# Group-Valued Regularization for Analysis of Articulated Motion

Guy Rosman[1*], Alex M. Bronstein[2], Michael M. Bronstein[3], Xue-Cheng Tai[4], and Ron Kimmel[5]

[1] Dept. of Computer Science
Technion - IIT
Haifa 32000 Israel
{rosman,ron}@cs.technion.ac.il
[2] School of Electrical Engineering
Faculty of Engineering
Tel Aviv University
Ramat Aviv 69978, Israel
bron@eng.tau.ac.il
[3] Institute of Computational Science, Faculty of Informatics
Universitá della Svizzera italiana
CH - 6904 Lugano, Switzerland michael.bronstein@usi.ch
[4] Dept. of Mathematics
University of Bergen
Johaness Brunsgate 12
Bergen 5007 Norway
tai@mi.uib.no

**Abstract.** We present a novel method for estimation of articulated motion in depth scans. The method is based on a framework for regularization of vector- and matrix- valued functions on parametric surfaces.
We extend augmented-Lagrangian total variation regularization to smooth rigid motion cues on the scanned $3D$ surface obtained from a range scanner. We demonstrate the resulting smoothed motion maps to be a powerful tool in articulated scene understanding, providing a basis for rigid parts segmentation, with little prior assumptions on the scene, despite the noisy depth measurements that often appear in commodity depth scanners.

**Key words:** Parameteric Surfaces, Motion Segmentation, Articulated Motion,

## 1 Introduction

Depth scanners are becoming an ever-more prevalent data source for computer vision applications. Interpretation of such data is still a challenge, especially

* This research was supported by Israel Science Foundation grant no.1551/09 and by the European Community's FP7- ERC program, grant agreement no. 267414.

given the wide range of environments and applications in which depth sensors are expected to be used. While many algorithms classify objects in range data (see [8, 23, 33, 17] and references therein), obtaining meaningful cues for the general settings in an efficient way is still a topic of intense research.

Motion cues are important for scene understanding [6, 14]. Obtaining a meaningful dense motion descriptor for $3D$ data is an important preprocessing step for the understanding of arbitrary scenes. Articulated motion is one specific type of motion, common in many natural and man-made scenes. Detection and understanding of articulated motion have therefore attracted the attention of numerous research efforts, see [1, 41, 2, 29, 26] for a few examples.

When treating rigid and articulated motion, Lie-group theory provides us with a well-motivated representation of motion. Lie-groups have been used extensively for motion interpretation [27, 38, 22, 29], tracking [37, 35], and modeling [7, 18], among other uses. It is only natural to use them as a dense motion descriptor to be used in the understanding process of $3D$ motions. While motion vectors alone can also be used to represent the *scene flow* [42], the *overparameterized* [25, 30] representation we favor leads naturally to interpretation of the scene in terms of piecewise rigid motions, as we demonstrate in this paper.

In this paper, we present a novel approach for characterizing articulated motion obtained from depth sensors, without assuming an explicit skeletal model. Instead, we favor an implicit approach, aggregating motion cues in local neighborhoods in a bottom-up manner. While a similar approach, could have been to perform diffusion of the Lie-group elements using the Lie-algebra, as suggested, for example in [29], this is done by an explicit approach to the evolution of the smoothing process, which is inherently slow. In this paper we try to alleviate the limitations associated with such approaches.

We do so using a regularization method that is based on an augmented-Lagrangian scheme for group-valued regularization on parameteric surfaces such as depth scans. This algorithm extends our recent work on fast total variation regularization of group-valued images [31], modifying it algorithm to handle smoothness on parameteric surfaces.

Using this regularization, we are able to create a scale-space re-interpreting the scene-flow in terms of local rigid transformations. Experiments shown in this paper demonstrate the usefulness of the resulting images for motion segmentation from noisy depth data.

The smoothing process we describe and associated preprocessing steps are highly parallelizable, lending themselves to real-time implementation on parallel hardware such as *graphics processing units* (GPUs), as shown for example, in [31]. The results obtained demonstrate the method's usefulness in identifying components of piecewise-rigid motion in real-life, noisy, range data.

