Journal of Mathematical Imaging and Vision (2020)

This is a preprint of an article published in Springer

Journal of Mathematical Imaging and Vision

The final version is available online at: https://link.springer.com/article/10.1007/s10851-021-01038-y

On Inversion-Free Mapping and Distortion Minimization: Supplementary Material

Alexander Naitsat[c,**], Gregory Naitzat[d], Yehoshua Y. Zeevi[c]

[c] Viterbi Faculty of Electrical Engineering, Technion - Israel Institute of Technology
[d] University of Chicago,IL, Department of Statistics

ARTICLE INFO

Article history:
Received December 10, 2020

ABSTRACT

This supplemental document presents some additional results and addresses alternative approaches to optimal mapping and to computation of the distortion gradient. The main text and its supplemental document use a coherent numbering for references and a common bibliography for citations. Likewise, we adopt here identical abbreviations and notations to those introduced in the main text.

I Variational Approach

Distortion minimization processes can be interpreted in two major ways: **(i)** As a problem $\operatorname{argmin} E(\boldsymbol{x})$ of optimizing coordinate vectors $\boldsymbol{x} \in \mathbb{R}^{n|\mathcal{V}|}$, analyzed in Sections 4 and 5; **(ii)** as a variational problem $\operatorname{argmin} E(f)$, where the objective variable is a vector function f.

The two problems are equivalent in the discrete settings, since target vertex coordinates $\boldsymbol{x}$ define simplicial map $f[\boldsymbol{x}]$ unambiguously (see optimization problems (56)-(58) and (53)-(55), defined in Section 4.2).

However, in more general case, it is useful to consider the variational interpretation of problem (1). Notably, the descending step, (62), can be reformulated via the Euler-Lagrange formula for gaining better theoretical understanding of the problem and extending distortion energies to more general weight functions.

To this end, we redefine distortion energies as functionals $E[\boldsymbol{f}]$[13], operating above the space of smooth deformations $\boldsymbol{f} \in \operatorname{Diff}(\mathbb{R}^n)$,

$$E[\boldsymbol{f}] \triangleq \int_S w(\boldsymbol{r})\,\mathcal{D}(\boldsymbol{f},\boldsymbol{r})\,d\,\boldsymbol{r},$$

where $S = \operatorname{Dom}(\boldsymbol{f})$, $w : S \to \mathbb{R}^+$ are continuous (positive) weight functions and $\mathcal{D}(\boldsymbol{f},\boldsymbol{r})$ is a distortion measure. Consequently, we consider the optimal mapping problem in the context of the calculus of variations as the following minimization:

$$\underset{\substack{\boldsymbol{f} \in \operatorname{Diff}(S,\mathbb{R}^n) \\ \boldsymbol{f}|_{S_0} = \boldsymbol{f}^0\big|_{S_0}}}{\operatorname{argmin}} \quad E[\boldsymbol{f}], \tag{98}$$

where $\boldsymbol{f}^0$ is a given initial deformation of $S \subset \mathbb{R}^n$ and S_0 is a subset of S. Since $\mathcal{D}(\boldsymbol{f},\boldsymbol{r})$ depends on n variables $\boldsymbol{r} = (\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n)$ and n functions $\boldsymbol{f} = (\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n)$, including their partial derivatives, we consider Euler-Lagrange equation for a vector function of multiple variables. Consider a general functional

$$\mathcal{L} = \mathcal{L}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_n;\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n;\boldsymbol{f}_{1,1},\ldots,\boldsymbol{f}_{1,n},\boldsymbol{f}_{2,1},\ldots,\boldsymbol{f}_{n,n}),$$

that depends on position $\boldsymbol{r}$, function $\boldsymbol{f}(\boldsymbol{r})$ and on Jacobian matrix $d\boldsymbol{f}$, represented as the list of the first order partial derivatives: $\boldsymbol{f}_{i,j} \triangleq \dfrac{\partial \boldsymbol{f}_i}{\partial \boldsymbol{r}_j}$, $1 \le i,j \le n$. Then, according to [CH65]

**Corresponding author
e-mail: anaitsat@campus.technion.ac.il (Alexander Naitsat)

[13] We write $E[\boldsymbol{f}]$ instead of $E(f)$ to distinguish between the variational formulation and the coordinate-based formulation, introduced in (56).

