Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Robust and accurate optimal transportation map by self-adaptive sampling*

Yingshi WANG[1], Xiaopeng ZHENG[2], Wei CHEN[2,3], Xin QI[4], Yuxue REN[3], Na LEI[†‡2,3], Xianfeng GU[†4]

[1]*Department of Computer Science, Inner Mongolia University of Finance and Economics, Hohhot 010010, China*
[2]*School of Software, Dalian University of Technology, Dalian 116620, China*
[3]*Beijing Advanced Innovation Center for Imaging Technology, Capital Normal University, Beijing 100048, China*
[4]*Department of Computer Science, Stony Brook University, Stony Brook, NY 11794, USA*
[†]E-mail: nalei@dlut.edu.cn; gu@cs.stonybrook.edu

Received May 25, 2020; Revision accepted Sept. 27, 2020; Crosschecked Aug. 19, 2021

Abstract: Optimal transportation plays a fundamental role in many fields in engineering and medicine, including surface parameterization in graphics, registration in computer vision, and generative models in deep learning. For quadratic distance cost, optimal transportation map is the gradient of the Brenier potential, which can be obtained by solving the Monge-Ampère equation. Furthermore, it is induced to a geometric convex optimization problem. The Monge-Ampère equation is highly non-linear, and during the solving process, the intermediate solutions have to be strictly convex. Specifically, the accuracy of the discrete solution heavily depends on the sampling pattern of the target measure. In this work, we propose a self-adaptive sampling algorithm which greatly reduces the sampling bias and improves the accuracy and robustness of the discrete solutions. Experimental results demonstrate the efficiency and efficacy of our method.

Key words: Optimal transportation; Monge-Ampère equation; Self-adaptive sampling
https://doi.org/10.1631/FITEE.2000250 **CLC number:** O242; TP391

1 Introduction

Optimal transportation theory studies transformations among probability measures. An optimal transportation map transforms one probability measure to another in the most economic way, and the cost of the map defines the Wasserstein distance between the two measures. Recently, optimal transportation theory has played an increasingly important role in many engineering and medical fields, such as area-preserving surface parameterization in computer graphics (Su KH et al., 2016, 2017), dynamic surface registration in computer vision (Su ZY et al., 2015), and shape comparison in medical imaging (Su ZY et al., 2013). In particular, the optimal transportation map has been used for generative models in deep learning (Arjovsky et al., 2017; Gulrajani et al., 2017; Lei et al., 2019). For example, in Wasserstein generative adversarial networks (Arjovsky et al., 2017), the Wasserstein distance was used to compute the distance between the data distribution and the generated distribution by the discriminator. The optimal transportation map was calculated by the generator to transform the white noise to the latent data distribution.

The study of computational algorithms for the optimal transportation map has become an active

‡ Corresponding author
* Project supported by the National Numerical Wind Tunnel Project, China (No. NNW2019ZT5-B13) and the National Natural Science Foundation of China (Nos. 61907005, 61772105, 61936002, and 61720106005)
ORCID: Na LEI, https://orcid.org/0000-0003-3361-0756

field. One major approach is based on the Brenier theory (Villani, 2008), which claims that the optimal transportation map is given as the gradient of a convex Brenier potential function when the cost function is strictly convex and the Brenier potential satisfies the classical Monge-Ampère equation. Yau-Gu-Luo's theory (Gu et al., 2016) shows that solving the Monge-Ampère equation is reduced to convex optimization. Because of the high non-linearity of the Monge-Ampère partial differential equation (PDE), during the solving process, the strict convexity of intermediate solutions needs to be preserved with great caution. In Yau-Gu-Luo's algorithm, the target probability measure needs to be discretized and approximated by a summation of Dirac measures. The accuracy of the computational results is heavily influenced by the sampling pattern of the target measure. If the target density function is highly non-uniform, as shown in our experiments, and the density has huge variations, then then conventional uniform sampling is inappropriate, the computational result has large approximation errors, and the optimization process is unstable and hardly convergent.

One of the intrinsic reasons for the instability is that the Brenier potential is represented by a convex polyhedral surface, on which each face corresponds to one sample. If the Dirac measures of samples differ drastically, then the areas of the faces will have large variations. Therefore, the construction of the combinatorial structure of the Brenier potential is prone to errors. Another reason is that during the optimization, the strict convexity of the Brenier potential has to be maintained all the time; namely, the search path cannot exceed the admissible space. Otherwise, if the convexity is violated in the process, the algorithm will collapse. If some faces are tiny, then the solution is very close to the boundary of the admissible space. Hence, great efforts need to be made to guarantee that the searching path is within the admissible space, which reduces the efficiency and robustness.

In the current work, we define the robustness as the computational stability and accuracy for solving optimal transportation maps for probability distributions with complicated supports and density functions with large variations.

This work tackles the challenging problem of the instability of the conventional optimal transportation mapping algorithm by proposing a novel self-adaptive sampling algorithm, such that the sampling density is consistent with the target probability density by the computational geometric method. This will ensure the uniformity of the face areas of the discrete Brenier potential, and keep the searching path far away from the the boundary of the admissible space. Experimental results demonstrate that our proposed method greatly improves the robustness and accuracy.

Fig. 1 shows one example of surface area-preserving parameterization using our proposed algorithm. The top row shows a three-dimensional (3D) human facial surface S viewed from different angles. The surface has the induced Euclidean metric g. The bottom left frame illustrates the Riemann mapping ($\varphi : S \to D$) which maps the facial surface onto the unit planar disk. The surface area element $\mathrm{d}A_g$ is pushed forward to a measure $\varphi_{\#}\mathrm{d}A_g$ on the disk. The optimal transportation map $T : D \to D$ from $\varphi_{\#}\mathrm{d}A_g$ to the Lebesgue measure on the disk is computed using the proposed method. The image of the optimal transport (OT) map is shown in the bottom right frame. The composition $T \circ \varphi : S \to D$ is an area-preserving parameterization from the top left frame to the bottom right frame.

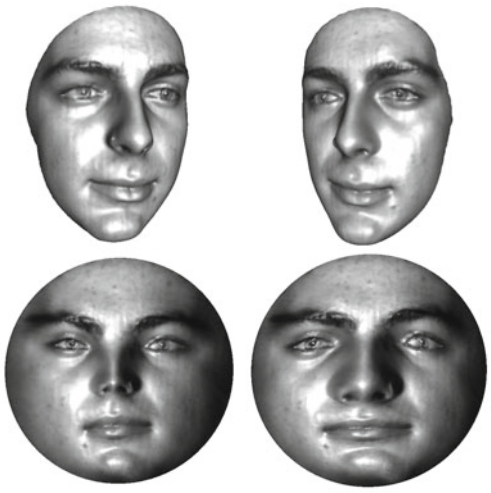

Fig. 1 The top row shows a three-dimensional (3D) human facial surface S viewed from different angles. The male facial surface (top left) S with a Riemannian metric g is conformally mapped onto the planar disk D (bottom left) by a Riemann mapping $\varphi : S \to D$. The mapping φ pushes forward the surface area element $\mathrm{d}A_g$ to a measure on the disk, denoted as $\varphi_{\#}\mathrm{d}A_g$. An optimal transportation map $T : D \to D$ from $\varphi_{\#}\mathrm{d}A_g$ to the Lebesgue measure on the disk is computed, and the image is shown in bottom right

2 Related works

In this section, we review the most related works.

2.1 Optimal transportation map

Monge-Kantorovich theory has been applied to solve the optimal transportation problem via a linear programming technique (Kantorovich, 2006). This method was intuitively applied in image registration and warping in early research works. This approach was proposed by ur Rehman et al. (2009). However, because of the high computational cost, the method can hardly handle the 3D image registration problem efficiently. The optimal transportation map was also applied for texture mapping purposes in Dominitz and Tannenbaum (2010), where the surface was initially mapped to the unit sphere conformally, and then the mapping was optimized by a gradient flow with multiple levels of resolutions to accelerate the convergence.

Recent research focuses more on the approach based on the Monge-Brenier theory (Brenier, 1991). Gu et al. (2016) used a geometric variational approach to prove the Alexandrov theory. This is equivalent to the discrete Brenier theory. The method leads to a constructive algorithm for computing the optimal transportation maps in general settings. de Goes et al. (2011) proposed to use OT for two-dimensional (2D) shape reconstruction and simplification. Later, they generalized this to use capacity-constrained Voronoi tessellation to deal with a blue noise processing problem (de Goes et al., 2012). Mérigot (2011) proposed a multi-scale approach to accelerate the computation for large-scale problems. Most of the early works focused on 2D image registration and processing, and more recent works generalized them to deal with 3D surfaces using computational geometric approaches. By incorporating with conformal mapping methods, optimal transportation maps were applied to obtain area-preserving maps in Dominitz and Tannenbaum (2010), Zhao et al. (2013), and Su ZY et al. (2015). This method can simultaneously balance the area and angle distortion.

