



école _____
normale _____
supérieure _____
paris—saclay _____

SPOT: SLICED PARTIAL OPTIMAL TRANSPORT

-

FAST PYTHON IMPLEMENTATION, ALGORITHM IMPROVEMENTS, AND PROBLEMS

Auteurs:
Quentin SPINAT

December 30, 2020

Contents

Introduction	1
1 Partial Transport in 1-D	4
1.1 Nearest Neighbor Assignment	4
1.2 Quadratic Time Algorithm	5
1.3 Quasilinear Time Problem Decomposition	9
2 Sliced Partial Transport	12
3 Results	13
3.1 Color Transfer	13
3.2 Point Cloud Registration	13
Conclusion	13
Appendix	13
3.3 Proof of the optimality of algorithm 2	13

Abstract

This paper is the project report for the 2020 MVA course *Computational Optimal Transport* taught by Gabriel Peyré. The purpose of it is to study the paper *SPOT: Sliced Partial Optimal Transport* [1].

Optimal Transport is a mathematical field that focus on finding the best way to match two sets with respect to a matching cost. Here we will focus on a subproblem which consists in finding the best injective assignment of a finite set of points to a second bigger set of points. Although it is a very precise problem, it is often encountered : color transfer from an Image to a bigger one, matching a small point clouds to a bigger one, ...

Nicolas Bonneel et al. proposed a method they called FIST : Fast Iterative Sliced Transport to solve this problem. It consists in solving the assignment problem considering only 1D optimal assignments, which are fast and easy to compute. Their algorithm is told to be quasilinear in time with respect to the datasets sizes.

In this project, I clarify the explanations of the original paper (which is full of typos, especially in the algorithm description), make a fast and user-friendly implementation of this algorithm in Python, and apply it to color transfer and point cloud registration.

Even though I wasn't able to parallelize my code due to python limitations, the implementation I made revealed to be faster than the original implementation in C++ for color transfer application. However I couldn't be able to reproduce good results for point cloud registration.

Keywords— Optimal Transport, Optimal assignment, SGD, Shape matching, Color transfer

Introduction

Problem

Optimal Transport is a mathematical field that focus on finding the best way to match two sets with respect to a matching cost. Here we will focus on a subproblem which consists in finding the best injective assignment of a finite set of points to a second bigger set of points. Although it is a very precise problem, it is often encountered : color transfer from an Image to a bigger one, matching a small point clouds to a bigger one.

Previous Work

Optimal transport is a computationally complex problem, often expressed as a linear program due to Kantorovich. The variables being optimized form what is known as a transport plan to describe the amount of mass from each bin of an input (possibly high-dimensional) histogram that needs to travel to each bin

of a target histogram. In this definition, it is assumed that both histograms are normalized. Directly solving this linear program can be very costly in time and memory, and when dealing with a large amount of points, algorithms become very greedy, especially because of the curse of dimensionality. Faster alternate approaches have been developed to tackle particular cases, but cannot be applied to the problem we try to overcome.

This is why the field of sliced optimal transport has emerged. It consists in projecting the point clouds onto random one-dimensional subspaces, and solving one-dimensional transport problems [Pitié et al. Rabin et al. [2], Bonneel et al. [3], [4]]. Slicing methods are very fast because of how simpler one-dimensional optimal transport is to solve. However, Sliced optimal transport is only an approximation and comes at the cost of being optimal only on the projections and not necessarily on the higher dimensional space. But optimal transport for non-normalized histograms has received little attention, even in one-dimension, which makes the sliced approach not possible for point clouds of different sizes yet.

When histograms are not normalized, the transport problem needs to be relaxed. Notably, the "unbalanced" optimal transport of Benamou et al. [5] replaces hard constraints by soft constraints in a Monge formulation to accommodate for densities of unequal masses within a PDE framework. However, it suffers from a big space complexity. Figalli introduces *partial optimal transport* that keeps all constraints of the linear program hard [6], but replaces equality with inequality constraints. The problem we are interested in is a special case of this linear program for the problem of matching point clouds.