Euler-Lagrange equation for the problem $\underset{f}{\operatorname{argmin}} \int_S \mathcal{L}\, d\boldsymbol{r}$, is

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_1} - \sum_{i=1}^{n} \dfrac{\partial}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_{1,i}} = 0\,, \\ \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_2} - \sum_{i=1}^{n} \dfrac{\partial}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_{2,i}} = 0\,, \\ \vdots \\ \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_n} - \sum_{i=1}^{n} \dfrac{\partial}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{f}_{n,i}} = 0\,. \end{cases} \tag{99}$$

Since, according to Theorem 2.1, $\mathcal{D}$ depends only on the map derivatives, substituting $\mathcal{L}(\boldsymbol{r}, \boldsymbol{f}, d\boldsymbol{f}) = w(\boldsymbol{r})\mathcal{D}(\boldsymbol{f}, \boldsymbol{r})$ into (99) yields

$$\begin{cases} \sum_{i=1}^{n} \dfrac{\partial w}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{D}}{\partial \boldsymbol{f}_{1,i}} = 0\,, \\ \sum_{i=1}^{n} \dfrac{\partial w}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{D}}{\partial \boldsymbol{f}_{2,i}} = 0\,, \\ \vdots \\ \sum_{i=1}^{n} \dfrac{\partial w}{\partial \boldsymbol{r}_i} \dfrac{\partial \mathcal{D}}{\partial \boldsymbol{f}_{n,i}} = 0\,. \end{cases} \tag{100}$$

Note that if the weight function $w(\boldsymbol{r})$ is constant, then the above Euler-Lagrange formula (100) is reduced to the trivial equality, $(0,\ldots,0) = (0,\ldots,0)$. Thus, we assume that, unlike piecewise constant weights, used in (47), $w(\boldsymbol{r})$ is a non-constant function of $\boldsymbol{r}$ and it is differentiable a.e.

According to (100) the functional (or the variational) derivative of $E[\boldsymbol{f}]$ is

$$\frac{\delta E}{\delta \boldsymbol{f}} \triangleq \frac{\delta E[\boldsymbol{f}(\boldsymbol{r})]}{\delta \boldsymbol{f}(\boldsymbol{r})} \tag{101}$$

$$= \left(\sum_{i=1}^{n} \frac{\partial w}{\partial \boldsymbol{r}_i} \frac{\partial \mathcal{D}}{\partial \boldsymbol{f}_{1,i}}, \ldots, \sum_{i=1}^{n} \frac{\partial w}{\partial \boldsymbol{r}_i} \frac{\partial \mathcal{D}}{\partial \boldsymbol{f}_{n,i}} \right). \tag{102}$$

Consequently, the variational approach can be employed for optimizing simiplicial maps with more general weight functions. For example, consider piecewise linear weights, defined on a triangulated domain $\boldsymbol{M}$. These weights can be expressed via barycentric coordinates (51) and they can be used in practice for interpolating data, sampled at vertices, and thereby extending the data to any point $\boldsymbol{r} \in \operatorname{conv}(\boldsymbol{M})$. For example, consider interpolation of texture coordinates on tessellated surfaces, interpolation of intensity values on tetrahedral meshes, reconstructed from MRI scans, and etc.

Practically, (100) and (101) define a variational GD process for obtaining an optimal mapping with respect to distortion measure $\mathcal{D}$. This process is similar to GD update of target coordinates $\boldsymbol{x}$; it starts with a deformation $\boldsymbol{f}^{(k)}$, $k = 0$, and computes the next deformation by

$$\boldsymbol{f}^{(k+1)} = \boldsymbol{f}^{(k)} - \Delta t \left. \frac{\delta E}{\delta \boldsymbol{f}} \right|_{\boldsymbol{f}^{(k)}}, \tag{103}$$

where Δt is a given time step and $\delta E / \delta \boldsymbol{f}$ is computed according to (102). We can simplify (100) and (102) by using SVD of Jacobian $d\boldsymbol{f}$, as follows:

Let $d\boldsymbol{f}_{\boldsymbol{r}} = U\Sigma V^T$ be an SVD decomposition (20) and let $\{\nu_1, \ldots, \nu_n\}$ and $\{u_1, \ldots, u_n\}$ be the corresponding right and left signualr vectors of $d\boldsymbol{f}$, respectively. We refer to the chosen frames as to *SVD coordinates*. The Jacobian matrix $d\boldsymbol{f}_{\boldsymbol{r}}$ in SVD coordinates is represented by

$$\widehat{d\boldsymbol{f}_{\boldsymbol{r}}} \triangleq [d\boldsymbol{f}_{\boldsymbol{r}}]_{V,U} = \operatorname{diag}(\sigma_1(d\boldsymbol{f}_{\boldsymbol{r}}), \ldots, \sigma_n(d\boldsymbol{f}_{\boldsymbol{r}}))\,. \tag{104}$$

Since distortions are functions of Jacobian singular values σ_i, $\mathcal{D}$ satisfies

$$\frac{\partial \mathcal{D}}{\partial \widehat{\boldsymbol{f}}_{i,j}} = \delta_{ij} \frac{\partial \mathcal{D}}{\partial \sigma_i}\,, \; 0 \le i, j \le n\,. \tag{105}$$

Combining this all together yields the following expression of the functional derivative in $\{V, U\}$ coordinate frames:

$$\frac{\delta E}{\delta \widehat{\boldsymbol{f}}} = \left(\frac{\partial w}{\partial \nu_1} \frac{\partial \mathcal{D}}{\partial \sigma_1}, \ldots, \frac{\partial w}{\partial \nu_n} \frac{\partial \mathcal{D}}{\partial \sigma_n} \right). \tag{106}$$

In the next section, we develop another expression for the energy gradient. In both expressions, energy gradients are functions of distortion derivatives, $\partial \mathcal{D} / \partial \sigma_i$. Therefore, both formulas can be interpreted as transformations from gradients, computed in SVD coordinate frames, to the gradients, computed with respect to the standard Euclidean bases.

II Distortion Energy Gradient

The remaining part is dedicated for developing an explicit expression for $\nabla_{\boldsymbol{x}} E$, with respect to an arbitrary distortion measure $\mathcal{D}(df_s[\boldsymbol{x}])$.

This method differs from the standard approach, presented in Appendix A.2, because it uses only the basic rules of partial derivatives. The standard method employs step-by-step computations for evaluating $\nabla_{\boldsymbol{x}} E$. In contrast to that approach, here we derive a single analytic expression for the gradient of $E(\boldsymbol{x})$, represented in a matrix form. In our computations, we rely on the two auxiliary functions that operate on indices that arise in the discrete case.

The task of the first auxiliary function, is to match between each simplex and its $n+1$ vertices to global indexing of vertices in $\mathcal{V}$. Assuming a global vertex order $\mathcal{V} = \{u_1, u_2, \ldots, u_{|\mathcal{V}|}\}$, we define $\boldsymbol{X}$ to be the matrix of all target coordinates $\boldsymbol{x}_{u_i}$ put in the matrix columns according to that vertex order. Then, we set the first auxiliary function to be

$$\Lambda : \{1, \ldots, |\mathcal{S}|\} \times \{1, \ldots, n+1\} \to \{1, \ldots, |\mathcal{V}|\}\,, \tag{107}$$

and use this function to rewrite target coordinate matrix (92) in the following form:

$$\boldsymbol{X}(s) = \left(\begin{array}{c|c|c} \boldsymbol{X}_{\Lambda(s,1)} & \cdots & \boldsymbol{X}_{\Lambda(s,n+1)} \end{array} \right). \tag{108}$$

Similarly, we provide a global order of hat functions $\{(h_s)_\nu : s \in \mathcal{S}, \nu \in s\}$ and match global indices of each $(h_s)_\nu$ with the number of row at which $(h_s)_v$ appears in (93). We denote this matching function by

$$\Omega : \{1,\ldots,|\mathcal{S}|\} \times \{1,\ldots,n+1\} \to \{1,\ldots,|\mathcal{S}|(n+1)\}. \quad (109)$$

