Second-Order Shape Optimization for Geometric Inverse Problems in Vision

Jonathan Balzer
University of California
Los Angeles, CA 90095
USA
balzer@cs.ucla.edu

Stefano Soatto
University of California
Los Angeles, CA 90095
USA
soatto@cs.ucla.edu

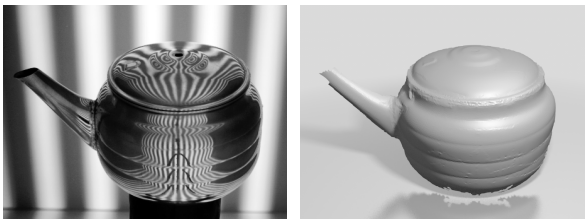
Abstract

We develop a method for optimization in shape spaces, i.e., sets of surfaces modulo re-parametrization. Unlike previously proposed gradient flows, we achieve superlinear convergence rates through an approximation of the shape Hessian, which is generally hard to compute and suffers from a series of degeneracies. Our analysis highlights the role of mean curvature motion in comparison with first-order schemes: instead of surface area, our approach penalizes deformation, either by its Dirichlet energy or total variation, and hence does not suffer from shrinkage. The latter regularizer sparks the development of an alternating direction method of multipliers on triangular meshes. Therein, a conjugate-gradient solver enables us to bypass formation of the Gaussian normal equations appearing in the course of the overall optimization. We combine all of these ideas in a versatile geometric *variation-regularized Levenberg-Marquardt-type method applicable to a variety of shape functionals, depending on intrinsic properties of the surface such as normal field and curvature as well as its embedding into space. Promising experimental results are reported.*

1. Introduction

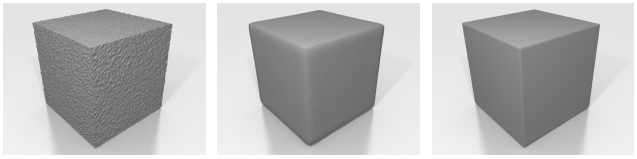
1.1. Motivation

Many inference tasks in vision amount to solving *inverse problems*, where a solution is understood to be an element x in a set X which, given some model $f : X \times M \mapsto M$, minimizes the *residual* $r(x) = f(x, s) - t$ between a *signal* s and its prediction under f. For instance, in optical flow, one wishes to minimize the distance between an image $\mathcal{I}_s$ and a warped version $\mathcal{I}_t \circ w$ of $\mathcal{I}_s$ w.r.t. w in the diffeomorphism group $X = \mathrm{diff}(D)$ of the image plane. In this paper, we are interested in the case where M is a linear space of functions, e.g., $BV(D)$ or $H^1(D)$, over some geometric domain D, but – quite similar to the example of $\mathrm{diff}(D)$ – the set of latent variables X is *not*. Instead, X is a *shape space* consisting of three-dimensional (3-d) surfaces up to re-parametrizations. The literature offers a wealth of first-order numerical methods. But despite their superior convergence properties, to this date there are no generally applicable second-order methods for shape optimization. This is explained by the difficulties in accurately and efficiently approximating the Hessian. We focus on a class of separable quadratic functionals to propose a second-order numerical method for solving visual inference problems on shape spaces. As shown in Sect. 2.2.2, the construction suppresses eigenspaces of the Hessian which are responsible for shrinking biases in traditional gradient flows. To ensure regularization, we suggest penalizing variations, not of the iterated surface itself, but *deformations* thereof. This leads to a variant of the classic Levenberg-Marquardt method which can be applied under weak assumptions on f by breaking down the nonlinear and possibly nonconvex global optimization problem into a sequence of convex subproblems. Depending on the choice of regularizer, one type of subproblem encountered is equivalent to the Rudin-Osher-Fatemi (ROF) model for image denoising and segmentation [20]. To solve it numerically, we develop an extension of the alternating direction method of multipliers (ADMM) to surfaces represented by trian-



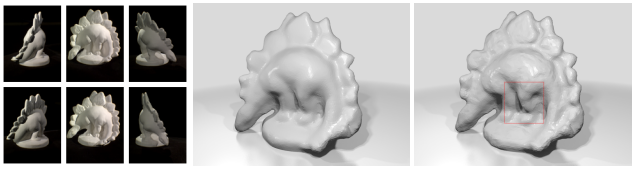
(a) (b)

Figure 1. Application *reconstruction of specular surfaces*: (a) Correspondences between camera pixels and scene points viewed along the surface are established by a structured-light coding technique. (b) The correspondences can be converted into normal information which is then integrated into a visible surface reconstruction by our method.



(a) Noisy cube (b) ML-LMTV denoising (c) ROF-LMTV denoising

Figure 2. Application *surface denoising*.



(a) Input images (b) Initialization (c) Optimized mesh

Figure 3. Application *photometric optimization*: the size of the shadows in the red-marked region in (c) suggests noticeable changes of the initial shape.

gular meshes (Sect. 2.3). We demonstrate that the chosen class of separable quadratic functions applies to a variety of problems relevant to vision, from mesh reconstruction from unorganized point clouds and deflectometric images (Fig. 1), surface denoising (Fig. 2), to photometric optimization (Fig. 3), which will all be explored in Sect. 3.

