DTU Library

Ultra Fast Optical Sectioning: Signal preserving filtering and surface reconstruction

Jensen, Rasmus Ramsbøl; Poel, Mike van der; Larsen, Rasmus; Paulsen, Rasmus Reinhold

Link back to DTU Orbit

Ultra Fast Optical Sectioning: Signal preserving filtering and surface reconstruction

Rasmus R. Jensen, Mike van der Poel, Rasmus Larsen, and Rasmus R. Paulsen

Informatics and Mathematical Modelling, Technical University of Denmark
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark
`raje@imm.dtu.dk`, `www.imm.dtu.dk`

Abstract. In 3D surface scanning it is desirable to filter away *bad* data without altering the quality of the remaining *good* data. Filtering of raw scanner data before surface reconstruction can minimize the induced error and improve on the probability of reconstructing the true surface. If outliers consist of actual data such as hair, and not just evenly distributed noise, these outliers tend to err smoothing algorithms away from the wanted result. We present a novel algorithm based on a Markov Random Field that uses a distance constraint to robustly classify a 3D scan volume. Through this classification a signal preserving filtering of the data set is done. The remaining data are used for a smooth surface reconstruction creating very plausible surfaces. The data used in our work comes from a newly developed hand held 3D scanner. The scanner is an Ultra Fast Optical Sectioning scanner, which is able to extract high quality 3D surface points from 2D images recorded at over 3000 fps. The scanner has been developed for digital impression taking in the dental area. Our work relates to future in-ear scanning for fitting custom hearing aids without impression taking.
-Keywords: *3D scanning, Markov Random Field, computer vision, surface reconstruction, noise filtering.*

1 Introduction

3D surface acquisition is an established and active research and development area. Novel applications and devices continues to emerge, where the data scale ranges from minuscule in microscopy optical sectioning [7] to large scale aerial surface laser scanning of the earth [13]. A variety of scanners and cameras exist; each with their own strengths and weaknesses. While some scanners produce 3D surface data, scanning an object from one direction is known as 2.5D scanning as it only portrays the object from one side and does not provide a full 3D model. To construct a 3D model several 2.5D scans need to be patched creating a full reconstruction [8].

We have worked with a new scanner, the Ultra Fast Optical Sectioning TRIOS scanner from 3Shape[1], which has been developed to facilitate 3D impression taking in the dental area. Our work is a study on how to reconstruct surfaces in the presence of structured noise. The study is a preliminary study,

which relates to fitting custom hearing aids, where the construction of an in-the-ear scanner would make the ear canal impression step obsolete.

Anticipating the problem of structured noise from hair in the ear canal we want an algorithm that filters the data and leaves only valid surface data. As no scanner has yet been produced that will actually go into the ear, we have used data of hairy arms and bearded chins recorded by the TRIOS scanner. The scanner produces high quality data with some very sparse *salt-and-pepper* type noise and also good scanning of actual hair strands. As the scanner is a dental scanner with both high precision and accuracy, we were faced with a specific problem of removing only the hair without degrading the remaining data.

Simple mean or median filtering can make any surface fair (if one smoothens enough) but these filters also distort the data set and are useable only when the noise is Gaussian (mean filtering) or when the number of outliers are few (median filtering). An adaptive application of such filters used only on outliers can remove noise without too much degrading of the data, but such an approach would break down in areas with a lot of outliers. Generally, local filtering does only preserve local structure; for areas with a lot of hair a global method is needed.

Implicit functions err towards outliers and noise. Splines [10] are a well known tool for creating both smooth curves and surfaces, but if noise is not Gaussian the smoothing will be skewed. This is the case in our data, where outliers are mainly found above the surface. Splines are also continuous and therefore do not handle discontinuity well.

A way of removing outliers in a data set is to use random sampling such as RANSAC [4] and fit to the random sample until a fit matches the data well. Even though this removes the influence of outliers it does not guarantee an optimal match. A RANSAC approach is described in [11], where an algorithm is created that finds basic shapes and structures in noisy data sets.

