Group-Valued Regularization for Analysis of Articulated Motion: Supplementary Material

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

This note supplements the paper "Group-Valued Regularization for Analysis of Articulated Motion", providing additional details on the exact setup and numerical algorithms used in our experiments. While this description is not critical for understanding the concepts, ideas, and algorithmic flow, we believe the additional details provided herein will aid understanding, implementation and and extension of the approach portrayed in the paper.

1 An Augmented Lagrangian Scheme for Parametric Surfaces

1.1 Group-Valued Regularization on Parametric Surfaces

The functional we wish to minimize should describe the irregularity of this motion field in terms of the scanned $3D$ surface with two or more poses. Since our input is a set of range images, it makes sense to use the $2D$ image domain as the integration domain. Our notion of smoothness of functions, however, should be defined in terms of the $3D$ surface tangent plane. Thus, the regularization term

* 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.

we seek is intimately linked to the problem of image processing for images defined on manifolds [22, 23, 13, 26]. The measure of smoothness of the associated locally-rigid motion should take into account the change in $3D$ coordinates. We therefore take the *total variation* (TV, [20])

$$E_S(u) = \int_{\Omega \in \mathbb{R}^2} \|\nabla_{\mathcal{M}} u\| d\Omega, \tag{1}$$

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, we can exchange the embedding space gradient norm $\|\nabla u\|$ for the intrinsic regularity measure $\|u^{-1}\nabla u\|$.

Additionally, we need to express the gradient on the surface in terms of the gradient in the parameterization plane. Given the local *first fundamental form* of the surface, the expression is relatively simple. Specifically, let $x_i, i = 1, 2$ be the image-plane coordinates, and $X_i, i = 1, 2$ be surface coordinates. The surface gradient in terms of a parametric domain gradient $\left(\frac{\partial u}{\partial x_i}\right)_i$ is given by ([10], page 102)

$$\left(\frac{G u_{x_1} - F u_{x_2}}{EG - F^2}, \frac{-F u_{x_1} + E u_{x_2}}{EG - F^2} \right), \tag{2}$$

where E, F, G are the components of the Riemmanian metric, given by

$$E = \langle X_1, X_1 \rangle, F = \langle X_1, X_2 \rangle, G = \langle X_2, X_2 \rangle. \tag{3}$$

The regularization term therefore becomes

$$\|\nabla_{\mathcal{M}} u\|^2 = \left\| \frac{1}{EG - F^2} \begin{pmatrix} G & -F \\ -F & E \end{pmatrix} \nabla u \right\|^2 = \|J_{\mathcal{M}} \nabla u\|^2, \tag{4}$$

where the Jacobian matrix $J_{\mathcal{M}}$ transforms the gradient in terms of the parametric plane into the gradient of the image as defined on the surface.

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

$$E_D = \int_{\Omega \in \mathbb{R}^2} \|u - u_0\|^2 d\Omega, \tag{5}$$

where u_0 is the given input function, for example a local motion estimate given by local *iterative closest point* search [5, 8], or by least-square fitting a rigid motion model to a deformation result based on other algorithms [6, 15]. We suggest an efficient non-rigid registration method in Subsection 2.4. A straightforward generalization that can be implemented using an iterative reweighted least-squares technique is a robust data term of the form

$$E_D = \int_{\Omega \in \mathbb{R}^2} \psi_D \left(\|u - u_0\|^2 \right) d\Omega, \tag{6}$$

where ψ_D is a robust fitting function, for example $\psi_D(s) = \sqrt{s+\epsilon^2}$ for some small ϵ.

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

$$\min_{u \in SE(3)} \quad E_S(u) + \lambda E_D(u), \tag{7}$$

where λ describes the relative strength of the data term. We describe in Section 2 a fast minimization algorithm for this cost function.

