

## 3D SUPER-RESOLUTION APPROACH FOR SPARSE LASER SCANNER DATA

S. Hosseinyalamdary<sup>a\*</sup>, A. Yilmaz<sup>a</sup>

<sup>a</sup> Photogrammetric Computer Vision (PCV) Lab  
2070 Neil avenue, Columbus, OH 43212,  
hosseinyalamdary.1@osu.edu ; yilmaz.15@osu.edu

Commission VI, WG VI/4

**KEY WORDS:** 3D Super-resolution, Geometric Surface Reconstruction, Diffusion Equations, isotropic and anisotropic

### ABSTRACT:

Laser scanner point cloud has been emerging in Photogrammetry and computer vision to achieve high level tasks such as object tracking, object recognition and scene understanding. However, low cost laser scanners are noisy, sparse and prone to systematic errors. This paper proposes a novel 3D super resolution approach to reconstruct surface of the objects in the scene. This method works on sparse, unorganized point clouds and has superior performance over other surface recovery approaches. Since the proposed approach uses anisotropic diffusion equation, it does not deteriorate the object boundaries and it preserves topology of the object.



Figure 1: The laser scanner point cloud may be noisy, sparse, prone to systematic errors. We propose an approach to overcome these problems and reconstruct surface of the objects in this paper.

## 1. INTRODUCTION

Surface of the objects can be observed by applying different techniques such as stereo images, laser scanner or structured light. These techniques provide 3D coordinates of the sampled points of a surface called point cloud. Depending on the technology applied to collect data, the noise in 3D coordinates can vary from less than millimeter in high quality laser scanners to few centimeters in stereo images. The resolution of the point cloud is also a factor of the applied technology. The stereo images and structured light provide dense point cloud, in contrast the laser scanner may collect sparse point cloud especially using inexpensive laser scanners. The point cloud may be affected by systematic errors. The stereo images and structured light may suffer from light overexposure or lack of texture may result the failure in these approaches. The laser scanner may be affected by the platform motion while the data is collected. In autonomous driving applications, the sensors' point of view is restricted to the installation of these sensors on the platform and therefore, objects are most likely to be observed from one point of view. If the object is seen from one point of view, back of the object will be self-occluded and therefore, the generated surface partially represents surface of the object. If the object is seen from different perspectives, surface of the object can be fully generated. Figure (1.a) shows the image of an object and the sampled points are projected to the image in Figure (1.b). The sparsity of the point cloud can be seen in this figure. This paper presents 3D super-resolution approach to overcome aforementioned problems. In order to 3D super-resolution We use a 3D generalization of anisotropic PDE

\*Corresponding author

based regularization which has been used for various computer vision applications including super-resolution, restoration, denoising, and inpainting (Tschumperlé and Deriche, 2005, Sapiro, 2006).

This paper is organized as follows. The surface reconstruction approaches are briefly described in section (2.). The implicit representation of surface and anisotropic diffusion equations using local patches are given in section (3.). The implementation aspect of the proposed approach is discussed in section (4.) and it is explained how the noise can be handled in laser scanner data. Experiments are described in section (5.), the results of our approach is compared to two popular approaches and conclusions are drawn in section (6.).

## 2. LITERATURE REVIEW

Laser scanner point cloud can be used for high level tasks in computer vision such as object tracking, object recognition and scene understanding. The observed depth of the object makes laser scanner more reliable to track objects rather than the image based object tracking. Held et al. used laser scanner and camera in multiple epochs to precisely track the objects (Held et al., 2014). In addition, it has been shown that the histogram of oriented depth can be employed to recognize people in the scene similar to the histogram of oriented gradient (Premebida et al., 2014). However, the use of the histogram of oriented depth may be adversely affected in sparse point clouds and the surface reconstruction approaches may be required to improve the density of point clouds. It has been shown that the success rate of the object recognition increases when the point cloud and images are used together (J. Dolson and Thrun, 2010). Surface of an object can be represented in different ways. In addition to the point based representation which is shown in Figure (1.b), surface can be explicitly approximated by triangular mesh, spline and its derivatives (Szeliski, 2010, Lee et al., 1997). Moreover, surface can be implicitly represented using level sets. In this approach the space is divided into inner object space and outer object space and an indicator function is defined as signed distance function with different signs for inner and outer object space and the surface is zero-crossing of this function (Osher and Fedkiw, 2003). Among many surface reconstruction algorithms marching cubes, moving least squares (MLS), and Poisson surface reconstruction may

be more popular. Marching cubes uses zero-crossing of the implicit surface representation to recover 3D model (Lorensen and Cline, 1987). MLS applies a moving window to fit a polynomial to the points in this window (Pauly et al., 2003, Alexa et al., 2003) and therefore, it is locally optimum. Kazhdan et al. (Kazhdan et al., 2006, Bolitho et al., n.d., Kazhdan and Hoppe, 2013) suggested to use implicit shape model to fit surface to the sparse sampled points. This approach obtains a global optimum surface which may not preserve topology and fill the holes in the object. Lim and Haron (Lim and Haron, 2014) have surveyed surface reconstruction techniques in more details. There are also some scientific endeavors to recover the normal vector field of an object. Tasdizen and Witaker have used an anisotropic diffusion tensor to smooth the normal vector field while it preserves the discontinuities (Tasdizen et al., 2002, Tasdizen and Whitaker, 2003). Anisotropic filtering of the normal vector field and curvature tensor is used to improve the quality of meshes of surface (Liu et al., 2007). Boulch and Marlet address the ambiguities of the normal vector field and edges and corners and provide a solution to handle it (Boulch and Marlet, 2012).

