Supplementary

Dongsheng An[1], Yang Guo[1], Min Zhang[2], Xin Qi[1], Na Lei[3], and Xianfang Gu[1]

[1] Stony Brook University
[2] Harvard Medical School
[3] Dalian University of Technology

1 Ideal sampling

In our context, *ideal sampling* of a manifold $\Sigma \subset \mathbb{R}^d$ is to find a dense enough point cloud S such that a triangulation T of S could be constructed by interpolation such that $|T|$ is homeomorphic to Σ, and the Hausdorff distance between T and Σ is small. In fact, in [2] the authors proved a theorem that by considering (ϵ, δ)-sampling, the weighted Delaunay triangulation of point cloud faithfully recovers the manifold.

In our model, we assume that the input real image dataset is an ideal sampling of real image manifold, and images are viewed as being i.i.d. uniformly sampled from the image manifold.

2 The construction of μ

Compute the SDOT map T In the first step, we compute the semi-discrete OT map $T : [0,1]^d \to \Omega$, with $T_{\#}Uni([0,1]^d) = \hat{\mu}_{gt}$. Under T, the continuous domain of $[0,1]^d$ is decomposed into cells $\{W_i\}$ with $T(w) = z_i \ \forall w \in W_i$, and the Lebesgue measure of each W_i is $\frac{1}{n}$. The cell structure is shown in the left frame of Fig. 1 (the orange cells).

According to [1], the computation of the semi-discrete OT map is equivalent to minimize the convex energy $E(h) = \int_0^h \sum_{i=1}^n \int_{W_i} d\eta - \frac{1}{n}\sum_{i=1}^n h_i$, where the height vector h is the unique optimizer under the condition that $\sum_i h_i = 0$ and W_i represents the i-th cell of the cell-decomposition induced by h [4]. By adaptive Monte-Carlo method introduced by [1], we can use $\hat{a}_i(h) = \#\{j \in \mathcal{J} \mid w_j \in W_i(h)\}/N$ to approximate the volume of W_i, thus the gradient $\nabla E(h) \approx (\hat{a}_i(h) - \nu_i)^T$ can be estimated. Here $w_j, j = 1, 2, \ldots, N$ is sampled from the uniform distribution. With the optimal weight h, the domain of the uniform distribution will be split into n cells, and the Lebesgue measure of each of the cell equals to $\hat{\mu}_{gt}(z_i) = 1/n$. Then we get the semi-discrete OT map $T : w \in W_i \to z_i$.

Piece-wise linear Extension of T Secondly, we extend the image domain of T from the discrete latent codes $\{z_i\}$ to a continuous neighborhood σ_ε, which serves as the supporting manifold of μ. Specifically, we construct a simplicial complex σ_ε from the latent codes $\{z_i\}$. Here $\varepsilon > 0$ is a constant. The 0-skeleton of σ_ε, represented by $\sigma_\varepsilon^{(0)}$, is the set of all latent codes $\{z_i\}$. The we define its k-skeletons $\sigma_\varepsilon^{(k)}$ by $\sigma_\varepsilon^{(k)} = \{[z_1, z_2, \ldots, z_k] \mid \|z_i - z_j\|_2 \leq \varepsilon, \forall 1 \leq i < j \leq k\}$ for $0 < k \leq d$. The right frame of Fig. 1 shows an example of σ_ε. By assuming that the latent code is densely sampled

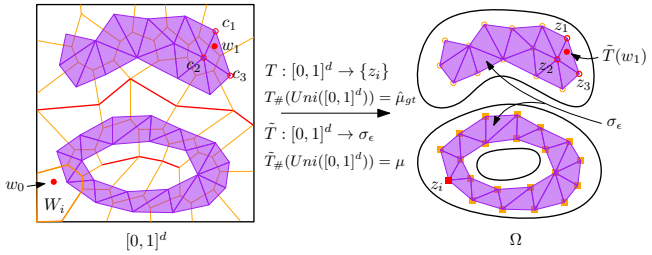

Fig. 1. OT map T and the extended OT map $\tilde{T}$ in 2D case. Here T maps points (e.g. w_0) in each polyhedral cell W_i (orange cells on the left) to the corresponding latent code z_i (circles and squares on the right). The piece-wise linear $\tilde{T}$ maps triangulated regions in $[0,1]^d$ to the simplicial complex σ_ε in the latent space (shown in purple). Given the barycenters c_i's of each W_i's, each triangle $\Delta c_i c_j c_k$ is mapped to the corresponding simplex $[z_i, z_j, z_k]$. For example, w_1 in the triangle $\Delta c_0 c_1 c_2$ is mapped to $\tilde{T}(w_1)$ in the simplex $[z_0, z_1, z_2]$. Red lines in $[0,1]^d$ illustrate the singular set of T, which corresponds to the pre-image of gaps or holes in Ω.

from the latent manifold Ω and with an appropriate ε, σ_ε will have consistent "hole" and "gap" structure with Ω, in the sense of homology equivalence, as explained in Chap. 3.