2 Numerical Methods

We now describe a minimization scheme for the cost function given in Equation (7), extending the scheme suggested in [19]. 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 the equality $v = u$ 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, as described in Subsection 2.3. This transforms the minimization problem into a saddle-point problem

$$\max_{\mu} \quad \min_{v \in SE(3), u} \int_{\Omega} \|\nabla_{\mathcal{M}} u\| + \frac{r}{2}\|v-u\|^2 + \mu^T(v-u) + \|u-u_0\|^2 d\Omega \tag{8}$$

Unlike previous approaches for augmented Lagrangian TV regularization [24], 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 [9], and can be made provably convergent with minor modifications, as shown in [18]. This is done by slightly modifying our algorithm, as suggested by Attouch et al. [2], changing the update steps for u, v into the minimization problems

$$\begin{aligned} u^k &= \operatorname*{argmin}_{u} \mathcal{F}(u, v^{k-1}, \mu) + \frac{1}{\theta_k}\|u - u^{k-1}\|^2 \\ v^k &= \operatorname*{argmin}_{v \in \mathcal{G}} \mathcal{F}(u^k, v, \mu) + \frac{1}{\theta_k}\|v - v^{k-1}\|^2, \end{aligned} \tag{9}$$

where $\mathcal{F}(u^k, v, \mu)$ is the function we minimize,

$$\mathcal{F}(u, v, \mu) = \int_{\Omega} \|\nabla_{\mathcal{M}} u\| + \frac{r}{2}\|v-u\|^2 + \mu^T(v-u) + \|u-u_0\|^2 d\Omega \tag{10}$$

The proof of convergence become quite easy, as shown by Attouch et al. [2, Lemma 5].

We now modify the augmented Lagrangian TV framework for the smoothness term described in (1).

2.1 Augmented Lagrangian TV Optimization of Vector Valued Functions on Parametric Surfaces

Let us put aside for a moment the matrix-valued nature of u, and formulate an efficient iterative scheme for smoothing (in the total variation sense) a function on a parametric surface. In our case, the surface is the visible surface obtained from a range scanner, and the parametrization domain is the image plane with its coordinates system. In order to regularize efficiently images given on parametric surfaces, we use the parametrization domain sampled on a Cartesian grid.

The scheme we present is based on the augmented Lagrangian TV optimization scheme [24]. 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 [24]. 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

$$\max_{\mu_2} \min_{u,p} \int_{\Omega} \|p\| + \frac{r_2}{2} \|J_{\mathcal{M}} \nabla u - p\|^2 + \mu_2^T (J_{\mathcal{M}} \nabla u - p) + \lambda \|u - u_0\|^2 d\Omega, \quad (11)$$

where μ_2 is our Lagrange multiplier for the gradient constraint. The optimization of (11) with respect to u is given by a diffusion equation

$$-r_2 \operatorname{div} \left(J_{\mathcal{M}}^T (J_{\mathcal{M}} \nabla u - p) \right) + \operatorname{div} J_{\mathcal{M}}^T \mu_2 + 2\lambda (u - u_0) = 0. \quad (12)$$

Optimization with respect to p can be expressed in closed form [25, 24] by a shrinkage operator,

$$p = \left(1 - \frac{1}{r_2} \frac{1}{\|w\|} \right) w, w = J_{\mathcal{M}} \nabla u - \frac{\mu_2}{r_2}, \quad (13)$$

where $\| \cdot \|$ is the Frobenius norm.

Finally, updating μ is given according to the augmented Lagrangian method [17, 11] by the update equation

$$\mu_2^k = \mu_2^{k-1} + r_2 \left(p^k - J_{\mathcal{M}} \nabla u^k \right). \quad (14)$$

We now turn to describe the

2.2 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

$$\max_{\mu,\mu_2} \min_{v \in SE(3)} \int_{\Omega} \left[\begin{array}{c} \|p\| + \frac{r}{2} \|u - v\|^2 + \mu^T (u - v) + \\ \frac{r_2}{2} \|J_{\mathcal{M}} \nabla u - p\|^2 + \mu_2^T (J_{\mathcal{M}} \nabla u - p) + \lambda \|u - u_0\|^2 \end{array} \right] d\Omega. \quad (15)$$

Minimization with respect to u is the same as in the scalar TV case, resulting in a diffusion equation. We now describe the update rule for v and p.

