TECHNISCHE
UNIVERSITÄT
DARMSTADT

Technical Report
13rp006-GRIS

Wavelet-based Surface Reconstruction from Multi-Scale Sample Points

Ronny Klowsky and Michael Goesele

January 2014
Originally written in 2013

Technische Universität Darmstadt, Germany
Department of Computer Science
Graphics, Capture and Massively Parallel Computing

Wavelet-based Surface Reconstruction from Multi-Scale Sample Points

Ronny Klowsky and Michael Goesele

TU Darmstadt

Abstract

Multi-view stereo reconstruction techniques yield inherently multi-scale point data typically fed into surface reconstruction algorithms. Following the intuition of scale space we assume that sample points originate from smoothed versions of the original surface. The smoothing can be characterized by a smoothing kernel that suppresses fine-scale structures. In this paper, we propose a surface reconstruction framework that correctly handles this multi-scale input data. We represent the surface using a multi-resolution analysis allowing us to reconstruct scales separately and to merge the sample points in frequency space. With an underlying wavelet basis we are able to locally model surface detail according to the surface properties or sample distribution. We first demonstrate the effectiveness of our method on a synthetic data set with known smoothing. For real-world data obtained by multi-view stereo we estimate the smoothing kernel and present reconstruction results with enhanced detail.

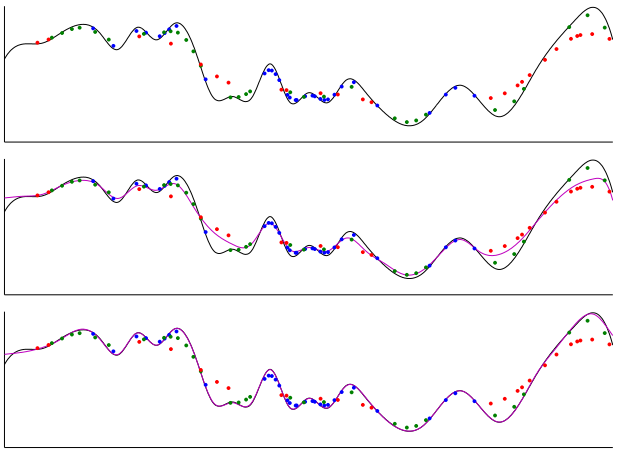


Figure 1: True surface (black) and multi-scale sample points (red–coarse, green–medium, blue–fine). *Top:* Input data. *Middle:* Reconstruction (magenta) treating all sample points equally. *Bottom:* Our reconstruction which takes scale into account and follows the true surface more clearly.

1 Introduction

Surface reconstruction from (unorganized) sample points is a well-researched area but also a continuous challenge. Popular methods include the pioneering work of Hoppe et al. [12], range image integration (VRIP) proposed by Curless and Levoy [8], and Poisson surface reconstruction by Kazhdan et al. [17]. Recent papers [10, 23, 26] give a detailed overview of the various methods available today. The focus of this paper lies on the multi-scale component inherent to many reconstruction techniques such as multi-view stereo. These approaches are able to deal with large scenes, for example comprising entire cities [1], and a mixture of various cameras ranging from mobile phones to digital SLRs. Drastically different object-to-camera distances and varying image resolutions automatically yield multi-scale sample points. When talking about scales of a surface we typically think of gradually removing detail structures of the original surface with a low-pass filter, which we model using a smoothing kernel. The main characteristic of multi-scale input data is that the samples are taken from successively smoothed versions (i.e., scales) in contrast to the simple case where all samples originate from the same scale (see the reconstruction in Fig. 1 top). In fact, it is commonly assumed that the input points are real point samples of the original surface implying that no or very little smoothing is involved (Fig. 1 middle). The first, and to our knowledge the only, to consider the multi-scale properties of sample points in a surface reconstruction algorithm are Fuhrmann and Goesele [10]. They essentially remove coarse-scale data points (originating from strongly smoothed versions of the origi-

nal surface) in areas where fine-scale points (less smoothed) with high confidence are available. Using this heuristic they are able to achieve impressive results on real world data sets. However, they rely on the correlation of resolution and scale suggesting that fine-scale sample points are usually present in higher resolution than coarse-scale samples. Also, discarding samples is a binary decision and information might be thrown away that could have been useful to close holes or even improve the fine-scale reconstruction. In summary, the fundamental problem of how to correctly merge multi-scale data points, i.e., combine the coarse- and fine-scale data instead of discarding the former, is still not convincingly solved.

In this paper we propose a reconstruction framework for 2.5D height field representations (Sec. 3) that explicitly models and incorporates the multi-scale properties of the input data (Fig. 1 bottom). We use the concept of multi-resolution analysis (multi-scale approximation) of the original surface. With the generating scaling functions and wavelets we are able to simultaneously decompose the surface in space and frequency domain. Given sample points with known or approximated smoothing kernel we show how the original surface can be recovered correctly. Hereby, our surface representation allows for locally varying degree of detail according to surfaces shape and sample point distribution. For practical application (Sec. 4) we add a regularization term to the surface recovery and integrate everything into one quadratic program. We further propose a specific wavelet representation and discuss the scale estimation in the context of multi-view stereo. Finally, we show results demonstrating the effectiveness of our method (Sec. 5) and conclude the paper with an outlook on future work (Sec. 6).

