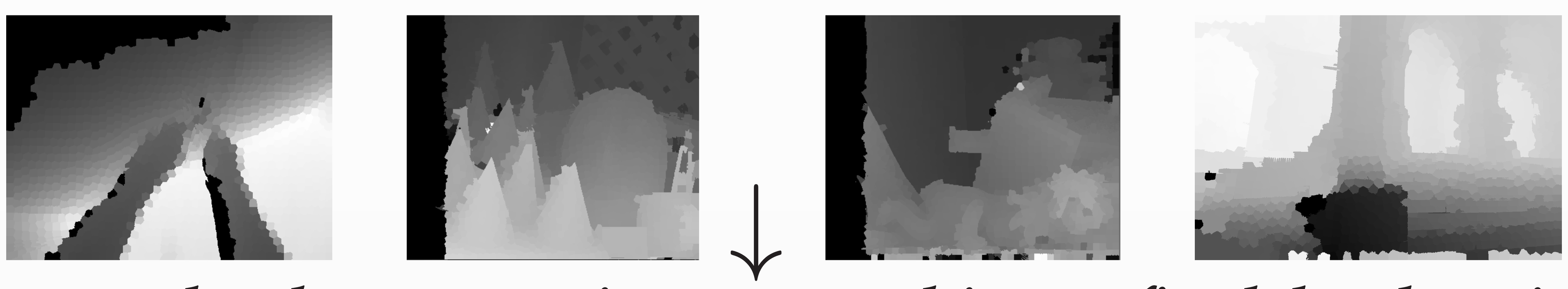
1. ***Input image + input depth*** (from any kind of source)

RGB-d depth map | Disparity map | Disparity map | Point cloud

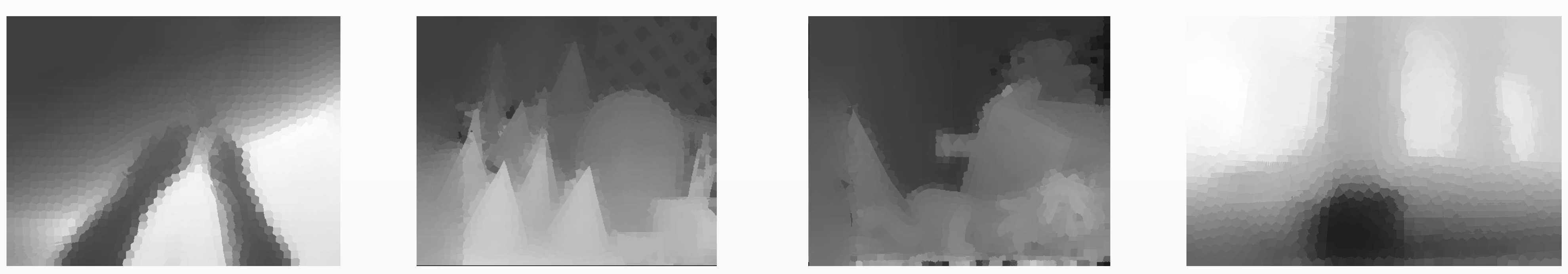


2. ***Superpixel segmentation*** [4]: reducing complexity of labeling problem

3. ***Input depth + segmentation = initial superpixel depth***: the whole superpixel gets a depth value according to its pixel depth distribution



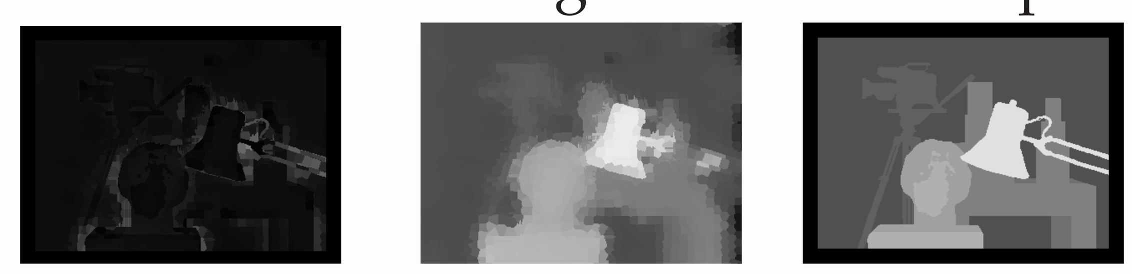
4. MRF based ***depth propagation + smoothing → final depth estimation***



Quantitative evaluation

- ***Error per pixel***: difference between the ground truth values and the MRF depth propagation