### 3. METHODOLOGY

In this section, we introduce mathematical foundation of the proposed approach. First, local coordinate system is defined and 3D surface is converted to depth scalar field. The foundations of the geometric anisotropic diffusion equation are described later.

#### 3.1 The Local coordinate system

Every surface can be considered as a two dimensional manifold embedded into three dimensional Euclidean space,  $S : \Lambda \subset \mathbb{R}^3 \rightarrow R^2$ .  $\Lambda$  is a subset of  $R^3$  that encompasses the surface. If we assume that the surface is smooth and consequently differentiable, it becomes a Riemannian manifold and tangent space can be defined. Let's assume that tangent space,  $T(u, v)$  is tangential to the surface at point  $\mathbf{X}$  and it is spanned by orthogonal vectors  $\mathbf{u}$  and  $\mathbf{v}$ , we define the local coordinate system in the way that its origin is located at point  $\mathbf{X}$  and  $\mathbf{u}$ ,  $\mathbf{v}$  and  $\mathbf{n}$  are the bases of this coordinate system.  $\mathbf{n}$  is normal vector of the surface at point  $\mathbf{X}$ . Therefore, every point  $\mathbf{X}'$  in the vicinity of point  $\mathbf{X}$  may be expressed as  $\mathbf{X}' = [u_{\mathbf{X}'}, v_{\mathbf{X}'}, 0]^T$  in local coordinate system. Point  $\mathbf{X}'$  may be also represented in three dimensional Euclidean coordinate system of sensors, which is called global coordinate system in this paper,  $\mathbf{X}_g' = [x, y, z]^T$ . Subscript  $g$  stands for global coordinate system and subscript  $l$  in local coordinate system is left for brevity. If the neighborhood of point  $\mathbf{X}$ ,  $\Omega_{\mathbf{X}}$ , becomes larger, the surface diverges from tangent space and  $\mathbf{X}' = [u_{\mathbf{X}'}, v_{\mathbf{X}'}, 0]^T$  is not valid anymore. Let's assume the distance of the every point on the surface from its projection in tangent space is  $W(u, v)$ . It is a scalar field referred as distance map or depth map. Therefore, every point  $\mathbf{X}'$  in the neighborhood of point  $\mathbf{X}$  can be represented as  $\mathbf{X}' = [u_{\mathbf{X}'}, v_{\mathbf{X}'}, W(u_{\mathbf{X}'}, v_{\mathbf{X}'})]^T$  in tangent space. In order to recover the surface, it suffices to estimate the scalar field  $W(u_{\mathbf{X}'}, v_{\mathbf{X}'})$  and convert the points in local coordinate system into global coordinate system. Figure (2) shows the profile of local coordinate system. Horizontal axis represents the tangent space where  $W(u, v) = 0$ , and vertical axis shows the normal direction to the surface. Point  $\mathbf{X}$  shows the origin of the coordinate system and every point  $\mathbf{X}'$  has  $(u, v)$  components in tangent space and  $W(u, v)$  component in normal direction.

Furthermore, The choice of the neighborhood is of importance: the large neighborhood may fill the holes in the objects and overextends the boundaries of the surface. In contrast, small neighborhood may not reconstruct the surface and does not overcome the

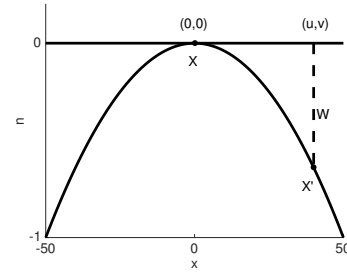


Figure 2: The local surface is shown by a quadratic curve. The coordinate system is defined in the way that point  $\mathbf{X}$  is the origin,  $(u, v)$  are the basis vectors in tangent space and  $\mathbf{n}$  is the basis vector in normal direction. Therefore, every point  $\mathbf{X}'$  can be represented as  $[u_{\mathbf{X}'}, v_{\mathbf{X}'}, W(u_{\mathbf{X}'}, v_{\mathbf{X}'})]^T$ .

sparsity of the sampled points. It is also preferable to change the neighborhood based on the geometry of the sampled points.

It should be noted that the proposed approach is condition to the convexity (or concavity) of the surface in  $\Omega_{\mathbf{X}}$ . The convexity (or concavity) of the surface in the neighborhood guarantees that  $W(u_{\mathbf{X}'}, v_{\mathbf{X}'})$  is a function of  $u_{\mathbf{X}'}$  and  $v_{\mathbf{X}'}$ . It is obviously invalid for realistic scenarios, but most of the objects may be assumed to be convex (or concave) in local neighborhood.