2 Related work

Classic surface reconstruction methods work on regularly sampled, some also on multi-resolution data points [8, 12, 17, 26]. The data is assumed to be single-scale which means that all points share the same noise model with the true surface as mean. A few recent approaches deviate from this paradigm. Klowsky et al. [20] use a Gaussian noise model but assign to each sample point a different standard deviation. They build a confidence volume represented in an octree and compute a minimum cut to reconstruct the surface (similar to other graph-cut based methods [3, 13, 27]). Fuhrmann and Goesele [10] integrate depth maps, similar to VRIP [8], into a hierarchical signed distance field (hSDF). They subsequently prune the hSDF removing coarse-scale data in regions where fine-scale data is available. The final surface is then extracted using a variant of the marching tetrahedra algorithm. Bailer et al. [2] handle the scale problem in a similar manner and also select locally the highest scale reconstruction available. Zach et al. [31] integrate range images into a global signed distance field and add a regularization term that minimizes the total variation (L1-regularization) of the SDF. Some of these methods support multi-resolution representations with locally varying level-of-detail and are capable of producing impressive results even on uncontrolled multi-view stereo data sets. However, none of them combines data from different scales while modeling the different degree of smoothing, i.e., sample points are still assumed to lie on the true surface when neglecting noise.

Pauly et al. [25] clarify the difference between multi-scale and multi-resolution surface representation. They use approximate low-pass filters to create a point-based multi-scale surface representation for the context of surface editing. Kazhdan [16] incorporates Fourier theory for surface reconstruction. The method aims at recovering the characteristic function of the solid by reconstructing its Fourier coefficients. While theoretically well founded the method requires summing over all input points to compute each single Fourier coefficient. This is computationally extremely expensive and implies that a single point influences the entire model which is counterintuitive. It also requires some heuristics to process non-uniformly sampled data. In a recent work, Digne et al. [9] propose a scale space meshing method that implements the mean curvature motion (MCM) on the raw point set. They reconstruct a smooth mesh first and then revert the MCM. It would be interesting to investigate handling of multi-scale data with this approach.

Several authors proposed surface reconstruction methods using smooth basis functions possibly integrated in a wavelet space. In the early work of Pastor and Rodríguez [24] spherical wavelets are used which naturally limits the application to objects that are topologically equivalent to a sphere.

Carr et al. [5] reconstruct smooth surfaces on the basis of smooth radial basis functions from noisy data. By computing the Fourier coefficients Kazhdan [16] actually represents the indicator function using dilations and translations of the sine function. Manson et al. [23] improve on this idea and apply wavelets instead, exploiting the local support to decrease complexity. A direct surface representation in Monge's form, as used in this paper, was proposed by Johnson et al. [15]. They use B-Splines and associated wavelets for scattered data reconstruction and give a theoretical error analysis. For better preserving depth discontinuities Ji et al. [14] seek for a piecewise smooth approximation in tight wavelet frames. In contrast to our work, all of these and other related methods in scattered data interpolation do not tackle the problem of multi-scale input data as we do in this paper. Also, the multi-scale structure of the basis functions is not exploited in order to adjust the granularity of the final reconstruction according to the input data.

3 Reconstruction framework

The basis of our reconstruction framework is a surface representation that allows us to operate on different scales of the surface. With that we can model surfaces with locally varying detail, either due to the surface itself or due to the distribution of the sample points. The classic Fourier transform is unsuited due to boundary handling issues and the missing locality. The latter also implies a constant frequency resolution over the entire space without taking into account the actual sample distribution. This involves the risk to hallucinate high frequency details in regions that are not sampled at all. In the following we first introduce our surface representation and describe afterwards how the surface can be recovered correctly from multi-scale sample points.

3.1 Surface representation

In this paper, we use an explicit surface representation assuming the surface can be parameterized as a height field $f(x)$. For simplicity the following derivation is for the 1D case $x \in \mathbb{R}$ but it can be easily extended to higher dimensions applying standard multi-dimensional wavelet construction as described by Mallat [22, Ch. 7.7]. We embed the surface in a multi-resolution analysis, written according to the notation of Stollnitz et al. [29, Ch. 7] as

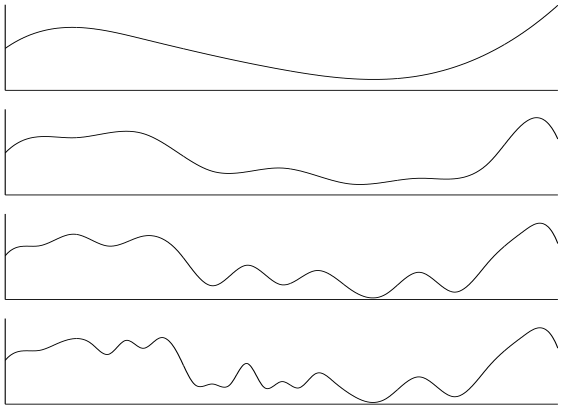


Figure 2: Multi-resolution analysis of a 1D surface. The detail level j increases from top to bottom and local surface details become visible.