New Contributions

Nicolas Bonneel et al. [1] proposed a Sliced Partial Optimal Transport (SPOT) method they called FIST : Fast Iterative Sliced Transport to solve the problem of optimal injective assignment between point clouds of different sizes. This is a sliced method which consists in minimizing the mean over all the one-dimensional projections of the Wasserstein distance between the two 1D projected point clouds. It hence makes it possible to solve the original problem by doing a Stochastic Gradient Descent and solving only the 1D projected subproblems. To solve these one-dimensional partial optimal transport subproblems efficiently, they propose an algorithm quasi-linear in time with respect to the datasets sizes.

Applications

In this project, we focus on two applications of this new method : color transfer problem and point cloud registration.

Color Transfer

Transferring colors between images has become a classical image processing problem. The purpose is to change the color distribution of an input image, without changing its content, to match the style of a target image. Numerous optimal transport solutions have already been developed [Bonneel et al. [4], [7]; Pitié et al. [2]; Pitié et al. [8]; Rabin et al. [9]]. In addition to their use of optimal transport, a

common point to these approaches is that they consider the problem of matching normalized histograms. A consequence of that is often acknowledged as a limitation: their content should not differ. For instance, matching an image with 80% trees and 20% sky to an image containing the opposite ratio will inevitably lead to trees becoming blue or sky becoming green. Several approaches also require the images to have exactly the same number of pixels – this is precisely the case for sliced transportation [Bonneel et al. [4]; Rabin et al. [9]].

Bonneel et al. proposed two solutions to overcome these problems, but I only focus on one of them : this solution simply enlarges the target images to give more freedom to each input pixel to be matched to target pixel values. This method is really fast and produces really good results

Point Cloud Registration

A common problem in point cloud processing is that of registering points under a given transformation model. For instance, one tries to match a point cloud with another by supposing the transformation between them is rigid, or constrained to a similarity, or affine, or homographic, etc. A well-known algorithm to solve this problem is the Iterative Closest Point (ICP) algorithm. Given a point cloud X_0 to be matched against X_1 , this algorithm proceeds by first matching points of X_0 to their nearest neighbors in X_1 , and, given this assignment, the best transformation in the class of allowed transformations is found by minimizing an energy. The algorithm then transforms the initial point cloud using the computed transformation. The process is repeated until convergence. However, this algorithm requires points clouds to be relatively close to start with and suffers from local minima. In practice, extremely bad behaviors may arise when considering similarity transforms (rotation, translation and scaling). In that case, the lack of injectivity of the nearest neighbor map tends to estimate progressively smaller scaling factors as iterations increase, occasionally leading to a trivial zero scale solution.

This motivates the use of a metric which accounts for an injective mapping, such as the sliced metric proposed by Bonneel et al. They thus propose to replace the nearest neighbor matching by a partial sliced optimal transport matching they call *Fast Iterative Sliced Transport* (FIST).

1 Partial Transport in 1-D

This section will explain the core of this new method, which is an efficient algorithm for Partial Transport in one dimension. It relies on the fact that the nearest neighbor assignment of two point sets can be found in linear time in one-dimension. However, this nearest neighbor assignment not being injective. However, we need to smartly modify it. This can easily be done in quadratic time. To go further and improve the algorithm performances, we need then to introduce a way to decompose the original problem into independent easier problems, which can be solved in parallel. With this problem decomposition, the whole algorithm becomes quasilinear in time.

In all the problem we consider that X is a set of 1-D points of cardinal m , Y a set of 1-D points of cardinal $n < m$, and that X and Y are sorted by increasing value (this can be done in quasilinear time). We also write $\alpha = \frac{m}{n}$.