1.2. Relation to prior work

The natural question arises of why we should develop *local* methods – even of second order – when globally optimizable convex programs for many reconstruction tasks have been proposed, cf. [7, 17]. The short answer is that first, in these models, convexity originates from embedding the unknown surface into a linear space through some implicit representation such as a distance or characteristic function. We would like to avoid such resource-hungry representations as much as possible and restrict their use to as-coarse-as-possible initialization. Second, as soon as visibility, which in turn depends on the optimization variable itself, is fully considered in these models, convexity will be lost. There are some analogies between the present paper and [19] in the sense that the problem of interest is decomposed into a sequence of nondifferentiable subproblems: The latter generalizes Candés' reweighted ℓ_1-algorithm, and the goal of the decomposition is to handle nonconvex regularizers. Sect. 2.2.2 is an extension of [4], where a regularization-free *Gauss-Newton* method was presented especially for normal field integration. The ADMM has been adapted to linear spaces over surfaces before, first by Wu et al. [24], later by Liu and Leung [18]. The authors of the latter paper are concerned with point-based surface models. The former approach is different from ours in that it explicitly forms Gaussian normal equations at every iteration. In comparison, we suggest executing a few preconditioned conjugate-gradient steps on the corresponding *overdetermined* linear system. A similar trick has been proposed previously for large-scale bundle adjustment [8]. The optimization framework developed in this paper is fairly general but applied to the sample problems in Sect. 3, it inherits some of the ideas found in the specialized literature: Similar to Avron et al. [1], we couple denoising of the normal field with subsequent integration for the purpose of surface fairing (Sect. 3.1) and reconstruction (Sect. 3.2). Inspired by [16], we consider orientation information for reconstruction but prefer explicit surface models and account for the nonlinearity of f, the Gauss map. Geometric applications of the Split Bregman method have been studied in [14], but different from Sect. 3.2 within a level set segmentation framework. The body of literature on our third sample application – photometric optimization from multiview stereo images – is too vast to do it justice here. Let us only explicitly mention the works [10, 15, 23] because they feature shape optimization albeit of first order.

2. Main result

2.1. From Gauss-Newton to Levenberg-Marquardt

With the notation introduced at the beginning of Sect. 1.1, any nonlinear least-squares (LS) problem takes the form

$$\min_{x\in X} \frac{1}{2}\|r(x)\|^2_{L^2(D)}. \tag{1}$$

The default optimization strategy is the Gauss-Newton algorithm, which exploits the fact that the Hessian of the L^2-energy at some $x_k \in X$ can be accurately approximated by the differential Dr of the residual: $D^2E_d|_{x_k} \approx (Dr|_{x_k})^\top Dr|_{x_k}$ in x_k. In combination with the classic Newton method, this gives the implicit time step

$$x_{k+1} = x_k + v \tag{2a}$$

in which the *update* v solves the linear equation system

$$Dr^\top Drv = -Dr^\top r. \tag{2b}$$

An illustrative interpretation of (2) is the following: Approximating r by its first order Taylor expansion $r(v) \approx r(x_k) + Dr|_{x_k} v$ results in a *local linear* LS problem

$$\min_v \frac{1}{2}\|r(x_k) + Dr|_{x_k} v\|^2_{L^2(D)} \tag{3}$$

whose normal equation is precisely (2b). Returning to the introductory example of optical flow, the Gauss-Newton method applied to the *nonlinear* gray value conservation law under a translational deformation model, in which w is assumed spatially constant, is – apart from the multiscale

strategy it adopts to avoid falling into local minima – equivalent to the Lucas-Kanade algorithm [3].

Occasionally, Eq. (2b) becomes underconstrained which causes ill-conditioning of $Dr^\top Dr$ by creating zeros among its eigenvalues. In the Levenberg-Marquardt method in its original form, the issue is addressed by adding $\frac{\lambda}{2}\|v\|^2_{L^2(D)}$ to (3). The new local approximation of the energy limits the step size (i.e., the size of the *trust region*) inversely proportional to the magnitude of λ. The regularizer appears in the normal matrix as λ-fold multiple of the identity, and hence, pushes the spectrum of the former towards positive values by an amount of λ. When the update step v possesses some form of spatial regularity, we can punish large *variations* in lieu of large magnitudes of v by replacing (3) with

$$\min_v \frac{1}{2}\|r(x_k) + Dr|_{x_k} v\|^2_{L^2(D)} + \frac{\lambda}{p}\|Dv\|^p_{L^p(D)}. \quad (4)$$

For $p = 2$, the regularizer equals the Dirichlet energy which maintains linearity of the local LS problem. For $p = 1$, Eq. (4) becomes the ROF functional. Bachmayr and Burger point out this connection in [2]. We will generalize the resulting variation-regularized Levenberg-Marquardt method consisting of Eqs. (2a) and (4) from vector spaces to sets of surfaces in 3-d. The necessary theoretical foundations are laid out in the following section before stating our main result in Sect. 2.2.2.

2.2. Formulation in shape space

2.2.1 Shape spaces, functions, and gradient flows

When we speak of *shape*, we mean the trace of a surface, i.e., the collection of its points in a set-theoretic sense modulo its symmetry group, which consists of all smoothness-preserving re-parametrizations. Let S_0 be the boundary of a smooth reference subdomain of $\mathbb{R}^3$. The set of all diffeomorphic embeddings $\operatorname{diff}(S_0, \mathbb{R}^3)$ becomes a *shape space* as soon as two embeddings $w, z \in \operatorname{diff}(S_0, \mathbb{R}^3)$ are considered equivalent if they differ by some $\tau \in \operatorname{diff}(S_0, S_0)$, i.e., $w = z \circ \tau$. Let us remark that in particular, all elements of $X = \operatorname{diff}(S_0, \mathbb{R}^3)/\operatorname{diff}(S_0, S_0)$ exhibit the same topology, namely that of S_0. A shape space has the structure of an infinite-dimensional manifold [11]. We will not worry about its intriguing geometrical and topological properties. For all practical purposes, it suffices to acknowledge that the tangent space of this manifold at a "point" S consists of all infinitesimal normal velocities v in $H^1(S)$ respectively $BV(S)$. This is quite intuitive: tangential deformations map surface points to surface points, do not alter shape, and hence preserve the equivalence class of S. We can also conduct analysis. An important example of a *shape function* is the surface integral