$$\mathbf{V}_0 \subset \mathbf{V}_1 \subset \mathbf{V}_2 \ldots \subset L^2(\mathbb{R}) \tag{1}$$

where $\mathbf{V}_0$ can be thought of containing very smooth surfaces and with increasing index j in V_j more detail can be added (see Fig. 2). Eventually all possible surfaces $f \in L^2(\mathbb{R})$ are included. The complements of $\mathbf{V}_j$ in $\mathbf{V}_{j+1}$ are denoted by $\mathbf{W}_j$ such that

$$\mathbf{V}_{j+1} = \mathbf{V}_j + \mathbf{W}_j, j > 0. \tag{2}$$

The $\mathbf{V}_j$ are spanned by shifted and dilated versions $\phi_{j,l} = \phi(2^j x - l)$ of the father wavelet (or scaling function) ϕ and the $\mathbf{W}_j$ by shifted and dilated versions $\psi_{j,l}$ of the mother wavelet ψ, respectively. With that the surface f can be represented by its wavelet decomposition

$$f(x) = \sum_l c_{0,l}\phi_{0,l}(x) + \sum_{j=0}^{\infty}\sum_l d_{j,l}\psi_{j,l}(x) \tag{3}$$

where the $c_{0,l}$ denote the scaling function and the $d_{j,l}$ the wavelet coefficients. One can think of modeling the rough shape through the $c_{0,l}$ and then adding more and more details with increasing j by activating the $d_{j,l}$. Typically, the (effective) support of the $\psi_{j,l}$ decreases with increasing j so that surface details can be modeled locally. Since $\mathbf{V}_j = \mathbf{V}_0 + \mathbf{W}_0 + \ldots + \mathbf{W}_{j-1}$ one could also start with scaling functions of higher level. Also, in practice one has to cut off somewhere resulting in the more general representation:

$$f(x) \approx \sum_l c_{j_0,l}\phi_{j_0,l}(x) + \sum_{j=j_0}^{j_{\max}}\sum_l d_{j,l}\psi_{j,l}(x). \tag{4}$$

Without loss of generality we will in the following assume $j_0 = 0$ and for convenience we will use the equal sign although we refer to the approximation.

3.2 Surface recovery from samples

Given ideal point samples $(x_i, y_i)_{i=1,\dots,N}$ from the surface with $y_i = f(x_i)$ we have a linear system of equations

$$y_i = \sum_l c_{0,l}\phi_{0,l}(x_i) + \sum_{j=0}^{j_{\max}} \sum_l d_{j,l}\psi_{j,l}(x_i) \quad (5)$$

and the coefficients $c_{0,l}, d_{j,l}, 0 \le j \le j_{\max}$ as unknown variables. We can rewrite Eq. (5) in matrix form using a short vector notation as per $\psi_j = [\dots, \psi_{j,l}, \dots]^{\mathrm{T}}$:

$$\begin{pmatrix} \phi_0^{\mathrm{T}}(x_1) & \psi_0^{\mathrm{T}}(x_1) & \dots & \psi_{j_{\max}}^{\mathrm{T}}(x_1) \\ \vdots & & & \vdots \\ \phi_0^{\mathrm{T}}(x_N) & \psi_0^{\mathrm{T}}(x_N) & \dots & \psi_{j_{\max}}^{\mathrm{T}}(x_N) \end{pmatrix} \begin{bmatrix} c_0 \\ d_0 \\ \vdots \\ d_{j_{\max}} \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (6)$$

For multi-scale samples, i.e., sample points from the gradually smoothed surface, we assume that for each sample (x_i, y_i) the convolution kernel g_i is known such that

$$y_i = (g_i * f)(x_i). \quad (7)$$

This is a very general setup since we do not commit ourselves to a particular smoothing kernel. In standard scale-space, with a Gaussian convolution, it is just the standard deviation σ_i that varies among the samples but here we allow for other kernels (e.g., Laplacians, splines, or box filters) as well. Note that ideal point samples are also covered by simply using the Dirac delta function $g_i(t) = \delta(t - x_i)$. With Eq. (7) the linear system changes to

$$\begin{aligned} y_i &= (g_i * f)(x_i) \quad (8) \\ &= \left(g * \left(\sum_l c_{0,l}\phi_{0,l} + \sum_{j=0}^{j_{\max}} \sum_l d_{j,l}\psi_{j,l} \right) \right)(x_i) \\ &= \sum_l c_{0,l}(g_i * \phi_{0,l})(x_i) + \sum_{j=0}^{j_{\max}} \sum_l d_{j,l}(g_i * \psi_{j,l})(x_i). \quad (9) \end{aligned}$$

Again, we can write Eq. (9) in matrix form similar to Eq. (6) replacing the basis functions in the matrix with the respective convolutions. In the following we will denote the resulting matrix by Ψ resulting in the linear system

$$\Psi \mathbf{d} = \mathbf{y}. \quad (10)$$

with $\mathbf{d}$ covering scaling function and wavelet coefficients.

By definition wavelets fulfill $\int \psi_{j,l} = 0$ and with increasing scale j the $\psi_{j,l}$ become narrower. As a consequence, the convolution with the smoothing kernel $(g * \psi_{j,l})$ will diminish towards zero as j increases. In other words, a sample point's significance on the wavelet coefficients $d_{j,l}$ decreases. At the same time, a coarse scale sample point has less influence on coefficient $d_{j,l}$ than a fine scale sample point at the same position because the convolution kernel g is broader. In this way, we respect all given samples but prevent coarse scale samples from interfering with fine scale surface structures.