**Contribution** The contribution of the paper is two-fold: a. We extend the framework of fast TV regularization to parametric surfaces, both for the scalar and the group-valued case. b. Based on this framework, we describe a fast method to obtain and segment rigid motion cues from depth-image videos, demonstrating its effectiveness on real depth videos.

In Section 2 we describe the model and functional behind our method. The algorithm and relevant numerical schemes are given in Section 3. Finally, we demonstrate the results of our method in Section 4, and discuss the results and potential uses in Section 5.

## 2 Model Description

We now describe our setting and model, and the resulting functional. We are given a depth video of an object undergoing articulated motion so that most of the object's visible surface is composed of rigidly moving parts. These parts are connected by joints, for which the motion is merely assumed to be smooth. At each frame in the sequence, we have a depth image of the object, with subsequent frames differing by small, piecewise-rigid motion. For an arbitrary depth video, a predetermined articulated skeleton cannot be assumed at this low-level vision phase. Instead, we propose to obtain a strong cue for detecting rigidly moving parts by spatially regularizing the observed motion in a piecewise-rigid manner. Rigid motions can be described by the elements of the *Lie-group* $SE(3)$, the Lie-group of rigid transformations in $\mathbb{R}^3$. We expect each point to be loosely associated with a local rigid motion, and wish to smooth the field of *motion particles* in the $3D$ scene. We therefore define our motion description by a group-valued function $u(x) \in SE(3)$, where $x$ denotes a point on the surface of the object at one frame. The resulting piecewise-smooth $SE(3)$ image is highly informative, and in many cases allows straightforward segmentation of the motion, without using data-driven classifiers [3] or growing the parts in a bottom-up manner and letting motion models compete [2, 12].

### 2.1 The special-Euclidean Group $SE(3)$

Lie-groups are algebraic groups endowed with a manifold structure. Their structure allows us to discuss smoothness of motion parameters in a well-defined manner. We refer the reader to standard literature on the topic for an in-depth discussion [15]. Because of the group nature of $SE(3)$, the tangent plane at each point on the Lie-group can be mapped onto the identity element's tangent plane, allowing us to define a linear space uniformly throughout the group. This space, the *Lie-algebra* associated with a Lie-group, allows us to define differentiation on the Lie-group.

Specifically, we look at the group of rigid $3D$ motions, the *special-Euclidean* group $SE(3)$,

$$SE(3) = \left\{ (\mathbf{R}\ \mathbf{t}) \,|\mathbf{R} \in \mathbb{R}_{3\times 3}, \mathbf{R}^T\mathbf{R} = \mathbf{I}, \mathbf{t} \in \mathbb{R}^3 \right\}, \tag{1}$$

and a map $u$ from points on the object surface onto an embedding of $SE(3)$ in $\mathbb{R}^{12}$. This map will be regularized by minimizing the functional we now describe.

### 2.2 Group-Valued Regularization on Parametric Surfaces

The functional we wish to minimize should describe the irregularity of the motion field in terms of the scanned $3D$ surface with two or more poses. Since our input is a range image, it makes sense to use the $2D$ image domain as the integration domain. Our notion of smoothness, however, should be defined in terms of the $3D$ surface tangent plane. Thus, the regularization term we seek is intimately linked to the problem of image processing for images defined on parametric surfaces [34, 19, 40]. The measure of smoothness of the associated locally-rigid motion should take into account the geometry of the $3D$ surface. We therefore take the *total variation* (TV, [32])

$$E_S(u) = \int_{\Omega \in \mathbb{R}^2} \|\nabla_{\mathcal{M}} u\| d\Omega, \tag{2}$$

to be our measure of regularity, defined in terms of $\nabla_{\mathcal{M}}$, the gradient of the function $u$ on the surface itself. For vector-valued functions, we denote by $\|\nabla u\|$ the Frobenius norm of the Jacobian matrix $\left(\frac{\partial u_i}{\partial x_j}\right)_{ij}$. In the case where $u$ is a Lie-group matrix whose elements are isometries, there is justification to use the embedding space gradient norm $\|\nabla u\|$ and not the group's intrinsic regularity measure $\|u^{-1}\nabla u\|$.