If a surface is an open surface, some of the points are located at the boundaries of the surface. The boundaries of an object can be represented as a curve and it can be assumed as one dimensional manifold embedded into two dimensional space,  $C : \tau \subset R^2 \rightarrow R^1$ , where  $\tau$  is the subset of  $R^2$  that encompasses the boundaries. The boundaries of the surface can be considered as the silhouette of an object in image processing and may be used to create the histogram of the oriented depth for object recognition (Premebida et al., 2014).

#### 3.2 Geometry based surface reconstruction

Laser scanner collects samples of points from surface of an object in the scene,  $\mathbf{X}^i$ . The surface can be reconstructed at this local neighborhood if the depth map,  $W(u, v)$  is estimated in this region. The surface reconstruction is an inverse problem and it is ill-conditioned. Therefore, we apply the smoothness assumption of the surface to regularize the surface reconstruction such that

$$\min_{W: \Omega \rightarrow R} E(W) = \int_{\Omega} \phi(\|\nabla W\|) dudv, \quad (1)$$

where  $\nabla W = [\frac{\partial W}{\partial u}, \frac{\partial W}{\partial v}] = [W_u, W_v]$  is the gradient of the depth map,  $\Phi(\cdot)$  is an arbitrary differentiable convex function and  $\|\nabla W\|$  is  $L_2$  norm of  $\nabla W$  at point  $\mathbf{X}'$ . The Euler-Lagrange equations can be applied to minimize (1) and it leads to isotropic diffusion equation

$$\begin{cases} \frac{\partial W}{\partial t} = \text{div}(\nabla W) = \Delta W & \text{if } \mathbf{X}' \notin \mathbf{X}^i \\ \frac{\partial W}{\partial t} = 0 & \text{if } \mathbf{X}' \in \mathbf{X}^i, \end{cases} \quad (2)$$

where  $\Delta$  is the Laplacian operator and (2) indicates to a famous problem in partial differential equation known as heat flow. Here, we considered  $\Phi(\cdot)$  as an  $L_2$  norm quadratic surface. Equation (2) can be solved using gradient decent minimization in few iterations. Obviously, we need an initial estimation of the surface and we can update it in every iteration such that

$$W_{k+1} = W_k - \frac{\partial W}{\partial t} \delta, \quad (3)$$

where  $\delta$  is the step size at each iteration. Equation (1) assures that the surface is smooth, but it may smooth the edges and corners. In order to preserve the edges and corners, a diffusion tensor can be applied to (1). A diffusion tensor,  $D$  can be multiplied by  $\nabla W$  in (1) to provide anisotropic diffusion filtering. Therefore,  $\phi(\nabla W)$  in (1) is replaced to  $\phi(D\nabla W)$ . Solution of (1) can be estimated using the gradient descent in an iterative scheme

$$\begin{cases} \frac{\partial W}{\partial t} = \text{div}(D\nabla W) & \text{if } \mathbf{X}' \notin \mathbf{X}^i \\ \frac{\partial W}{\partial t} = 0 & \text{if } \mathbf{X}' \in \mathbf{X}^i. \end{cases} \quad (4)$$

The term  $\frac{\partial W}{\partial t} = 0$  guarantees that the diffusion equation maintains the sampled points intact and smoothes other points on the surface. In order to preserve the edges and corners,  $D$  should be selected in the way that it suppresses the smoothing in the direction of  $\nabla W$ . Perona and Malik suggested to use inverse of exponential function of the squared gradient in their work (Perona and Malik, 1990). If the sampled points  $\mathbf{X}_i$  in the neighborhood of point  $\mathbf{X}$  are converted to the local coordinate system, the local coordinates of these points become  $\mathbf{X}_i = [u_{\mathbf{X}_i}, v_{\mathbf{X}_i}, W(u_{\mathbf{X}_i}, v_{\mathbf{X}_i})]^\top$ .  $\nabla W$  can be approximated as  $W_u \simeq \frac{W(u_{\mathbf{X}_i}, v_{\mathbf{X}_i}) - W(u_{\mathbf{X}}, v_{\mathbf{X}})}{u_{\mathbf{X}_i} - u_{\mathbf{X}}}$  and  $W_v \simeq \frac{W(u_{\mathbf{X}_i}, v_{\mathbf{X}_i}) - W(u_{\mathbf{X}}, v_{\mathbf{X}})}{v_{\mathbf{X}_i} - v_{\mathbf{X}}}$ . We can generate the structure tensor such that

$$\mathbf{G} = \nabla W^\top \nabla W. \quad (5)$$

The structure tensor represents local geometry of the surface and its eigenvectors show the maximum and minimum gradient directions  $W$ . The eigenvalues,  $\lambda_1 \geq \lambda_2 \geq 0$ , and eigenvectors  $\theta_1, \theta_2$  of structure tensor show how the points are distributed in the neighborhood. In order to preserve the edges and corners, the smoothing filter should be panelized where the gradient is large. The eigenvalues of the structure tensor indicate to the magnitude of the gradient and the eigenvectors are the direction of the gradients. Let's define the diffusion tensor,  $D$  such that

$$D = f_1(\lambda_1, \lambda_2)\theta_1\theta_1^\top + f_2(\lambda_1, \lambda_2)\theta_2\theta_2^\top. \quad (6)$$