4 Surface reconstruction

Samples given in a real application are disturbed by noise and regions are irregularly sampled regarding density and scale. The consequence is that the linear system (10) cannot be solved exactly and we have to formulate an optimization problem. We introduce and discuss a regularization to avoid over-fitting and formulate the entire problem as a quadratic program. Thereafter we discuss how the smoothing kernel g_i can be estimated or even influenced in the context of multi-view stereo sample points and examine whether an optimal kernel exists. At the end of this section we review a particular wavelet family which we use in our experiments.

4.1 Optimization

The main problem we face when fitting a function to sample points is to reconstruct a smooth surface while still modeling the details. Besides the presence of noise and sparse sampling our model has a more inherent problem of over-fitting. When trying to recover fine scale details that are not sufficiently supported by the data, the entries of an entire row of the matrix Ψ vanish, and there is almost no control on the corresponding wavelet coefficients $d_{j,l}$. One way to counteract this is to decrease the maximum scale $j_{\max}$ but this effect might just be local and we do not want to decrease the overall detail level according to the worst represented region.

Consequently, a regularization is necessary that prevents all kinds of over-fitting. We add a penalty on the second order derivatives similar to Calakli and Taubin [4] and solve the following optimization problem

$$\underset{\mathbf{d}}{\text{minimize}} \quad \frac{1}{N}\|\Psi\mathbf{d}-\mathbf{y}\|^2+\lambda\int\|Hf(x)\|^2\,dx \tag{11}$$

where f denotes the final surface represented as in Eq. 4. $Hf(x)$ is the Hessian containing the second-order partial derivatives of f and $\|Hf(x)\|$ is the Frobenius norm of the matrix $Hf(x)$. Note that the smoothing term automatically affects regions with low-scale samples more than regions where high-scale samples are present because the corresponding coefficients are less restricted. We can reformulate the problem into a quadratic program

$$\underset{\mathbf{d}}{\text{minimize}} \quad \mathbf{d}^{\mathrm{T}}[\frac{1}{N}\Psi^{\mathrm{T}}\Psi+\lambda Q^s]\mathbf{d}-\frac{2}{N}\mathbf{y}^{\mathrm{T}}\Psi\mathbf{d} \tag{12}$$

where the matrix Q^s is the contribution of the second order derivative term. It consists of

$$Q^s_{\alpha,\beta}=\int<H\chi_\alpha(x),H\chi_\beta(x)>\,dx. \tag{13}$$

where we used the indices α and β to consecutively number the basis functions χ_α which are either scaling functions or wavelets. The matrix $Q=\frac{1}{N}\Psi^{\mathrm{T}}\Psi+\lambda Q^s$ is symmetric and positive definite, so problem (12) can be solved using a large-scale quadratic program or linear system solver.

4.2 Scale estimation

Until now we assumed that the convolution kernels g_i are known. However, it is not clear how to determine the kernel for given sample points in a real world application. All we can do is look for a good approximation of g_i. Klowsky et al. [18] give an approximation of the smoothing kernel for patch-based depth reconstruction. They show that the window based photo-consistency optimization between images leads to sample points that lie on a box filtered version of the original surface. The width of the box filter can be computed from the pixels footprint, i.e., the projected size of the pixel spacing in world space, multiplied with the window size in pixels. In a follow up paper [19] they apply a weighted photo-consistency optimization for depth reconstruction and show that the convolution kernel is equal to the applied weighting function (accordingly scaled to match the world-coordinate system). This not only allows us to estimate the convolution kernel g_i for the samples but to actively influence it during creation of the sample points. We will exploit this in our experiments in Sec. 5.

4.3 Optimal smoothing kernel

Before presenting the results of our method we want to spend some extra thought on choosing the optimal smoothing kernel. Ideally, the way the samples are generated matches the multi-resolution analysis used for the surface representation. In other words the significance of a sample point vanishes completely for all wavelet coefficients $d_{j,l}$ with j larger than the sample's scale. How can this be modeled? In the case of (semi-)orthogonal wavelets we have

$$<\phi_{0,k},\psi_{j,l}>=0,\ \text{for all } j\geq 0. \tag{14}$$

If we further assume symmetric scaling functions we can establish the following relationship between the inner product and the convolution

$$<\phi_{0,k},\psi_{j,l}>=\int\phi(t-k)\psi_{j,l}(t)\,dt \tag{15}$$

$$=(\phi*\psi_{j,l})(k)=0. \tag{16}$$

That is, if we had $g_i(t)=\phi(t)$ as the convolution kernel and samples at the integer positions $x_i\in\mathbb{Z}$ we would get

$$y_i=(\phi*f)(x_i)=\sum_l c_{0,l}(\phi*\phi)(x_i+l). \tag{17}$$

Having this kind of sample points we could solely solve for the scaling function coefficients $c_{0,l}$. Following this path, with $g_i(t)=\phi(2^j t)$ and sampling positions $x_i\in\{2^{-j}k,k\in\mathbb{Z}\}$ one could obtain the wavelet coefficients up to $d_{j-1,l}$. Note that in such a scenario the inherent over-fitting discussed in Sec. 4.1 is removed to a large extent.

Unfortunately, due to obvious reasons this is not achievable in practice: Firstly, we are very likely to not exactly hit the desired sampling positions and secondly we are incapable to (exactly) control the dilation of the smoothing kernel. In addition, we lose the possibility to exploit redundancy by sampling more positions than actually required. Therefore it remains a thought experiment and in practice we prefer to choose a smoothing kernel that behaves well and simplifies computations.

