Supplementary Material

Dongsheng An[1], Na Lei[2], and Xianfeng Gu[1]

[1] Stony Brook University
[2] Dalian University of Technology
{doan, gu}@cs.stonybrook.edu, nalei@dlut.edu.cn

1 Proof of Thm. 6

Theorem 6 Given $\nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^d)$, the auxiliary measure is μ, $T_k : \mu \to \nu_k$ are the OT maps with k=1,2. Suppose the distance from μ to the geodesic connecting ν_1 and ν_2 is d, then $T_2 \circ T_1^{-1} : \nu_1 \to \nu_2$ is measure preserving and its transport cost $\mathcal{C}$ is bounded by

$$\mathcal{W}_c(\nu_1, \nu_2) \le \mathcal{C}^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \le \mathcal{W}_c(\nu_1, \nu_2) + 2d \tag{1}$$

Proof. Suppose the geodesic connecting ν_1 and ν_2 is γ, μ^* is the closest point to μ on γ. By definition, $(T_k)_\# \mu = \nu_k$, then we have

$$(T_2 \circ T_1^{-1})_\# \nu_1 = (T_2)_\# (T_1^{-1})_\# \nu_1 = (T_2)_\# \mu = \nu_2. \tag{2}$$

Thus, $T_2 \circ T_1^{-1}$ is measure preserving, but it may not be optimal. Since here we assume that the cost function is given by the L^2 distance, we have $\mathcal{C}(T_k) = W_c^2(\mu, T_k)$. Then

$$\mathcal{C}(T_2 \circ T_1^{-1}) \ge \mathcal{W}_c^2(\nu_1, \nu_2). \tag{3}$$

T_k's are the optimal transport maps, according to the triangle inequality, we have

$$\mathcal{C}^{\frac{1}{2}}(T_1) + \mathcal{C}^{\frac{1}{2}}(T_2) \le \mathcal{W}_c(\nu_1, \nu_2) + 2d. \tag{4}$$

Assume the cell decomposition of T_1 and T_2 is given by $\{W_i^1\}$ and $\{W_j^2\}$, and the refined cell decomposition of $\{W_i^1\}$ and $\{W_j^2\}$ is $\{W_{ij}\}$ with $W_{ij} := W_i^1 \cap W_j^2$. If we set $d(x, y) = \|x - y\|_2$ and by Minkowski inequality,

$$
\begin{aligned}
&\mathcal{C}^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \\
&= \left[\sum_{i,j=1}^{m,n} \int_{W_{ij}} d(y_i^1, y_j^2)^2 d\mu(x) \right]^{\frac{1}{2}} \\
&\leq \left[\sum_{i,j=1}^{m,n} \int_{W_{ij}} (d(x, y_i^1) + d(x, y_j^2))^2 d\mu(x) \right]^{\frac{1}{2}} \\
&\leq \left[\sum_{i,j=1}^{m,n} \int_{W_{ij}} d(x, y_i^1)^2 d\mu(x) \right]^{\frac{1}{2}} + \left[\sum_{i,j=1}^{m,n} \int_{W_{ij}} d(x, y_j^2)^2 d\mu(x) \right]^{\frac{1}{2}} \\
&= \left[\sum_{i=1}^{m} \int_{W_i^1} \|x - y_i^1\|^2 d\mu(x) \right]^{\frac{1}{2}} + \left[\sum_{j=1}^{n} \int_{W_j^2} \|x - y_j^2\|^2 d\mu(x) \right]^{\frac{1}{2}} \\
&= \mathcal{C}^{\frac{1}{2}}(T_1) + \mathcal{C}^{\frac{1}{2}}(T_2)
\end{aligned}
\tag{5}
$$

Thus,

$$\mathcal{C}^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \leq \mathcal{W}_c(\nu_1, \nu_2) + 2d. \tag{6}$$

Combining the above estimates, we obtain the bounds

$$\mathcal{W}_c(\nu_1, \nu_2) \leq \mathcal{C}^{\frac{1}{2}}(T_2 \circ T_1^{-1}) \leq \mathcal{W}_c(\nu_1, \nu_2) + 2d \tag{7}$$

2 Proof of Proposition 7

Proposition 7 Given $\mu = \sum_{i=1}^m \nu_i^1 N(x_i, \sigma^2 I_d)$ and $\nu_1 = \sum_{i=1}^m \nu_i^1 \delta(x - x_i)$, then we have $\mathcal{W}_c(\mu, \nu_1) \leq \sigma$ under the quadratic Euclidean cost. Moreover, if σ is small enough, then the cell W_i of the cell decomposition induced by the semi-discrete OT map from μ to ν_1 should cover x_i itself.