The method based on the Monge-Brenier theory has been applied for area-preserving brain mapping in Su ZY et al. (2013), where the cortical surface was mapped onto the unit sphere conformally, and then onto the extended complex plane by the stereographic projection. The Monge-Brenier approach has been improved by Nadeem et al. (2017) using the conformal welding method. Since the exact evaluation of the Wasserstein distance is expensive, the heat kernel method was applied to approximate it in Cuturi (2013) and Solomon et al. (2014, 2015). While most research works deal with the optimal transportation problem with the Euclidean metric, Wang (2004) and Cui et al. (2019) focused on solving the optimal transportation problem in the spherical domain.

Recent research also introduced the optimal transportation theory in the optical design field. Reflector design problems were summarized as a group of Monge-Ampère equation problems in Wang (1996, 2004) and Guan et al. (1998). The correspondence of the Monge-Ampère equation reflector design problem was listed as one of the open problems in Yau (1998), and can further be related to the optimal transportation theory. Similar research in lens design was introduced in Gutiérrez and Huang (2009). Numerical methods and simulations of optical design problems were proposed in Meyron et al. (2018).

2.2 Adaptive mesh generation

In mesh generation field, Delaunay refinement algorithms were originally designed for meshing planar domains, and later were generalized for meshing surfaces and volumes. Chew (1989)'s algorithm split any triangle whose circum-radius was greater than the prescribed shortest edge length parameter e, and hence generated triangulation of uniform density and with no angle smaller than 30°. However, the number of triangles produced is not optimal. Chew (1993)'s algorithm split any triangle whose circum-radius-to-shortest-edge ratio was greater than one, and hence in practice, produced graded meshes. A similar split criterion was used in Ruppert (1995)'s algorithm, which had the theoretical guarantee of the minimal angle of no less than 20.7°. Shewchuk (2002) unified the pioneering mesh generation algorithms of Chew (1989) and Ruppert (1995), improved the algorithms in several ways, and helped solve the difficult problem of meshing non-manifold domains with small angles. Dey (2007) developed a series of algorithms for surface meshing and remeshing based on volumetric Delaunay refinement (Cheng et al., 2007; Dey and Levine, 2007;

Dey and Ray, 2010). Mesh adaption algorithms used variational approaches to improve meshing qualities by optimizing various functionals (Cheng et al., 2012).

Mesh adaption was proposed to use variational approaches which are based on the optimization of mesh related functionals to achieve the best adapted meshes. The mesh adaption methods include centroidal Voronoi tessellation (CVT) (Du and Wang, 2005; Liu et al., 2009; Rong et al., 2011; Yan et al., 2011), interpolation error minimization (Kunert, 2002; Alauzet et al., 2006), optimal Delaunay triangulations (ODTs) (Chen and Xu, 2004), mesh refinement (Apel and Lube, 1996; Bossen and Heckbert, 1996; Löhner and Cebral, 2000; Berndt and Shashkov, 2003; Apel et al., 2004), monitor functions (Huang, 2006), mesh smoothing (Sirois, et al., 2006), and a branch of other methods (Bottasso, 2004; Dobrzynski and Frey, 2008). Many of these methods were generalized to produce anisotropic meshes by incorporating a metric tensor into the functional.

3 Theoretical background

In this section, we briefly introduce the basic concepts and theorems. For details, readers can refer to Villani (2008) and Gu et al. (2018).

3.1 Brenier theory

Suppose that Ω is a domain in the Euclidean space $\mathbb{R}^2$ with probability measure μ, $\mathrm{d}\mu(x) = f(x)\mathrm{d}x$, where the density function $f(x)$ is absolutely continuous positive, and that $D \subset \mathbb{R}^2$ is another domain with probability measure ν with a density function, $\mathrm{d}\nu(y) = g(y)\mathrm{d}y$. The total measures are equal, $\mu(\Omega) = \nu(D)$, namely,

$$\int_{\Omega} f(x)\mathrm{d}x = \int_{D} g(y)\mathrm{d}y.$$

Mapping $T : \Omega \to D$ induces a push-forward measure $T_{\#}\mu$, i.e., for any Borel set $B \subset D$,

$$T_{\#}\mu(B) := \mu(T^{-1}(B)) = \int_{T^{-1}B} f(x)\mathrm{d}x.$$

We say that mapping T is measure-preserving if $T_{\#}\mu = \nu$. If T is with C^1 continuity, then the mapping satisfies the Jacobian equation:

$$\det(\mathbf{DT}(x)) = \frac{f(x)}{g(T(x))}, \tag{1}$$

where $\mathbf{DT}$ is the Jacobian matrix of map T.

Supposing that $c : \Omega \times D \to \mathbb{R}$ is a cost function which measures the cost to transport the unit mass from $x \in \Omega$ to $y \in D$, the optimal transportation problem is to find a measure-preserving mapping that minimizes the total transportation cost:

$$\min_{T_{\#}\mu=\nu} \mathcal{C}(T) = \min_{T_{\#}\mu=\nu} \int_{\Omega} c(x, T(x)) f(x)\mathrm{d}x.$$

The minimizer is called the optimal transportation map. Suppose that the cost is the quadratic Euclidean distance:

$$c(x, y) = \frac{1}{2}|x - y|^2.$$

Then, the square root of the total transportation cost $\mathcal{C}(T)$ of the optimal transportation map is called the Wasserstein distance between μ and ν:

$$\mathcal{W}_2^2(\mu, \nu) = \min_{T_{\#}\mu=\nu} \frac{1}{2} \int_{\Omega} |x - T(x)|^2 f(x)\mathrm{d}x.$$

Theorem 1 (Brenier theory (Villani, 2008)) Suppose that Ω is strictly convex, that f is absolutely continuous, and that $g \in L^1(D)$ is bounded and measurable. Then, there exists a convex function $u : \Omega \to \mathbb{R}$ whose gradient map $\nabla u : x \mapsto \nabla u(x)$ gives the unique optimal transportation map. Such kind of u is unique up to a constant.

The convex function claimed in the Brenier theory is the so-called Brenier potential whose gradient maps give the optimal transportation map $T(x) = \nabla u(x)$. By Jacobian equation (1), we obtain the Monge-Ampère equation:

$$\det(D^2 u(x)) = \frac{f(x)}{g(\nabla u(x))}. \tag{2}$$

Hence, solving the optimal transportation problem is reduced to solving the Monge-Ampère equation.

3.2 Geometric variation theory

Suppose that $u : \Omega \to \mathbb{R}$ is a convex function, and its subdifferential at $x \in \Omega$ is defined as

$$\partial u(x) := \{\boldsymbol{q} \in \mathbb{R}^2 | u(y) \geq \langle \boldsymbol{q}, y - x \rangle + u(x), \forall y \in \Omega\}. \tag{3}$$

Then

$$\partial u(E) := \cup_{x \in E} \partial u(x).$$

We can treat ∂u as a set-valued mapping. The function u solves the problem

$$(\partial u)_{\#}\mu = \nu,$$

with the boundary condition $\partial u(\Omega) = D$, and is called the Alexandrov solution to Monge-Ampère equation (2). According to the Monge-Ampère PDE theory, the Alexandrov solutions converge to the classical solution under appropriate regularity conditions of the densities $f(x)$ and $g(y)$, the convexity of Ω, and the regularity of $\partial\Omega$ (Villani, 2008).