4.4 Spline wavelets on the interval

We now further specify the surface representation. Because the observed surface will always be of finite extent we can only identify corresponding coefficients. Consequently, there is no point in describing the surface using wavelets on the entire $\mathbb{R}^2$ (or $\mathbb{R}$) which would lead to border handling problems. Therefore we employ wavelets on bounded intervals, w.l.o.g. on $[0,1]$.

For our implementation we decided to use spline wavelets. From a variety of good reasons to do so (see Unser et al. [30]) we out point two: First, closed form solutions exist, not only for the basis functions but also for the convolution with, e.g., a Gaussian. Second, the basis functions are smooth allowing us to easily represent smooth surfaces. In the following we will shortly review the semi-orthogonal spline wavelets on $L^2([0,1])$ which were initially introduced by Chui and Quak [6] (see also Stollnitz et al. [29]). They are a natural extension of the semi-orthogonal spline wavelets on $L^2(\mathbb{R})$ developed by Chui and Wang [7].

A basis for V_j is given by the B-splines $B_{i,m,j}$ with $i = -m+1,\ldots,2^j-1$ which are defined as follows:

$$B_{i,m,j} = (t^{(j)}_{i+m} - t^{(j)}_i)[t^{(j)}_i,\ldots,t^{(j)}_{i+m}]_t\,(t-x)^{m-1}_+ \tag{18}$$

$$t^{(j)}_k = \begin{cases} 0, & k = -m+1,\ldots,0 \\ k2^{-j}, & k = 1,\ldots,2^j-1 \\ 1, & k = 2^j,\ldots,2^j+m-1 \end{cases} \tag{19}$$

where m denotes the spline order and the term $[\cdot,\ldots,\cdot]_t$ refers to the m-th divided difference of $(t-x)^{m-1}_+$ with respect to t. The inner scaling functions $B_{i,m,j}$, for $i = 0,\ldots,2^j-m$, are equal to the scaling functions for $L^2(\mathbb{R})$ which are just dilations and translations of the cardinal B-spline $N_m(x) = m[0,1,\ldots,m]_t\,(t-x)^{m-1}_+$:

$$\phi_{j,i}(x) = B_{i,m,j}(x) = N_m(2^j x - i),\ i = 0,\ldots,2^j-m. \tag{20}$$

The inner wavelets are equal to the Chui–Wang wavelets of order m:

$$\psi_{j,i}(x) = \frac{1}{2^{2m-1}} \sum_{k=0}^{2m-2} (-1)^k N_{2m}(k+1)\, B^{(m)}_{2i+k,2m,t^{(j+1)}_m}(x). \tag{21}$$

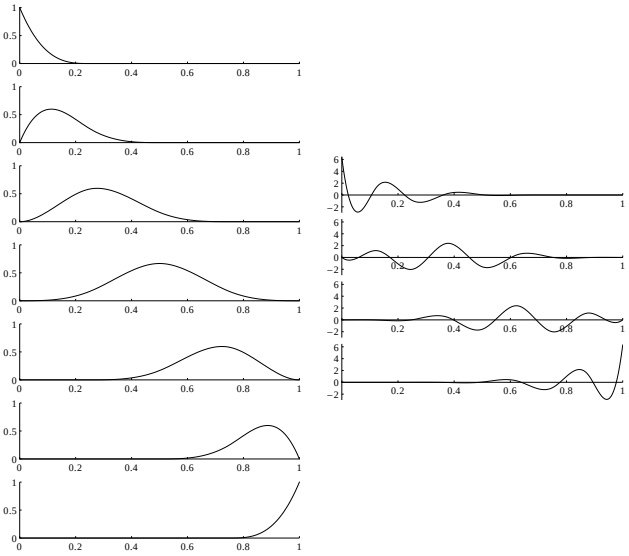


Figure 3: The seven scaling functions (*left*) and four wavelets (*right*) on the interval for $j = 2$ spanning $\mathbf{V}_3$.

We refer to Chui and Quak [6] on how to construct the border wavelets in the general case. For cubic splines ($m = 4$) the coefficients of the refinement equation are given in [29, App. B]). Figure 3 shows the scaling functions and wavelets for $j = 2$. For the results presented in next section we use the tensor product wavelets on $[0,1]^2$.

5 Results

To apply our technique in practice, all we have to do is set up the matrices in Equation (12) and solve the quadratic program in order to determine the coefficient vector d for the wavelets. In all experiments we assume that the final surface can be described as a height field $z = f(x,y)$ with $(x,y) \in [0,1]^2$. This is realized using a rigid transformation plus an additional scaling, thus easily invertible after reconstruction. Instead of comparing our method to some arbitrary other method we decided to use our framework and neglect scale setting the convolution kernel g to be the Dirac delta function. This corresponds to the commonly used assumption of other methods that sample points lie on the true surface (plus zero-mean noise).