1.1 Nearest Neighbor Assignment

In one dimension, the nearest neighbor assignment $t : \llbracket 1, m \rrbracket \rightarrow \llbracket 1, n \rrbracket$ from X to Y can be found in linear time. Since $t(i) \leq t(i+1)$, the algorithm consists in scanning X and Y simultaneously from left to right and compare their value.

Algorithm 1 Nearest Neighbor Assignment

Input: sorted X, Y

Output: t

```
1:  $i \leftarrow 1$ 
2:  $j \leftarrow 1$ 
3: while  $i < m$  do
4:   if  $x_i \leq y_j$  then
5:      $t[i] \leftarrow j$ 
6:      $i \leftarrow i + 1$ 
7:   else if  $j = n$  then:
8:      $t[i] \leftarrow n$ 
9:      $i \leftarrow i + 1$ 
10:  else if  $y_{j+1} < x_i$  then
11:     $j \leftarrow j + 1$ 
12:  else if  $|x_i - y_j| < |x_i - y_{j+1}|$  then:
13:     $t[i] \leftarrow j$ 
14:     $i \leftarrow i + 1$ 
15:  else
16:     $t[i] \leftarrow j + 1$ 
17:     $j \leftarrow j + 1$ 
18:     $i \leftarrow i + 1$ 
19:  end if
20: end while
21: return  $t$ 
```

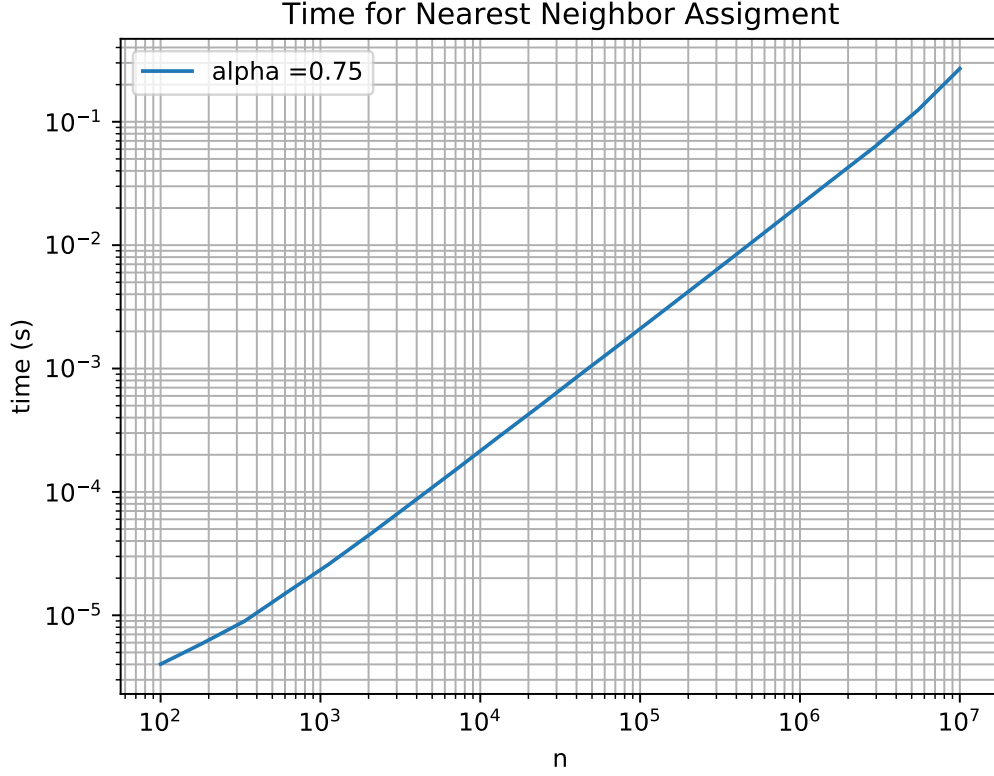


Figure 1: Mean time over 100 simulations for Nearest Neighbor Assignment with respect to n , $\alpha = 0.75$, logscale

1.2 Quadratic Time Algorithm

As the nearest neighbor assignment may not be injective, we need to change it resolving issues around points where $t[i] = t[j]$ for $i \neq j$. To do so, we scan X from left to right and consider the sub-problem of matching $X' = \{x_i\}_{i \in [1; m']}$ with Y , progressively increasing m' from 1 to m .