If  $f_1(\cdot)$  and  $f_2(\cdot)$  are decreasing functions. There are many variations of  $f_1(\cdot)$  and  $f_2(\cdot)$ , but it should be defined in the way that the large gradient should be suppressed. Therefore, the diffusion process does not pass through the discontinuities such as edges and corners in the surface. In this paper, we chose  $f_1(\cdot)$  and  $f_2(\cdot)$  as follows

$$\begin{cases} f_1(\lambda_1, \lambda_2) = \frac{1}{\sqrt{1+\lambda_1+\lambda_2}} \\ f_2(\lambda_1, \lambda_2) = \frac{1}{1+\lambda_1+\lambda_2}. \end{cases} \quad (7)$$

For more details, readers are referred to (Tschumperlé and Deriche, 2005, Sapiro, 2006). Solution of (4) is equivalent to an oriented Gaussian smoothing. Therefore, anisotropic diffusion equations are equivalent to oriented Gaussian kernel,  $K$ ,  $w(\text{Tschumperlé and Deriche, 2005, Sapiro, 2006})$  such that

$$\begin{cases} K(\mathbf{t}) = \frac{1}{4\pi t} \exp(-\frac{\mathbf{x}^\top D^{-1} \mathbf{x}}{4t}) \\ W(t) = K(\mathbf{t}) * W(0), \end{cases} \quad (8)$$

where  $t$  stands for the number of iterations,  $*$  is convolution operator,  $W(0)$  indicates to the depth scalar field at initial surface, and  $\mathbf{x}$  is the location of the location of the element in the kernel. This method is initially applied for image enhancement and the use of it in surface reconstruction may require some modifications and we will address it later.

## 4. IMPLEMENTATION

A simple method to convert local coordinate system to global coordinate system is presented. This discussion is followed by the initial estimation of the surface and the numerical conditioning of the problem.

### 4.1 Transformation between local and global coordinates

The local coordinate system is defined by three vector bases  $\mathbf{u}$ ,  $\mathbf{v}$ , and  $\mathbf{n}$  and it is originated at point  $\mathbf{X}$ . Although  $\mathbf{u}$  and  $\mathbf{v}$  are chosen in the directions of minimum and maximum surface curvature in many applications, their directions are not important for the proposed approach as long as they are orthonormal and the transformation between local and global coordinate systems is known. Therefore, it is sufficient if the rotation between  $z$  axis (in global coordinates system) and normal to the surface at point  $\mathbf{X}$  (in local coordinate system) is calculated. This transformation can be estimated such that

$$\mathbf{n} = \mathbf{R}_1(\alpha)\mathbf{R}_2(\beta)\mathbf{z}. \quad (9)$$

If  $\mathbf{n} = [n_1, n_2, n_3]$  and  $\mathbf{z} = [0, 0, 1]$ , it leads to  $\alpha = -\arctan(\frac{n_2}{n_3})$  and  $\beta = \arccos \sqrt{n_2^2 + n_3^2}$ . Obviously, the translation vector is the coordinates of point  $\mathbf{X}$  and for every point  $\mathbf{X}'$  in the neighborhood in local coordinate system can be converted to global coordinate system

$$\mathbf{X}'_g = \mathbf{R}_1(\alpha)\mathbf{R}_2(\beta)\mathbf{X}'_l + \mathbf{X}_g. \quad (10)$$

### 4.2 Initial surface

Geometry of the surface in a local neighborhood can be represented by structure tensor. Unlike (5), where the structure tensor is created in local neighborhood, the structure tensor for the sampled points can be calculated in global coordinate system. The structure tensor for 3D Euclidean coordinate system (global coordinate system) is created such that

$$\mathbf{G}_{3 \times 3} = (\mathbf{X}_g^i - \mathbf{X}_g)^\top (\mathbf{X}_g^i - \mathbf{X}_g). \quad (11)$$

Now, the structure tensor can be decomposed into its ( $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$ ) and eigenvectors ( $\theta_1, \theta_2, \theta_3$ ). The eigenvector corresponding to the least eigenvalue is aligned to the normal vector of the surface. It can be used to generate tangent space which has been defined in section (3.). The other eigenvalues and eigenvectors indicate to the shape of the object in the neighborhood of point  $\mathbf{X}_g$ . The other two eigenvectors ( $\theta_1$  and  $\theta_2$ ) to define the neighborhood that can be adjusted based on the geometry of the object. The eigenvalues  $\lambda_1$  and  $\lambda_2$  provides the major and minor axes of an ellipse in tangent space to approximate the neighborhood. If the points are located at a plane, therefore  $\lambda_1 \geq \lambda_2 \gg 0$  and ellipse and consequently neighborhood become large (if  $\lambda_1 \simeq \lambda_2$ , it becomes circle). If the points are located at a narrow part of the object (Roof of the car is seen like a narrow object in Figure (1.a)),  $\lambda_2 \simeq 0$  and the ellipse elongates the narrow part. Therefore, it enforces maintaining the geometry of the object and prevents any overgrowing region. Later, it is shown that this property maintains topology of the surface. In order to dispense the points in the neighborhood of point  $\mathbf{X}$ , a  $100 \times 100$  grid is created around the point  $\mathbf{X}$  and the depth map of the surface is estimated for every point on the grid. Assuming the depth of each point on the grid is zero, the surface is approximated with pieces of planes. However, it is preferable if higher order polynomial is fit to sampled points  $\mathbf{X}^i$  in the neighborhood of point  $\mathbf{X}$ . A quadratic surface is used here to approximate the depth for the points on the grid. Let's assume the quadratic form