5.1 Synthetic data

We start with a synthetic data set where we know both the ground truth surface (see Fig. 4 (*left*)) and

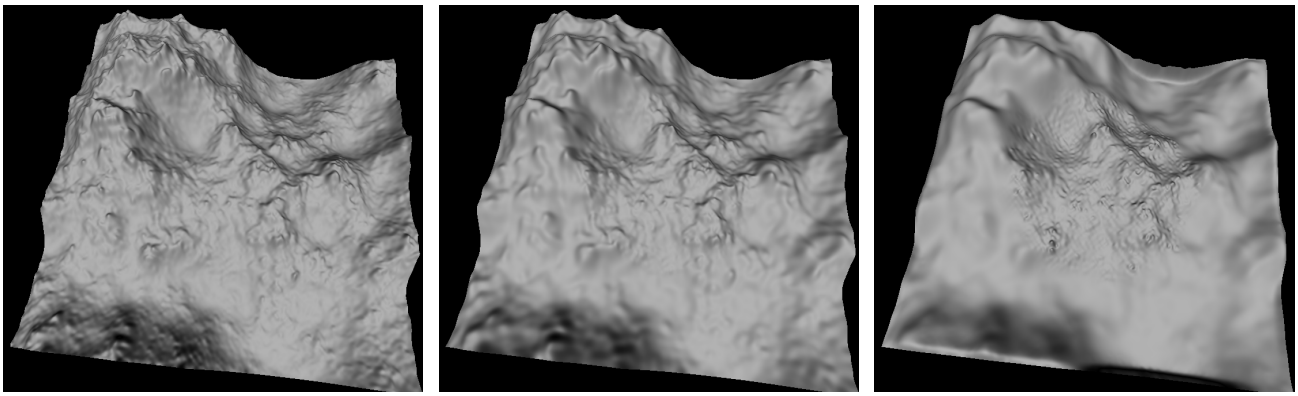


Figure 4: *Left:* Ground truth surface from which we generate low- and high-scale samples. *Middle:* Our reconstruction taking scale into account. *Right:* Treating all samples as real point samples neglecting the scale.

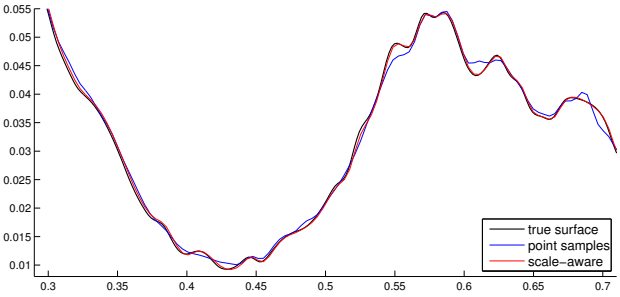


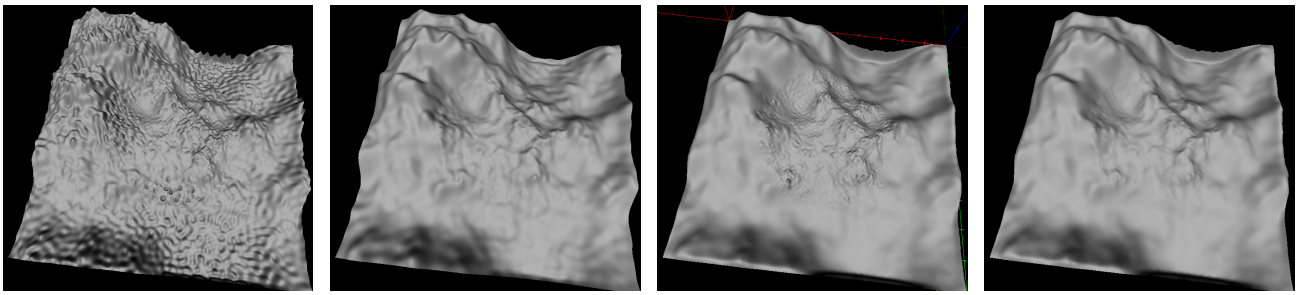
Figure 5: A segment of the central horizontal scanline through the geometry in Fig. 4 showing that our scale-aware reconstruction accurately follows the ground truth.

its wavelet decomposition. The input to our method are sample points from the convolved version of this surface using a Gaussian with known standard deviation σ. We generate 20,000 sample points from which $\frac{4}{5}$ are uniformly sampled over $[0,1]^2$ with $\sigma = 0.01$ (low-scale), and $\frac{1}{5}$ are uniformly sampled on a centered circle with radius 0.25 with $\sigma = 0.002$ (high-scale). For the reconstruction we use $j_0 = 4$ and $j_{\max} = 6$, i.e., using scaling functions $\phi_{4,\cdot}$ and wavelets $\psi_{4,\cdot}, \psi_{5,\cdot}, \psi_{6,\cdot}$. The smoothness weight is $\lambda = 10^{-12}$. The result of our method can be seen in Fig. 4 (*middle*) compared to neglecting scale in Fig. 4 (*right*). The benefit of taking the scale into account, even in the areas with only low-scale sample points, is clearly visible. Fig. 5 shows a segment from the center horizontal scanline that confirms this impression.

In Fig. 6 we demonstrate the effect of the smoothness weight. We reconstruct effectively on the same scale, that is in V_7, but using scaling functions $\phi_{6,\cdot}$ and wavelets $\psi_{6,\cdot}$. Now, the smoothing kernel is roughly as big as the basis function and there is only very small or no data force on the basis function coefficients leading to "ripple" artifacts. The same effect can be caused by under-sampling. Then the smoothness weight λ has to be chosen accordingly to prevent introducing high-frequency artifacts.