To create this assignment a from t , we start by initializing $a[1] = t[1]$. Once $\{a[k]\}_{k \in [0; i]}$ is created, there is two options. If $t[i+1] > a[i]$, we put $a[i+1] = t[i+1]$. Else, there are two possible cases : we can either put $a[i+1] = a[i] + 1$ (case 1), or put $a[i+1] = a[i]$ and translate every previous $a[k]$ from one to the left (case 2). The second option writes $a[i+1] = a[i]$ and $a[r : i] = a[r : i] - 1$, where r is such that $y_{a[r]-1}$ is the last available free spot of Y . We chose the solution that minimizes the distance between points (i.e $\sum_{k=1}^{i+1} (x_k - y_{a[k]})^2$).

Proof that the partial assignment given by this algorithm is optimal is given in appendix 3.3.

This algorithm can be optimized to run faster. First, there trivial subcases :

- If $m = n$, we can directly put $a[i] = i$ for i from 1 to m .
- If $m = n - 1$, there exists some k to be determined such that $a[1 : k - 1] = [1 : k - 1]$ and $a[k : m] = [k + 1 : n]$. Such k minimizes $\sum_{i=1}^{k-1} (x_i - y_i)^2 + \sum_{i=k}^m (x_i -$

Algorithm 2 Quadratic Partial Optimal Assignment

Input: sorted X, Y

Output: a

```

1: compute  $t$ 
2:  $a[1] \leftarrow t[1]$ 
3: for  $i$  from 1 to  $m - 1$  do
4:   if  $t[i + 1] > a[i]$  then
5:      $a[i + 1] \leftarrow t[i + 1]$ 
6:     update  $r$ 
7:   else
8:      $w_1 \leftarrow \sum_{k=r}^i (x_k - y_{a[k]})^2 + (x_{i+1} - y_{a[i]+1})^2$ 
9:      $w_2 \leftarrow \sum_{k=r}^i (x_k - y_{a[k]-1})^2 + (x_{i+1} - y_{a[i]})^2$ 
10:    if  $w_1 \leq w_2$  then ▷ Case 1
11:       $a[i + 1] \leftarrow a[i] + 1$ 
12:    else ▷ Case 2
13:       $a[i + 1] \leftarrow a[i]$ 
14:       $a[r : i] \leftarrow a[r : i] - 1$ 
15:      update  $r$ 
16:    end if
17:  end if
18: end for
19: return  $a$ 

```

$y_{i+1})^2$. Denoting $C = \sum_{i=1}^m (x_i - y_{i+1})^2$, the above minimization problem can be written $\min_k \sum_{i=1}^{k-1} (x_i - y_i)^2 + C - (x_i - y_{i+1})^2$ which can be obtained within a single linear search, as C is a constant and need not be computed.

- If t is injective, we put $a = t$.

Second, we can simplify the problem before computing a .

- If there exists i such that $X[1 : i] \leq Y[1 : i]$, we can put $a[1 : i] = \llbracket 1 : i \rrbracket$
- If there exists i such that $Y[i : n] \leq X[i : m]$, we can put $a[i : m] = \llbracket i : m \rrbracket$
- If p is the number of non-injective values of t , $p = \text{card}\{t[i] = t[i + 1], \forall i < m\}$, then, it is enough to consider the sub-problem of matching X to $Y' = \{y_j\}_{j \in \llbracket \max(1, t[1]-p); \min(t[m]+p, n) \rrbracket}$