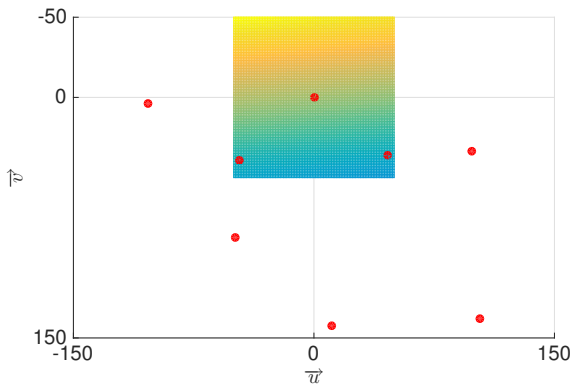


Figure 3: A  $100 \times 100$  grid is created to generate the surface at the local neighborhood. This local neighborhood is shown in Figure (1.a) with a red rectangle. Depth map should be estimated for the points on the grid to reconstruct the surface in this neighborhood.

is like  $ax^2 + by^2 + cxy = z$  for a point  $p = [x, y, z]^T$ . If the sample points are  $\mathbf{X}_i = [u_{\mathbf{X}^i}, v_{\mathbf{X}^i}, W(u_{\mathbf{X}^i}, v_{\mathbf{X}^i})]^T$ , the quadratic form of the surface can be estimated such that

$$\begin{bmatrix} W_{\mathbf{X}^1} \\ \vdots \\ W_{\mathbf{X}^n} \end{bmatrix} = \begin{bmatrix} u_{\mathbf{X}^1} u_{\mathbf{X}^1} & v_{\mathbf{X}^1} v_{\mathbf{X}^1} & u_{\mathbf{X}^1} v_{\mathbf{X}^1} \\ \vdots & \vdots & \vdots \\ u_{\mathbf{X}^n} u_{\mathbf{X}^n} & v_{\mathbf{X}^n} v_{\mathbf{X}^n} & u_{\mathbf{X}^n} v_{\mathbf{X}^n} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}. \quad (12)$$

The constant term in (12) is considered to be zero to guarantee that the surface passes through the origin. The first order terms are also assumed to be zero in order to keep the slope of the surface at the origin zero and maintain the normal vector. The second order coefficients can be estimated in (12) using least squares and the quadratic form is used to estimate the depth of the grid points. Figure (3) illustrates a  $100 \times 100$  grid points and the neighboring sampled points. The red dots are the sampled points in the local coordinate system. The color of the surface shows the depth of every point in the grid. This patch which is used in this paper to explain the proposed approach is shown in Figure (1.a) with a red rectangle. This patch is located at the boundary of the object and the sampled points are distributed in half space of local coordinate system.

### 4.3 Normalization

Solution to (2) is equivalent to the convolution of the depth map with a Gaussian kernel based filter,  $K$ . The exponential function in (8) may lead to very large numbers especially if the points are not evenly distributed in space. Therefore, it should be normalized to overcome the numerical ill-condition problem. The maximum distance of the sampled points,  $\mathbf{X}^i$ , in the neighborhood to the point,  $\mathbf{X}$ ,  $d_{max}$ , is calculated and scale is defined as  $s = \frac{\sqrt{(3)}}{d_{max}}$ . This scale maps the distance such that  $[0, \infty] \rightarrow [0, 1]$ . The Gaussian kernel is also mapped such that  $\frac{1}{4\pi t}[0, 1] \rightarrow \frac{1}{4\pi t}[\frac{1}{e}, 1]$ . The similar scheme is commonly used in Photogrammetry and computer vision where the interest points are not evenly distributed over the image space. For instance, the significant improvement has been shown in homography estimation using numerical conditioning (Hartley and Zisserman, 2004, Hartley, 1997). Here, we

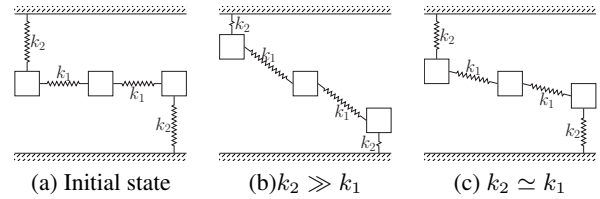


Figure 4: The mechanical system which can be interpreted as the diffusion equation and noise suppression in geometric surface reconstruction. If the boxes move vertically and initial state of this mechanical system is like (a), The final state of the system depends of the constant factor characteristic of the springs. If  $k_2 \gg k_1$  the boxes move toward walls and if  $k_2 \simeq k_1$ , the final state suggests that the boxes should be almost at the same level.

generalized the numerical conditioning in 2D image space for the sparse points in 3D Euclidean space.