2.3 Minimization with respect to v

Optimization with respect to v is using the same projection operator per-pixel as in [18]. Looking at optimization with respect to v, we obtain

$$\underset{v \in SE(3)}{\operatorname{argmin}} \frac{r}{2}\|v-u\|^2 + \langle \mu, u-v \rangle = \underset{v \in SE(3)}{\operatorname{argmin}} \frac{r}{2}\left\|v - \left(\frac{\mu}{r}+u\right)\right\|^2 = \underset{SE(3)}{\operatorname{Proj}}\left(\frac{\mu}{r}+u\right), \quad (16)$$

where $\operatorname{Proj}_{SE(3)}(\cdot)$ denotes the orthogonal projection operator onto $SE(3)$, given by a *singular value decomposition*.

The optimization of the Lagrange multipliers is done according to the update rule

$$\mu^k = \mu^{k-1} + r\left(v^k - u^k\right); \quad \mu_2^k = \mu_2^{k-1} + r_2\left(p^k - J_{\mathcal{M}}\nabla u^k\right). \quad (17)$$

An algorithmic description of the resulting scheme is given as Algorithm 1.

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

1: **for** $k = 1, 2, \dots,$ until convergence **do**
2: Update $u^k(x)$, according to (12).
3: Update $p^k(x)$, according to (13).
4: Update $v^k(x)$, by projection onto the matrix group, using SVD.
5: Update $\mu^k(x), \mu_2^k(x)$, according to (17).
6: **end for**

2.4 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 non-rigid registration process, followed by estimation of the locally-rigid motion that takes place between the two point-clouds. In order to obtain a correspondence between two time-frames of a depth video, we employ a version of a non-rigid ICP algorithm, similar to the approach suggested by Li et al. [14].

We model the motion between the two point clouds using a simple additive model,

$$\mathbf{y}(\mathbf{x}) = \mathbf{x} + \mathbf{w}(\mathbf{x}), \quad (18)$$

where $\mathbf{x}, \mathbf{w}(\mathbf{x}) \in \mathbb{R}^3$ denote object points and the corresponding motion vector, and $\mathbf{y}(\mathbf{x})$ denotes the point corresponding to $\mathbf{x}$. While in our examples use a point-to-point distance, other, more accurate, distance functions [16] can also be used.

Enforcing the local smoothness of the motion field, we obtain a simple deformation energy, similar in a sense to optical flow functionals [12],

$$E_{REG} = \int_{\Omega} \lambda_D \psi_D \left(\|\mathbf{x} + \mathbf{w} - \mathbf{y}(\mathbf{x})\|^2 \right) + \psi_S \left(\|\nabla \mathbf{w}\|^2 \right) d\Omega, \tag{19}$$

where λ_D defines the relative importance of the data fitting term.

In order to obtain the deformation we iteratively update $\mathbf{y}(\mathbf{x})$ and solve for $\mathbf{w}(\mathbf{x})$ using the linearized equation resulting from the Euler-Lagrange condition,

$$\mathrm{div}_S \left(\psi_S' (\cdot) \nabla_S \mathbf{w} \right) = 2\lambda_D \psi_D' (\cdot) \left(\mathbf{y}(\mathbf{x} + \mathbf{w}) - \mathbf{x} \right), \tag{20}$$

where div_S and ∇_S denote respectively the divergence and Laplacian operators on a point cloud, and $\psi_S' (\cdot), \psi_D' (\cdot)$ are the derivatives of the robust functions with respect to their argument. This linearized equation is solved iteratively in a Gauss-Seidel manner, as in optical flow algorithms (see for example, [7]). At each linearization step we compute a corredponding point $\mathbf{y}(\mathbf{x})$ for each point $\mathbf{x}$, as is often the case for non-rigid ICP algorithms. While we used an approximate nearest neighbor search tree in $3D$ [1], the small changes between frames allow to use a 2D registration for a more efficient implementation. Finally, we note that as in ICP algorithms, various measures can be used to remove inconsistent point correspondences [21] in order to obtain robustness to outliers in the fitting process. This was not necessary in our case, and is left for future work. In order to compute the divergence, gradient, and Laplacian operators we use a local polynomial model for first derivatives and Belkin et al. approximation [3] for the Laplacian. The deformation scheme is described as Algorithm 2.

