P-03

Towards parameter free denoising

Mickaël Buchet[1*], Tamal. K. Dey[2], Jiayuan Wang[2], and Yusu Wang[2]

[1]AIMR – Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577 Japan
[2] Department of Computer Science and Engineering – The Ohio State University, Columbus, OH 43210, USA

In many data analysis applications, we are given a point set that is supposedly sampling an unknown object called the ground truth. This sample is corrupted by noise some of the points may lie far away from the ground truth, creating outliers also termed as ambient noise. One of the main goals of denoising techniques is to remove those points so that the error becomes bounded with respect to the Hausdorff distance. Popular denoising approaches such as deconvolution and thresholding require the input of several parameters by the user, either to define the algorithm or the noise model. We aim at lightening the burden put on the user by reducing the number of parameters to one in the general case and zero in the case of a uniform sampling.

We propose a one-step algorithm that, given a scale defining parameter, removes points that are outliers at this scale. The presence of this parameter cannot be avoided for general point sets as the definition of a scale is indispensable. However, if we add some uniformity in the data, then this uniformity can replace the notion of scale. In this case, we propose an algorithm alternating decluttering steps and resampling steps that denoise the data without knowing the correct scale.

This algorithm presents three interesting features. First, the decluttering is made in a way that is dual to classic algorithms. Usually, one select good points and then removed all points that are not closed to a good point as being irrelevant. In our approach, we look at all points and keep only those that are far from a good point, the notion of closeness being defined by the point currently being investigated. A good choice on the order in which points are investigated yield theoretical guarantees. Second, we introduce a new technique of resampling. The set of points obtained through the decluttering step is too usually too sparse. We hence add again some of the points we removed in order to obtain a denser sampling before iterating. Finally, our algorithm does not need any parameter and returns a theoretically guaranteed result without computing the correct scale, which is a very difficult question. Instead, we guarantee that after passing the unknown optimal scale, we will not worsen our sampling.

[1] Declutter and Resample: Towards parameter free denoising, M. Buchet, T. K. Dey, J. Wang, Y. Wang, preprint