$$E(S) = \int_S \phi(S)\,\mathrm{d}S. \quad (5)$$

We admit *costs* $\phi(\boldsymbol{x}, \boldsymbol{n})$ depending on $\boldsymbol{x} \in \mathbb{R}^3$ as well as the unit surface normal $\boldsymbol{n} \in \mathbb{S}^2$, but generalizing what follows to higher-order differential surface properties (e.g, the Willmore energy, cf. Appx. A in the appending technical report) is possible. The *shape differential* of E at some S in the tangential direction of v is given by

$$DE(S; v) = \int_S \underbrace{\left(\kappa\phi + \langle \nabla\phi, \boldsymbol{n}\rangle - \nabla_S^\top \nabla_{\mathbb{S}^2}\phi\right)}_{g_E} v\,\mathrm{d}S, \quad (6)$$

where κ denotes the mean curvature and g_E the *shape gradient* of E. A derivation of this formula can be found in several places, cf. [11, 13, 15, 22], its application in many more, cf. [9, 10, 23]. Note that the domain of ϕ may extend to the embedding space $\mathbb{R}^3 \times \mathbb{R}^3 \supset S \times \mathbb{S}^2$. Correspondingly, ∇ is the *Euclidean* nabla operator, whereas ∇_S and $\nabla_{\mathbb{S}^2}$ denote the *intrinsic* or *surface gradient* on S respectively the unit sphere $\mathbb{S}^2$. Also note that because $\ker \nabla_S^\top = \operatorname{span}(\boldsymbol{n})$ pointwise, it is sufficient to calculate the Euclidean derivative of ϕ w.r.t. $\boldsymbol{n}$ without reprojecting onto $\mathbb{S}^2$. By evolving some S_0 in the steepest descent direction $-g_E$ according to

$$S_{k+1} = S_k - (\kappa\phi + \langle \nabla\phi, \boldsymbol{n}\rangle - \nabla_S^\top \nabla_{\mathbb{S}^2}\phi)\boldsymbol{n}(S_k), \quad (7)$$

we can decrease (5) in two ways: either by reducing the surface area[1] via (weighted) mean curvature motion (MCM) in the direction $-\kappa\phi$; alternatively, we let each point follow the direction of greatest decrease of the cost function $\nabla\phi$ respectively $\nabla_{\mathbb{S}^2}\phi$. The stationary point of the descent, at which $g_E = 0$, will be determined by the equilibrium between these two forces. This equilibrium is responsible for a phenomenon called *minimal surface bias*: First, whenever the descent direction w.r.t. ϕ is uninformative in the sense that $\nabla\phi = \mathbf{0}$ while $\phi > 0$, the evolution will locally favor surfaces of minimal area. Second, due to the counterforce, the limit surface cannot fully account for the regularity of ϕ leading to visible oversmoothing and retraction of boundaries if present. Finally, where both ϕ *and* its derivatives w.r.t. $\boldsymbol{x}$ and $\boldsymbol{n}$ vanish[2], the evolution (7) will stagnate.

2.2.2 Hessian-free superlinear optimization

While the computation of first-order shape differentials like (6) is relatively straightforward, nonzero geodesic curvature of shape spaces significantly aggravates this process for second-order derivatives [11]. So far, the lack of symmetry and positive-definiteness have defeated any attempt to implement a pure Newton method for (5). Our key insight is that this problem can be circumvented under the

[1] The area integral measures – up to some material properties inherent in ϕ – the *tangential strain* or *membrane energy* of a surface.

[2] As an example, consider the re-projection error of multiple views onto a homogeneously textured surface region.

condition that ϕ is *separable* and *quadratic*:

$$\phi(\boldsymbol{x}, \boldsymbol{n}) = \frac{1}{2}(\|\boldsymbol{r}_x(\boldsymbol{x})\|^2 + \|\boldsymbol{r}_n(\boldsymbol{n})\|^2). \tag{8}$$

The residual $\boldsymbol{r}_x$ of ϕ over $\mathbb{R}^3$ arises from the (dis)location of surface points in space. Note that $\boldsymbol{r}_x$ is generally vector-valued, e.g., to account for multi-channel images or distances to known points (Sect. 3.2 and 3.3). The shape differential of $\boldsymbol{r}_x$, describing the impact of infinitesimal normal deformations v on the value of $\boldsymbol{r}_x$, is directly given by $D\boldsymbol{r}_x(v\boldsymbol{n})$. In perfect analogy, the normal error $\boldsymbol{r}_n$ is a map taking $\boldsymbol{n} \in \mathbb{S}^2$ to the embedding space[3] $\mathbb{R}^3$ with Jacobian $D_{\mathbb{S}^2}\boldsymbol{r}_n : T\mathbb{S}^2 \to T\mathbb{R}^3$. Invocation of the chain rule yields $-D_{\mathbb{S}^2}\boldsymbol{r}_n\nabla_S v$ for the shape differential of $\boldsymbol{r}_n$. Here, we have used the fact that pure infinitesimal rotations of the normal are related to the velocity v by its negative surface gradient $-\nabla_S v$, cf. [4, Prop. 1]. The shape differentials of $\boldsymbol{r}_x$ and $\boldsymbol{r}_n$ enable a local quadratic approximation

$$E_d(v) := \frac{1}{2}\|\boldsymbol{r}_x(\boldsymbol{x}) + D\boldsymbol{r}_x(v\boldsymbol{n})\|^2_{L^2(S)} + \frac{1}{2}\|\boldsymbol{r}_n(\boldsymbol{n}) - D_{\mathbb{S}^2}\boldsymbol{r}_n\nabla_S v\|^2_{L^2(S)} \tag{9}$$

of (5) around S. The equivalence of (4) and (2b) then immediately implies a shape space analogue of (2):

$$S_{k+1} = S_k + v\boldsymbol{n}(S_k) \tag{10a}$$

where the normal velocity v is the unique minimizer of

$$E_d(v) + \frac{\lambda}{p}\|\nabla_S v\|^p_{L^p(S_k)}. \tag{10b}$$

As shown in Sect. 2.3, there are efficient ways of minimizing this function (for fixed S_k and $p = 1, 2$).