In practice, the target domain is sampled and represented by the discrete samples $D = \{y_1, y_2, \dots, y_n\}$, and the target measure is approximated by a summation of Dirac measures:

$$\nu = \sum_{i=1}^{n} \nu_i \delta(y - y_i),$$

where $\delta(\cdot)$ is the Dirac function. Each sample y_i corresponds to a linear equation:

$$l_i(x) := \langle x, y_i \rangle - h_i,$$

where h_i is a constant. Let the height vector be $\boldsymbol{h} = (h_1, h_2, \dots, h_n)$. The Brenier potential is given by the upper envelope of these planes:

$$u_{\boldsymbol{h}(x)} := \sup_{x \in \Omega} \{\langle y_i, x \rangle - h_i, \forall y_i \in D\}. \tag{4}$$

The graph of the Brenier potential is a convex polyhedral surface. Its face is defined as

$$W_{\boldsymbol{h}}^i := \{x \in \Omega | u_{\boldsymbol{h}}(x) = l_i(x) \geq l_j(x), \forall j = 1, 2, \dots, n\}.$$

Then μ-area of face $W_{\boldsymbol{h}}^i$ is defined as

$$w_{\boldsymbol{h}}^i := \mu(W_{\boldsymbol{h}}^i) = \int_{W_{\boldsymbol{h}}^i} f(x) \mathrm{d}x.$$

We define the admissible space of height vectors as

$$\Sigma := \left\{\boldsymbol{h} \in \mathbb{R}^n | w_{\boldsymbol{h}}^i > 0, \forall i = 1, 2, \dots, n\right\}. \tag{5}$$

By the Brunn-Minkowski inequality, we can show that Σ is convex. Furthermore, we define the hyperplane as

$$\Pi := \{\boldsymbol{h} \in \mathbb{R}^n | h_1 + h_2 + \cdots + h_n = 0\}.$$

The optimal transportation map is given by the following theorem:

Theorem 2 (Yau-Gu-Luo theory (Gu et al., 2016)) Suppose that Ω is convex, that f is absolutely continuous, and that the total measures are equal:

$$\int_{\Omega} f(x) \mathrm{d}x = \sum_{i=1}^{n} \nu_i.$$

Then, on the set $\Sigma \cap \Pi$, the following energy is strictly convex:

$$E(\boldsymbol{h}) = \int^{\boldsymbol{h}} \sum_{i=1}^{n} \left(\nu_i - w_{\boldsymbol{h}}^i\right) \mathrm{d}h_i, \tag{6}$$

and has a unique optimizer $\boldsymbol{h}_0$, giving the optimal transportation map under the quadratic distance cost:

$$T : \Omega \to D, \quad W_{\boldsymbol{h_0}}^i \mapsto y_i, \quad i = 1, 2, \dots, n.$$

The convex energy in the Yau-Gu-Luo theory is called the Alexandrov potential. Given a convex function $u : \Omega \to \mathbb{R}$, we can define its Legendre dual as

$$u^*(y) := \sup_{x \in \Omega} \{\langle x, y \rangle - u(x)\}, \tag{7}$$

which is also convex. The discrete convex function $u_{\boldsymbol{h}}$ in Eq. (4) also has Legendre dual $u_{\boldsymbol{h}}^*$, as shown in Fig. 2.

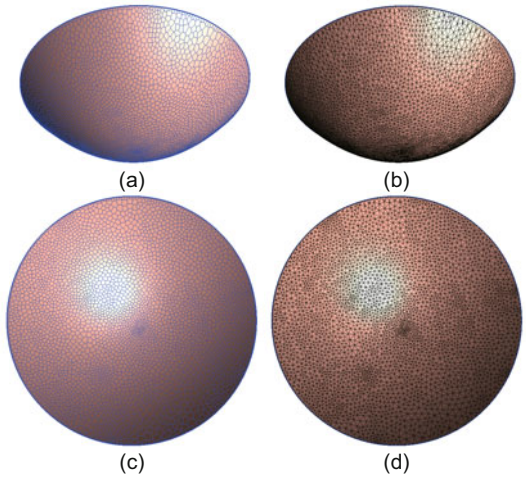

Fig. 2 Discrete Brenier potential $u_{\boldsymbol{h}}$ (a), its Legendre dual $u_{\boldsymbol{h}}^*$ (b), and their projections: power diagram (c) and the weighted Delaunay triangulation (d)

Definition 1 (Power diagram and weighted Delaunay triangulation) The projection of $u_{\boldsymbol{h}}$ induces all cell decomposition of Ω:

$$\Omega = \cup_{i=1}^{n} W_{\boldsymbol{h}}^i, \tag{8}$$

which is called the power diagram induced by $\boldsymbol{h}$. The projection of $u_{\boldsymbol{h}}^*$ on $D = \{y_1, y_2, \dots, y_n\}$ induces a triangulation of D, which is called the weighted Delaunay triangulation induced by $\boldsymbol{h}$.

The gradient and Hessian matrix of the Alexandrov potential in Eq. (6) can be directly derived from the power diagram and the weighted Delaunay triangulation, respectively. The gradient is

$$\nabla E(\boldsymbol{h}) = (\nu_1 - w^1_{\boldsymbol{h}}, \nu_2 - w^2_{\boldsymbol{h}}, \ldots, \nu_n - w^n_{\boldsymbol{h}})^{\mathrm{T}}. \quad (9)$$

The Hessian matrix is given by

$$\frac{\partial^2 u(\boldsymbol{h})}{\partial h_i \partial h_j} = \frac{\partial w^i_{\boldsymbol{h}}}{\partial h_j} = \frac{\mu(w^i_{\boldsymbol{h}} \cap w^j_{\boldsymbol{h}})}{|y_i - y_j|} \quad (10)$$

and

$$\frac{\partial^2 u(\boldsymbol{h})}{\partial h_i^2} = -\sum_{j=1}^{n} \frac{\partial^2 u(\boldsymbol{h})}{\partial h_i \partial h_j}. \quad (11)$$

4 Computational algorithms

The variational Theorem 2 leads to practical computational methods. The theorem itself is geometric, and can be translated to the computational geometry algorithm. The whole algorithm is equivalent to a non-linear convex optimization problem. At each step, one needs to construct the Brenier potential. The Brenier potential can be computed by the upper envelope of supporting planes. By the Legendre dual, the envelope can be converted to a convex hull of the dual points of the planes. The theory is general to any dimension, but in the current work, we demonstrate only the 2D cases. This section focuses on the computational algorithm for solving the 2D optimal transportation problem.

4.1 Geometric variational algorithm

The geometric variational algorithm is based on the Yau-Gu-Luo theory (Theorem 2).

4.1.1 Legendre dual

Given the target samples $D = \{y_1, y_2, \ldots, y_n\}$, associated with the target Dirac measures $\{\nu_1, \nu_2, \ldots, \nu_n\}$, for each height vector $\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$, we compute the convex hull which gives the Legendre dual of the Brenier potential:

$$u^*_{\boldsymbol{h}} := \mathrm{Conv}\big(\{(y_1, h_1), (y_2, h_2), \ldots, (y_n, h_n)\}\big).$$

The Brenier potential $u_{\boldsymbol{h}}$ can be carried out by the Legendre dual of $u^*_{\boldsymbol{h}}$:

$$u_{\boldsymbol{h}} = u^{**}_{\boldsymbol{h}}.$$

Combinatorially, the connectivities of $u_{\boldsymbol{h}}$ and $u^*_{\boldsymbol{h}}$ are Poincarè dual to each other. In details, each vertex $v \in u_{\boldsymbol{h}}$ is dual to a face $v^* \in u^*_{\boldsymbol{h}}$, each edge $[v_i, v_j] \in u_{\boldsymbol{h}}$ corresponds to an edge $v^*_i \cap v^*_j \in u^*_{\boldsymbol{h}}$, and each face $[v_i, v_j, v_k] \in u_{\boldsymbol{h}}$ corresponds to a vertex $v^*_i \cap v^*_j \cap v^*_k \in u^*_{\boldsymbol{h}}$. Note that $u_{\boldsymbol{h}}$ and $u^*_{\boldsymbol{h}}$ are open polyhedral surfaces. We need to add a vertex at infinity, and connect it with all boundary vertices of $u^*_{\boldsymbol{h}}$. The Poincarè dual still holds between them.

Geometrically, each face f is supported by a plane in $\mathbb{R}^3$, with representation $z = ax + by - c$. Then its dual vertex is with the position (a, b, c), and vice versa. Special attention needs to be paid to boundary vertices. Each of them corresponds to an open planar face. In practice, we clip all the faces on the Brenier potential $u_{\boldsymbol{h}}$ by a cylinder, generated by vertical lines through the boundary of Ω. The projection of $u_{\boldsymbol{h}}$ induces the power diagram of Ω as defined in Eq. (8). The projection of $u^*_{\boldsymbol{h}}$ induces a weighted Delaunay triangulation. Fig. 3 shows that the discrete Brenier potential $u_{\boldsymbol{h}}$ projects to the power diagram, and that its Legendre dual $u^*_{\boldsymbol{h}}$ projects to the weighted Delaunay triangulation. Furthermore, the figure shows the Poincarè dual between the power diagram and the weighted Delaunay triangulation.

