



**HAL**  
open science

## Bundle Adjustment Constrained Smoothing for Multi-view Point Cloud Data

Kun Liu, Rhaleb Zayer

► **To cite this version:**

Kun Liu, Rhaleb Zayer. Bundle Adjustment Constrained Smoothing for Multi-view Point Cloud Data. SVC 2012, 8th International Symposium on Visual Computing, Jul 2012, Rethymnon, Crete, Greece. pp.126-137, 10.1007/978-3-642-33179-4\_13 . hal-00763442

**HAL Id: hal-00763442**

**<https://inria.hal.science/hal-00763442>**

Submitted on 10 Dec 2012

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Bundle Adjustment Constrained Smoothing For Multi-View Point Cloud Data

Kun Liu, Rhaleb Zayer

Inria, Villers-lès-Nancy, F-54600, France

**Abstract.** Direct use of denoising and mesh reconstruction algorithms on point clouds originating from multi-view images is often oblivious to the reprojection error. This can be a severe limitation in applications which require accurate point tracking, e.g., metrology.

In this paper, we propose a method for improving the quality of such data without forfeiting the original matches. We formulate the problem as a robust smoothness cost function constrained by a bounded reprojection error. The arising optimization problem is addressed as a sequence of unconstrained optimization problems by virtue of the barrier method. Substantiated experiments on synthetic and acquired data compare our approach to alternative techniques.

## 1 Introduction

Over the last decade, *bundle adjustment* (BA) has become one of the key steps in multi-view reconstruction. It intervenes as a single nonlinear optimization which simultaneously fine-tunes the 3D structure and the viewing parameter estimates [1]. BA requires a set of feature correspondences which can be sparse, quasi-dense or dense in order to control the reprojection error and yields a refined visual reconstruction. In an ideal setting the resulting point cloud data would reflect the exact geometry of the original object. In practice however, several factors such as ill-textured objects, spatial discretization, structured noise, and lighting conditions contribute toward matching errors. These errors cannot be fully fixed by BA and the point cloud generally exhibits noise to varying degrees. A commonly adopted solution is the construction of an approximating surface using existing meshing algorithms, e.g. [2, 3]. These geometric algorithms operate mainly in the three dimensional domain and do not necessarily maintain correspondences between the scene and image features. As a result, the cross-image correspondences are lost and can only be approximated by reprojection on the surface. Although recent approaches in multiview reconstruction, e.g. [4], can improve the visual appearance tremendously, they cannot be readily used in applications such as a metrology or non-contact shape and deformation measurement where an accurate and consistent tracking of surface points over time is crucial for gathering information such as strain or parameter estimation. Furthermore, we are not aware of geometry processing methods which enforce bounds on the reprojection error in the literature.

In this paper, we propose a point cloud smoothing approach tailored for multi-view point cloud data. We formulate the problem as the minimization of a smoothness measure constrained by a bound on the reprojection error. For the former, we propose a measure which favors local flatness of the point cloud data and for the latter, we adopt a formulation similar to standard sparse bundle adjustment. Both measures are combined into a constrained nonlinear optimization formulation. A barrier approach is used to drive the numerical optimization towards a smooth point cloud where the bounds on the reprojection error are enforced. In order to overcome numerical problems related to the densely populated nature of the arising matrix equations, we take advantage of the Sherman-Morisson formula. This allows for addressing relatively large data sets while keeping reasonable memory requirements.

We evaluate the quality of resulting point cloud data by means of ground truth data generated synthetically. Tests on real data acquired and reconstructed using existing methods e.g. [5] confirm the quality of our results. We demonstrate the robustness of our approach to irregular data sampling, to sharp features and to shrinkage. Our approach does not make any assumption on the nature of the noise in the data and does not require any additional input, e.g., visual hulls. The only assumption made is the geometric smoothness, which is often a property of the original model. Although, we do not perform any further matching computations on the underlying images, experiments on synthetic data sets suggest that our approach moves existing matches closer towards to the exact matches.