Remarkably, while the steepest descent (7) used in previous approaches strives to reduce surface area, the solution of the local subproblem (10b) does not. The simple explanation is that minimization is performed w.r.t. the velocity field v and coupled with the properties of the surface only through the shape differentials of $\boldsymbol{r}_x$ and $\boldsymbol{r}_n$. Unfortunately, problems arise if the cost $\phi = \phi(\boldsymbol{x})$ is independent of the normal, like in the applications discussed in Sects. 3.2 and 3.3. When $\boldsymbol{r}_n = \boldsymbol{0}$, the minimizer of (9) can be obtained in closed form:

$$v\boldsymbol{n} = (D\boldsymbol{r}_x)^{-1}\boldsymbol{r}_x.$$

This, however, requires $D\boldsymbol{r}_x$ to be of full rank, a condition which can never hold in the vicinity of a stationary point where $D\boldsymbol{r}_x$ should be identically zero. A more intuitive explanation is the following: Loss of the mean curvature term in the descent rule cannot remain without consequences. Surface area correlates with surface smoothness.

[3] The example of the difference between two unit vectors shows that clearly the image of $\boldsymbol{r}_n$ is not necessarily contained in $\mathbb{S}^2$.

(a) (b) $\lambda = 10$

Figure 4. (a) The cameraman test image with additive Gaussian noise of standard deviation $\sigma = 20$ texture-mapped onto the Stanford bunny. (b) Result of ROF denoising on the surface.

Without the binding influence of κ, points on the surface will be able to move around *separately*, quickly compromising its integrity unless $\boldsymbol{r}_x$ is unrealistically smooth. The regularization term in (10b) comes to the rescue by enforcing either harmonic ($p = 2$) or piecewise constant ($p = 1$) *descent directions* or *deformations* in (10a). A convenient side effect is that the regularizer will inpaint nonzero values of v to regions where both $\boldsymbol{r}_x$ and $D\boldsymbol{r}_x$ vanish and a gradient descent would come to a complete halt (as discussed at the end of Sect. 2.2.1).

2.3. Conjugate gradient ADMM on triangular meshes

To minimize (10b) efficiently, we now describe a variation of the ADMM on surfaces. Since, in the end, we are interested in designing a numerical algorithm, let us consider finite-dimensional representations of S and the function spaces on it. In particular, let us assume we have a triangulation S_h of S but emphasize that the continuous formulation in Sect. 2.2.2 equally admits other kinds of discretizations, like e.g. with zero-sets of a scalar-valued function on $\mathbb{R}^3$. The precise details, in particular the lengthy derivation of the mass matrices $\mathbf{W}_x$ and $\mathbf{W}_n$ as well as the discrete nabla operator ∇_{S_h} on S_h, must be deferred to the technical report accompanying this paper[4]. We collect the Jacobians and residuals from (9) in

$$\mathbf{A} = \begin{pmatrix} \mathbf{W}_x \operatorname{diag}(D\boldsymbol{r}_x) \\ -\mathbf{W}_n \operatorname{diag}(D_{\mathbb{S}^2}\boldsymbol{r}_n)\nabla_{S_h} \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} \mathbf{W}_x\mathbf{r}_x \\ \mathbf{W}_n\mathbf{r}_n \end{pmatrix}.$$

The upper block-diagonal matrix is assembled from the values that the corresponding continuous quantities take at the vertices, the lower half respectively from the values on the faces. With these abbreviations in place, starting from $\mathbf{v}_0 = \mathbf{d}_0 = \mathbf{b}_0 = \mathbf{0}$, the ADMM for minimization of (10b) iterates the following three steps:

$$(\mathbf{A}^\top\mathbf{A} - \lambda\mu\nabla_{S_h}^\top\nabla_{S_h})\mathbf{v}_{k+1} = \mathbf{A}^\top\mathbf{f} + \lambda\mu\nabla_{S_h}^\top\mathbf{d}_k, \tag{11a}$$

$$\mathbf{d}_{k+1} = \operatorname{prox}_{\|\cdot\|_{\ell^1}}(\nabla_{S_h}\mathbf{v}_{k+1} + \mathbf{b}, \mu^{-1}), \tag{11b}$$

$$\mathbf{b}_{k+1} = \mathbf{b}_k + \nabla_{S_h}\mathbf{v}_{k+1} - \mathbf{d}_{k+1}. \tag{11c}$$

Note that the roles of μ and λ have switched opposed to the

[4] http://arxiv.org/abs/1311.2626

	Cube ML	Cube ROF	Teapot	Sphere	MVS
n	$24,578$	$24,578$	$17,974$	$7,842$	$34,834$
GD	0.94	0.89	0.56	0.43	30.9
LMD	0.68	0.67	1.98	0.48	28.4
LMTV	1.52	1.5	7.2	1.41	36.8

Table 1. Execution time for a single step in seconds. The first row contains the number n of vertices in the optimized mesh.

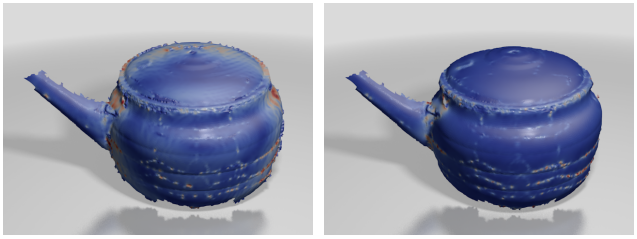
canonical notation in [14]. The reason is that our focus is on the Levenberg-Marquardt method here, in which the parameter discounting the step length is conventionally referred to by λ. We make the following modification to the original algorithm and its surface-based variant proposed in [24]: First, note that (11a) is the Gaussian normal equation of the LS problem associated with (9). The only advantage of working with the normal equation is that therein, $\mathbf{A}^\top\mathbf{A}$ becomes symmetric and strictly diagonally-dominant. This is exploited in [14] by invoking a simple and very efficient Jacobi solver. At the same time, small eigenvalues will become even smaller with deteriorating influence on the condition number. Additionally, discrete divergence and Laplace-Beltrami operators defined by $\operatorname{div}_{S_h} := \nabla_{S_h}^\top$ respectively $\Delta_{S_h} := \nabla_{S_h}^\top \nabla_{S_h}$ are inconsistent with discrete conservation laws, which may lead to numerical instabilities [12]. Last but not least, there is the cost of computing the matrix product. Therefore, we propose to rearrange the normal equation of (9) as follows:

$$\begin{pmatrix} \mathbf{A} \\ \lambda\mu\nabla_{S_h} \end{pmatrix} \mathbf{v}_{k+1} = \begin{pmatrix} \mathbf{f} \\ \lambda\mu\mathbf{d}_k \end{pmatrix}.$$

This linear system is now overdetermined but amenable to the Conjugate Gradients Least-Squares method [6], which avoids explicit formation of the normal equation. Its iterative nature allows us to preserve the inexactness of the original ADMM. When $p = 2$ and hence (9) is differentiable, setting $\mu = 1$ and $\mathbf{d}_k = \mathbf{0}$, we immediately obtain the update $\mathbf{v}_{k+1}$ from (11a) without the need for shrinkage (11b) and executing Bregman steps (11c).

3. Applications

All algorithms discussed in the following section were implemented in C++ and executed on a single 3.4 GHz core of a commodity computer with 12 GB of main memory. The source code is available for download on the first author's website. To begin with, we showcase the viability of our ADMM variant at hand of texture denoising (Fig. 4). Here, the surface remains static so we achieve essentially the same as [24, 18]. Applications of the method introduced in Sect. 2.2, in which the surface itself plays the role of the optimization variable, will be presented in the following sections. Thereby, we abbreviate the Levenberg-Marquardt method with a TV-regularizer ($p = 1$) by LMTV



(a) GD (b) Our method

Figure 5. Result from Fig. 1(b) overlayed with the colormapped error distribution in the interval $[0, 0.5]$.

respectively LMD when (10b) contains the Dirichlet energy ($p = 2$). We compare LMTV and LMD with the existing gradient descent (GD) scheme. Let us remark that its ad-hoc formulation (7) does not directly lend itself to implementation because it suffers from numerical stiffness due to the MCM term. Noticing that $\kappa\boldsymbol{n} = \Delta_S(S)$, i.e., the *mean curvature vector* $\kappa\boldsymbol{n}$ is just the Laplace-Beltrami operator $\Delta_S = \nabla_S^\top\nabla_S$ applied to the functions that embeds S into $\mathbb{R}^3$, we arrive at the backward Euler scheme

$$S_{k+1} + (\lambda+\phi)\Delta_S(S_{k+1}) = S_k - (\langle\nabla\phi, \boldsymbol{n}\rangle - \nabla_S^\top\nabla_{\mathbb{S}^2}\phi)\boldsymbol{n}. \quad (12)$$

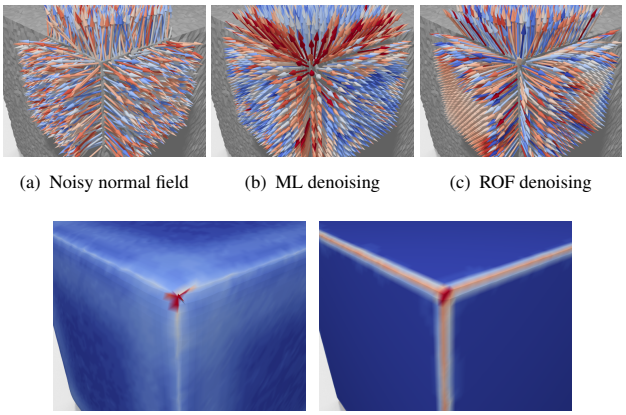
Additionally, a regularization weight λ has been introduced as a factor of κ, amplifying the smoothing effect of MCM if necessary. The price to pay for stability is the inversion of the matrix $\mathrm{id} + (\lambda+\phi)\Delta_S$ at each iteration. Consequently, the number of floating point operations per GD step is not significantly smaller than in LMD, see Tab. 1.

3.1. Normal field integration and denoising

Let $\boldsymbol{n}_d$ denote some *desired* normal field. Integration we understand as finding a surface S such that $\boldsymbol{n}(S) = \boldsymbol{n}_d(S)$. This is an inverse problem in the spirit of Sect. 1.1: noise in the data prevents the integrability of $\boldsymbol{n}$ and hence the existence of such a strong solution. Instead, we look for a minimizer of

$$E_n(S) = \int_S \frac{1}{2}\|\boldsymbol{n} - \boldsymbol{n}_d\|^2 \,\mathrm{d}S. \quad (13)$$

This energy constitutes a special case of (8) in which $\boldsymbol{r}_n(\boldsymbol{n}) = \boldsymbol{n} - \boldsymbol{n}_d$ and $\boldsymbol{r}_x(\boldsymbol{x}) = \mathbf{0}$. It is useful in a variety of applications which are classified by how they define the target normal field $\boldsymbol{n}_d$. Take for instance the deflectometric reconstruction of specular surfaces. In deflectometry, one measures the correspondence between pixels on the image plane and points in the scene they see *via* the specular surface, cf. Fig. 1(a) and [5]. Reconstructions are shown in Fig. 5. As seen in the convergence plot in Fig. 7(a), despite



(a) Noisy normal field (b) ML denoising (c) ROF denoising

(d) (e)

Figure 6. Integral surfaces of the normal fields in (b) and (c) obtained by LMD are shown in (d) respectively (e) together with the local residual, also see Fig. 2.