We note that expressing the gradient on the surface in terms of the parameterization plane is relatively simple given the *first fundamental form* of the surface (see for example, [13], page 102, or the supplementary notes).

In addition, we also require a data term. The simplest data term in use is the least squares fitting term,

$$E_D = \int_{\Omega \in \mathbb{R}^2} \|u - u_0\|^2 d\Omega = \int_{\Omega \in \mathbb{R}^2} \|u_0^{-1} u - Id\|^2 d\Omega, \tag{3}$$

where $\|\cdot\|$ is the Frobenius norm, and $u_0$ is the given input function, for example a local motion estimate given by local *iterative closest point* search [4, 9], or by least-square fitting a rigid motion model to a deformation result based on other algorithms [5, 24]. We note that by inverting $u_0$, this distance is the same as the distance often used between $SE(3)$ elements [20]. We suggest an efficient non-rigid registration method in Subsection 3.3, which can be easily extended to include robust data terms.

The overall cost function we intend to minimize will be of the form

$$\min_{u \in SE(3)} \quad E_S(u) + \lambda E_D(u), \tag{4}$$

where $\lambda$ describes the relative strength of the data term. We describe in Section 3 a fast minimization algorithm for this cost function.

## 3 Numerical Methods

We now try to minimize the overall cost function (4). In order to enforce the constraint $u \in SE(3)$ we use an auxiliary variable $v$ such that $v = u, v \in$

$SE(3)$. We obtain $v = u$ using an augmented Lagrangian term added to the cost function. The resulting constraint causes the optimization with respect to $v$ to become a projection operation per-pixel, using *singular value decomposition* (SVD). This transforms the minimization problem into a saddle-point problem

$$\max_{\mu} \min_{v \in SE(3), u} \int_{\Omega} \|\nabla_{\mathcal{M}} u\| + \frac{r}{2}\|v-u\|^2 + \mu^T(v-u) + \|u-u_0\|^2 d\Omega \quad (5)$$

Unlike previous approaches for augmented Lagrangian TV regularization [36], our approach differs in the measure of smoothness we use. We note that while the update step for $v$ comes from minimizing the cost function, it is highly linked to the intuitive choice of updating $u$ and then projecting it, as well as to optimization by proximal operators [11], and can be made provably convergent with minor modifications, as shown in [31]. We now modify the augmented Lagrangian TV framework for the smoothness term described in (2).

### 3.1 Augmented Lagrangian TV Optimization of Vector Valued Functions on Parametric Surfaces

Let us start with the simpler case of a general vectorial function $u$, and formulate an efficient iterative scheme for smoothing (in the TV sense) functions on a parametric surface. In our case, this surface will be obtained from a range scanner, and the parametrization domain will be the image plane with its coordinates system. In order to efficiently regularize images on parametric surfaces, we sample the parametrization domain on a Cartesian grid.

The scheme we present is based on the augmented Lagrangian TV optimization scheme [36]. We use an auxiliary variable $p$ to describe the surface-domain gradient, rather than the image-domain gradient. That is, we add an auxiliary variable $p = J_{\mathcal{M}} \nabla u$, and optimize with respect to it using a shrinkage operator, similar to the image-domain TV case [36]. $J_{\mathcal{M}}$ denotes the Jacobian relating image coordinates and surface coordinates. We enforce the gradient constraint by adding an augmented Lagrangian term with Lagrange multipliers. We update $u$,$p$, and the Lagrange multiplier iteratively. For the vector-valued TV case, minimizing the functional now becomes a solution of the saddle-point problem

$$\max_{\mu_2} \min_{u,p} \int_{\Omega} \|p\| + \frac{r_2}{2}\|J_{\mathcal{M}}\nabla u - p\|^2 + \mu_2^T(J_{\mathcal{M}}\nabla u - p) + \lambda\|u-u_0\|^2 d\Omega, \quad (6)$$

where $\mu_2$ is our Lagrange multiplier for the gradient constraint. The optimization of (6) with respect to $u$ is given by a diffusion equation

$$-r_2 \operatorname{div}\left(J_{\mathcal{M}}^T \left(J_{\mathcal{M}}\nabla u - p\right)\right) + \operatorname{div} J_{\mathcal{M}}^T \mu_2 + 2\lambda\left(u-u_0\right) = 0. \quad (7)$$