Finally, we don't have to re-evaluate the sums w_1 and w_2 each time. If we write $S_{i,1} = \sum_{k=r}^i (x_k - y_{a[k]})^2$ and $S_{i,2} = \sum_{k=r}^i (x_k - y_{a[k]-1})^2$, then in the loop i , we have $w_1 = S_{i,1} + (x_{i+1} - y_{a[i]+1})^2$ and $w_2 = S_{i,2} + (x_{i+1} - y_{a[i]})^2$. There is then three cases :

- If we put $a[i + 1] > a[i] + 1$, then $S_{i+1,1} = (x_{i+1} - y_{a[i+1]})^2$ and $S_{i+1,2} = (x_{i+1} - y_{a[i+1]-1})^2$
- As long as we put $a[i + 1] = a[i] + 1$, we just have $S_{i+1,1} = S_{i,1} + (x_{i+1} - y_{a[i+1]})^2$ and $S_{i+1,2} = S_{i,2} + (x_{i+1} - y_{a[i+1]-1})^2$

- In the last case (case 2), we need to re-compute almost entirely $S_{i+1,1}$ and $S_{i+1,2}$

Algorithm 3 Optimized Quadratic Partial Optimal Assignment

Input: sorted X, Y

Output: a

```

1: compute  $t$ 
2: simplify the problem
3:  $a[1] \leftarrow t[1]$ 
4:  $S_1 \leftarrow (x_1 - y_{a[1]})^2$ 
5:  $S_2 \leftarrow (x_1 - y_{a[1]-1})^2$ 
6: for  $i$  from 1 to  $m - 1$  do
7:   if  $t[i + 1] > a[i]$  then
8:      $a[i + 1] \leftarrow t[i + 1]$ 
9:     if  $a[i + 1] > a[i] + 1$  then
10:       $r \leftarrow i + 1$ 
11:       $S_1 \leftarrow (x_{i+1} - y_{a[i+1]})^2$ 
12:       $S_1 \leftarrow (x_{i+1} - y_{a[i+1]-1})^2$ 
13:    else ▷  $r$  does not change
14:       $S_1 \leftarrow S_1 + (x_{i+1} - y_{a[i+1]})^2$ 
15:       $S_2 \leftarrow S_2 + (x_{i+1} - y_{a[i+1]-1})^2$ 
16:    end if
17:  else
18:     $w_1 \leftarrow S_1 + (x_{i+1} - y_{a[i]+1})^2$ 
19:     $w_2 \leftarrow S_2 + (x_{i+1} - y_{a[i]})^2$ 
20:    if  $w_1 \leq w_2$  then ▷ case 1
21:       $a[i + 1] \leftarrow a[i] + 1$ 
22:       $S_1 \leftarrow S_1 + (x_{i+1} - y_{a[i+1]})^2$ 
23:       $S_2 \leftarrow S_2 + (x_{i+1} - y_{a[i+1]-1})^2$ 
24:    else ▷ case 2
25:       $a[i + 1] \leftarrow a[i]$ 
26:       $a[r : i] \leftarrow a[r : i] - 1$ 
27:      update  $r$ 
28:       $S_1 \leftarrow \sum_{k=r}^{i+1} (x_k - y_{a[k]})^2$ 
29:       $S_2 \leftarrow \sum_{k=r}^{i+1} (x_k - y_{a[k]-1})^2$ 
30:    end if
31:  end if
32: end for
33: return  $a$ 