Let $\boldsymbol{\eta}(s)_i$ be the normal vector $\boldsymbol{\eta}_i$ of the i-th face of $s \in \mathcal{S}$ and define $\boldsymbol{N}(s)$ to be a matrix of negative normal vectors

$$\{-\boldsymbol{\eta}(s)_i |\, i = 1,\ldots,n+1\},$$

put in the order of hat functions. We express the rescaled Jacobian $n\,\mathrm{Vol}(s)d\boldsymbol{h}_s$ by means of the indexing function (109) as

$$\boldsymbol{N}(s) = \begin{pmatrix} \boldsymbol{N}_{\Omega(s,1)} \\ \hline \vdots \\ \hline \boldsymbol{N}_{\Omega(s,n+1)} \end{pmatrix}, \quad (110)$$

and rewrite the Jacobian df_s in the following form:

$$df_s = \frac{1}{n\,\mathrm{Vol}(s)} \boldsymbol{X}(s)\boldsymbol{N}(s). \quad (111)$$

We then differentiate energy E with respect to j-th coordinate of target vertex $\boldsymbol{x}_{v_i}$

$$\frac{\partial E}{\partial(\boldsymbol{x}_{v_i})_j} = \sum_{s\in\mathcal{S}} w(s) \frac{\partial \mathcal{D}(df_s)}{\partial(\boldsymbol{x}_{v_i})_j} \quad (112)$$

$$= \sum_{s\in\mathcal{S}(v_i)} w(s) \frac{\partial \mathcal{D}(df_s)}{\partial(\boldsymbol{x}_{v_i})_j}, \quad (113)$$

where $\mathcal{S}(v_i)$ denotes the set of simplices sharing vertex v_i. The derivative of the distortion measure $\mathcal{D}$, with respect to target coordinate in (113), is obtained via the chain rule and the singular value formulation of Theorem 2.1:

$$\frac{\partial \mathcal{D}(df_s)}{\partial(\boldsymbol{x}_{v_i})_j} = \left[\frac{\partial \mathcal{D}}{\partial \sigma}\right]^\top \frac{\partial \sigma(df_s)}{\partial(\boldsymbol{x}_{v_i})_j} \quad (114)$$

$$= \sum_{k=1}^{n} \frac{\partial \mathcal{D}}{\partial \sigma_k} \frac{\partial \sigma_k(df_s)}{\partial(\boldsymbol{x}_{v_i})_j}. \quad (115)$$

By employing the indexing of (107) and (109), we rewrite (111) in the element-wise form:

$$df_s^{qp} = \frac{1}{n\,\mathrm{Vol}(s)} \sum_{k=1}^{n+1} (\boldsymbol{X}_{\Lambda(s,k)})_q (\boldsymbol{N}_{\Omega(s,k)})_p, \quad (116)$$

where the superscript, df_s^{pq}, denotes the p-row, q-column element of Jacobian df_s. Differentiating (116) with respect to $(\boldsymbol{x}_{v_i})_j$ yields

$$\frac{\partial(df_s^{qp})}{\partial(\boldsymbol{x}_{v_i})_j} = \frac{1}{n\,\mathrm{Vol}(s)} \frac{\partial\left[(\boldsymbol{x}_{v_i})_q \left(\boldsymbol{N}_{\Omega(s,\Omega_s^{-1}(i))}\right)_p\right]}{\partial(\boldsymbol{x}_{v_i})_j} \quad (117)$$

$$= \delta_{qj} \frac{\left(\boldsymbol{N}_{\Omega(s,\Omega_s^{-1}(i))}\right)_p}{n\,\mathrm{Vol}(s)}, \quad (118)$$

where δ denotes the Kronecker delta and Ω_s^{-1} is the inverse of the indexing function (109) restricted to $\{s\}\times\{1,\ldots,n+1\}$, that is,

$$\forall s \in \mathcal{S}, i \in \Omega(\{s\}\times\{1,\ldots,n+1\}) : \ \Omega\left(s,\Omega_s^{-1}(i)\right) = i.$$