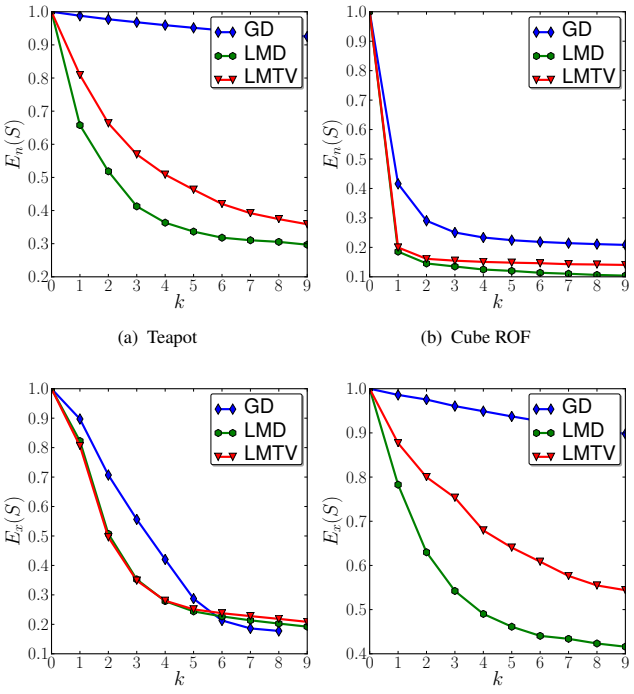
the implicit Euler integration (12), the gradient descent suffers from severe step size restrictions and terminates prematurely in a local minimum. Another example is fourth-order surface denoising [1]. The idea is that instead of smoothing the surface itself, which would involve a second-order diffusion equation, one first applies the smoothing to the normal field of the surface and in a second step integrates the result[5] $\boldsymbol{n}_d$. We obtain $\boldsymbol{n}_d$ either in terms of a maximum likelihood (ML) estimate (Fig. 6(b)) or from the output of our ADMM variant applied to the ROF-functional (Fig. 6(c)) of the original normal field (Fig. 6(a)). Remarkably, the convergence rate of GD at integration becomes competitive again given that the input data has undergone the initial smoothing (Fig. 7(b)).

3.2. Surface reconstruction from point clouds

Here, we are given a set of discrete points in space $\mathcal{P} = \{\boldsymbol{p}_l \in \mathbb{R}^3 \mid l \in \mathbb{N}\}$ (Fig. 9(a)), which are the representation of choice for many reconstruction methodologies embracing the triangulation principle. Our goal is to find a surface S with minimal average distance

$$E_x(S) = \int_S \frac{1}{2}\|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 \,\mathrm{d}\boldsymbol{x} \tag{14}$$

to $\mathcal{P}$ where $\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{p}_l \in \mathcal{P}} \|\boldsymbol{x} - \boldsymbol{p}_l\|$. Making the substitution $\boldsymbol{r}_x(\boldsymbol{x}) = \boldsymbol{x} - \hat{\boldsymbol{x}}$ respectively $\phi(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{r}_x\|^2$, this energy can be brought into the form (8) with $\boldsymbol{r}_n$ vanishing. Supposing that $\mathcal{P}$ is sufficiently dense, the global minimum with value 0 is given by the zero-set $\phi^{-1}(0)$ of $\phi(\boldsymbol{x})$. This direct approach requires a representation of the squared distance function over a Cartesian grid. Regularity and dimensionality of such a representation imply a trade-off between reconstruction quality and computational efficiency[6], making it difficult to take advantage of the full resolution of the raw data. While it lacks fine geometric details, $\phi^{-1}(0)$ generally captures the topology of the surface we wish to infer, thus providing an adequate initial guess S_0 for refinement by (10). The Jacobian $D\boldsymbol{r}_x|_{\boldsymbol{x}}$ is given by the vector that connects $\boldsymbol{x}$ with its closest point $\hat{\boldsymbol{x}} \in \mathcal{P}$. If the point cloud is oriented such that for each $\boldsymbol{p}_l \in \mathcal{P}$, we have a desired orientation $\boldsymbol{n}_d$, we can combine (14) and (13) with $\boldsymbol{r}_n(\boldsymbol{x}) = \boldsymbol{n}(\boldsymbol{x}) - \boldsymbol{n}_d(\hat{\boldsymbol{x}})$ similar to a *screened Poisson reconstruction* [16]. To test the performance of the different algorithms under ideal circumstances, we synthesized the toy example shown in Fig. 8(a). We observe at hand of Fig. 8 that LMTV is the only method that achieves a stable stationary state, which justifies the TV as a regularizer. Again, the error decay in the GD method is satisfactory, which can be attributed to the ideal circumstances and that the Euler steps are backwards. A more realistic scenario is depicted in Fig. 9. Here, the point cloud stems from an RGBD sensor. Given the known and regular topology of the image lattice, one can easily obtain a normal map for each depth image by finite-differencing (Fig. 9(b)). We refine the level of detail of an initial Poisson reconstruction (Fig. 9(c)) by minimizing a blend of the functionals in Eqs. (13) and (14).



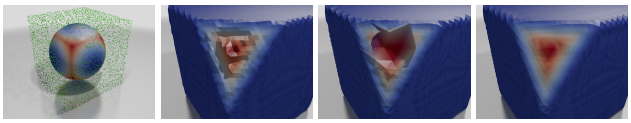
(a) Teapot (b) Cube ROF

(c) Sphere (d) Photometric optimization

Figure 7. Convergence rates of first- vs. second-order methods.

[5] In both steps, a second-order partial differential equation has to be solved, hence, we have a scheme of total order four.

[6] Insufficient spatial resolution is known to cause so-called *staircasing artifacts* at the numerical extraction of the level set.



(a) Initialization (b) GD (c) LMD (d) LMTV

Figure 8. (a) To interpolate a box-shaped point cloud with one corner cut off, we evolve a sphere towards minimal average distance. (b) GD suffers from a catastrophic loss of stability after 10 steps, (c) LMD develops folds after 30, while (d) LMTV remains stable.