```

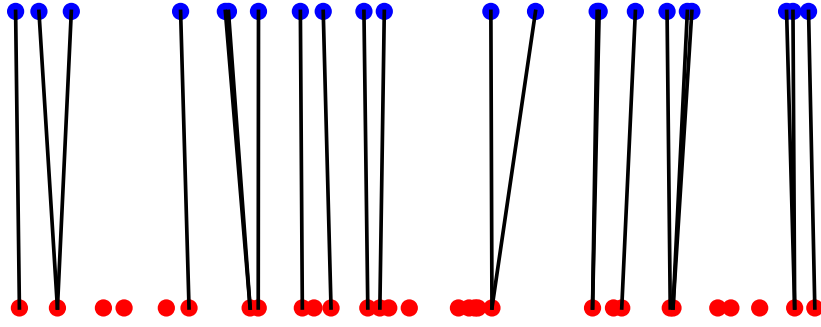


Figure 2: Nearest Neighbor Assignment

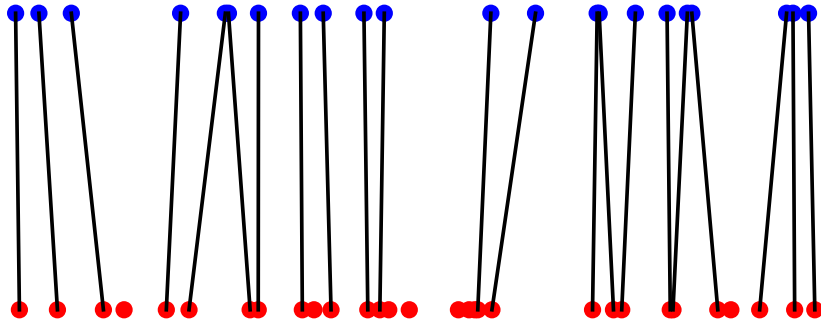


Figure 3: Partial Optimal Assignment

Surprisingly, after performances analysis, Numba accelerates loops and succeed in making the quadratic assignment linear with n for n not too big. This is better than expected, and problem decomposition may not be necessary for n smaller than one million. However it could still enables loop parallelization, and reduce complexity for bigger n .

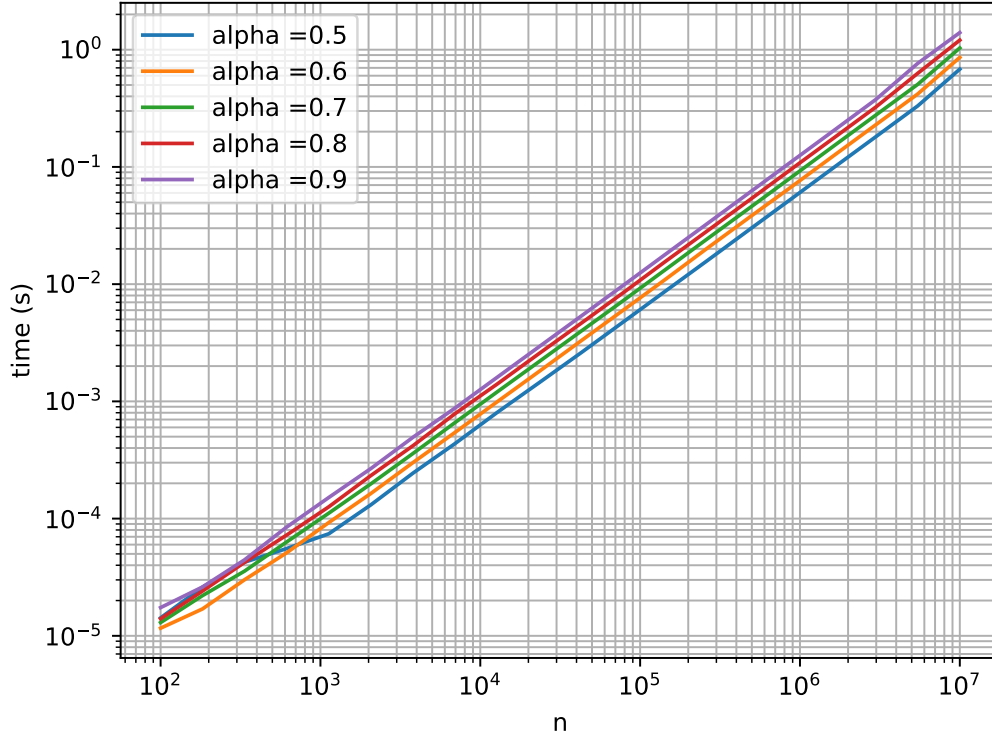


Figure 4: Mean time over 100 simulations for Quadratic Partial Optimal Assignment with respect to n , for different values of α , logscale