In summary, this paper makes the following contributions:

- Formulate a constrained optimization for smoothing multi-view point clouds with bounded reprojection error
- Develop robust and efficient numerical solution procedure

Our approach can be regarded as a post-processing tool and could be used in conjunction with existing reconstruction algorithms. The rest of this paper is organized as follows: Section 2 covers the most related work, section 3 lays out the general setup and the notation, and section 4 reviews bundle adjustment and discusses using the Laplacian operator as smoothing regularizer. In section 5, we introduce our smoothing cost function and show how to set up the constrained optimization. Numerical aspects of our approach are discussed in section 6 and the results are summarized in section 7.

## 2 Related work

The prior art on multi-view model acquisition is extensive. In order to keep this exposition succinct, we restrict ourselves to the most related work and refer the reader to [6, 7] for a general overview.

Despite its long history, most of research effort on BA has been dedicated to numerical optimization strategies [1]. Subjects such as fusing it or enhancing it with additional input has been studied less. For instance, [8, 9] introduce model-based constraints as a regularizer within the bundle adjustment formulation in

the context of head reconstruction. The authors of [10] propose using GPS (global positioning system) data as a penalty for the reprojection error and optimize the problem in the least squares sense. More closely related to our work is the approach proposed in [11], where GPS and Structure-from-Motion data are fused within a constrained optimization formulation. The approach is applied to a setting which combines monocular image sequence with GPS data. In our case, we are more concerned about the geometric smoothness of the multi-view data and we cannot take advantage of readily available sensor data.

Smoothing surface meshes [12–15] and point clouds [16, 17], are well studied topics in geometry processing. Traditionally, these approaches are tailored for input data obtained from scanners. Unfortunately, many of these methods do not perform well on point clouds originating from multi-view stereo reconstruction. Furthermore they enforce the correspondences between 3D structure and image feature points.

### 3 Problem setting and notation

In the following, we assume the input data consists of a point cloud originating from standard multi-view acquisition [6] along with the camera parameters. The data can stem from dense [18] or quasi dense [5] matching approaches.

The point cloud will be represented as  $x = (x_1^T, \dots, x_n^T)$ , where  $x_i$  represent the three dimensional coordinates of the  $i$ -th point. For a set of  $m$  views, the camera matrices can be conveniently assembled as  $p = (p_1, \dots, p_m)$  where  $p_j$  be the vector of parameter for camera  $j$ . The point corresponding to  $x_i$  on an image  $j$  will be denoted  $a_{ij}$ .

### 4 Bundle adjustment with smoothing regularization

The overall goal of this work is to determine a smooth geometric model and the configuration of cameras that are maximally consistent with the observations. A good starting point is the bundle adjustment.

#### 4.1 Bundle adjustment

BA is a nonlinear least square problem [6] where the cost function penalizes reprojection error with respect to 3D structure and viewing parameters.

Concretely, this can be formulated as:

$$h(x, p) = \sum_i^n \sum_j^m \delta_{ij} \|Q(p_j, x_i) - a_{ij}\|^2 \quad (1)$$

where  $Q(p_j, x_i)$  is the predicted projection of point  $i$  on image  $j$ , and  $\delta_{ij}$  flags 1 if point  $i$  is visible in image  $j$  and 0 otherwise.

The best known algorithm to solve this kind of nonlinear least-squares problems is the the Levenberg-Marquardt (LM) algorithm [19].

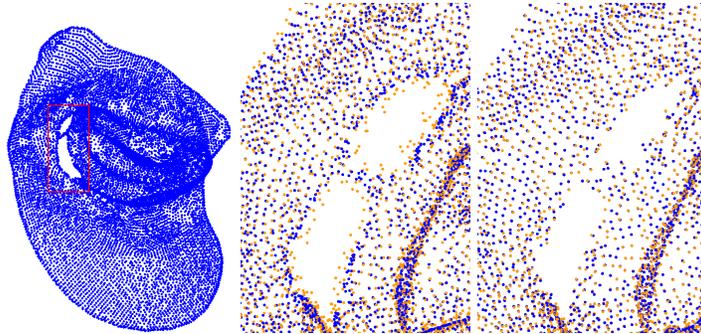
## 4.2 Bundle adjustment with smoothing regularization