There exist a large body of literature covering noise properties and handling in direct surface scanners [3]. However, the used scanning device is not directly comparable with devices previously investigated.

To maintain as much of the good data as possible in our scans, we have solved our problem based a *Markov Random Field* (MRF) formulation on a 3D voxel grid. An early description of MRFs in 2D image analysis is on the noise removal in dirty pictures addressed in [2]. The novelty of our work is the use of a 3D MRF with a distance based smoothness prior that classifies the data set into surface data and not surface data. The classification allows for a signal preserving filtering of the data set before any actual surface reconstruction.

2   Data

The data come from an Ultra Fast Optical Sectioning TRIOS scanner, which records a stream of 2D images. Approximately 130 images are collected into a set that constitutes a voxel volume. From known changes in the scanner during the volume acquisition a scan surface is constructed which can be converted

into a real world coordinate system through a calibration step. The scanner computes the voxels that defines the interface between air and solid material as seen from the scanners viewpoint. Since the scanner is known to be above the surface it is possible to label the voxels along the depth axis as either *above* or *below* the sought surface. When the scanner firmware has determined which voxels belong to the interface, the row, column and depth volume is transformed into real world 3D points using the calibrating parameters. The resulting 2.5D point cloud represents the interface between air and solid material as seen from the scanners viewpoint. To perform a full 3D surface scan, the scanner is moved around the object and the partial scans are merged together using a proprietary algorithm.

The scanner is normally used for digital impression taking in dental work, which requires a very high accuracy. Therefore, the quality of scans are very high and the noise levels due to the scanner hardware is minimal. However, real physical objects as for example hair will also be captured by the scanner. In the current application (direct ear-scanning) we are interested in the true surface of the ear and therefore hair is considered noise and should be removed from the scan. In the following we consider hair as being structured noise.

A 2.5D scan can also be considered a depth map, where the pixel value reflect the distance to the object. Figure 1 shows four depth maps of scans from different surfaces. The scans constitute a range from very bad to perfect, and the algorithm should be able to handle all examples. When there is a lot of structured noise in the scan, we need an algorithm that does not break down but does a good filtering leaving only the actual surface data (even if it is very little).
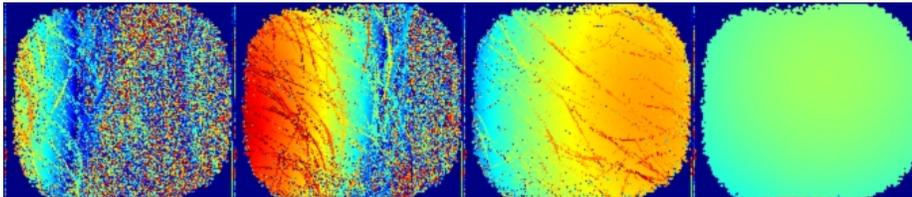

Fig. 1. Four depthmaps of scan surfaces. From left to right: Little surface coverage with hair, half coverage with hair, full coverage with hair, full coverage without hair.

3 Markov Random Field volume classification

The raw output from the scanner (and the scanner firmware) is a voxel set where the scanner is virtually placed above the voxel volume looking down the depth direction. Each voxel is labeled as being either *above* or *under* the *scan surface*. The scan surface, S_{scan}, is the initial surface that can be extracted as the interface between the *above* and *under* voxel sets. However, S_{scan} is noisy (in the

sense that hair is present in the scan) and does not represent the true underlying skin surface. We aim to produce a consistent and locally smooth skin surface, S_{skin}, from the data set. In order to re-label the voxel set and thereby implicitly producing, S_{skin}, a Markov Random Field (MRF) classification/regularisation approach is chosen. In the following, a short introduction to MRFs and a description of the chosen models are given.