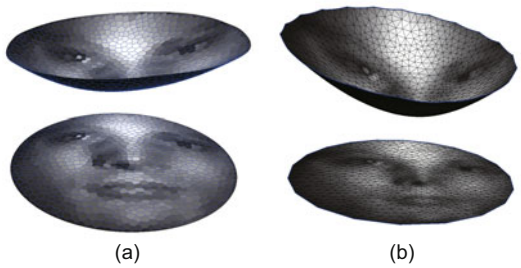

Fig. 3 The first male face example for the optimal transportation map: Brenier potential and power diagram (a) and its Legendre dual and the weighted Delaunay triangulation (b)

4.1.2 Damping Newton's method

Initially, we set $\boldsymbol{h}^0 = \frac{1}{2}(u_1^2 + v_1^2, u_2^2 + v_2^2, \ldots, u_n^2 + v_n^2)$, where (u_i, v_i) are the coordinates of the i^{th} sample y_i $(i = 1, 2, ..., n)$. The initial power diagram and the weighted Delaunay triangulation are a conventional Voronoi diagram and a Delaunay triangulation, respectively. This guarantees that the initial Brenier potential and its Legendre dual are strictly

convex; namely, the initial height vector belongs to the admissible space, $\boldsymbol{h}^0 \in \Sigma$.

Assume that at the k^{th} step, we have $\boldsymbol{h}^k$, the Brenier potential $u_{\boldsymbol{h}^k}$ and its Legendre dual $u^*_{\boldsymbol{h}^k}$, and the power diagram $\{W^i_{\boldsymbol{h}^k}\}_{i=1}^n$. We compute the gradient of the Alexandrov potential in Eq. (9) and Hessian matrix $\boldsymbol{H}$ as described in Eqs. (10) and (11). Then we solve the linear system as

$$\nabla E(\boldsymbol{h}^k) = \text{Hess}(\boldsymbol{h}^k)\boldsymbol{d}.$$

Next, we need to determine the step length λ. We initialize λ as one, and compute the convex hull of the points

$$\{(y_1, h_1^k + \lambda d_1), (y_2, h_2^k + \lambda d_2), \cdots, (y_n, h_n^k + \lambda d_n)\}.$$

If the convex hull misses any point, then $\boldsymbol{h}^k + \lambda \boldsymbol{d}$ is outside the admissible space and the corresponding Brenier potential is not strictly convex. Then we reduce the step length λ by half, i.e., $\lambda \leftarrow \frac{1}{2}\lambda$, and repeat the trial. We repeat this procedure and find the minimal l such that

$$\min_l \{\boldsymbol{h}^k + 2^{-l}\boldsymbol{d}\} \in \Sigma.$$

By iterating this procedure, we reduce the Alexandrov potential monotonously, until the difference between the target measure and the current one (measured by the norm of the gradient of the Alexandrov potential in Eq. (9)) is less than a prescribed threshold $\varepsilon > 0$. The algorithm is summarized in Algorithm 1.

Algorithm 1 Geometric variational method for the optimal transportation map

Input: convex domain Ω with measure μ; discrete samples $Y := \{y_1, y_2, \ldots, y_n\}$ with measures $\nu_1, \nu_2, \ldots, \nu_n$, with equal measure $\mu(\Omega) = \sum_{i=1}^n \nu_i$

Output: optimal transportation map $T : \Omega \to Y$

Initialize $\boldsymbol{h}^0 = (h_1, h_2, \ldots, h_n) \leftarrow 1/2(|y_1|^2, |y_2|^2, \cdots, |y_n|^2)$
while true **do**
  Compute the Brenier potential $u_{\boldsymbol{h}^k}$ and its Legendre dual $u^*_{\boldsymbol{h}^k}$
  Project $u_{\boldsymbol{h}^k}$ and $u^*_{\boldsymbol{h}^k}$ to obtain the power diagram and the weighted Delaunay triangulation
  Compute the gradient $\nabla E(\boldsymbol{h}^k)$ of the Alexandrov potential in Eq. (9)
  if $\|\nabla E(\boldsymbol{h}^k)\|$ is less than ε **then**
    Return $T = \nabla u_{\boldsymbol{h}^k}$
  end if
  Compute the Hessian matrix of Alexandrov potential in Eqs. (10) and (11)
  Solve the linear system $\nabla E(\boldsymbol{h}^k) = \text{Hess}(\boldsymbol{h}^k)\boldsymbol{d}$
  Set the step length $\lambda \leftarrow 1$
  repeat
    $\lambda \leftarrow \lambda/2$
    Construct the convex hull of $\{(y_i, h_i^k + \lambda d_i)\}_{i=1}^n$
  until all sample points are on the convex hull
  Update the height vector $\boldsymbol{h}^{k+1} \leftarrow \boldsymbol{h}^k + \lambda \boldsymbol{d}$
end while

4.2 Self-adaptive sampling

The initial input target measure ν is with a density function $g(y)$. We generate random samples $\{y_i\}_{i=1}^n$ inside D, compute a cell decomposition $D = \bigcup W_i$ with each cell centered at one sample $y_i \in W_i$, and define the Dirac measure $\nu_i = \int_{W_i} g(y)\mathrm{d}y$, $\nu = \sum_{i=1}^n (\nu_i \delta(y - y_i))$.

The sampling of the target probability measure ν is crucial for the robustness of the algorithm and the accuracy of the computational result. If the density function $g(y)$ has a large L^∞ norm, then ν_i has large variations, and therefore the sizes of cells in the power diagram vary a lot as well. Some cells are relatively small, and therefore the desired solution is very close to the boundary of the admissible space. This makes the optimization process highly unstable. Furthermore, the convex hull algorithm in computational geometry heavily depends on the numerical accuracy of visibility testing. A tiny cell means that the visibility testing for the vertex is difficult and that the computational result is inaccurate. To conquer these difficulties, we propose to use a self-adaptive sampling scheme, aiming at making the target Dirac measures ν_i's as uniform as possible. Namely, we want to minimize the variation as

$$\min_{y_1, y_2, \ldots, y_n} \sum_{i=1}^n |\nu_i - \bar{\nu}|^2, \quad \bar{\nu} = \frac{1}{n}\int_D g(y)\mathrm{d}y. \tag{12}$$

The direct optimization of Eq. (12) is difficult. Instead we optimize the sampling pattern using a mesh generation algorithm using density $g(y)$ as the grading function.

In the mesh generation field, given a geometric model, a grading function determines the element density in each region of the model based on the complexity of the geometry. More complex geometry with more angles and cuts will have a higher density of elements to better capture the topographical features. The mesh density grades away from the

complex regions. In our scenario, we use the target density as the grading function for mesh generation.

We use the weighted Delaunay refinement algorithm with grading function g. Supposing that D is a planar convex domain, we want to generate samples $\{y_i\}$ inside D and a triangulation $\mathcal{T}$ with y_i as vertices, such that the ν-area of each triangle is uniform, close to ε^2.

Given a planar triangle, with edge lengths a, b, and c, we measure the shape quality as the ratio of its inner circle radius r to the circum-radius R:

$$q := \frac{2r}{R} = \frac{(b+c-a)(c+a-b)(a+b-c)}{abc}.$$

If the triangle is equilateral, then q is 1; if the triangle is degenerated, then q is 0. In practice, we require the minimal corner angle to be greater than $\pi/6$. Also, we compute its circum-radius R as the size quality measure. The ν-length of the circum-radius is R multiplied by $g(y)$ at the center of the triangle. Based on Brandts et al. (2008), the shape quality q is equivalent to the minimal corner angle criterion. In our algorithm, we set $q > 0.366$, such that the minimal angle is greater than $\pi/6$.

First, we sample the boundary of D, $\{y_1, y_2, \dots, y_k\} \subset \partial D$, and partition the boundary ∂D into boundary segments, $\gamma_1, \gamma_2, \dots, \gamma_k$, such that the ν-length of each segment is close to ε. We compute the Delaunay triangulation for the current boundary sample set. Then we select the triangle with the worst quality; i.e., either q is too small or μ-length value of the circum-radius is too big. We compute the circum center of the triangle, insert it into the sample set, and update the Delaunay triangulation. We repeat this process, until all the faces are with ν-length of circum-radius less than ε and the shape quality is greater than 0.366. Based on Edelsbrunner (2001), the weighted Delaunay refinement algorithm can produce meshes with a uniform circum-radius. Combining the minimal angle and the uniform circum-radius, we can guarantee the uniformity of the samples $\{y_i\}$.

Based on Edelsbrunner (2001), the complexity for computing the power diagram/weighted Delaunay triangulation of n points is $O(n \ln n)$. By Kitagawa et al. (2019), the convergence rate of Newton's method for the optimal transport is linear. If the density function of μ is $C^{0,\alpha}$, $\alpha \in [0, 1)$, then $\frac{|x_{k+1}-x^*|}{|x_k-x^*|} \leq \frac{1}{1+\alpha}$, where x^* is the optimum and x_k is the k^{th} iteration solution. Supposing that the error threshold is τ, the number of iterations is $O(-\ln \tau)$. Hence, the total computational complexity is $O(-(\ln \tau) n \ln n)$. The algorithm is summarized in Algorithm 2.