As we seek to smooth point cloud data while minimizing BA, it seems natural to consider reformulating the problem as a minimization of the following objective function

$$h(p, x) + \alpha \|L(x)\|^2 \quad (2)$$

where  $h$  is the BA function from equation 1,  $L$  is the Laplacian operator discretized locally using the  $k$ -nearest neighbors, and  $\alpha$  is a weighting parameter. This formulation blends naturally within the standard Levenberg-Marquardt algorithm as the Laplacian operator can be simply considered as regularizer.

Unfortunately such an approach would suffer from over-smoothing as well as limitations known to Laplacian operator especially with respect to sharp features. Additionally due to the sparsity of the point cloud and the discretization based on nearest neighbors, shrinking effects can appear around holes or areas where data is missing as illustrated in figure 1. In the following section, we propose an alternative approach which remedies such artifacts as illustrated in the aforementioned figure.

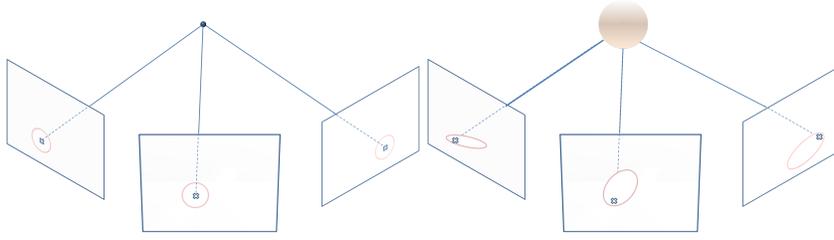


**Fig. 1.** A zoom on the ear model (left) illustrates the shrinking effect of Laplacian regularization (middle, blue). Constrained smoothing (right, blue) is more robust to such artifacts. In both results, the original data is shown in orange.

## 5 Bundle adjustment constrained smoothing

We regard the problem as searching for a smooth surface such that reprojection error is minimal. In order to account for possible errors in the cross-image feature correspondences, we can allow the matches to evolve in a small disk around their initial location as illustrated in figure 2. This would allow searching for a smooth surface while maintaining image feature correspondences in the vicinity of their initial positions.

In the context of an optimization for the whole point cloud data, we do not need to enforce the radius constraint for the individual points. Instead, we enforce



**Fig. 2.** Starting from a converged bundle adjustment, our approach (left) searches for new spatial position of the 3d point while guaranteeing that the reprojection error is bounded i.e. the matches are maintained within a disk around the input matches. On the other hand, constraining the smoothing within a ball around the initial spatial position (right) can lead to larger reprojection errors as the shape of the corresponding projection (planar ellipses) is not taken into account (please refer to the text for details).

it as a global constraint. In the following, we define our smoothness measure and we show how it can be combined with the reprojection error.

It would be possible to tailor a smoothing approach which restricts the displacements within a small ball around the initial spatial point locations. This kind of smoothing however, does not take into consideration the reprojection into image space and can lead to large errors. This is illustrated in figure (2-right) which shows the spatial search domain (ball around the initial point) and its counterpart in image space (planar ellipses). As the planar ellipses can be elongated, the reprojected point position can lay far from the initial match and hence such an approach would corrupt the initial matching results. This effect can be further amplified when dealing with wide base-line views. In contrast, we formulate the constraints in image space. The spatial position is then forced to lay at the intersection of the fat bundle-lines (small cylinders around the bundle lines) and thus a tight bound on reprojection error is guaranteed.

### 5.1 Smoothness measure

In order to define smoothness for point cloud data, we endow the points with local adjacency relations. We use k-nearest neighbor algorithm to construct a directed graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ , where  $\mathcal{V}$  is the point set we add an edge  $(x_i, x_j)$  in  $\mathcal{E}$  if  $x_i$  is one of k-nearest neighbor of  $x_j$ . In all our experiments  $k$  is set to 10, we note that using higher values hardly changes the results. Additionally, we endow each point with a normal direction. We estimate the point normal  $n_i$  ( $\|n_i\| = 1$ ) for every point  $i$  by the method proposed in [20], which uses principal component analysis on the local neighborhoods.