### 4.4 Noise in sampled points

Up to this point it is assumed that the sampled points are accurate and the surface should be constructed in the way it passes through these sampled points. It may arguably be considered acceptable since these points are the only observations from the surface. As it will be discussed later, the laser scanner range error may be significant and the sampled points may be noisy. In addition, other systematic errors such as platform motion can adversely affect the quality of point cloud. Therefore, the noise of the sampled points should be suppressed in surface recovery. In (2) and (4) the second term assures that the surface passes through the sampled points. If the second term in these equations is replaced with  $\frac{\partial W}{\partial t} = \epsilon$  for the sampled points ( $\mathbf{X}' \in \mathbf{X}^i$ ), the surface does not pass through the sampled points anymore. The value of  $\epsilon$  which depends on the laser scanner range error balances how much the surface should follow the sampled points and how smooth it should be. Figure (4) shows a schematic mechanical system that represents the concept of diffusion and noise suppression. If the boxes in Figure (4) can move vertically and their final position depend on only the force created by springs, there will be an analogy between the surface recovery and the final state of the springs. The upper and lower walls are similar to the observed sampled points and the boxes are similar to the generated surface. The springs can be interpreted as the reliability of the sampled points. If the sample points are not noisy, these points are reliable and the surface should be forced to pass through the sampled points. If the sampled points are noisy, the sampled points may be smoothed. Analogously, if the springs attached to the walls are much stronger than the springs connecting the boxes,  $k_2 \gg k_1$ , the first and last boxes will get close to the walls (b). If we consider that the springs attached to the walls are not stronger than the springs connecting the boxes,  $k_2 \simeq k_1$ , and the boxes will be positioned in a smooth curve (c).

## 5. EXPERIMENTS AND RESULTS

The objects are segmented from dataset 5 in KITTI dataset (Geiger et al., 2012, Geiger et al., 2013). We focused on the moving objects but the results are valid for the static objects since the motion cue is not applied. Velodyne HDL-64E sensor is used to reconstruct surface of the object and the PointGray Flea2 color camera is used to color the point cloud and provide a better visualization. Figure (1.b) shows the point cloud superimposed on the image. The sampled points are 2 centimeters accurate in range, they are sampled in  $0.09^\circ$  angular resolution and laser scanner has 120 meters range (Geiger et al., 2012, Geiger et al., 2013). The images are 1.4 Megapixels,  $90^\circ$  opening angle and global shutter.

The cameras and laser scanner are calibrated and external and internal calibration parameters of the sensors are known. The image content is not used for surface reconstruction and we have used the images for only visualization purposes. To the knowledge of the authors there are not quantitative evaluation metrics for surface reconstruction. Also, there is no benchmark for 3D recovery in realistic scenario. Since laser scanner is one of the most accurate 3D surface reconstruction methods, there are not many options left to be used as the ground truth. As mentioned earlier, the generated surface can be applied for object recognition. The boundaries of the generated surface are very important in object recognition. Therefore, we discuss how well different approaches can maintain the boundaries. In order to evaluate the quality of each method, we project the generated surface into image and compare the boundaries of the generated surface to the boundaries of the object in image. In addition, topology of the object may be important. Topology of an object may be represented as the number of holes in the object. We have considered three object in our experiments: car, cyclist and pedestrian. Car has two (or more) windows that laser passes through them. Therefore, topology of the object is two. For the cyclist case, topology of the object is very complicated. The wheels of bike, handles of bike and hands of cyclist may create holes. For the pedestrian, surface do not have hole. It is arguably assumed that topology should be preserved in this paper and explained if topology of the surface can be preserved. Among many 3D surface recovery approaches, we chose moving least squares and Poisson surface reconstruction approaches and compared the generated surfaces with the results of the proposed approach in this paper. These approaches are popular in surface reconstruction and have been commonly used to up-sample point clouds. Our approach and MLS use local estimation of the surface, in contrast to Poisson surface reconstruction that uses a global optimization to recover the surface. All of these three approaches implicitly represent the surface, but Poisson surface reconstruction uses level sets concept to obtain 3D surface. Poisson surface reconstruction originally assumes that the surface is closed and some modifications may be required to apply this approach to open surfaces. Furthermore, the proposed method and MLS populate the point cloud and create more accurate point based representation of the surface. Poisson surface reconstruction is mesh based surface representation approach and surface is constructed from faces and vertices here.