Optimization with respect to $p$ can be expressed in closed form [39, 36] by a shrinkage operator,

$$p = \max\left(0, 1 - \frac{1}{r_2}\frac{1}{\|w\|}\right) w, w = J_{\mathcal{M}}\nabla u - \frac{\mu_2}{r_2}, \quad (8)$$

Finally, updating $\mu_2$ is given according to the augmented Lagrangian method [28, 16].

### 3.2 Augmented Lagrangian Regularization of Group-Valued Maps on Parametric Surfaces

Using an augmented Lagrangian term in order to enforce the constraint of $u = v \in SE(3)$, the overall functional reads

$$\max_{\mu,\mu_2} \min_{\substack{v \in SE(3) \\ u}} \int_{\Omega} \left[ \begin{array}{c} \|p\| + \frac{r}{2}\|u - v\|^2 + \mu^T(u - v) + \\ \frac{r_2}{2}\|J_{\mathcal{M}}\nabla u - p\|^2 + \mu_2^T(J_{\mathcal{M}}\nabla u - p) + \lambda\|u - u_0\|^2 \end{array} \right] d\Omega. \tag{9}$$

Minimization with respect to $u$ is done as in the same as in subsection 3.1. Minimization with respect to $p$ is given by equation 8. We update $\mu$ according to the augmented Lagrangian method [28, 16].

Optimization with respect to $v$ is using the same projection operator per-pixel as in [31]. Looking at optimization with respect to $v$, we obtain

$$\operatorname*{argmin}_{v \in SE(3)} \frac{r}{2}\|v - u\|^2 + \langle \mu, u - v \rangle = \operatorname*{argmin}_{v \in SE(3)} \frac{r}{2} \left\| v - \left( \frac{\mu}{r} + u \right) \right\|^2 = \operatorname*{Proj}_{SE(3)} \left( \frac{\mu}{r} + u \right), \tag{10}$$

where $\text{Proj}_{SE(3)}(\cdot)$ denotes the orthogonal projection operator onto $SE(3)$, given by an SVD operation, setting the singular values of $v$ to all-ones.

An algorithmic description of the resulting scheme is given as Algorithm 1. For further numerical details, the reader is referred to the supplementary material.

**Algorithm 1** Fast TV regularization of group-valued images on parametric surfaces

```
for k = 1, 2, ..., until convergence do
    Update u^k(x), according to (7).
    Update p^k(x), according to (8).
    Update v^k(x), by projection onto the matrix group, using SVD.
    Update μ^k(x), μ_2^k(x), according to the augmented Lagrangian scheme.
end for
```

### 3.3 Estimating Non-Rigid Motion in Depth Videos

In order to estimate the non-rigid motion occuring between two subsequent time-frames of a depth video, we first apply a simple non-rigid registration process, similar to the approach suggested by Li et al. [21], followed by the estimation of a locally-rigid motion model, as described in the supplementary material. In general, any motion estimation method can be used.

Since the overall motion field can involve both piecewise rigid and non-rigid motion components, and because of the noisy scan results often obtained from commodity depth scanners, the estimated instantaneous motion is quite noisy, as can be seen in Figure 2. The motion field should be post-processed so as to obtain locally-rigid interpretation. This can be obtained by the regularization process described in Section 3.2. The overall algorithm is summarized as Algorithm 2.

During the third step of the algorithm different $\lambda$ values can be used so as to obtain a scale-space of motion interpretation, for detecting salient candidates for rigid parts, or as features for learning-based motion segmentation.