Given computed normal directions, we define the local planarity for each edge  $(x_i, x_j)$  in  $\mathcal{E}$  as

$$(n_{ij} \cdot (x_i - x_j))^2, \quad (3)$$

where  $n_{ij}$  is the average normal associated with the mid-point of edge  $(x_i, x_j)$ . It penalizes the deviation of the points from the average plane defined by the

midpoint and its normal. We sum the contribution of all edges in  $\mathcal{E}$  to define the global cost function

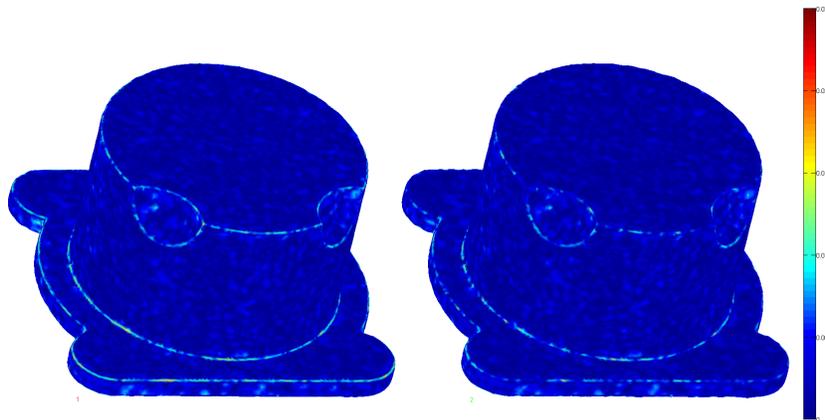
$$f(x) = \sum_{(x_i, x_j) \in \mathcal{E}} (n_{ij} \cdot (x_i - x_j))^2 \quad (4)$$

This function acts in two ways, it tends to improve the local flatness by minimizing the scalar product and second, since the edge vector is not normalized, it tends to pull neighboring points together. This local flatness measure is commonly used in the context of mesh simplification[21].

A similar formulation in the  $\mathbf{L}^1$ -norm has been proposed in [22]. In order to avoid differentiability difficulties raised by lower order norms we use the  $\mathbf{L}^2$  in view of coupling this measure with the reprojection error. In general,  $\mathbf{L}^2$  responds strongly to outliers. We avoid this shortcoming by means of a robust norm

$$\Psi(s^2) = \sqrt{s^2 + \epsilon^2}; \quad \epsilon = 1e^{-6} \quad (5)$$

This function can be regarded as a differentiable norm of the absolute norm function and its impact is illustrated in figure 3.



**Fig. 3.** The result of our approach on a noisy input data using  $\mathbf{L}^2$  norm (left) and the robust norm (right). The color coding shows the Hausdorff distance to the ground truth point cloud. (Best viewed in the electronic version).