For surface reconstruction of car, we had around 500 sampled points of the surface which are shown by red dots in Figure (5). The density of the points was around 20 centimeters which could change depending on the depth. As mentioned before, a grid of  $100 \times 100$  is generated for every sampled point, the points outside the elliptic mask are removed and the surface is generated with more than 1 million points. For MLS approach a surface with 1 centimeter resolution which has more than 500,000 points is generated. A mesh with more than 6000 faces is created in Poisson surface reconstruction. The vertices of the faces were extracted from every face, they were colored and demonstrated in Figure (5.c). Although mesh representation of the surface makes Poisson approach useful for many applications in computer graphics, we converted it to the point-based representation and the results of this method may be affected by this conversion. In order to preserve the quality of images, we have save the results in Encapsulated PostScript vector graphics (eps) format. Obviously, illustration of these points are not possible and the generated surfaces are down-sampled to be visible on the screen. The same procedure has been performed for cyclist and pedestrian in the scene. The results shows that the projected point cloud of car properly fit the image of car in our approach. In contrast, MLS and Poisson surface reconstruction methods are not able to recover the

boundaries of car properly and the generated surface exceeds the boundaries. Topology of a surface may be described with the number of holes in the surface. Our method is the only method that preserves topology and MLS and Poisson approaches fill in the holes (windows) in car. The surface of the cyclist has been reconstructed using three approaches. Similar to car, the boundaries of the cyclist cannot properly preserved in MLS and Poisson. In contrast, our approach shows better performance and the generated surface is similar to the cyclist. There are some miniature holes in cyclist using our method which it comes from the down-sampling for visualization purposes. Also, there are some sampled points on the leg of the cyclist and generated surface does not cover them. Since we use noise suppression, the generated surface does not follow the sampled points. However, the boundaries are still well-recovered in most regions. MLS and Poisson surface reconstruction fail to recover the boundaries of the pedestrian. Compared to car and cyclist, we believe our method fail to recover accurate boundaries of pedestrian. Our method connects the stretched legs and hands of the pedestrian to the body and creates a surface that may not be useful for applications such as object recognition.

## 6. CONCLUSION

In this paper, we proposed a novel method that uses local implicit representation of the surface and applied 3D generalization of super-resolution to populate the point cloud and generate a dense surface. The proposed approach has superior performance over the existing approaches and it generates an accurate surface that can be used in histogram of oriented depth for object recognition. However, the proposed approach may connect different parts of an object and fail to preserve the boundaries.

This paper can be extended in different ways: The image content can be used to sharpen the edges and improve the quality of the generated surface; A global term can be added to the regularizer to assure the global smoothness of the surface; The generated surface using this method can be used to recognize the objects and the accuracy of object recognition using this method may be studied.

## REFERENCES

- Alexa, M., Behr, J., Cohen-Or, D., Fleishman, S., Levin, D. and Silva, C. T., 2003. Computing and rendering point set surfaces. *IEEE Transactions on Visualization and Computer Graphics* 9(1), pp. 3–15.
- Boltho, M., Kazhdan, M., Burns, R. and Hoppe, H., 2009. Parallel poisson surface reconstruction. In: *Proceedings of the 5th International Symposium on Advances in Visual Computing: Part I, ISVC '09, Las Vegas, Nevada*, pp. 678–689.
- Boulch, A. and Marlet, R., 2012. Fast and robust normal estimation for point clouds with sharp features. *Comp. Graph. Forum* 31(5), pp. 1765–1774.
- Geiger, A., Lenz, P. and Urtasun, R., 2012. Are we ready for autonomous driving? the kitti vision benchmark suite. In: *Computer Vision and Pattern Recognition (CVPR)*, IEEE, Providence, Rhode Island, US, pp. 3354–3361.
- Geiger, A., Lenz, P., Stiller, C. and Urtasun, R., 2013. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)* 32(11), pp. 1231–1237.
- Hartley, R. I., 1997. In defense of the eight-point algorithm. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(6), pp. 580–593.