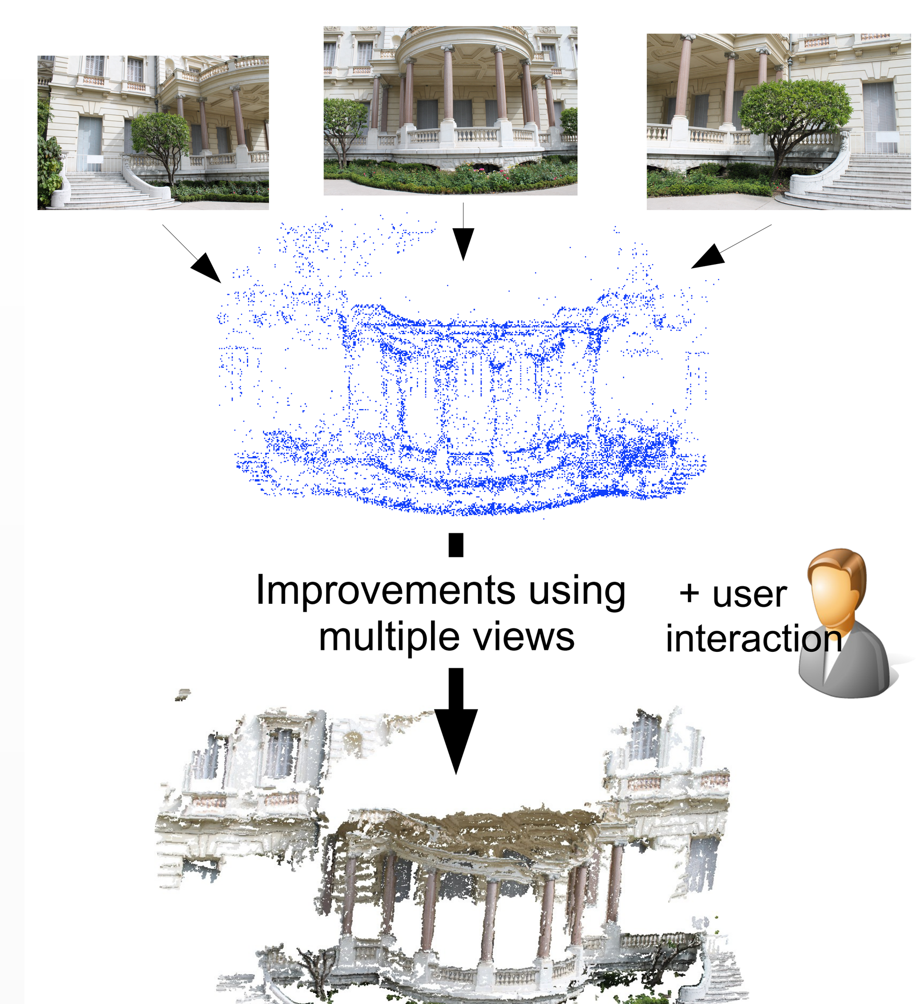
error = gt - final depth



Input image [a]	Input depth	Initial superpixel	Depth propagation
Tsukuba	23.4555	21.4349	20.7857
Venus	17.9249	13.3355	8.8710
Cones	31.9362	25.0465	8.9230
Teddy	32.0781	25.6495	9.6206
Sawtooth	16.0922	13.2377	9.51
Bull	12.7467	10.3099	7.4231
Poster	15.7758	11.4537	9.3676
Barn1	15.6588	12.3351	9.5196
Barn2	15.3817	12.8297	10.2926
Map	21.5471	23.2944	21.7738

[a] Datasets: `http://vision.middlebury.edu/stereo/data/`

Application - Future work



Analysis of MRF cost functions

Energy Function:

$$E = \sum_{p \in N} C(l_p) + \sum_{\{p,q\} \in V} C(l_p, l_q)$$

Unary cost

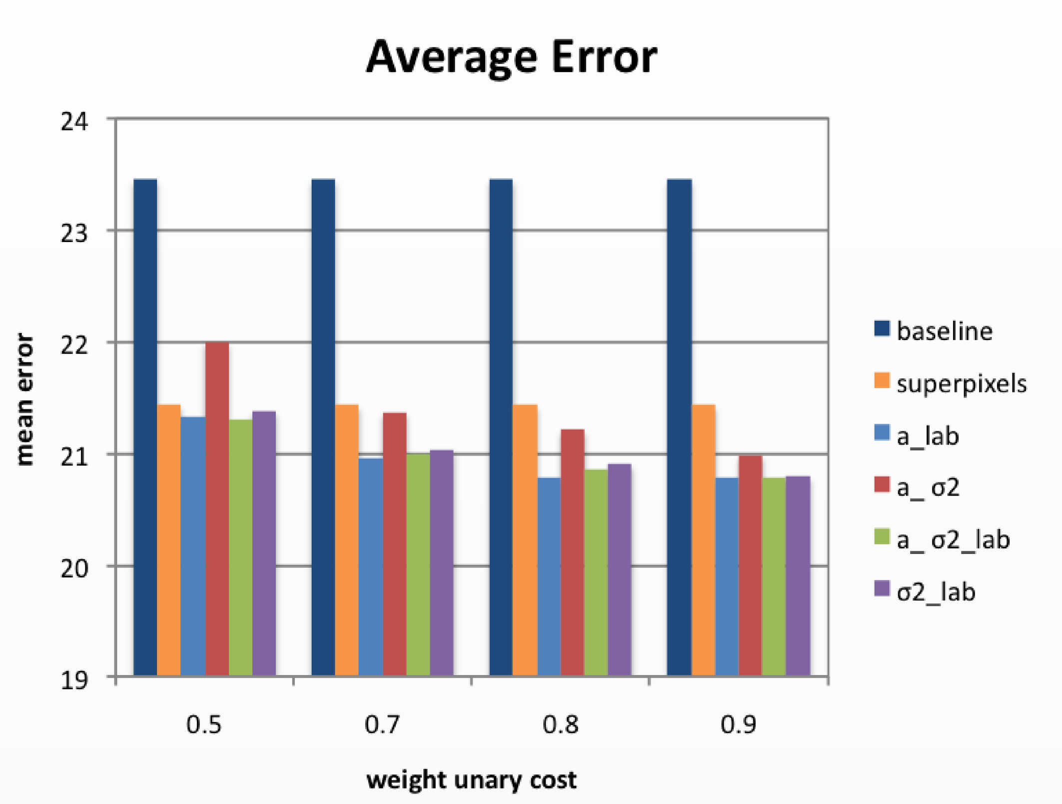
(per superpixel)