Finally, we define the piece-wise linear extended OT map $\tilde{T} : [0,1]^d \to \sigma_\varepsilon$. Representing the cells W_i by their mass centers c_i, we can get the point-wise map $t : c_i \mapsto z_i$. Also, $\{c_i\}$ gives a dual cell decomposition of Ω, as shown in the purple triangles of the left frame of Fig. 1. Given a random sample w sampled from $Uni([0,1]^d)$, we can approximate the dual cell (the purple triangles in the left frame of Fig. 1) containing it by its nearby mass centers c_j. Computing the barycentric parameters λ_j's with respect to the nearby c_js of w, i.e. compute λ_j's such that $w = \Sigma\lambda_j c_j$ with $0 \leq \lambda_j \leq 1$ and $\Sigma\lambda_j = 1$. Then w is mapped to $\tilde{T}(w) := \Sigma\lambda_j T(c_j) = \Sigma\lambda_j z_j$ if the corresponding z_j's form a simplex of σ_ε. Otherwise we only map w to z_i, i.e. $\tilde{T}(w) := T(c_i) = z_i$ where c_i is the nearest neighbour of w in $\{c_i\}$. As illustrated in Fig. 1, compared to the many-to-one semi-discrete OT map T, $\tilde{T}$ maps samples within the triangular areas (the purple triangles on the left frame) in $[0,1]^d$ *piecewise linearly* to the corresponding simplices in σ_ε (the purple triangles on the right frame) in a bijective manner. We denote the pushed forward distribution under $\tilde{T}$ as $\mu_\varepsilon := \tilde{T}_\# Uni([0,1]^d)$.

Theorem 1. *The 2-Wasserstein distance between μ_ε and $\hat{\mu}_{gt}$ satisfies $W_2(\mu_\varepsilon, \hat{\mu}_{gt}) \leq \varepsilon$, where ε is a given constant to build μ. Moreover, if the latent codes are densely sampled from the latent manifold Ω, we have $W_2(\mu_\varepsilon, \mu_{gt}) \leq 2\varepsilon$, μ-almost surely.*

Proof. In fact, by the construction of μ_ϵ, samples in a cell W_i in $[0,1]^d$ is mapped to the neighboring simplices of $z_i = T(W_i)$ under $\tilde{T}_\epsilon$, or z_i itself. We could construct a transport map U mapping $\tilde{T}_\epsilon(w)$ to z_i if $w \in W_i$. To see U is a transport map, we note that since $U \circ \tilde{T} = T$ and $T_\# Uni([0,1]^d) = \hat{\mu}_{gt}$, we have $U_\# \mu = \hat{\mu}_{gt}$ where $\mu = \tilde{T}_\# Uni([0,1]^d)$. The total cost of this transport map under cost function $c(x,y) = |x-y|^2$ is less than ϵ^2. This is because $\tilde{T}(w)$ is mapped into adjacent simplices of z_i, and these simplices are contained in the ball of radius ϵ centered at z_i. Therefore we have $|z_i, \tilde{T}_\epsilon(w_i)|^2 \leq \epsilon^2$. Moreover, since the Wasserstein distance involves the minimal

l^2 cost among all transport plans, we have $W_2(\mu_\epsilon, \hat{\mu}_{gt}) \leq \epsilon$. This proves the first part of the theorem.

The second part of the theorem comes from the fact that the empirical distribution $\hat{\mu}$ converges to μ weakly a.s. and that W_2 is a metric in probability space.

To avoid confusion, we omit the subscript ε and denote μ_ε as μ in the paper.