**Algorithm 2** Regularized estimation of rigid motion from depth video

1: **for** $k = 1, 2, \ldots,$ until convergence **do**
2: Estimate motion field between depth frames according to a non-rigid ICP.
3: Estimate $u_0(x)$ at each point using least median squares fitting.
4: Regularize $u_0(x)$ using Algorithm (1).
5: **end for**

## 4 Results

We now demonstrate the results of our algorithm on both synthetic and real data. In Figure 1, we demonstrate results based on a synthetic hand model undergoing motion. We used the non-rigid registration model to track the surface over several frames so as to obtain a sufficiently large motion. While the detected motion is not completely piecewise-rigid due to skinning artifacts, occlusions, etc, the fingers are detected quite well. Using a standard mean-shift algorithm on the log-coordinates of the rotation matrices, we obtain segmentation of the fingers and the phalanges that undergo motion.

In Figure 2, we demonstrate TV regularization of $SE(3)$ for several frame pairs in a depth sequence by a Kinect sensor. Visualization is done using log-coordinates of the rotation matrix. The resulting estimated rigid motion allows segmentation of body parts and finger phalanges. In Figure 3 we use the mean-shift clustering algorithm [10] on the $SE(3)$ images' projection onto the small rotations standard linearization basis in order to segment the main moving parts. Despite the simple choice of the segmentation algorithm, that does not take into account the geometry of the surface and linearizes the Lie-group in the simplest possible manner, the segmentation of the moving parts is clear. It is expected that utilizing geometric prior on the regions size will prevent artifacts such as oversegmentation. In the examples shown here, QVGA resolution ($320 \times 240$) was used. The estimation of motion coefficients takes in Matlab about 5 seconds on an Intel i3 CPU. The regularization is similar algorithmically to [31], which took about a tenth of a second to compute on GPU. Preliminary results support this efficiency claim.

## 5 Conclusions

In this paper we extended the notion of fast $SE(3)$-valued image regularization to parametric surfaces. This allowed us to selectively smooth articulated motion based on noisy depth data. In many cases this regularized representation can provide a partition of an object into rigid parts, as shown in several examples. In future work, we intend to further explore the resulting scale-space of $SE(3)$

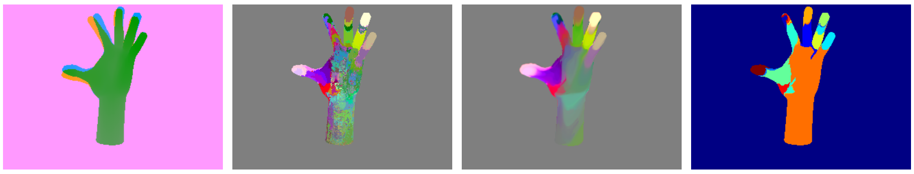

**Fig. 1.** TV regularization based on Equation (9) of an $SE3$-valued image placed on a rendered depth surface. Left-To-Right, Top-To-Bottom: An overlay of the two consecutive time-frames used to obtain motion estimation, the estimated and regularized $SE(3)$ images, and a resulting segmentation using mean-shift.

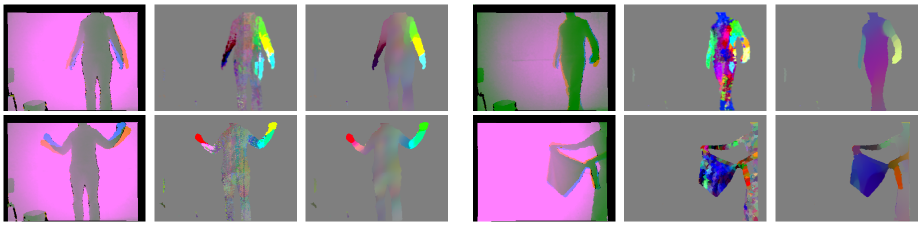

**Fig. 2.** TV regularization based on Equation (9) of an $SE3$-valued image given on a scanned depth surface. . Each row represents results on two different frames from a depth sequence. Left-To-Right, for each frame: An overlay of the two consecutive time-frames used to obtain motion estimation, the estimated $SE(3)$ measurement, and regularized image. Raw depth data is used to estimate the motion. The regularized $SE(3)$ image hints at joint locations for parts that were moving at the time the depth frames were taken. Note in the last example, using a slightly stronger regularization, a nonrigid object (a shirt) is still separated clearly from the arms.