Proof. If we transport all the mass corresponding to $N(x_i, \sigma I_d)$ to x_i of ν_1, then we get a transport plan from μ to ν_1. By defining $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\{\frac{\|x\|2}{2\sigma^2}\}$, the transport cost of such a transport plan is given by

$$\mathcal{C} = \sum_{i=1}^{m} \nu_i^1 \int \|x\|^2 f(d) dx = \sigma^2 \tag{8}$$

Thus, the optimal transport cost from μ to ν_1, namely $\mathcal{W}_c^2(\mu, \nu_1)$, should be no more than σ^2. This gives

$$\mathcal{W}_c(\mu, \nu_1) \leq \sigma \tag{9}$$

When $\sigma \ll \min_{i \neq j} \|x_i - x_j\|_2$, the cell W_is of the cell decomposition induced by the semi-discrete OT map from μ to ν_1 should cover the corresponding x_is, namely nearly all mass of $\nu_i^1 \mathcal{N}(x_i, \sigma^2 I_d)$ should be transported to x_i. If W_i does not cover x_i, some mass of $\mathcal{N}(x_j, \sigma^2 I_d)$ with $x_j \neq x_i$ will be transported to x_i, as a result $\mathcal{W}_c(\mu, \nu_1)$ will be larger than σ. This corresponds to the cyclical monotonicity of the optimal transport (Chapter 5 of [4]).

Algorithm 1 Semi-discrete OT Map

Input: the absolutely continuous source measure μ and the discrete target measure $\nu = \sum_{i=1}^{n} \nu_i \delta(x - x_i)$, number of Monte Carlo samples N, positive integer s and the measure accuracy θ.
Output: Optimal transport map $T(\cdot)$.
Initialize $h = (h_1, h_2, \ldots, h_{|\mathcal{I}|}) \leftarrow (0, 0, \ldots, 0)$.
repeat
  Sample N samples $\{z_j\}_{j=1}^{N} \sim \mu$.
  Calculate $\nabla h = (\hat{w}_i(h) - \nu_i)^T$.
  $\nabla h = \nabla h - mean(\nabla h)$.
  Update h by Adam algorithm with $\beta_1 = 0.9, \beta_2 = 0.5$.
  if $E(h)$ has not decreased for s steps **then**
    $N \leftarrow N \times 2$.
  end if
until $\sum_{i=1}^{n} abs(\hat{w}_i(h) - \nu_i) < \theta$
OT map $T(\cdot) \leftarrow \nabla(\max_i \langle \cdot, x_i \rangle + h_i)$.

Algorithm 2 Construct the sparse matrix

Input: the absolutely continuous source measure μ, the computed h_1 for ν_1, and the computed h_2 for ν_2.
Output: Sparse matrix S of the transport plan.
Initialize $S = 0_{m \times n}$.
repeat
  Sample $z \sim \mu$.
  Find the cell W_i^1 in $\{W_i^1\}$ that contains z.
  Find the cell W_j^2 in $\{W_j^2\}$ that contains z.
  Set $S(i, j) = 1$
until converge

3 Algorithm Pipeline for the SDOT algorithm

Based on [1], we summarize the whole pipeline of the SDOT (semi-discrete optimal transport) algorithm in Alg. 1.

4 Algorithm Pipeline for constructing the spare matrix

We also summarize the whole pipeline of constructing and extending the sparse matrix S in Alg. 2.

5 Algorithm for Discrete OT plan with continuous μ where the source measure is sampled from

In the section, we give the algorithm pipeline for computing the discrete OT plan with the continuous μ where the source measure ν_1 is sampled from, as shown in Alg. 3.

Algorithm 3 Discrete Optimal Transport Plan

1: **Input:** The absolutely continuous source measure μ, $\nu_1 = \sum_{i=1}^{m} \nu_i^1 \delta(x - x_i)$ and $\nu_2 = \sum_{j=1}^{n} \nu_j^2 \delta(y - y_j)$, the μ-volume distortion θ and the number k of the nearest neighbours.
2: **Output:** The approximate OT plan.
3: Compute the semi-discrete OT map T_1 and T_2 from μ to ν_1 and ν_2 with the parameter θ.
4: Initialize the sparse matrix S according to Alg. 2.
5: Extend S according to its k nearest neighbours.
6: Solve the sparse LP problem Eqn. (7).