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 instantenous motion components are quite noisy. 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 2.2. The overall algorithm is summarized as Algorithm 3. During the third step of the algorithm, when smoothing the motion field, different λ 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 [4].

Algorithm 2 Regularized estimation of rigid motion from depth video

SELECT($\{\mathbf{x}\}$), a set of representative points.

1: **for** $k = 1, 2, \dots,$ until convergence **do**
2: Compute corresponding points $\mathbf{y}(\mathbf{x})$ for each point in the point cloud.
3: Compute IRLS weights, $\psi_S' (\cdot), \psi_D' (\cdot)$.
4: Update $\mathbf{w}(\mathbf{x})$ by solving (via Gauss-Seidel iterations) Equation 20.
5: **end for**

Algorithm 3 Regularized estimation of rigid motion from depth video

1: **for** $k = 1, 2, \ldots,$ until convergence **do**
2: Estimate smooth motion field between depth frames according to (20).
3: Estimate $u_0(x)$ at each point using least median squares fitting.
4: Regularize $u_0(x)$ using Algorithm (1).
5: **end for**

References

1. S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu. An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM*, 45(6):891–923, 1998.
2. H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Lojasiewicz inequality. *Math. Oper. Res.*, 35:438–457, 2010.
3. M. Belkin, J. Sun, and Y. Wang. Constructing Laplace operator from point clouds in rd. In *Symposium on Discrete Algorithms*, pages 1031–1040, Philadelphia, PA, USA, 2009. SIAM.
4. H. Benhabiles, G. Lavoué, J.-P. Vandeborre, and M. Daoudi. Learning boundary edges for 3D-mesh segmentation. *Comp. Graphics Forum*, 2011.
5. P. J. Besl and N. D. McKay. A method for registration of 3D shapes. *IEEE Trans. PAMI*, 14(2):239–256, 1992.
6. 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.
7. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *ECCV*, pages 25–36, Prague, Czech Republic, May 2004. Springer Verlag.
8. Y. Chen and G. Medioni. Object modelling by registration of multiple range images. *Image Vision Comput.*, 10:145–155, April 1992.
9. P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. May 2010.
10. M. P. do Carmo. *Differential Geometry of Curves and Surfaces*. Prentice-Hall, 1976.
11. M. R. Hesteness. Multipliers and gradient methods. *J. of Optimization Theory and Applications*, 4:303–320, 1969.
12. B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.
13. R. Lai and T. F. Chan. A framework for intrinsic image processing on surfaces. *Comput. Vis. Image Underst.*, 115:1647–1661, Dec. 2011.
14. 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.
15. A. Myronenko and X. B. Song. Point-set registration: Coherent point drift. *CoRR*, abs/0905.2635, 2009.
16. H. Pottmann and M. Hofer. Geometry of the squared distance function to curves and surfaces. In H. C. Hege and K. Polthier, editors, *Visualization and Mathematics III*, pages 223–244. Springer, 2003.

17. M. J. Powell. *Optimization*, chapter A method for nonlinear constraints in minimization problems, pages 283–298. Academic Press, 1969.
18. G. Rosman, Y. Wang, X.-C. Tai, R. Kimmel, and A. M. Bruckstein. Fast regularization of matrix-valued images. Technical Report CAM11-87, UCLA, 2011.
19. G. Rosman, Y. Wang, X.-C. Tai, R. Kimmel, and A. M. Bruckstein. Fast regularization of matrix-valued images. In *ECCV*, 2012. Accepted.
20. L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D Letters*, 60:259–268, 1992.
21. S. Rusinkiewicz and M. Levoy. Efficient variants of the ICP algorithm. In *Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, June 2001.
22. A. Spira and R. Kimmel. An efficient solution to the Eikonal equation on parametric manifolds. In *INTERPHASE 2003 meeting*, volume 3, pages 315–327, 2003.
23. A. Spira and R. Kimmel. Geometric curve flows on parametric manifolds. *J. Comput. Phys.*, 223:235–249, April 2007.
24. X.-C. Tai and C. Wu. Augmented Lagrangian method, dual methods and split Bregman iteration for ROF model. In *SSVM*, pages 502–513, 2009.
25. 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.
26. 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.