3.1 The volume random field

We define a random field with spatial voxel positions $\{v_1, v_2, \ldots, v_n\}$ in the volume V with the index set I. In this set each voxel v_i takes a value x_i from the binary label set $L = \{under, above\}$, where $under$ is under and $above$ is over the skin surface. Notice that we make a distinction between scan surfaces S_{scan} and skin surfaces S_{skin}; the classification relates to the latter. All values of x_i are represented by the vector $\mathbf{x}$, which is the configuration of the random field.

A neighbourhood system to v_i is defined as $N = \{N_i | i \in I\}$ for which it holds that $i \notin N_i$ and $i \in N_j \Leftrightarrow j \in N_i$. A random field is said to be a Markov field, if the probability P of any configuration of $\mathbf{x}$ satisfies the positivity property:

$$P(\mathbf{x}) > 0 \qquad \forall \mathbf{x} \in L \tag{1}$$

And the Markovian property:

$$P(x_i | \{x_j : j \in I \backslash \{i\}\}) = P(x_i | \{x_j : j \in N_i\}) \tag{2}$$

Or in other words the probability of x_i given the index set $I \backslash \{i\}$ is the same as the probability given the neighbourhood of i. Our use of neighbourhood is limited to the direct 6-neighbours in the volume. The goal is to compute the configuration of the field that maximizes the probability.

3.2 Defining the Markov Random Field

We aim to produce a labelling $\mathbf{x}$ of the voxel volume V, such that the boundary of the labelling coincides with the skin surface. A likelihood term and two priors are defined on the following:

- Under scan surface is likely to be under skin surface ($under$), while above scan surface is likely to be above skin surface ($above$).
- Skin surface points are in the vicinity of scan points.
- Skin surface is locally smooth.

Here $surface\ points$ and $scan\ points$ are defined as the local interface between $above$ and $under$ labelled voxels. Using these priors a MRF is created for which a minimal energy problem is defined using the following:

Likelihood term: This term is based on the relative position of a voxel and the scan surface:

$$\Phi(v_i | x_i) = -\log P(v_i | x_i) \tag{3}$$

Generally, we expect the scan and skin surfaces to coincide and therefore make a simplification of the likelihood function, such that voxels below the scan surface S_{scan} have low energy if labelled *under* and high energy if labelled *above* and vice versa for voxels above the scan surface. The energy then becomes:

$$\Phi(v_i < S_{\text{scan}}|x_i) = \begin{cases} 1 & x_i = under \\ 0 & x_i = above \end{cases}$$
$$\Phi(v_i > S_{\text{scan}}|x_i) = \begin{cases} 0 & x_i = under \\ 1 & x_i = above \end{cases} \tag{4}$$

This somewhat loosely defined term alone would just produce the scan surface.

Vicinity prior: The skin surface should be close to scan surface points. This is induced by adding an energy penalty to changes in label, which relates to the distance from the voxel v_i to the nearest scan point S_{scan}:

$$\lambda(x_i, x_j) = \begin{cases} K_{\text{dist}} \cdot \text{dist}(v_i, S_{\text{scan}}) & x_i \neq x_j \\ 0 & x_i = x_j \end{cases} \tag{5}$$

The distance is approximated using an Euclidean distance transform (EDT) [5]. This is a fast linear time algorithm that approximates distance in a number of sweeps. This penalty should only affect voxels that are not in the immediate neighbourhood of scan surface points, which is why 1 is subtracted from the distances, such that both an actual surface point and its direct neighbours have distance 0. The vicinity constraint mainly effects areas with high discontinuity but it also forces the resulting surface to be true to the data in areas with continuity.

Smoothness prior: Neighbouring voxels are expected to have the same label with higher probability than having different labels, therefore an energy penalty is given to adjacent voxels with different label:

$$\psi(x_i, x_j) = \begin{cases} K_{ij} & x_i \neq x_j \\ 0 & x_i = x_j \end{cases} \tag{6}$$

This is a general smoothness constraint.