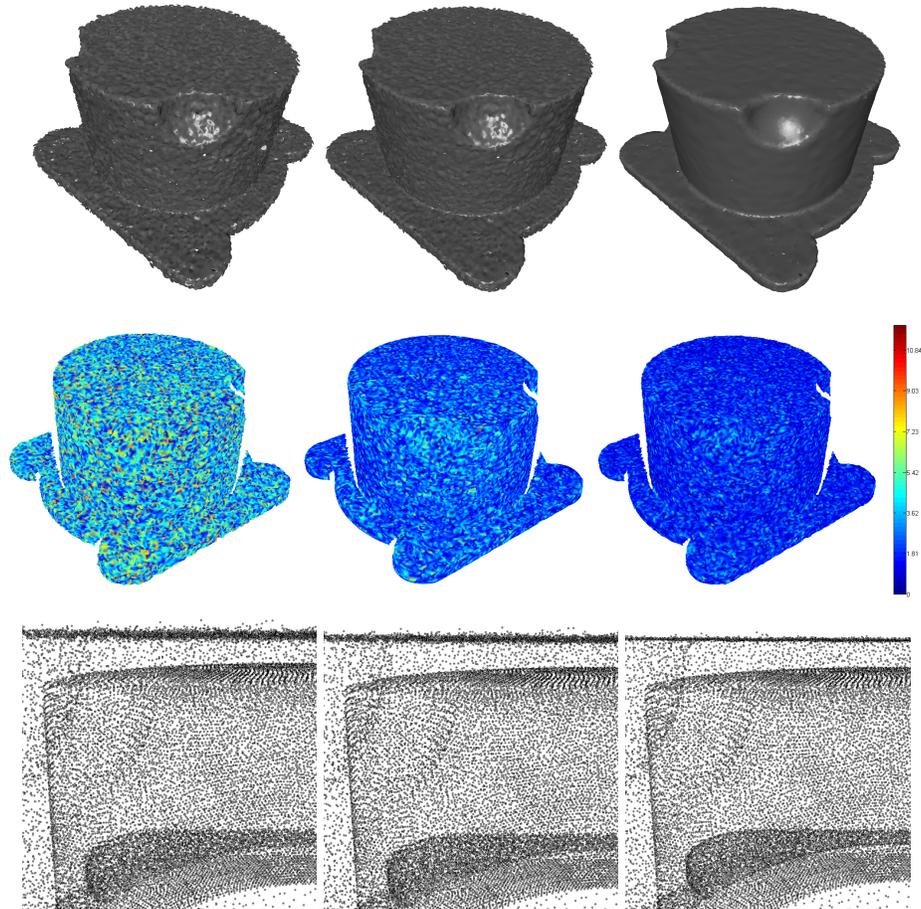
## 5.2 Reprojection error constraint

Let's assume we apply BA to the input point cloud data and let  $\epsilon_0$  be the residual reprojection. We can then define the our reprojection constraint as

$$h(x) < \lambda \epsilon_0 \quad (6)$$

with relaxing parameter  $\lambda > 1$ . The constraint defined above caps the reprojection error. It also resolves cases where minima of the cost function  $f(x)$  is not unique by restricting the search within a very close range.

## 6 Constrained Optimization



**Fig. 4.** A noisy point cloud (left-top) is processed using BA with Laplacian regularization (middle-top) smoothing and BA constrained smoothing (right-top), all views are shown in splating mode. The middle row shows the reprojection error for the same view. The bottom row shows a zoom on the corresponding point cloud data.

At this stage, we have all the ingredients necessary for the problem setup and we can formulate the smoothing procedure as the following constrained optimization

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && g(x) \leq 0 \end{aligned} \tag{7}$$

where  $g(x) = h(p, x) - \lambda\epsilon_0$ . The constraint function  $g(x)$  depends only on the structure as we keep the camera parameters fixed in what follows.

This problem is more intricate than BA alone as the cost function and the constraint are both nonlinear. We address it using the barrier method [23], which is a procedure for approximating constrained optimization problems by unconstrained ones. The smoothing problem can be solved then as a series of nonlinear minimization problems of the form

$$f(x) + B(g(x), c) \tag{8}$$

where  $B(g(x), c)$  is the barrier function and  $c$  is a positive constant. The barrier function operates by introducing a singularity along the constraint boundary. Probably, the popular choice is the logarithmic barrier function which tends to infinity at the constraint boundary. Furthermore, Its simplicity w.r.t subsequent derivative computations and its satisfaction of the self-concordance criterion makes it an attractive choice [19]. In the current paper, we use a logarithmic function

$$B(g(x), c) = -c \cdot \ln(-g(x)) \tag{9}$$

Concretely, solving the problem in equation. 7 amount to minimizing a series of function in the form of Eq. 8 with different constant  $c = c_k$ , where  $\{c_k\}$  is a decreasing sequence tending to 0, i.e. for each  $k$ ,  $c_k \geq 0$ ,  $c_{k+1} < c_k$ .