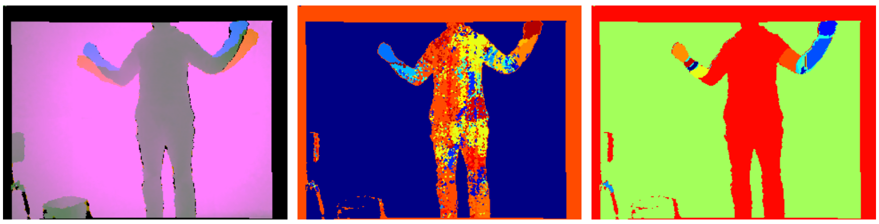

**Fig. 3.** Segmentation based on mean-shift clustering of the $SE(3)$ image. Left-to-right:

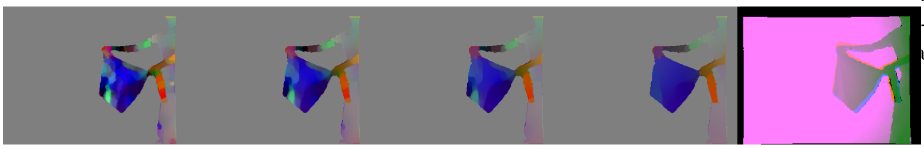

**Fig. 4.** A scale-space obtained by changing the fidelity coefficient through the values $\lambda = 5, 2.5, 1.5, 0.8$.

images and their use for scene understanding and $3D$ object segmentation as well as investigate several priors and data terms.

## References


1. G. Adiv. Determining three-dimensional motion structure from optic flow generated by several moving object,. *IEEE Trans. PAMI*, 7(4):384–401, 1985.
2. D. Anguelov, D. Koller, H.-C. Pang, P. Srinivasan, and S. Thrun. Recovering articulated object models from 3D range data. In *Proc. Conf. on Uncertainty in Artificial Intelligence*, pages 18–26. AUAI Press, 2004.
3. H. Benhabiles, G. Lavoué, J.-P. Vandeborre, and M. Daoudi. Learning boundary edges for 3D-mesh segmentation. *Comp. Graphics Forum*, 2011.
4. P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. PAMI*, 14(2):239–256, 1992.
5. A. M. Bronstein, M. M. Bronstein, and R. Kimmel. Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proc. Natl. Acad. Sci. USA*, 103(5):1168–1172, January 2006.
6. T. Brox, M. Rousson, R. Deriche, Dr., and J. Weickert. Colour, texture, and motion in level set based segmentation and tracking. *Image and Vision Computing*, 28(3):376–390, Mar. 2010.
7. E. Celledoni and B. Owren. Lie group methods for rigid body dynamics and time integration on manifolds. *Computer Methods in Applied Mechanics and Engineering*, 19:421–438, 1999.
8. H. Chen and B. Bhanu. 3D free form object recognition in range images using local surface patches. *Pattern Recognition Letters*, 28:1252–1262, 2007.
9. Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image Vision Comput.*, 10:145–155, April 1992.
10. D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Trans. PAMI*, 24:603–619, May 2002.
11. P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. May 2010.
12. D. Cremers and S. Soatto. Motion competition: A variational framework for piecewise parametric motion segmentation. *IJCV*, 62(3):249–265, 2005.
13. M. P. do Carmo. *Differential Geometry of Curves and Surfaces.* Prentice-Hall, 1976.
14. J. Fayad, C. Russell, and L. de Agapito. Automated articulated structure and 3D shape recovery from point correspondences. In *ICCV*, pages 431–438, 2011.
15. B. C. Hall. *Lie Groups, Lie Algebras,and Representations, An Elementary Introduction.* Springer Verlag, 2004.
16. M. R. Hesteness. Multipliers and gradient methods. *J. of Optimization Theory and Applications*, 4:303–320, 1969.
17. E. Kim and G. G. Medioni. 3D object recognition in range images using visibility context. In *IROS*, pages 3800–3807, 2011.
18. M. Kobilarov, K. Crane, and M. Desbrun. Lie group integrators for animation and control of vehicles. *ACM Trans. Graph.*, 28(2):1–14, 2009.
19. R. Lai and T. F. Chan. A framework for intrinsic image processing on surfaces. *Comput. Vis. Image Underst.*, 115:1647–1661, Dec. 2011.
20. P. M. Larochelle, A. P. Murray, and J. Angeles. A Distance Metric for Finite Sets of Rigid-Body Displacements via the Polar Decomposition. *Journal of Mechanical Design*, 129, 2007.