Algorithm 4 Discrete Optimal Transport Plan by GM model

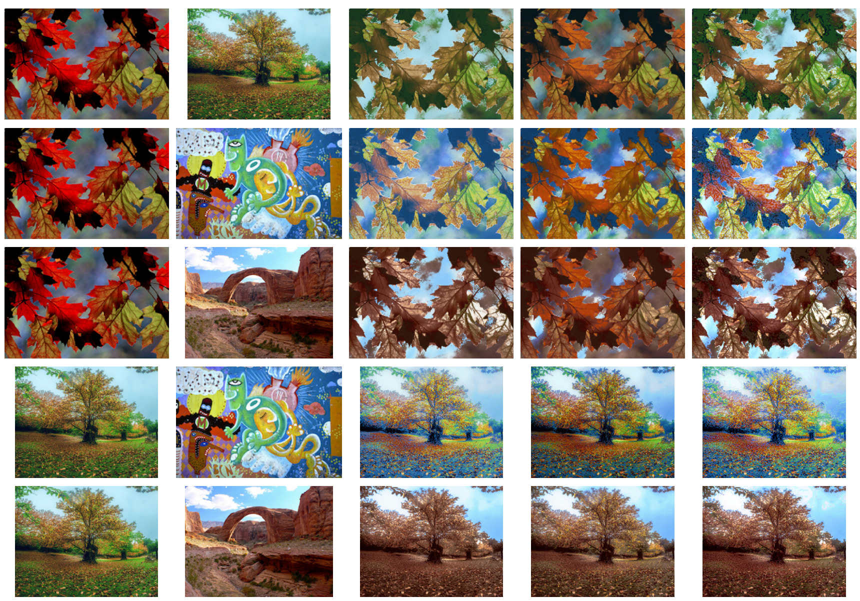
1: **Input:** $\nu_1 = \sum_{i=1}^{m} \nu_i^1 \delta(x - x_i)$ and $\nu_2 = \sum_{j=1}^{n} \nu_j^2 \delta(y - y_j)$, the measure accuracy θ and the nearest number of k.
2: **Output:** The transport plan.
3: Construct $\mu = \sum_{i=1}^{m} \nu_i^1 \mathbb{N}(x_i, \sigma I_d)$, with $\sigma = 0.1 \min_{i \neq k} d(x_i, x_k)$.
4: Compute the semi-discrete OT map T_2 from μ to ν_2 with the parameter θ based on Alg. 1.
5: Initialize the sparse matrix S: for each sample x_i, find the cell W_j^2 covering it. Then set $S(i, j) = 1$.
6: Extend S according to the k nearest neighbours.
7: Solve the sparse LP problem of Eqn. (7).

6 Algorithm for Discrete OT plan with Gaussian Mixture μ defined by the source measure

In this section, we introduce the algorithm to compute the discrete OT plan with μ being Gaussian mixture model defined by the source measure ν_1, as shown in Alg. 4.

7 More results of Color Transfer

In Fig. 1, we show the additional color transfer results of **(i)** autumn to comunion; **(ii)** autumn to graffiti; **(iii)** autumn to rainbow-bridge; **(iv)** comunion to graffiti; and **(v)** comunion to rainbow-bridge. It is obvious that the results of the proposed method are sharper than those of Sinkhorn [3]. And though the color transferred images of SOT [2] are sharp, the color spaces of them are problematic, as shown in the first three images of the 4th column.



(a) Source (b) Target (c) Sinkhorn (d) SOT (e) Ours.

Fig. 1. Additional comparison of the results on color transfer tasks.

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. Blondel, M., Seguy, V., Rolet, A.: Smooth and sparse optimal transport. In: Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics. pp. 880–889 (2018)
3. Cuturi, M.: Sinkhorn distances: Lightspeed computation of optimal transportation distances. In: International Conference on Neural Information Processing Systems. vol. 26, pp. 2292–2300 (2013)
4. Villani, C.: Optimal transport: old and new, vol. 338. Springer Science & Business Media (2008)