Each nonlinear minimization Eq.8 is solved iteratively using Newton's method. The associated Hessian matrix is

$$H = H_{f(x)} + B' H_{g(x)} + B'' \nabla g \nabla g^T \tag{10}$$

where  $H_{f(x)}$  and  $H_{g(x)}$  are the Hessian matrices of  $f(x)$  and  $g(x)$  respectively,  $B'$  and  $B''$  are the the first and second derivatives,  $\nabla g$  is the gradient of  $g(x)$ . In our implementation we approximate  $H_{g(x)}$  by  $J^T J$ , where  $J$  is the Jacobian of  $g$ .

The last term in equation 10 is a densely populated matrix and turns out to be problematic when solving the linear systems involved at each Newton iteration. We avoid this issue by means of the Sherman-Morrison formula [24] which reads

$$(A + uv^T)^{-1} = (I - \frac{A^{-1}uv^T}{1 + v^T A^{-1}u})A^{-1} \tag{11}$$

and holds for arbitrary invertible square matrix  $A$  and vectors  $u$  and  $v$  such that  $1 + v^T A^{-1}u$  is non zero.

Writing  $\hat{H} = H_{f(x)} + B' H_{g(x)}$  and  $\hat{g} = \sqrt{B''} \nabla g$ , we have by virtue of equation 11

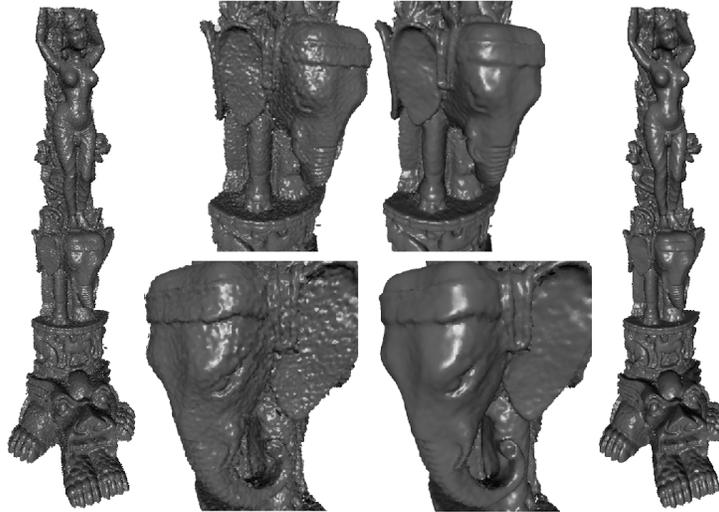
$$H^{-1} = (I - \frac{\hat{H}^{-1} \hat{g} \hat{g}^T}{1 + \hat{g}^T \hat{H}^{-1} \hat{g}}) \hat{H}^{-1} \tag{12}$$

Therefore, linear systems of the form  $(\hat{H} + \hat{g}\hat{g}^\top)x = b$  can be converted into

$$x = \left(I - \frac{\hat{H}^{-1}\hat{g}\hat{g}^\top}{1 + \hat{g}^\top\hat{H}^{-1}\hat{g}}\right)\hat{H}^{-1}b \quad (13)$$