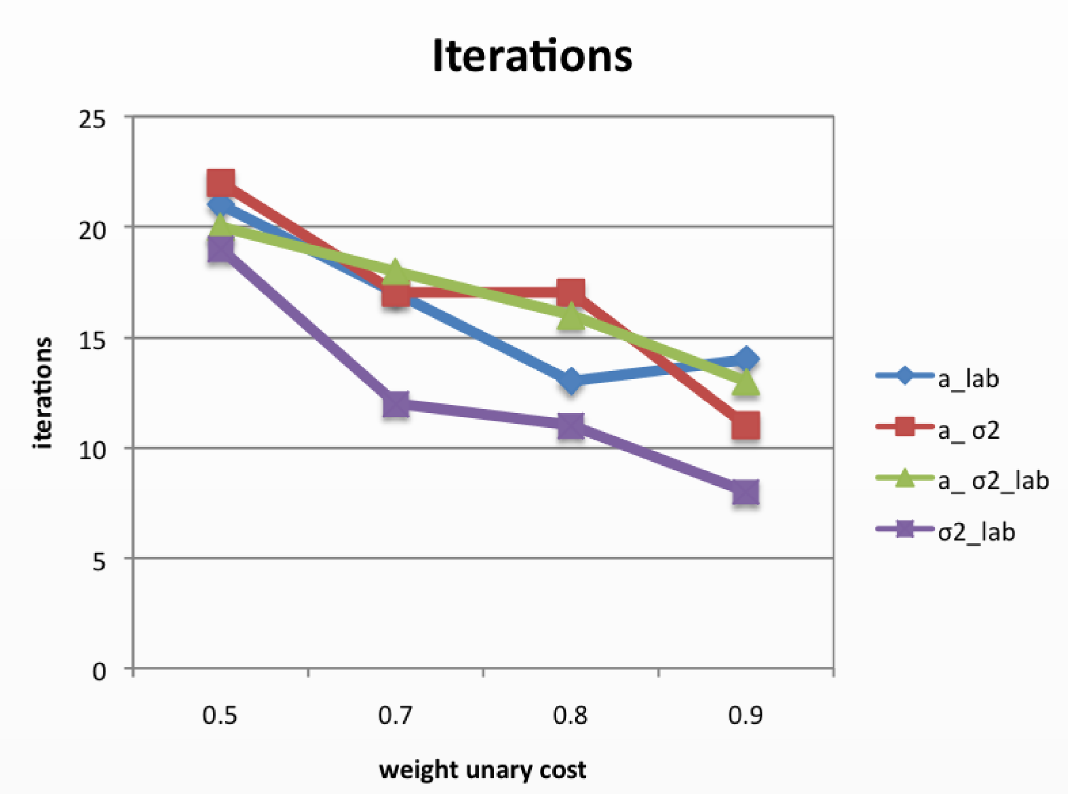
$$C(l_p) = \begin{cases} 0 & : z_p = 0 \\ w_u \cdot a_p \cdot (1-\sigma^2) \cdot (z_p - l_p)^2 & : z_p > 0 \end{cases}$$

Binary cost

(for each couple p,q of neighbouring superpixels)

$$C(l_p, l_q) = (1 - w_u)(1 - d_{lab})(l_p - l_q)^2$$





References

[1] D., Scharstein, R., Szeliski, High-accuracy stereo depth maps using structured light, in *CVPR*, 2003

[2] D., Scharstein, R., Szeliski, A taxonomy and evaluation of dense two-frame stereo correspondence algorithms, in *IJCV*, 2002

[3] M., Van den Bergh, M., Carton, D., Carton, L.J., Van Gool, Depth SEEDS: Recovering incomplete depth data using superpixels, in *WACV*, 2013

[4] R., Achanta, A., Shaji, K., Smith, A., Lucchi, P., Fua, S., Süsstrunk, Slic superpixels, in *EPFL*, 2002

[5] R., Szeliski, R., Zabih, D., Scharstein, O., Veksler, V., Kolmogorov, A., Agarwala, M., Tappen, C., Rother, A comparative study of energy minimization methods for markov random fields with smoothness-based priors in *Pattern Analysis and Machine Intelligence*, 2008

Acknowledge

This work was funded by project TAMA (DGA).