3 Homology property of constructed latent simplex σ_ϵ

Suppose $\{z_i\}$ are densely sampled on its supporting manifold Ω. Here dense sampling means the number of samples is larger than a constant related to the condition number of Ω (see e.g. [5]). We show that σ_ϵ is a subcomplex of a simplicial complex that has the same homology as Ω. In fact, given the point cloud $\{z_i\}$, we consider Euclidean balls $B_\epsilon(z_i)$ that centers at z_i and has radius ϵ. By Niyogi-Smale-Weinberger theorem [5], if the data $\hat{\mu}_{gt}$ is densely sampled on Ω, then with high probability and sufficiently small ϵ, the ball union $U_{2\epsilon} = \cup_{z_i} B_{2\epsilon}(z_i)$ has equal homology. By the nerve theorem, the ball union $U_{2\epsilon}$ retracts to the Cech complex $C_{2\epsilon}$. This means $C_{2\epsilon}$ has the same homology as Ω. On the other hand, by construction, σ_ϵ is the Vietoris-Rips complex of weight ϵ on point cloud $\{z_i\}$, which can be embedded between $C_{\epsilon/2}$ and $C_{2\epsilon}$ [3].

In our model, we want to choose a suitable supporting manifold for μ. By this we mean that when the support χ of the ground-truth distribution ν_{gt} in the image space has several connected components or if there are “holes" in the components, we want our latent manifold to have similar structure. This is guaranteed by the fact that σ_ϵ is a subcomplex of $C_{2\epsilon}$ which has the same homology as Ω, and the assumption that Ω is a nice latent representation of χ.

In practice, the computation of homology in high dimensional space is complicated and time consuming. We treat ϵ as a model hyperparameter without explicitly computing the homology groups.

4 The network architecture for CelebA-HQ

In this section, we introduce the architecture used to train the CelebA-HQ dataset in Tab. 1. We also made a video to show the evolution of our generator trained on CelebA-HQ during the training process. From the video it can be noted that as the epochs increase, the performance of the generator is significantly enhanced, as both of the reconstructed and generated images show better visual qualities.

Encoder	Kernel Size	Resampling	Output Shape
Conv	3 x 3		32 x 128 x 128
Residual block	[3 x 3] x 2	Down	64 x 64 x 64
Residual block	[3 x 3] x 2	Down	128 x 32 x 32
Residual block	[3 x 3] x 2	Down	256 x 16 x 16
Residual block	[3 x 3] x 2	Down	512 x 8 x 8
Residual block	[3 x 3] x 2	Down	1024 x 4 x 4
Linear			128
Decoder			
Linear			1024 x 4 x 4
Residual block	[3 x 3] x 2	Up	512 x 8 x 8
Residual block	[3 x 3] x 2	Up	256 x 16 x 16
Residual block	[3 x 3] x 2	Up	128 x 32 x 32
Residual block	[3 x 3] x 2	Up	64 x 64 x 64
Residual block	[3 x 3] x 2	Up	32 x 128 x 128
Conv, tanh	3 x 3		3 x 256 x 256
Discriminator			
Conv	3 x 3		32 x 128 x 128
Residual block	[3 x 3] x 2	Down	64 x 64 x 64
Residual block	[3 x 3] x 2	Down	128 x 32 x 32
Residual block	[3 x 3] x 2	Down	256 x 16 x 16
Residual block	[3 x 3] x 2	Down	512 x 8 x 8
Residual block	[3 x 3] x 2	Down	1024 x 4 x 4
Linear			1

Table 1. The architecture used to train the CelebA-HQ dataset.

References

1. An, D., Guo, Y., Lei, N., Luo, Z., Yau, S.T., Gu, X.: Ae-ot: A new generative model based on extended semi-discrete optimal transport. In: International Conference on Learning Representations (2020)
2. Cheng, S.W., Dey, T.K., Ramos, E.A.: Manifold reconstruction from point samples. In: SODA. vol. 5, pp. 1018–1027 (2005)
3. De Silva, V., Ghrist, R.: Coverage in sensor networks via persistent homology. Algebraic & Geometric Topology **7**(1), 339–358 (2007)
4. Gu, D.X., Luo, F., Sun, j., Yau, S.T.: Variational principles for minkowski type problems, discrete optimal transport, and discrete monge-ampère equations. Asian Journal of Mathematics (2016)
5. Niyogi, P., Smale, S., Weinberger, S.: Finding the homology of submanifolds with high confidence from random samples. Discrete & Computational Geometry **39**(1-3), 419–441 (2008)