Algorithm 2 Self-adaptive sampling

Input: convex domain D with measure density function $g : D \to \mathbb{R}$, and the size threshold ε

Output: a set of samples $Y := \{y_1, y_2, \dots, y_n\}$ and a triangulation $\mathcal{T}$, such that all the triangles are of uniform ν-area $O(\varepsilon^2)$

Generate a set of boundary samples $Y := \{y_1, y_2, \dots, y_m\}$, which divides the boundary into segments, and each segment has ν-length ε
Compute the Delaunay triangluate $\mathcal{T}$ of the current sample set Y
while true **do**
  Compute the shape quality and the size quality of all triangles in $\mathcal{T}$
  Select the triangle f with the worst shape quality or the largest ν-length circum-radius
  if the shape quality is greater than 8.75 and the ν-length circum-radius of f is less than ε **then**
    Return Y and $\mathcal{T}$
  end if
  Compute the circum center $\tilde{y}$ of f, and insert $\tilde{y}$ into the sample set Y
  Update the Delaunay triangulation $\mathcal{T}$
end while
Initialize $\boldsymbol{h}^0 = (h_1, h_2, \dots, h_n) \leftarrow 1/2(|y_1|^2, |y_2|^2, \dots, |y_n|^2)$
while true **do**
  Compute the Brenier potential $u_{\boldsymbol{h}^k}$ and its Legendre dual $u^*_{\boldsymbol{h}^k}$
  Project $u_{\boldsymbol{h}^k}$ and $u^*_{\boldsymbol{h}^k}$ to obtain the power diagram and the weighted Delaunay triangulation
  Compute the gradient $\nabla E(\boldsymbol{h}^k)$ of Alexandrov potential in Eq. (9)
  if $\|\nabla E(\boldsymbol{h}^k)\|$ is less than ε **then**
    Return $T = \nabla u_{\boldsymbol{h}^k}$
  end if
  Compute the Hessian matrix of Alexandrov potential in Eqs. (10) and (11)
  Solve the linear system $\nabla E(\boldsymbol{h}^k) = \text{Hess}(\boldsymbol{h}^k)\boldsymbol{d}$
  Set the step length $\lambda \leftarrow 1$
  repeat
    $\lambda \leftarrow \lambda/2$
    Construct the convex hull of $\{(y_i, h_i^k + \lambda d_i)\}_{i=1}^n$
  until all sample points are on the convex hull
  Update the height vector $\boldsymbol{h}^{k+1} \leftarrow \boldsymbol{h}^k + \lambda \boldsymbol{d}$
end while

5 Experimental results

We have conducted numerical experiments to test the performance of our algorithm. We used real geometric surfaces, either scanned from human faces or organ surfaces reconstructed from medical images. The algorithms were implemented using generic C++ on the Windows 10 platform with our hand-made halfedge mesh library and Eigen linear solver CGAL library for computing the convex hull. All the experiments were tested on a laptop with Intel Xeon CPU@3.00 GHz and 64.0 GB RAM.

As shown in Fig. 4, given a genus zero surface with a single boundary S, it has an induced Euclidean metric g, which induces the surface area element $\mathrm{d}A_g$. After normalization, the total surface area is π. The Riemann mapping $\varphi : (S, g) \to (D, \mathrm{d}u^2 + \mathrm{d}v^2)$ maps the surface onto the unit disk and pushes the area element to the disk, denoted as $\varphi_{\#}\mathrm{d}A_g$. Since Riemann mapping is conformal, the surface area element can be written as

$$\mathrm{d}A_g(u, v) = \mathrm{e}^{2\lambda(u,v)}\mathrm{d}u\mathrm{d}v,$$

where $\mathrm{e}^{2\lambda(u,v)}$ is the area distortion function and can be treated as the target density function.

On the disk, the Lebesgue measure, or equivalently the Euclidean metric $\mathrm{d}u^2 + \mathrm{d}v^2$, induces the Euclidean area element $\mathrm{d}u\mathrm{d}v$. We compute the optimal transportation $T : (D, \mathrm{d}u\mathrm{d}v) \to (D, \varphi_{\#}\mathrm{d}A_g)$ using the geometric variational method. The optimal mapping result is shown between the two planar images. The composition between the Riemann mapping φ and the inverse of the optimal transportation map T^{-1} gives an area-preserving mapping:

$$T^{-1} \circ \varphi : (S, g) \to (D, \mathrm{d}u\mathrm{d}v),\ (T \circ \varphi)_{\#}\mathrm{d}A_g = \mathrm{d}u\mathrm{d}v.$$

To visualize that mapping $T^{-1} \circ \varphi$ is area-preserving, we put circle packing texture on the planar unit disk, and pull it back to the original surface as shown in the top right frame of Fig. 4. We can see that the small circles are mapped to ellipses with similar area.

As shown in Fig. 5, we compute the histograms to measure the distortions. The top row shows the histograms of the conformal mapping of Fig. 4, and the bottom row show those of the optimal transportation map. The left column shows the angle distortion histograms, and the right column shows the area distortion histograms. The angle distortion histogram is calculated as follows: the triangle mesh

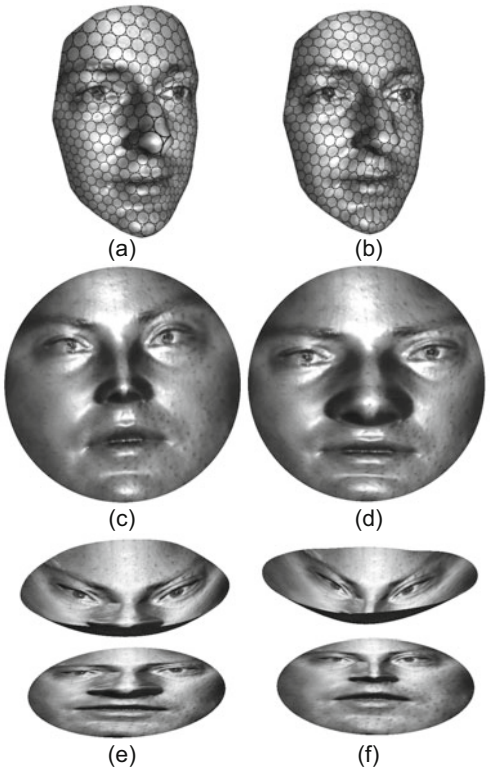

Fig. 4 The optimal transportation map for a male face: (a) conformal parameterization; (b) area-preserving parameterization; (c) conformal mapping; (d) optimal transportation map; (e) Brenier potential; (f) Legendre dual

S in $\mathbb{R}^3$ and its planar image share the same triangulation, and each corner angle in S corresponds to a planar corner angle. We compute the logarithm of the ratio between the corresponding corner angles, and construct the histograms. From the left column of Fig. 5, it is obvious that the angle distortion histograms of the conformal mapping highly concentrate in the zero point, showing that the conformal mapping induces very small angle distortions. In contrast, the optimal transportation map induces large angle distortions. The right column shows the area distortion histograms. These were obtained by computing the logarithm of the ratio between the corresponding face areas. It can be seen that the optimal transportation map induces very small area distortions, whereas the conformal mapping induces large area distortions.

Challenging examples are given by the cortical surfaces, as shown in Figs. 6 and 7. The density function induced by the conformal map has large variations, as shown in Fig. 7a. We use self-adaptive sampling to generate samples which are uniform with