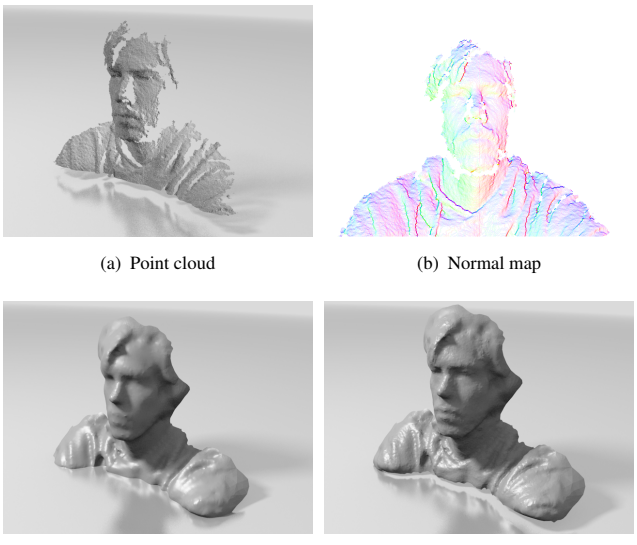
Thereby, the trade-off between point and normal fidelity can be steered by a scalar weight. The outcome is shown in Fig. 9(d).

3.3. Photometric optimization

Suppose we have a Lambertian surface of which we capture a set of gray value images from different vantage points. Multiview stereo is concerned with the inverse problem of converting the data into a geometric model of the surface. We cannot delve into the details of this highly sophisticated process. We limit the discussion to another application of our algorithm in a stage at the very end of the reconstruction pipeline, i.e., after an initial approximation of the surface as a set of (oriented) points has been armed with the topological structure of a surface. *Photometric optimization* seeks a minimizer of the shape functional

$$E_x(S) = \int_S \frac{1}{2}\rho^2 \, \mathrm{d}\boldsymbol{x}$$

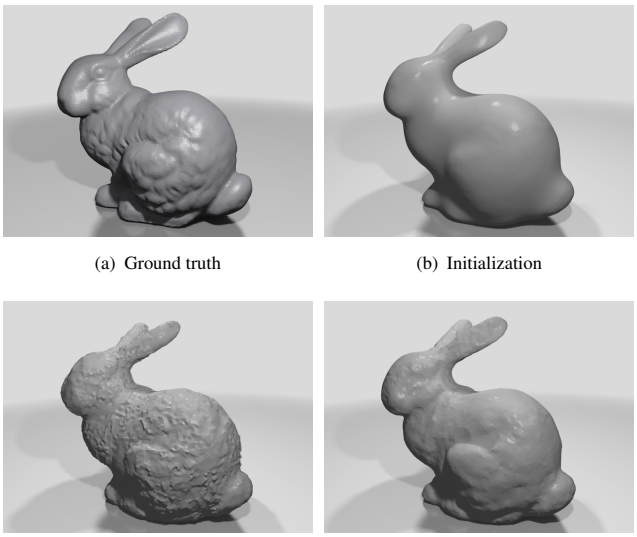
where ρ measures the instantaneous photoconsistency between pairs of images. Generally, ρ depends on shape and radiometry of the unknown surface as well as the set of vantage points (see the technical report for more details). Let us remark that we do not exploit additional knowledge on the location of the contour generators. We call ρ *instantaneous* because ρ also depends on *visibility*, which can only be modelled numerically, but not analytically. In local shape optimization, this typically happens at each iteration. To study our algorithm in a controllable scenario, we rendered synthetic images of the Stanford bunny (Fig. 10(a)) outfitted with a random texture. Needless to say, even under these conditions, perfect recovery of the ground truth model is all but impossible as it crucially depends on sufficient texture and sampling. As it is standard, we estimated the visual hull for an initialization but applied aggressive Laplacian smoothing to it to obtain a surface further away from the minimizer (Fig. 10(b)). Fig. 7(d) confirms the superior convergence rate of LMTV and LMD. The behavior of GD is similarly bad as in the teapot example, although the synthetic input data should be far less challenging. The stiffness of Eq. (12) and thus the maximal attainable step size is determined by the value of the regularization weight λ. In all our experiments, we were forced to set λ to extremely high values for GD to maintain stability yet encountered reconstructions such as in Fig. 10(c), unusually rugged compared to Fig. 10(d). We also ran our algorithm on the dino sparse ring data set described in [21], see Figs. 3 and 12. Plotting the surface evolved by gradient descent relative to the ground truth model (green) reveals the undesired effects of MCM during GD (Fig. 11). Even without contour constraints, these are absent in the second-order evolution. In fact, the opposite is true: we observe in Fig. 11(b) a local expansion relative to the initial guess (green). Finally, let



(a) Point cloud (b) Normal map

(c) Initialization (d) LMTV refinement

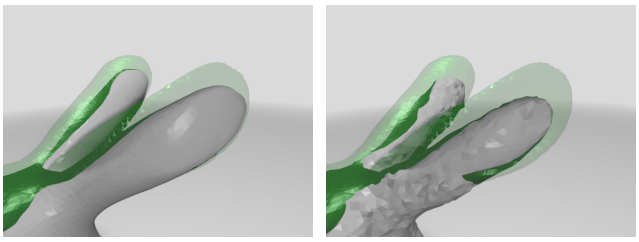
Figure 9. Hermite interpolation of points from an RGBD camera.



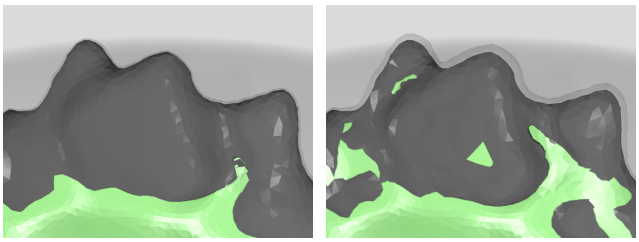
(a) Ground truth (b) Initialization

(c) GD after 48 steps (d) LMTV at $k = 5$

Figure 10. Photometric optimization based on synthetic input data.