Combining the likelihood term with the vicinity and smoothness prior, we get the following energy minimization problem for the scan volume:

$$E(\mathbf{x}) = \sum_{i \in I} \left(\Phi(v_i|x_i) + \sum_{j \in N_i} (\lambda(x_i, x_j) + \psi(x_i, x_j)) \right) \tag{7}$$

Where $\mathbf{x}$ is the classification of the whole volume. The vicinity constant K_{dist} and the smoothness constant K_{ij} relates to each other and the likelihood term, which we defined to be 0 or 1. The solution to the *MRF* ensures maximum probability with the constraint that the labelling is both highly consistent with the data and smooth. To solve the minimization problem the *Graph Cut* algorithm [6] is used. This algorithm is an efficient way to find the optimal solution for such binary problems. From the resulting MRF classification a new surface S_{MRF} is extracted as the interface between the *under* and *above* labelled voxels.

4 Filtering based on the MRF solution

With a solution to the MRF and a new surface estimate S_{MRF} the changes compared to the original surface S_{scan} can be analyzed. As the MRF is set up so create a surface that coincides with the skin surface S_{skin}, we set up a filtering based on the following:

$$S_{\mathrm{skin}} = S_{\mathrm{scan}} \cap S_{\mathrm{MRF}} \quad \wedge \quad S_{\mathrm{hair}} = S_{\mathrm{scan}} \setminus S_{\mathrm{MRF}}$$

This results in a pure skin surface estimate and an estimate of the structured noise (hair) removed from the original scan to create the skin surface. We have used the strict definition but to give some flexibility one might add a threshold on how much is considered a change when comparing the original data with the MRF surface. Figure 2 shows examples of the resulting depthmaps of the *MRF* based filtering. Even though it is difficult to quantify the result, clearly both hair and noise are removed leaving only the smooth skin surface data.
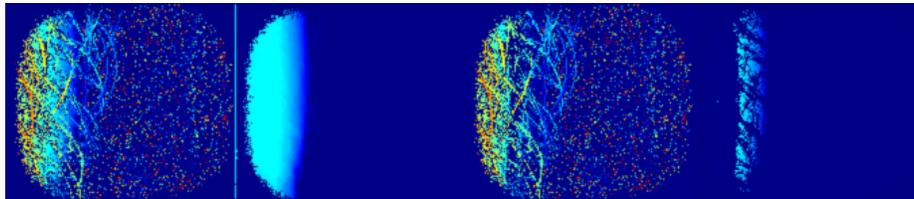
(a) Very little skin with hair
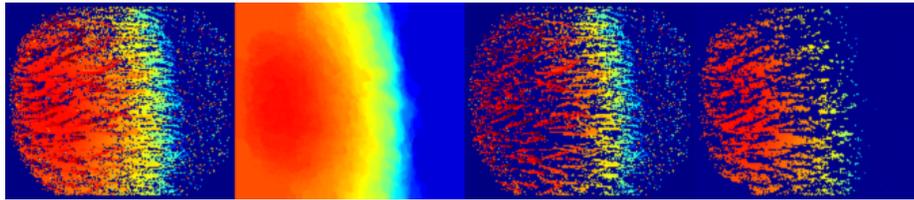
(b) Half skin coverage with hair

Fig. 2. From left to right: depth map of original scan, surface based on volume classification, changes made in new surface, unchanged data in new surface. It is especially worth noting that the first scan is of very poor quality. The scanner is only focusing on very little skin surface and on top of that there is a lot of hair. In spite of this the algorithm returns the small amount of actual skin surface present in the scan.

Setting the relations between the likelihood term, the vicinity and smoothness prior is not trivial and based on *trial-and-error*. However, a reasonable approach is (at least for this type of data) to set the smoothness prior K_{ij} to 0, while incrementing the vicinity prior K_{dist} until a good result is achieved (Fig. 3(a)). This parameter effects areas with discontinuity such as skin to hair,

while it actually forces the algorithm to be true to the input data in areas with continuity such as skin to skin. Setting the parameter removes most of the hair strands leaving only a little stubble, which can then be removed by adding the smoothness constraint. Figure 3(b) shows the difference between surface data found using only the vicinity prior and using both vicinity and smoothness, the difference is seen as green, and the blue is the filtered surface data.