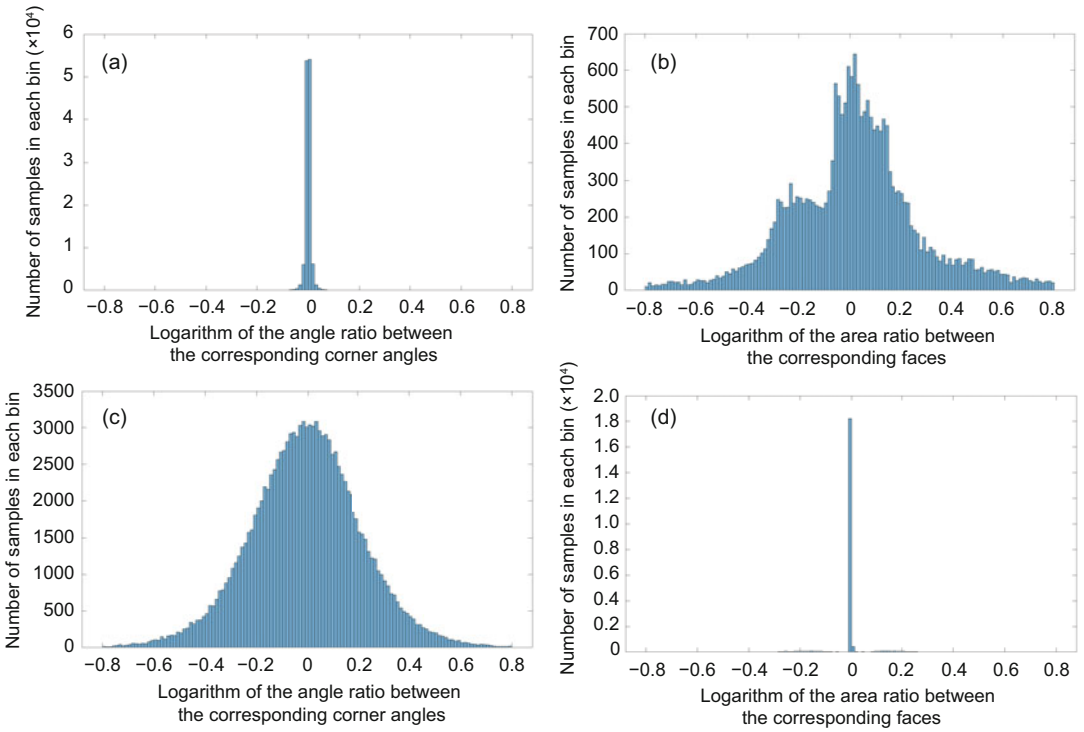

Fig. 5 Angle distortion and area distortion histograms of the male surface in Fig. 4: (a) angle distortion of conformal mapping; (b) area distortion of conformal mapping; (c) angle distortion of optimal transportation map; (d) area distortion of optimal transort map

respect to this density function, as shown in Figs. 6c and 6d. The sampling ensures the stablity of the optimization and obtains a precise optimal transportation map, as shown in Fig. 7b.

The most challenging example is the hand surface as shown in Figs. 8 and 9. In Fig. 8, the hand surface is conformally mapped onto the unit disk. The conformal factor at each finger tip is exponential with respect to the length of the finger. Therefore, it has a huge variance. The samples produced by conventional sampling methods lead to collapse during the optimization process. We apply our self-adaptive sampling to the conformal image, and concentrate most samples near the finger tip regions, as shown in Fig. 8d. The samples produced by our proposed method guarantee the stability of the computation, and produce a highly accurate optimal transportation map as shown in Fig. 9. In contrast, conventional sampling methods cannot produce meaningful results. Fig. 10 shows the area preserving parameterization for an oldman face model with complicated wrinkles. Fig. 11 shows the similar computational process for a female facial surface scanned by our team.

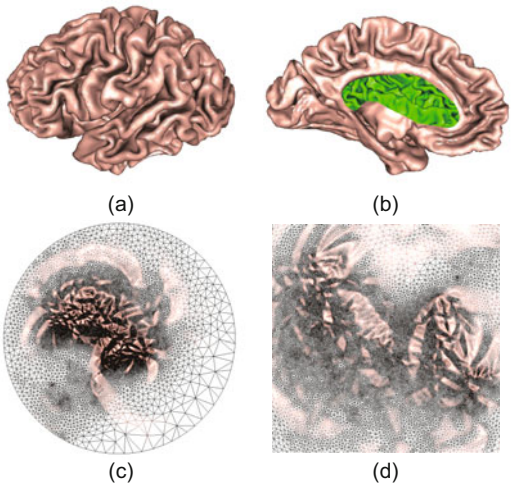

Fig. 6 The cortical surface is conformally mapped onto the unit disk. The conformal factor is treated as the density function of the target measure ν, and has large variation. The self-adaptive sampling produces uniform sampling with respect to ν, as shown in the bottom row: (a) cortical surface front view; (b) cortical surface back view; (c) self-adaptive sampling; (d) zoomed in of (c)

The statistics is reported in Table 1, including the numbers of vertices, edges, and faces. The

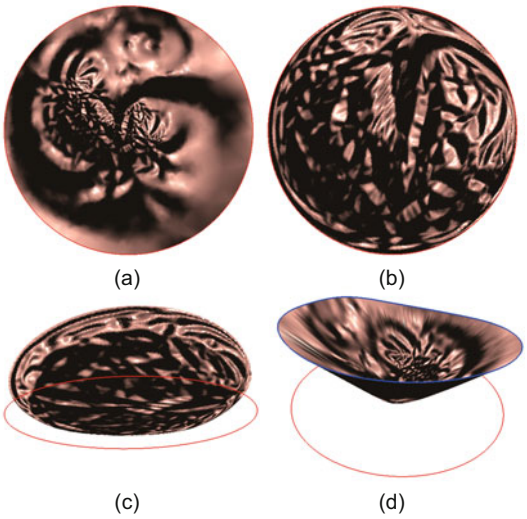

Fig. 7 Images of the conformal mapping (a), the optimal transportation map (b), Brenier potential (c), and Legendre dual of the Brenier potential (d)

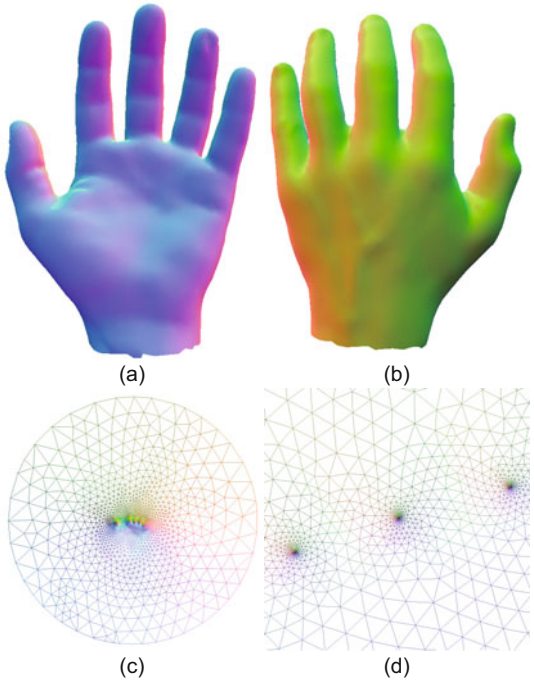

Fig. 8 The hand surface is conformally mapped on the unit disk. The self-adaptive sampling produces samples which are uniform with respect to the surface area element: (a) front view of the hand surface; (b) back view of the hand surface; (c) self-adaptice sampling; (d) zoomed in of (c)

number of iterations in Newton's method and the final relative error are shown. From the table, we can see that the number of iterations depends more on the complexity of the geometry of the model than on the combinatorial complexity of the triangulation. Hence, although the hand model in Figs. 8 and 9 has

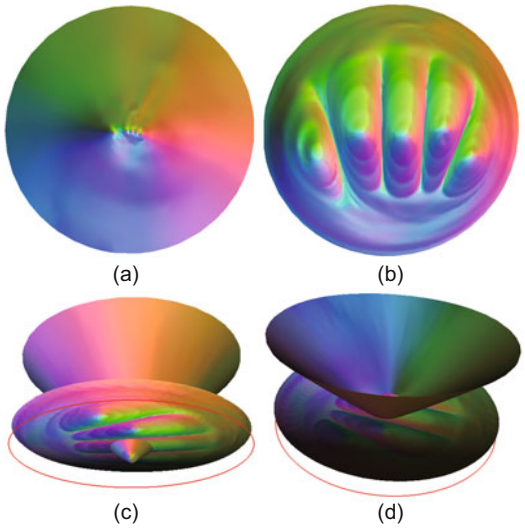

Fig. 9 Images of the conformal mapping (a), the optimal transportation map (b), Brenier potential (c), and its Legendre dual (d)

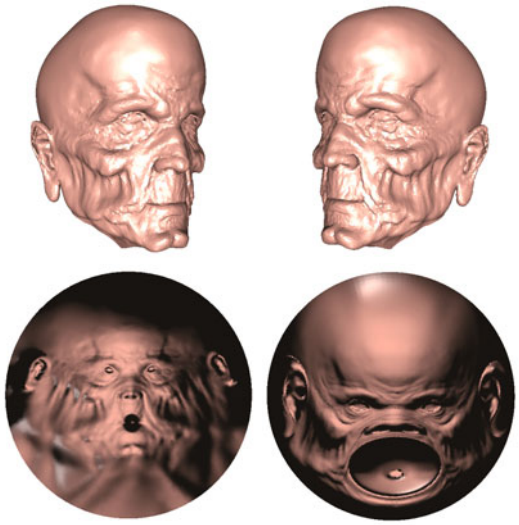

Fig. 10 Oldman

Table 1 Number of iterations and final relative error of different models

Model	Vert	Edge	Face	Fig.	Iter	Error
Buddha	32 765	97 897	65 133	12	11	4.305e-11
Brain	25 909	77 602	51 699	6, 7	35	5.425e-11
Hand	3706	11 072	7367	8, 9	28	6.212e-10
Male 1	24 848	74 032	49 185	1	17	7.563e-11
Male 2	24 977	74 418	49 442	4	54	6.555e-11
Oldman	25 627	76 554	50 928	10	16	3.963e-11
Female	24 979	74 458	49 480	11	54	9.310e-11