21. H. Li, R. W. Sumner, and M. Pauly. Global correspondence optimization for non-rigid registration of depth scans. *Computer Graphics Forum*, 27(5), July 2008.
22. D. Lin, W. Grimson, and J. Fisher. Learning visual flows: A Lie algebraic approach. In *CVPR*, pages 747–754, 2009.
23. T.-W. R. Lo and J. P. Siebert. Local feature extraction and matching on range images: 2.5D SIFT. *Comput. Vis. Image Underst.*, 113:1235–1250, December 2009.
24. A. Myronenko and X. B. Song. Point-set registration: Coherent point drift. *CoRR*, abs/0905.2635, 2009.
25. T. Nir, A. M. Bruckstein, and R. Kimmel. Over-parameterized variational optical flow. *IJCV*, 76(2):205–216, 2008.
26. M. Paladini, A. Del Bue, J. a. Xavier, L. Agapito, M. Stošić, and M. Dodig. Optimal metric projections for deformable and articulated Structure-from-Motion. *IJCV*, pages 1–25, July 2011.
27. F. C. Park, J. E. Bobrow, and S. R. Ploen. A Lie group formulation of robot dynamics. *Int. J. Rob. Res.*, 14:609–618, December 1995.
28. M. J. Powell. *Optimization*, chapter A method for nonlinear constraints in minimization problems, pages 283–298. Academic Press, 1969.
29. G. Rosman, M. M. Bronstein, A. M. Bronstein, and R. Kimmel. Articulated motion segmentation of point clouds by group-valued regularization. In *Eurographics Workshop on 3D Object Retrieval*, 2012.
30. G. Rosman, S. Shemtov, D. Bitton, T. Nir, G. Adiv, R. Kimmel, A. Feuer, and A. M. Bruckstein. Over-parameterized optical flow using a stereoscopic constraint. In *SSVM*, volume 6667 of *LNCS*, pages 761–772, 2011.
31. G. Rosman, Y. Wang, X.-C. Tai, R. Kimmel, and A. M. Bruckstein. Fast regularization of matrix-valued images. In *ECCV*, 2012. Accepted.
32. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D Letters*, 60:259–268, 1992.
33. J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-Time human pose recognition in parts from single depth images. June 2011.
34. A. Spira and R. Kimmel. Geometric curve flows on parametric manifolds. *J. Comput. Phys.*, 223:235–249, April 2007.
35. R. Subbarao and P. Meer. Nonlinear mean shift over Riemannian manifolds. *IJCV*, 84(1):1–20, 2009.
36. X.-C. Tai and C. Wu. Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model. In *SSVM*, pages 502–513, 2009.
37. O. Tuzel, F. Porikli, and P. Meer. Learning on Lie groups for invariant detection and tracking. In *CVPR*, 2008.
38. M. Žefran, V. Kumar, and C. Croke. On the generation of smooth three-dimensional rigid body motions. *Robotics and Automation, IEEE Transactions on*, 14(4):576 –589, Aug. 1998.
39. Y. Wang, J. Yang, W. Yin, and Y. Zhang. A new alternating minimization algorithm for total variation image reconstruction. *SIAM J. Imag. Sci.*, 1(3):248–272, 2008.
40. C. Wu, J. Zhang, Y. Duan, and X.-C. Tai. Augmented lagrangian method for total variation based image restoration and segmentation over triangulated surfaces. *J. Sci. Comput.*, 50(1):145–166, 2012.
41. Y. Yacoob and L. S. Davis. Learned models for estimation of rigid and articulated human motion from stationary or moving camera. *IJCV*, 36:5–30, January 2000.
42. Y. Zhang and C. Kambhamettu. Integrated 3D scene flow and structure recovery from multiview image sequences. *CVPR*, 2:2674, 2000.