We proceed by applying the chain rule and the SVD derivative formula from matrix calculus [PL00], to expand the right-side multiplier in (115):

$$\frac{\partial \sigma_k(df_s)}{\partial(\boldsymbol{x}_{v_i})_j} = \sum_{q,p} \frac{\partial \sigma_k(df_s)}{\partial(df_s^{qp})} \frac{\partial(df_s^{qp})}{\partial(\boldsymbol{x}_{v_i})_j} \quad (119)$$

$$= \sum_{q,p} U_s^{qk} V_s^{qk} \frac{\partial(df_s^{qp})}{\partial(\boldsymbol{x}_{v_i})_j}, \quad (120)$$

where U_s and V_s are orthonormal matrices of left and-right singular vectors from the SVD of df_s, defined in (20).

Finally, we expand (113) by (115) and then substitute the resulting expression back in (115). This leads to the following expression:

$$\frac{\partial E}{\partial(\boldsymbol{x}_{v_i})_j} = \sum_{s\in\mathcal{S}(v_i)} w(s) \sum_{k=1}^{n} \frac{\partial \mathcal{D}}{\partial \sigma_k} \frac{\partial \sigma_k(df_s)}{\partial(\boldsymbol{x}_{v_i})_j} \quad (121)$$

$$= \sum_{s\in\mathcal{S}(v_i)} w(s) \sum_{k=1}^{n} \frac{\partial \mathcal{D}}{\partial \sigma_k} \sum_{p} U_s^{jk} V_s^{pk} \frac{\left(\boldsymbol{N}_{\Omega(s,\Omega_s^{-1}(i))}\right)_p}{n\,\mathrm{Vol}(s)} \quad (122)$$

$$= \sum_{s\in\mathcal{S}(v_i)} \frac{w(s)}{n\,\mathrm{Vol}(s)} \left[\sum_{k=1}^{n} \frac{\partial \mathcal{D}}{\partial \sigma_k} U_s^{jk} \left(\sum_{p} V_s^{pk} \left(\boldsymbol{N}_{\Omega(s,\Omega_s^{-1}(i))}\right)_p\right)\right]. \quad (123)$$

For obtaining a simpler expression of the distortion gradient, we rewrite (123) in a vector form:

$$\frac{\partial E}{\partial \boldsymbol{x}_{v_i}} = \left[\frac{\partial E}{\partial(\boldsymbol{x}_{v_i})_1}, \cdots, \frac{\partial E}{\partial(\boldsymbol{x}_{v_i})_m}\right] \quad (124)$$

$$= \sum_{s\in\mathcal{S}(v_i)} \frac{w(s)\boldsymbol{N}_{\Omega(s,\Omega_s^{-1}(i))} V_s}{n\,\mathrm{Vol}(s)} U_s^\top \odot \nabla_\sigma \mathcal{D}^{[n]}, \quad (125)$$

where $\odot$ denotes the Hadamard (element-wise) product, and $\nabla_\sigma \mathcal{D}^{[n]}$ is the stack of n copies of $\nabla_\sigma \mathcal{D}$,

$$\nabla_\sigma \mathcal{D}^{[n]} \triangleq \left(\begin{array}{c|c|c} \nabla_\sigma \mathcal{D} & \cdots & \nabla_\sigma \mathcal{D} \end{array} \right)_{n\times n}. \quad (126)$$

Summarizing the above, we attain the distortion Jacobian:

$$d_{\boldsymbol{x}}E \triangleq \begin{pmatrix} \partial E/\partial \boldsymbol{x}_{v_1} \\ \hline \vdots \\ \hline \partial E/\partial \boldsymbol{x}_{v_{|\mathcal{V}|}} \end{pmatrix}. \quad (127)$$

The matrix form $d_{\boldsymbol{x}}E$ is readily available for GD and BGD optimization algorithms. For more general 1st and 2nd order solvers of (56), we need to lay down rows of (127) to obtain a vector form of the gradient $\nabla_{\boldsymbol{x}}E$. The column vector form of the gradient is used in the linear system (63) and in the quadratic proxy equation (64).