Even though the *MRF* solution actually creates a surface without holes, this surface tends to have bumps where the surface is closed below the hair strands. This is why only the unchanged part of the MRF surface should be kept. These remaining data are a much better starting point for a smooth surface reconstruction, as these data belong to the actual surface. Figure 3(c) shows a surface reconstruction based on the filtered point set. The surface has been reconstructed with the Markov Random Field Surface Reconstruction [9]. This algorithm uses a Markov Random Field to regularize a point distance field and it creates plausible hole filling, where data is missing. If computation time is a factor a simple 2D Delaunay triangulation [12] (considering the data as a 2D height map) would produce a reasonably fair surface.

5   Conclusion

In this paper data from a novel surface scanner has been used in an approach to remove structured noise from raw scanner data. The approach is based on re-labelling a voxel set using a Markov Random Field classification and extracting the sought surface as the interface between two voxel labels.

We have achieved good results on filtering noise and hair from our surface scans. Even though it is hard to quantify the ability to filter both noise and hair, qualitative visual inspections of the results are very promising. Because of the high quality expectance of the scans it is a strong point that the data is filtered such that the remaining data is unchanged.

Our algorithm has been implemented in Matlab with crucial parts in C++ `.mex`-files and it filters in a matter of a few seconds on a 2.8 GHz Intel processor laptop. There is good reason to believe that the surface filtering could be done in realtime with a full implementation in a precompiled programming language.

One could argue that a full volume classification as part of the filtering is not necessary. This is meant in the sense that large parts of the scan volume will be *far* from the actual surface and therefore a smarter way of classification close to the surface could speed up the process. This would however complicate the approach and is a topic for future research.
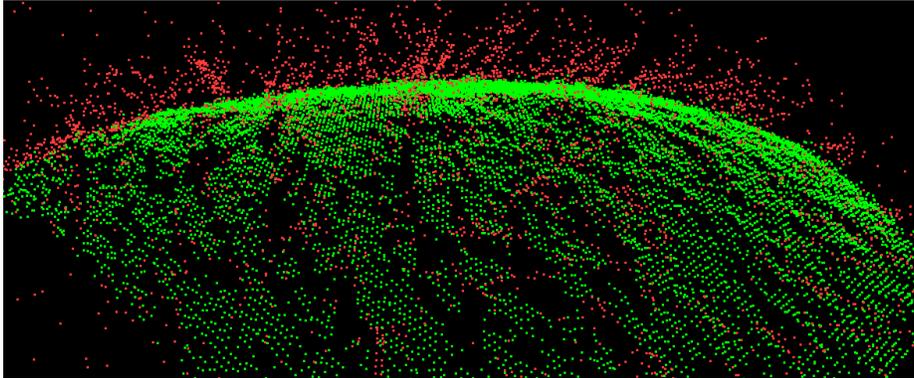
(a) Filtering done using only the vicinity prior
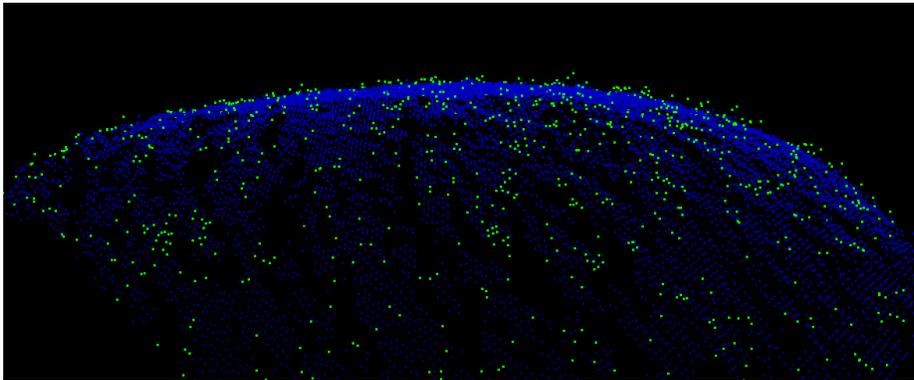
(b) Filtering after adding the smoothness prior
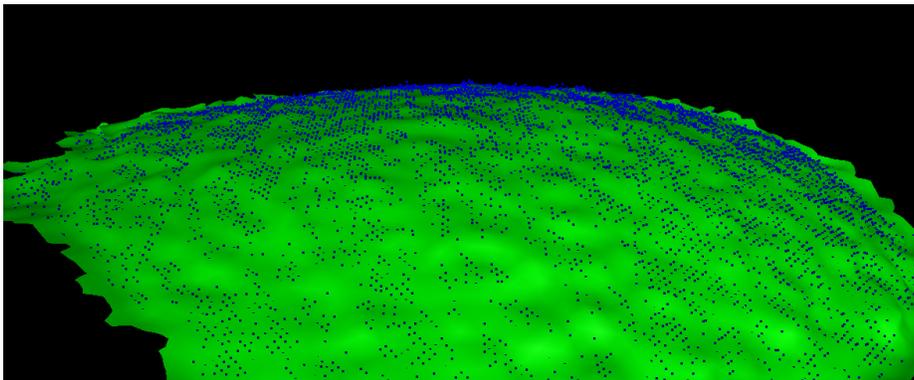
(c) Filtered data and surface reconstruction

