ABSTRACT

Nowadays the enhanced capabilities of commercially available active depth sensing devices are of great interest as 3D structure of scene can be efficiently represented by depth maps. Accurate depth maps with good resolution are essential in many applications such as robot navigation, intelligent transportation, object tracking and recognition, human computer interaction, virtual reality, 3D TV and so on. This has been facilitated by depth sensing devices such as time-of-flight (ToF) cameras, Microsoft Kinect sensor, ASUS Xtion PRO and Intel RealSense which provide range maps in real time. The depth sensors specified above, project infrared illumination on to the scene and from the reflected signal depth information is estimated. Despite advances in the technology of depth sensing devices, active sensors cannot work with shiny or self-illuminated surfaces and sometimes foreground objects occlude the background scene.

Passive methods for depth/disparity map estimation such as multi-view stereo, binocular stereo, shape from focus (SFF), depth from defocus (DFD), structure from motion (SfM) and photometric stereo also suffer from challenges such as occlusions, low textured regions, poor or varying illumination and inaccurate correspondence matching. In general depth maps suffer from low spatial resolution and are affected by missing data regions of irregular shapes and sizes especially around depth discontinuities due to poor reflectivity and occlusions. Hence, hole-filling/disocclusion and super-resolution (SR) of depth maps are challenging topics of research.

Depth map completion is an inverse and ill-posed problem, which means that infinite number of solutions may exist. Hence, to choose the optimal solution from the set of possible solutions and to resolve ambiguities apriori information can be utilized. Regularization with non-local self-similar depth patches, low-rank, sparsity and deep learning priors have recently become popular.

This thesis investigates non-local and low-rank priors for completion of depth profiles. In the literature, several works exist which utilize first order neighbourhood Markov random fields for inpainting missing regions in depth maps. However, local interaction based spatial regularization does not yield accurate results for depth map completion. Particularly, it is observed that depth discontinuities are not preserved and smearing or blurring distortions occur. Initially, in the second chapter of the thesis a convex optimization framework is proposed along with nonlocal extension of Gauss-Markov random field (GMRF) prior, to inpaint small missing regions using a single degraded depth profile as input. We observe that this simple extension is able to preserve depth discontinuities and prevent distortions.

As depth maps consist of many low-textured regions with occasional discontinuities, to exploit the correlated nature of depth data we are motivated to use the recently popular lowrank regularization techniques such as nuclear norm minimization (NNM). However, one basic drawback of NNM is that all the singular values are given equal importance and each of them are shrunk with the same value of threshold. As meaningful information is mostly associated

with higher singular values, they should be given more weightage as compared to lower singular values during rank minimization. Hence, a weighted nuclear norm minimization (WNNM) approach is proposed in the third chapter of the thesis to inpaint the degraded depth map.

It has been demonstrated in recent literature, that the Schatten p-norm ($0 < p \leq 1$) which is a non-convex relaxation of rank minimization yields better approximation to the solution in comparison to nuclear norm minimization Use of this metric can mitigate the over-shrinking drawback of NNM without adding much computational cost. Therefore, in chapter 4, we propose to use weighted Schatten p-norm as a penalty function to complete the missing depth data by taking as input only a single degraded depth map. The proposed method estimates depth values more accurately, particularly, at depth discontinuities or boundaries of objects.

In the real world, depth maps captured by sensors such as Microsoft Kinect or time-of-flight devices suffer from large missing regions especially at depth edges. Filling-in such large holes is difficult for unguided depth inpainting algorithms. Hence, in the literature, many works have appeared which utilize the information embedded in the corresponding RGB image of the same scene. Therefore in chapter 5 of the thesis to handle large missing regions in the input depth map and to increase the accuracy of filling-in by reducing blur or smoothness at the depth edges the corresponding RGB frame is over-segmented using superpixels at multiple scales. These superpixels are overlaid on the degraded depth map wherein large missing regions are inpainted by optimization of an objective function formulated using the non-local GMRF prior. In contrast to typical non-local image processing works in the literature instead of using a rectangular window, the non-local search window is restricted to the superpixel containing the reference patch.

We provide comparisons in each chapter with recent state-of-the-art algorithms to demonstrate the effectiveness of non-local and low-rank priors investigated in this thesis.

***Keywords**: Depth map completion, maximum a-posteriori, Markov random field, non-local means, superpixels, RGB-D data, low-rank regularization, weighted nuclear norm, weighted Schatten p-norm.*