Vert: number of vertices; Edge: number of edges; Face: number of faces; Iter: number of iterations; Error: relative error

fewer vertices than the Buddha model in Fig. 12, it requires more than twice the iterations of those of the Buddha model. The number of iterations also

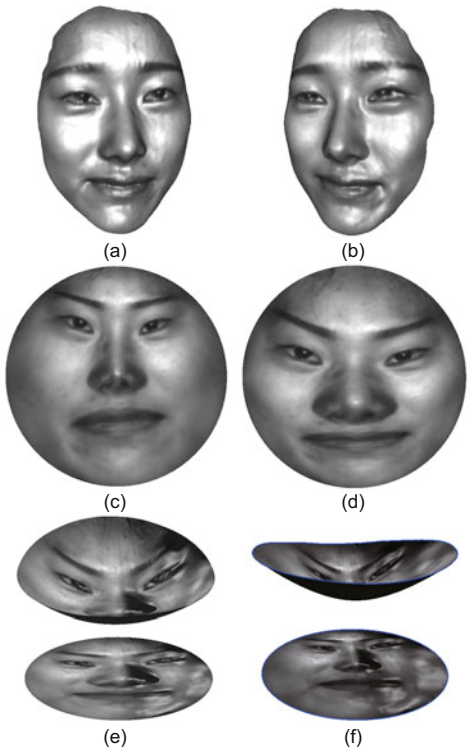

Fig. 11 Optimal transportation map of a female facial surface: (a) left view of the face; (b) right view of the face; (c) conformal mapping; (d) the optimal transportation map; (e) Brenier potential; (f) Legendre dual

depends on the triangulation quality. The second male face in Fig. 4 has roughly the same number of triangular faces as that of the first face in Fig. 1, but with a lower quality triangulation. Hence, it takes much more iterations than the first male face.

We further compare the robustness and accuracy of the method using adaptive sampling and those of the uniform sampling. Using uniform sampling and the weighted samples as the target Dirac measure, some cells are close to being degenerated, with the searching process wandering about the boundary of the admissible space Σ. The hand model in Figs. 8 and 9 and the brain model in Figs. 6 and 7 do not converge. The Buddha model in Fig. 12 converges with the maximal relative error of 2.03e-4 with the same number of iterations, and the facial models have a relative error about the 3e-7 level.

To further verify the robustness, we add two new experiments to test the target measure with complicated supports. Unlike previous experiments, the optimal transportation maps for these models are discontinuous, and we accurately computed the discontinuity singularities, as shown in Fig. 13.

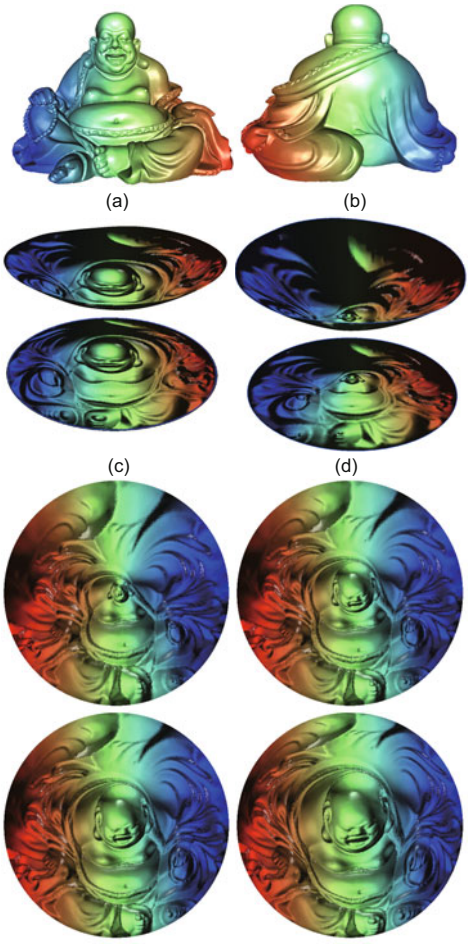

Fig. 12 Buddha surface: (a) front view ot the Buddha surface; (b) back view of the Buddha surface; (c) Brenier potential; (d) Legendre dual (The last two rows show the intermediate computational results during the optimization)

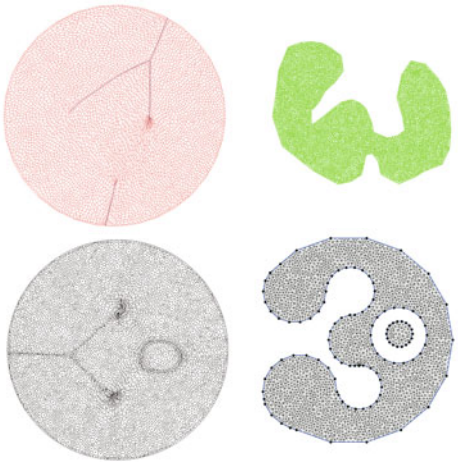

Fig. 13 Optimal transportation maps to target measures with complicated supports. The OT maps are discontinuous at the singularity sets, which are the dark curves in the disks

6 Conclusions

In this study, we proposed a novel self-adaptive sampling method which is uniform with respect to the density of the target measure, and this greatly improved the robustness and accuracy of the geometric variational algorithm for computing optimal transportation maps. Experimental results demonstrated the effectiveness of this method. In the future, we will generalize this method to higher-dimensional situations and compute the optimal transportation maps with more general cost functions.

Contributors

Yingshi WANG, Na LEI, and Xianfeng GU offered the insights and proposed the algorithms. Xiaopeng ZHENG, Wei CHEN, Xin QI, and Yuxue REN implemented the algorithms and visualized the data. All the authors have collaboratively organized the materials, conducted the experiments, drafted the manuscript, and revised and finalized the paper.

Compliance with ethics guidelines

Yingshi WANG, Xiaopeng ZHENG, Wei CHEN, Xin QI, Yuxue REN, Na LEI, and Xianfeng GU declare that they have no conflict of interest.

References

Alauzet F, Loseille A, Dervieux A, et al., 2006. Multi-dimensional continuous metric for mesh adaptation. Proc 15th Int Meshing Roundtable, p.191-214. https://doi.org/10.1007/978-3-540-34958-7_12

Apel T, Lube G, 1996. Anisotropic mesh refinement in stabilized Galerkin methods. *Num Math*, 74(3):261-282. https://doi.org/10.1007/s002110050216

Apel T, Grosman S, Jimack PK, et al., 2004. A new methodology for anisotropic mesh refinement based upon error gradients. *Appl Num Math*, 50(3-4):329-341. https://doi.org/10.1016/j.apnum.2004.01.006

Arjovsky M, Chintala S, Bottou L, 2017. Wasserstein generative adversarial networks. Proc 34th Int Conf on Machine Learning, p.214-223.

Berndt M, Shashkov MJ, 2003. Multilevel accelerated optimization for problems in grid generation. Proc 12th Int Meshing Roundtable, p.351-359.

Bossen FJ, Heckbert PS, 1996. A pliant method for anisotropic mesh generation. Proc 5th Int Meshing Roundtable, p.115-126.

Bottasso CL, 2004. Anisotropic mesh adaption by metric-driven optimization. *Int J Num Methods Eng*, 60(3):597-639. https://doi.org/10.1002/nme.977

Brandts J, Korotov S, Křížek M, 2008. On the equivalence of regularity criteria for triangular and tetrahedral finite element partitions. *Comp Math Appl*, 55(10):2227-2233. https://doi.org/10.1016/j.camwa.2007.11.010

Brenier Y, 1991. Polar factorization and monotone rearrangement of vector-valued functions. *Comm Pure Appl Math*, 44(4):375-417. https://doi.org/10.1002/cpa.3160440402

Chen L, Xu JC, 2004. Optimal Delaunay triangulations. *J Comp Math*, 22(2):299-308.

Cheng SW, Dey TK, Ramos EA, et al., 2007. Sampling and meshing a surface with guaranteed topology and geometry. *SIAM J Comp*, 37(4):1199-1227. https://doi.org/10.1137/060665889

Cheng SW, Dey TK, Shewchuk JR, 2012. Delaunay Mesh Generation. CRC Press, Boca Raton, USA.

Chew LP, 1989. Guaranteed-Quality Triangular Meshes. Technical Report TR 89-983. Computer Science Department, Cornell University, USA.