5.2 Real-world data

To test our algorithm on real-world data we took 174 images of a relief on a stone wall (see Fig. 7). We registered the images using structure-from-motion [28] and reconstructed depth maps per view using a multi-view stereo implementation similar to Goesele et al. [11]. In contrast to them we use a weighted photo-consistency optimization. More precisely we use a patch of size 21×21 pixels in image space and apply a Gaussian with $\sigma = 4$. We use such a big patch to get less noise in the reconstruction and to achieve a reasonably sized smoothing kernel to better visualize the effect of our method. The input images have a resolution of about 1000×666 pixels. According to Klowsky et al. [19] we can then estimate the smoothing kernel g to be a Gaussian as well with a scaled standard deviation depending on the internal camera parameters and the estimated depth. In order to meet the height field assumption we fit a plane to the feature points obtained by structure-from-motion and compute a transformation that maps it on the $x,y-$plane. As input to our method we merge the reconstructed



(a) (b) (c) (d)

Figure 6: The starting scale $j_0 = 6$ is chosen higher than in Fig. 4 resulting in less supported coefficients of the scaling function. *(a)* A small smoothness weight ($\lambda = 10^{-12}$) can lead to artifacts. *(b)* Choosing a larger weight ($\lambda = 10^{-10}$) fixes this problem. *(c)+(d)* Using the same smoothness weights ($\lambda = 10^{-12}$ and $\lambda = 10^{-10}$, respectively) but assuming all samples are real point samples. This variant is naturally less sensitive to the smoothness weight but also preserves less detail.

Figure 7: Example input images of the Relief data set.

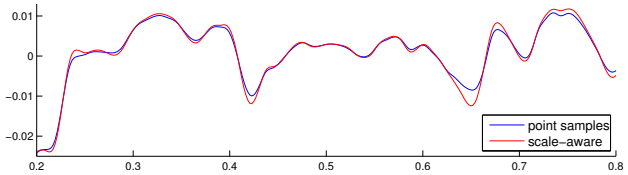


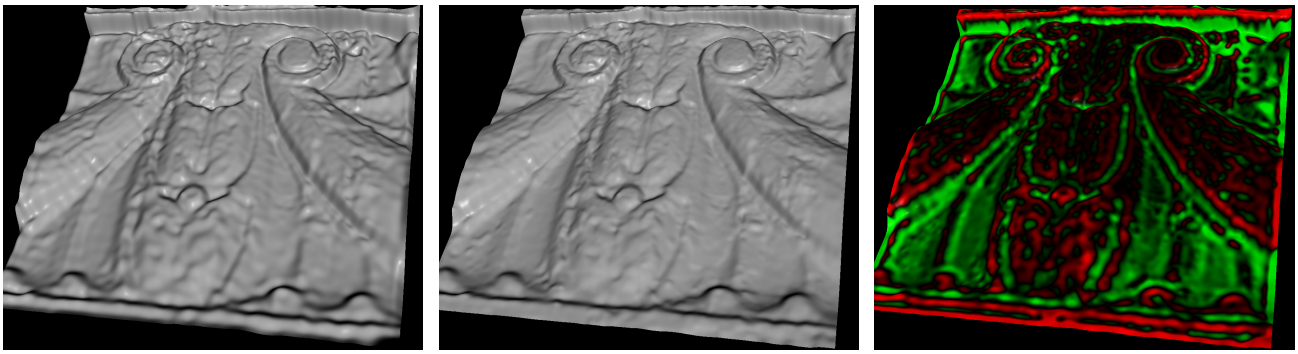
Figure 9: A profile of the Relief reconstruction (see Fig. 8) showing that our scale-aware reconstruction preserves more detail than treating all samples as real point samples.

points from 6 depth maps covering a range of about factor 3 in scale, i.e., $\sigma_{\max} \approx 3\sigma_{\min}$. This yields a total of about 1.6 million points.

We reconstruct a surface using $j_0 = 5$ and $j_{\max} = 6$, i.e., using $35^2 = 1,225$ scaling functions $\phi_{5,\cdot}$ spanning V_5, $3,264$ wavelets $\psi_{5,\cdot}$ spanning W_5, and $12,672$ wavelets $\psi_{6,\cdot}$ spanning W_6. In total we optimize for $17,161$ basis function coefficients. Fig. 8 shows the comparison between our scale-aware (*left*) reconstruction and using the same setup but ignoring scale (*center*), i.e., treating all samples as real point samples with zero-mean noise. Detail in the middle and lower part of the rendering is emphasized while some artifacts from multi-view stereo become more visible.

6 Conclusion and future work

We present a general surface reconstruction framework that incorporates the (multi-)scale property of the samples points. To our knowledge we are the first to dissolve the paradigm of point samples that lie on the true surface but still incorporate all data in the reconstruction process. Using the concept of multi-resolution analysis we can merge the sample points in frequency space while still maintaining locality due to the wavelet basis. On synthetic data we demonstrate clearly that our method correctly integrates the multi-scale input data. The real-world example indicates the improvement of our method as well, however, we have to struggle with registration errors and multi-view stereo artifacts. As pointed out by Klowsky et al. [19] the modeling of the multi-view stereo reconstruction is imperfect and thus the estimated smoothing kernel is not accurate. Experience from the image domain (e.g. Levin et al. [21]) suggests that a better kernel estimate will



(a) (b) (c)

Figure 8: Reconstruction using $j_0 = 5$ and $j_{\max} = 6$. *(a)* Taking scale into account preserves more detail compared to treating all samples as real point samples in *(b)*. The colored mesh *(c)* has vertex positions identical to *(b)* and the vertex colors encode the differences in height compared to *(a)*. Changes mainly affect the edges since we amplify high frequencies.

likely improve reconstruction quality. The biggest limitation of our method is probably the current restriction to height fields. Using an implicit surface representation, e.g., the signed distance field, it would be possible to extend the method to a more general class of surfaces. We do, however, face the problem that it is still unclear how reconstruction techniques affect the signed distance field.