(a) GD achieves regularity by favoring smaller surface area leading to shrinkage.



(b) Our method punishes large variations of the deformation at each step.

Figure 11. Contrary regularization strategies

us remark that evaluating ρ generates the majority of computational cost at each time step (Tab. 1). Since, independent of the optimization order, this cost scales quadratically in the number of views, superlinear convergence becomes critically important in the present application.

4. Conclusion and future work

We have presented a general second-order optimization method for shape functionals with several applications in the realm of visual reconstruction. We hope to pave the way for second-order methods in shape optimization but wish to further investigate their practical relevance in future work ourselves.

Acknowledgements

This research was supported by grants AFSOR FA9550-12-1-0364:P00001 and NGA HM02101310004.

References

[1] H. Avron, A. Sharf, C. Greif, and D. Cohen-Or. ℓ^1-Sparse reconstruction of sharp point set surfaces. *ACM T. Graphic.*, 29(5):1–12, 2010. 2, 6

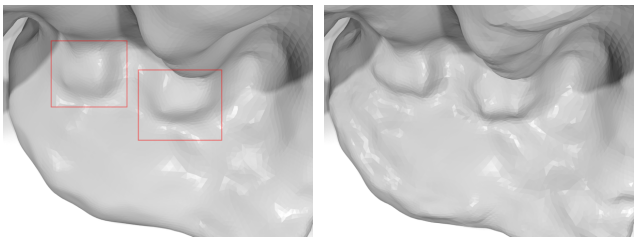
[2] M. Bachmayr and M. Burger. Iterative total variation schemes for nonlinear inverse problems. *Inverse. Probl.*, 25:1–26, 2009. 3

[3] S. Baker and I. Matthews. Lucas-Kanade 20 Years On: A Unifying Framework. *Int. J. Comput. Vision*, 56(3):221–255, 2004. 3

[4] J. Balzer. A Gauss-Newton Method for the Integration of Spatial Normal Fields in Shape Space. *J. Math. Imaging Vis.*, 441(1):65–79, 2012. 2, 4

[5] J. Balzer and S. Werling. Principles of Shape from Specular Reflection. *Measurement*, 43(10):1305–1317, 2010. 5

[6] A. Björck. *Numerical Methods for Least Squares Problems*. SIAM, 1996. 5



(a) $k = 0$ (b) $k = 5$

Figure 12. Additional LMD iterations following adaptive mesh refinement enhance small details like the dino's toes.

[7] Y. Boykov, V. Kolmogorov, D. Cremers, and A. Delong. An integral solution to surface evolution PDEs via geo-cuts. *Proc. ECCV*, 3953:409–422, 2006. 2

[8] M. Byröd and K. Åström. Conjugate gradient bundle adjustment. *Proc. ECCV*, pages 114–127, 2010. 2

[9] J. Chang, K. Lee, and S. Lee. Multiview normal field integration using level set methods. *Proc. CVPR IEEE*, 1:1–8, 2007. 3

[10] A. Delaunoy and E. Prados. Gradient Flows for Optimizing Triangular Mesh-based Surfaces: Applications to 3D Reconstruction Problems Dealing with Visibility. *Int. J. Comput. Vision*, 95(2):100–123, 2010. 2, 3

[11] M. Delfour and J. Zolésio. *Shapes and Geometries*. SIAM, 2001. 3

[12] M. Desbrun, A. Hirani, M. Leok, and J. Marsden. Discrete exterior calculus. Technical report, Caltech, 2005. 5

[13] B. Goldlücke, I. Ihrke, C. Linz, and M. Magnor. Weighted Minimal Surface Reconstruction. *IEEE T. Pattern Anal.*, 29(7):1194–1208, 2007. 3

[14] T. Goldstein and S. Osher. The Split Bregman Method for L^1-Regularized Problems. *SIAM J. Imag. Sci.*, 2(2):323, 2009. 2, 5

[15] H. Jin, A. Yezzi, and S. Soatto. Multi-View Stereo Reconstruction of Dense Shape and Complex Appearance. *Int. J. Comput. Vision*, 63(3):175–189, 2005. 2, 3

[16] M. Kazhdan and H. Hoppe. Screened poisson surface reconstruction. *ACM T. Graphic.*, 32(3):1–13, 2013. 2, 6

[17] K. Kolev, M. Klodt, T. Brox, and D. Cremers. Continuous Global Optimization in Multiview 3D Reconstruction. *Int. J. Comput. Vision*, 84(1):80–96, 2009. 2

[18] J. Liu and S. Leung. A Splitting Algorithm for Image Segmentation on Manifolds Represented by the Grid Based Particle Method. *Journal of Scientific Computing*, 56(2):243–266, 2012. 2, 5

[19] P. Ochs, A. Dosovitskiy, T. Brox, and T. Pock. An Iterated L^1-Algorithm for Non-smooth Non-convex Optimization in Computer Vision. *Proc. CVPR IEEE*, 1:1759–1766, 2013. 2

[20] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1):259–268, 1992. 1

[21] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A comparison and evaluation of multi-view stereo reconstruction algorithms. *Proc. CVPR IEEE*, 1:519–528, 2006. 7

[22] J. Solem and N. Overgaard. A Geometric Formulation of Gradient Descent for Variational Problems with Moving Surfaces. *Lect. Notes Comput. Sc.*, 3459:419–430, 2005. 3

[23] R. Tyleček and R. Šará. Refinement of Surface Mesh for Accurate Multi-View Reconstruction. *International Journal of Virtual Reality*, 9(1):45–54, 2010. 2, 3

[24] 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, 2011. 2, 5