Chew LP, 1993. Guaranteed-quality mesh generation for curved surfaces. Proc 9th Annual Symp on Computational Geometry, p.274-280. https://doi.org/10.1145/160985.161150

Cui L, Qi X, Wen CF, et al., 2019. Spherical optimal transportation. *Comp Aided Des*, 115:181-193. https://doi.org/10.1016/j.cad.2019.05.024

Cuturi M, 2013. Sinkhorn distances: lightspeed computation of optimal transportation distances. https://arxiv.org/abs/1306.0895

de Goes F, Cohen-Steiner D, Alliez P, et al., 2011. An optimal transport approach to robust reconstruction and simplification of 2D shapes. *Comp Graph Forum*, 30(5):1593-1602. https://doi.org/10.1111/j.1467-8659.2011.02033.x

de Goes F, Breeden K, Ostromoukhov V, et al., 2012. Blue noise through optimal transport. *ACM Trans Graph*, 31(6):171. https://doi.org/10.1145/2366145.2366190

Dey TK, 2007. Curve and Surface Reconstruction: Algorithms with Mathematical Analysis. Cambridge University Press, NY, UK.

Dey TK, Levine JA, 2007. Delaunay meshing of isosurfaces. Proc Shape Modeling Int, p.241-250. https://doi.org/10.1007/s00371-008-0224-1

Dey TK, Ray T, 2010. Polygonal surface remeshing with Delaunay refinement. *Eng Comp*, 26(3):289-301. https://doi.org/10.1007/s00366-009-0162-1

Dobrzynski C, Frey P, 2008. Anisotropic Delaunay mesh adaptation for unsteady simulations. Proc 17th Int Meshing Roundtable, p.177-194. https://doi.org/10.1007/978-3-540-87921-3_11

Dominitz A, Tannenbaum A, 2010. Texture mapping via optimal mass transport. *IEEE Trans Vis Comput Graph*, 16(3):419-433. https://doi.org/10.1109/TVCG.2009.64

Du Q, Wang DS, 2005. Anisotropic centroidal Voronoi tessellations and their applications. *SIAM J Sci Comput*, 26(3):737-761. https://doi.org/10.1137/S1064827503428527

Edelsbrunner H, 2001. Geometry and Topology for Mesh Generation. Cambridge University Press, Cambridge, England, UK.

Gu XF, Luo F, Sun J, et al., 2016. Variational principles for Minkowski type problems, discrete optimal transport, and discrete Monge-Ampère equations. *Asian J Math*, 20(2):383-398. https://doi.org/10.4310/AJM.2016.v20.n2.a7

Gu XF, Luo F, Sun J, et al., 2018. A discrete uniformization theorem for polyhedral surfaces. *J Diff Geom*, 109(2):223-256. https://doi.org/10.4310/jdg/1527040872

Guan P, Wang XJ, et al., 1998. On a Monge-Ampère equation arising in geometric optics. *J Diff Geom*, 48(2):205-223. https://doi.org/10.4310/jdg/1214460795

Gulrajani I, Ahmed F, Arjovsky M, et al., 2017. Improved training of Wasserstein GANs. Proc 31st Int Conf on Neural Information Processing Systems, p.5769-5779.

Gutiérrez CE, Huang Q, 2009. The refractor problem in reshaping light beams. *Archive Ration Mech Anal*, 193(2):423-443. https://doi.org/10.1007/s00205-008-0165-x

Huang WZ, 2006. Mathematical principles of anisotropic mesh adaptation. *Commun Comput Phys*, 1(2):276-310.

Kantorovich LV, 2006. On a problem of Monge. *J Math Sci*, 133(4):1383. https://doi.org/10.1007/s10958-006-0050-9

Kitagawa J, Mèrigot Q, Thibert B, 2019. Convergence of a Newton algorithm for semi-discrete optimal transport. *J Eur Math Soc*, 21(9):2603-2651. https://doi.org/10.4171/JEMS/889

Kunert G, 2002. Toward anisotropic mesh construction and error estimation in the finite element method. *Num Meth Partial Diff Equ*, 18(5):625-648. https://doi.org/10.1002/num.10023

Lei N, Su KH, Cui L, et al., 2019. A geometric view of optimal transportation and generative mode. *Comp Aided Geomet Des*, 68:1-21. https://doi.org/10.1016/j.cagd.2018.10.005

Liu Y, Wang WP, Lévy B, et al., 2009. On centroidal Voronoi tessellation-energy smoothness and fast computation. *ACM Trans Graph*, 28(4):1-17. https://doi.org/10.1145/1559755.1559758

Löhner R, Cebral J, 2000. Generation of non-isotropic unstructured grids via directional enrichment. *Int J Numl Meth Eng*, 49(1-2):219-232.

Mérigot Q, 2011. A multiscale approach to optimal transport. *Comp Graph Forum*, 30(5):1583-1592. https://doi.org/10.1111/j.1467-8659.2011.02032.x

Meyron J, Mérigot Q, Thibert B, 2018. Light in power: a general and parameter-free algorithm for caustic design. *ACM Trans Graph*, 37(6):224. https://doi.org/10.1145/3272127.3275056

Nadeem S, Su ZY, Zeng W, et al., 2017. Spherical parameterization balancing angle and area distortions. *IEEE Trans Vis Comput Graph*, 23(6):1663-1676. https://doi.org/10.1109/TVCG.2016.2542073

Rong G, Liu Y, Wang W, et al., 2011. GPU-assisted computation of centroidal Voronoi tessellation. *IEEE Trans Vis Comp Graph*, 17(3):345-356. https://doi.org/10.1109/TVCG.2010.53

Ruppert J, 1995. A Delaunay refinement algorithm for quality 2-dimensional mesh generation. *J Algorithms*, 18(3):548-585. https://doi.org/10.1006/jagm.1995.1021

Shewchuk JR, 2002. Delaunay refinement algorithms for triangular mesh generation. *Comput Geom*, 22(1-3):21-74. https://doi.org/10.1016/S0925-7721(01)00047-5

Sirois, Y, Dompierre J, Vallet MG, et al., 2006. Mesh smoothing based on Riemannian metric non-conformity minimization. Proc 15th Int Meshing Roundtable, p.271-288. https://doi.org/10.1007/978-3-540-34958-7_16

Solomon J, Rustamov R, Guibas L, et al., 2014. Earth mover's distances on discrete surfaces. *ACM Trans Graph*, 33(4):67. https://doi.org/10.1145/2601097.2601175

Solomon J, de Goes F, Peyré G, et al., 2015. Convolutional Wasserstein distances: efficient optimal transportation on geometric domains. *ACM Trans Graph*, 34(4):66. https://doi.org/10.1145/2766963

Su KH, Cui L, Qian K, et al., 2016. Area-preserving mesh parameterization for poly-annulus surfaces based on optimal mass transportation. *Comp Aided Geom Des*, 46:76-91. https://doi.org/10.1016/j.cagd.2016.05.005

Su KH, Chen W, Lei N, et al., 2017. Volume preserving mesh parameterization based on optimal mass transportation. *Comp Aided Des*, 82:42-56. https://doi.org/10.1016/j.cad.2016.05.020

Su ZY, Zeng W, Shi R, et al., 2013. Area preserving brain mapping. Proc IEEE Conf on Computer Vision and Pattern Recognition, p.2235-2242. https://doi.org/10.1109/CVPR.2013.290

Su ZY, Wang YL, Shi R, et al., 2015. Optimal mass transport for shape matching and comparison. *IEEE Trans Patt Anal Mach Intell*, 37(11):2246-2259. https://doi.org/10.1109/TPAMI.2015.2408346

ur Rehman T, Haber E, Pryor G, et al., 2009. 3D nonrigid registration via optimal mass transport on the GPU. *Med Image Anal*, 13(6):931-940. https://doi.org/10.1016/j.media.2008.10.008

Villani C, 2008. Optimal Transport: Old and New. Springer Science & Business Media, Berlin, Germany.

Wang XJ, 1996. On the design of a reflector antenna. *Inverse Probl*, 12(3):351. https://doi.org/10.1088/0266-5611/12/3/013

Wang XJ, 2004. On the design of a reflector antenna II. *Calc Var Part Diff Equ*, 20(3):329-341. https://doi.org/10.1007/s00526-003-0239-4

Yan DM, Wang K, Levy B, et al., 2011. Computing 2D periodic centroidal Voronoi tessellation. 8th Int Symp on Voronoi Diagrams in Science and Engineering, p.177-184. https://doi.org/10.1109/ISVD.2011.31

Yau ST, 1998. S.S. Chern: a Great Geometer of the Twentieth Century. International Press of Boston, p.366.

Zhao X, Su ZY, Gu XD, et al., 2013. Area-preservation mapping using optimal mass transport. *IEEE Trans Vis Comp Graph*, 19(12):2838-2847. https://doi.org/10.1109/TVCG.2013.135