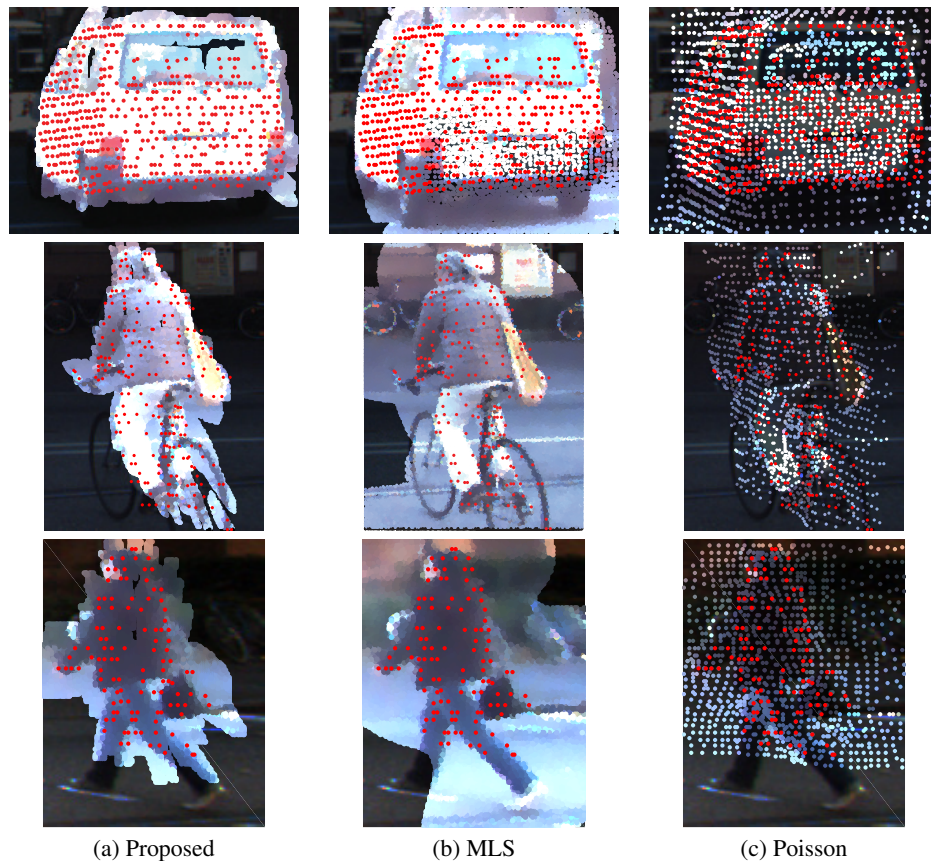


Figure 5: The proposed approach (a) is compared with MLS (b) and Poisson (c) surface reconstruction. The generated surface in our approach does follow boundaries of object in the image. However the MLS and Poisson approaches fail to maintain the boundaries of the object.

Hartley, R. I. and Zisserman, A., 2004. *Multiple View Geometry in Computer Vision*. 2 edn, Cambridge University Press, ISBN: 0521540518, New York, NY, USA.

Held, D., Levinson, J., Thrun, S. and Savarese, S., 2014. Combining 3d shape, color, and motion for robust anytime tracking. In: *Proceedings of Robotics: Science and Systems*, Berkeley, USA.

J. Dolson, J. Baek, C. P. and Thrun, S., 2010. Upsampling range data in dynamic environments. In: *Computer Vision and Pattern Recognition (CVPR)*, IEEE, San Francisco, US, pp. 1141–1148.

Kazhdan, M. and Hoppe, H., 2013. Screened poisson surface reconstruction. *ACM Trans. Graph.* 32(3), pp. 29:1–29:13.

Kazhdan, M., Bolitho, M. and Hoppe, H., 2006. Poisson surface reconstruction. In: *Proceedings of the fourth Eurographics symposium on Geometry processing*, Sardinia, Italy, pp. 61–70.

Lee, S., Wolberg, G. and Shin, S. Y., 1997. Scattered data interpolation with multilevel b-splines scattered data interpolation with multilevel b-splines scattered data interpolation with multilevel b-spline. *IEEE Transactions on Visualization and Computer Graphics* 3(3), pp. 228–244.

Lim, S. P. and Haron, H., 2014. Surface reconstruction techniques: a review. *Artificial Intelligence Review* 42(1), pp. 59–78.

Liu, M., Liu, Y. and Ramani, K., 2007. Anisotropic filtering on normal field and curvature tensor field using optimal estimation theory. In: *International Conference on Shape Modeling and Applications*, IEEE, Lyon, France, pp. 169–178.

Lorensen, W. E. and Cline, H. E., 1987. Marching cubes: A high resolution 3d surface construction algorithm. In: *Conference*

on Computer Graphics and Interactive Techniques (SIGGRAPH), SIGGRAPH '87, ACM, New York, NY, USA, pp. 163–169.

Osher, S. and Fedkiw, R. P., 2003. *Level set methods and dynamic implicit surfaces*. Applied mathematical science, Springer, New York, N.Y.

Pauly, M., Keiser, R., Kobbelt, L. P. and Gross, M., 2003. Shape modeling with point-sampled geometry. In: *Special Interest Group on Computer Graphics and Interactive Techniques (SIGGRAPH)*, San Diego, US, pp. 641–650.

Perona, P. and Malik, J., 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* 12(7), pp. 629–639.

Premevida, C., Carreira, J., Batista, J. and Nunes, U., 2014. Pedestrian detection combining RGB and dense LIDAR data. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, Chicago, US.

Sapiro, G., 2006. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, New York, NY, USA.

Szeliski, R., 2010. *Computer Vision: Algorithms and Applications*. 1 edn, Springer-Verlag New York, Inc., New York, NY, USA.

Tasdizen, T. and Whitaker, R., 2003. Anisotropic diffusion of surface normals for feature preserving surface reconstruction. In: *Proceedings of Fourth International Conference on 3-D Imaging and Modeling*, pp. 353–360.

Tasdizen, T., Whitaker, R., Burchard, P. and Osher, S., 2002. Geometric surface smoothing via anisotropic diffusion of normals.

In: Proceedings of the Conference on Visualization '02, VIS '02, IEEE Computer Society, Washington, DC, USA, pp. 125–132.

Tschumperlé, D. and Deriche, R., 2005. Vector-valued image regularization with pdes: A common framework for different applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(4), pp. 506–517.