Acknowledgements This work was supported in part by the DFG Emmy Noether fellowship GO 1752/3-1.

References

[1] Sameer Agarwal, Noah Snavely, Ian Simon, Steven M. Seitz, and Richard Szeliski. Building Rome in a day. In *ICCV*, 2009.

[2] Christian Bailer, Manuel Finckh, and Hendrik P.A. Lensch. Scale robust multi view stereo. In *ECCV*, 2012.

[3] Yuri Boykov and Vladimir Kolmogorov. Computing geodesics and minimal surfaces via graph cuts. In *ICCV*, 2003.

[4] Fatih Calakli and Gabriel Taubin. SSD: Smooth signed distance surface reconstruction. In *Pacific Graphics*, 2011.

[5] J. C. Carr, R. K. Beatson, J. B. Cherrie, T. J. Mitchell, W. R. Fright, B. C. McCallum, and T. R. Evans. Reconstruction and representation of 3D objects with radial basis functions. In *SIGGRAPH*, pages 67–76, New York, NY, USA, 2001. ACM.

[6] Charles K. Chui and Ewald Quak. Wavelets on a bounded interval. In *Numerical Methods in Approximation Theory*, 1992.

[7] Charles K. Chui and J. Z. Wang. On compactly supported spline wavelets and a duality principle. In *Trans. of the American Mathematical Society*, 1992.

[8] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *SIGGRAPH*, 1996.

[9] Julie Digne, Jean Michael Morel, Charyar-Mehdi Souzani, and Claire Lartigue. Scale space meshing of raw data point sets. *Computer Graphics Forum*, 30(6):1630–1642, 2011.

[10] Simon Fuhrmann and Michael Goesele. Fusion of depth maps with multiple scales. In *SIGGRAPH Asia*, 2011.

[11] Michael Goesele, Noah Snavely, Brian Curless, Hugues Hoppe, and Steven M. Seitz. Multi-view stereo for community photo collections. In *ICCV*, 2007.

[12] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. Surface reconstruction from unorganized points. In *SIGGRAPH*, 1992.

[13] Alexander Hornung and Leif Kobbelt. Robust reconstruction of watertight 3D models from non-uniformly sampled point clouds without normal information. In *EG SGP*, 2006.

[14] Hui Ji, Zuowei Shen, and Yuhong Xu. Wavelet frame based scene reconstruction from range data. *Journal of Computational Physics*, 229(6):2093–2108, 2010.

[15] Michael J. Johnson, Zuowei Shen, and Yuhong Xu. Scattered data reconstruction by regularization in b-spline and associated wavelet spaces. In *Approximation Theory*, 2009.

[16] Michael Kazhdan. Reconstruction of solid models from oriented point sets. In *EG SGP*, 2005.

[17] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *EG SGP*, 2006.

[18] Ronny Klowsky, Arjan Kuijper, and Michael Goesele. Modulation transfer function of patch-based stereo systems. In *CVPR*, 2012.

[19] Ronny Klowsky, Arjan Kuijper, and Michael Goesele. Weighted patch-based reconstruction: linking (multi-view) stereo to scale space. In *SSVM*, volume 7893 of *LNCS*, pages 234–245, 2013.

[20] Ronny Klowsky, Patrick Mücke, and Michael Goesele. Hierarchical surface reconstruction from multi-resolution point samples. In *LNCS*, volume 7474, 2012.

[21] A. Levin, Y. Weiss, F. Durand, and W.T. Freeman. Understanding blind deconvolution algorithms. *PAMI*, 33(12):2354–2367, 2011.

[22] Stéphane Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 3rd edition, 2008.

[23] Josiah Manson, G Petrova, and Scott Schaefer. Streaming surface reconstruction using wavelets. In *EG SGP*, 2008.

[24] Luis Pastor and Angel Rodríguez. Surface approximation of 3D objects from irregularly sampled clouds of 3D points using spherical wavelets. In *ICIAP*, 1999.

[25] Mark Pauly, Leif P. Kobbelt, and Markus Gross. Point-based multiscale surface representation. In *ACM TOG*, 2006.

[26] Shy Shalom, Ariel Shamir, Hao Zhang, and Daniel Cohen-Or. Cone carving for surface reconstruction. In *SIGGRAPH Asia*, 2010.

[27] Sudipta N. Sinha, Philippos Mordohai, and Marc Pollefeys. Multi-view stereo via graph cuts on the dual of an adaptive tetrahedral mesh. In *ICCV*, 2007.

[28] Noah Snavely, Steven M. Seitz, and Richard Szeliski. Skeletal sets for efficient structure from motion. In *CVPR*, 2008.

[29] Eric J. Stollnitz, Tony D. DeRose, and David H. Salesin. *Wavelets for Computer Graphics: Theory and Applications*. Morgan Kaufman, 1996.

[30] Michael Unser. Ten good reasons for using spline wavelets. In *SPIE Conf. on Mathematical Imaging*, volume 3169, 1997.

[31] Christopher Zach, Thomas Pock, and Horst Bischof. A globally optimal algorithm for robust TV-L1 range image integration. In *ICCV*, 2007.