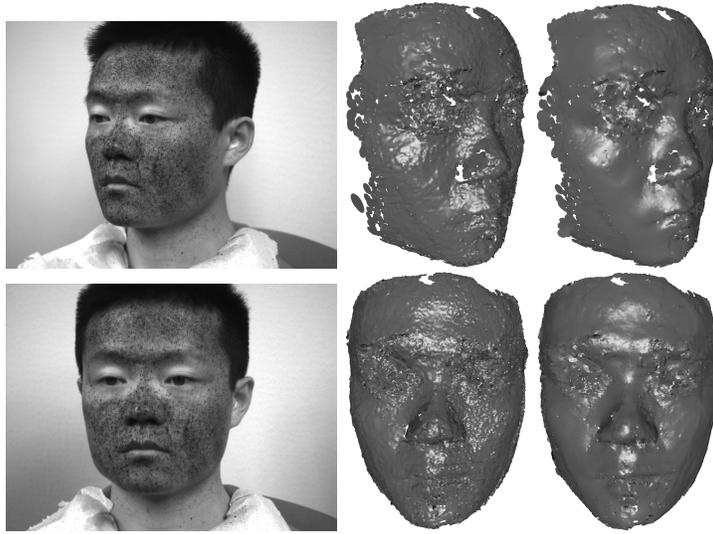
Since  $\hat{H}$  is sparse and not densely populated, the system can be handled using standard linear solvers. It is imperative to note that the inverse of  $\hat{H}$  need not be computed. Instead, equation 13 is split into two subsystems  $\hat{H}y = b$  and  $\hat{H}z = \hat{g}$ . The results are then plugged back into equation 13.

## 7 Results



**Fig. 5.** Illustration of our method on a large data set ( 200K points). Image correspondences across 56 views were perturbed by a gaussian noise with a unit variance and a peak of 3 which yields the noisy reconstruction (left). The result of our approach is shown to the right. Middle image show a zoom on the elephant head. All views are shown in splating mode.

We tested the proposed method on a set of synthetic and real world data. For synthetic data, camera captures were generated using existing 3D models with the help of Blender [25]. In this way, we have all the ground truth data necessary for evaluation. Noise was added in two different ways. In the first scenario, noise was directly added to the image correspondences (figures 3, 5). In the second scenario, noise was added to the original model and the image correspondences were obtained as weighted average of the noiseless and noisy projections (figure 4). In both scenarios, a gaussian function was used to generate the noise.



**Fig. 6.** Sample images (left) out of a set of 6 wide base-line images were used to generate a quasi-dense point cloud (middle) using the propagation approach in [5]. Our result (right) shows an overall quality improvement of the point cloud. Point clouds are shown in splating mode.

Figure 4 shows a comparison of Laplacian regularized BA of section 4 and the BA constrained smoothing of section 5. Although the former approach is easier to implement, it under-performs in comparison to latter approach. Figure 5 shows the performance of our approach on a large data set where more than 50 views were used. A typical result of our approach on real world data is shown in figure 6. In this example, 6 views were combined using the quasi-dense propagation approach of [5] to generate the initial point cloud. Our approach reduces the asperities in the point cloud and yields a smoother result. Even in regions such the eye area which are generally difficult to construct, our approach helps smoothing out the noise to a great extent.

**Limitations and discussion:** Our approach requires a good initial estimate e.g. results from bundle adjustment and also a sufficient sampling of the data in order to yield optimal results. However, when sampling points are missing in some region, as illustrated in figure 1, our approach does not suffer from shrinkage and still produces coherent results. When the initial data suffers from large noise, our approach can improve the results but only within the limits allowed by the reprojection error control. In this scenario, the noise reflects extensive errors in the matching, projection matrices or both and therefore enforcing reprojection errors based on such corrupt data might not be a viable goal. Nevertheless, relaxing the reprojection error bound would induce smoother geometric results.

## 8 Conclusion

In this paper, we studied the effect of combining a smoothness measure with bundle adjustment. To overcome limitations of using simple strategies such as Laplacian regularization or constrained spatial smoothing, we developed a robust and efficient approach based on containing the reprojection error while enforcing the smoothness of the point cloud data. In future work, we plan to investigate the use of our approach on time-dependent data.

**Acknowledgments and Credits:** The author would like thank, Alejandro Galindo for help with acquiring the face data, Nicolas Ray, Dmitry Sokolov, and the anonymous reviewers for their feedback on the paper. This work was funded by the ANR (Agence Nationale de la Recherche) under grant (PhysiGrafix ANR-09-CEXC-014-01).