Fig. 3. The figure shows a side view of a hairy surface. In the first image the effect of adding the vicinity prior alone is shown. Hair strands are coloured red and the surface green. The second image shows the difference between the surface using only vicinity prior and the surface, where both vicinity and smoothness prior has been applied. The difference is shown in green, while the resulting surface points of both vicinity and smoothness prior filtering are shown in blue. The third image shows a surface reconstruction using the filtered point set.

Bibliography

[1] 3Shape. Trios scanner. Website, 2011. http://www.3shapedental.com/restoration/dentist/digital-impression-taking.aspx.

[2] Julian Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 48(3):259–302, 1986. ISSN 00359246.

[3] B. Curless and M. Levoy. A volumetric method for building complex models from range images. *Proceedings of ACM SIGGRAPH*, pages 303–312, 1996.

[4] M.A. Fischler and R.C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981. ISSN 0001-0782.

[5] M.W. Jones et al. 3D distance fields: A survey of techniques and applications. *IEEE Transactions on Visualization and Computer Graphics*, pages 581–599, 2006. ISSN 1077-2626.

[6] V. Kolmogorov and R. Zabin. What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159, 2004. ISSN 01628828.

[7] MAA Neil, R. Juskaitis, and T. Wilson. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Optics Letters*, 22(24):1905–1907, 1997. ISSN 1539-4794.

[8] R. R. Paulsen and R. Larsen. Anatomically plausible surface alignment and reconstruction, sep 2010. URL http://www2.imm.dtu.dk/pubdb/p.php?5893.

[9] R.R. Paulsen, J.A. Baerentzen, and R. Larsen. Markov random field surface reconstruction. *IEEE Transactions on Visualization and Computer Graphics*, pages 636–646, 2010. ISSN 1077-2626.

[10] C.H. Reinsch. Smoothing by spline functions. *Numerische Mathematik*, 10 (3):177–183, 1967. ISSN 0029-599X.

[11] R. Schnabel, R. Wahl, and R. Klein. Efficient RANSAC for Point-Cloud Shape Detection. In *Computer Graphics Forum*, volume 26, pages 214–226. Wiley Online Library, 2007.

[12] J. Shewchuk. Triangle: Engineering a 2D quality mesh generator and Delaunay triangulator. *Applied Computational Geometry Towards Geometric Engineering*, pages 203–222, 1996.

[13] A. Wehr and U. Lohr. Airborne laser scanning–an introduction and overview. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2-3): 68–82, 1999. ISSN 0924-2716.