## References

1. Triggs, B., McLauchlan, P., Hartley, R., Fitzgibbon, A.: Bundle adjustment – a modern synthesis. *Vision algorithms: theory and practice* (2000) 153–177
2. Kazhdan, M., Bolitho, M., Hoppe, H.: Poisson surface reconstruction. In: *Proceedings of the fourth Eurographics symposium on Geometry processing. SGP '06*, Eurographics Association (2006) 61–70
3. Lempitsky, V., Boykov, Y.: Global optimization for shape fitting. In: *IEEE Conference on Computer Vision and Pattern Recognition, CVPR '07*. (2007) 1–8
4. Furukawa, Y., Ponce, J.: Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **32** (2010) 1362–1376
5. Lhuillier, M., Quan, L.: Match propagation for image-based modeling and rendering. *IEEE Trans. Pattern Anal. Mach. Intell.* **24** (2002) 1140–1146
6. Hartley, R., Zisserman, A.: *Multiple view geometry in computer vision*. Volume 2. Cambridge Univ Press (2000)
7. Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1. CVPR '06* (2006) 519–528
8. Fua, P.: Using model-driven bundle-adjustment to model heads from raw video sequences. In: *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999*. Volume 1. (1999) 46–53
9. Fua, P.: Regularized bundle-adjustment to model heads from image sequences without calibration data. *International Journal of Computer Vision* **38** (2000) 153–171
10. Kume, H., Taketomi, T., Sato, T., Yokoya, N.: Extrinsic camera parameter estimation using video images and gps considering gps positioning accuracy. In: *Proceedings of the 2010 20th International Conference on Pattern Recognition. ICPR '10*, IEEE Computer Society (2010) 3923–3926
11. Lhuillier, M.: Fusion of gps and structure-from-motion using constrained bundle adjustments. In: *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE (2011) 3025–3032
12. Taubin, G.: A signal processing approach to fair surface design. In: *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, ACM (1995) 351–358

13. Desbrun, M., Meyer, M., Schröder, P., Barr, A.: Implicit fairing of irregular meshes using diffusion and curvature flow. In: Proceedings of the 26th annual conference on Computer graphics and interactive techniques. (1999) 317–324
14. Fleishman, S., Drori, I., Cohen-Or, D.: Bilateral mesh denoising. In: ACM Transactions on Graphics (TOG). Volume 22., ACM (2003) 950–953
15. Jones, T.R., Durand, F., Desbrun, M.: Non-iterative, feature-preserving mesh smoothing. ACM Transactions on Graphics (TOG) **22** (2003) 943–949
16. Weyrich, T., Pauly, M., Keiser, R., Heinzle, S., Scandella, S., Gross, M.: Post-processing of scanned 3d surface data. In: Symposium on Point-Based Graphics. Volume 1. (2004)
17. Schall, O., Belyaev, A., Seidel, H.: Robust filtering of noisy scattered point data. In: Point-Based Graphics, 2005. Eurographics/IEEE VGTC Symposium Proceedings, IEEE (2005) 71–144
18. Ferrari, V., Tuytelaars, T., Gool, L.V.: Simultaneous object recognition and segmentation by image exploration. In Pajdla, T., Matas, J., eds.: European Conference on Computer Vision (ECCV). Volume I., Springer (2004) 40–54
19. Nocedal, J., Wright, S.: Numerical optimization. Springer verlag (1999)
20. Hoppe, H., DeRose, T., Duchamp, T., McDonald, J., Stuetzle, W.: Surface reconstruction from unorganized points. SIGGRAPH Comput. Graph. **26** (1992) 71–78
21. Heckbert, P.S., Garland, M.: Optimal triangulation and quadric-based surface simplification. Comput. Geom. Theory Appl. **14** (1999) 49–65
22. Avron, H., Sharf, A., Greif, C., Cohen-Or, D.:  $l^1$ -sparse reconstruction of sharp point set surfaces. ACM Transactions on Graphics (TOG) **29** (2010) 135:1–135:12
23. Luenberger, D., Ye, Y.: Linear and nonlinear programming. Volume 116. Springer Verlag (2008)
24. Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: Numerical Recipes: The Art of Scientific Computing(3rd ed.). Cambridge University Press (2007)
25. : Blender. (<http://www.blender.org>)