1.3 Quasilinear Time Problem Decomposition

In order to save more computational time, and decrease the complexity of the algorithm Bonneel et al. [1] proposed an algorithm to decompose the original problem into many easier subproblems, which can be solved independently with the previous partial optimal assignment algorithm. By doing this, it is possible to parallelize each subproblem solving and gain a consequent amount of time. However, Numba [?] the Python library for fast calculation I used, encounters problems with efficient parallelization and I wasn't able to parallelize problem solving after subproblem decomposition. An efficient implementation would require either to debug the Numba library, or to re-implement everything in a compiled language as C++.

The pseudo-algorithm provided by Bonneel et al. seemed to be false (or full of typo), and easy examples where it doesn't work are easy to find. (For example on their Fig. 3, f_{s_2} and f_{s_3} are True so problems Y_2 and Y_3 are not supposed to merge). I hence took the liberty to rewrite it differently.

The main intuition to the decomposition algorithm is that when considering a new point assignment, we need to either create a new subproblem (if $t[i+1] > a[i]$), or increase the size of the current problem. When increasing the size of the current problem, it is important to note that we just need to consider one more point on the right (case where $a[i+1] = a[i] + 1$) and one more point on the left (case where $a[i+1] = a[i]$ and translation).

Specifically, we again consider an increasing sub-problem involving $\{x_i\}_{i \in \llbracket 1; m' \rrbracket}$ and keep track of the last available free spot s as well as the currently last value l considered by each subproblem. The sub-problem A_k thus involves $\{y_j\}_{j \in \llbracket s_k+1; l_k \rrbracket}$. We also store a flag $\{f_j\}_{j \in \llbracket 1; n \rrbracket}$ where $f_j = \text{false}$ indicates that y_j has been considered by a subproblem at some point in the algorithm, or $f_j = \text{true}$ otherwise.

Algorithm 4 Decomposition of the assignment problem

Input: sorted X, Y

Output: A

```

1: compute  $t$ 
2: initialize Boolean flags  $\{f_j\}_{j \in \llbracket 1; n \rrbracket}$  to true;
3: for  $m'$  from 1 to  $m$  do
4:   if  $f_{t[m']}$  is True then ▷ create new subproblem
5:     create new subproblem  $A_k = (\{x_{m'}\}, \{y_{t[m']}\})$ 
6:      $f_{t[m']} \leftarrow \text{False}$ 
7:   else
8:     Consider  $A_{k'}$  last subproblem
9:     if  $t[m'] \neq t[m' - 1]$  then ▷ expand on the right only
10:       $A_{k'} \leftarrow (X_{k'} \cup \{x_{m'}\}, Y_{k'} \cup \{y_{t_{k'}+1}\})$ 
11:    else ▷ expand on the right and on the left
12:      while  $f_{s_{k'}}$  is false do ▷ merge with previous subproblems
13:        Consider  $A_{k'-1}$  subproblem before  $A_{k'}$ 
14:         $A_{k'-1} \leftarrow A_{k'} \cup A_{k'-1} = (X_{k'-1} \cup X_{k'}, Y_{k'-1} \cup Y_{k'})$ 
15:         $k' \leftarrow k' - 1$ 
16:      end while
17:       $A_{k'} \leftarrow (X_{k'} \cup \{x_{m'}\}, Y_{k'} \cup \{y_{s_{k'}}, y_{t_{k'}+1}\})$ 
18:    end if
19:  end if
20: end for
21: return  $A$ 

```

It is important to note that the algorithm use the fact that X and Y are sorted, and is hence made in such a way that consecutive subproblems A_k and A_{k+1} consider consecutive values of X and increasing values of Y , which means that $l_k \leq s_{k+1}$.

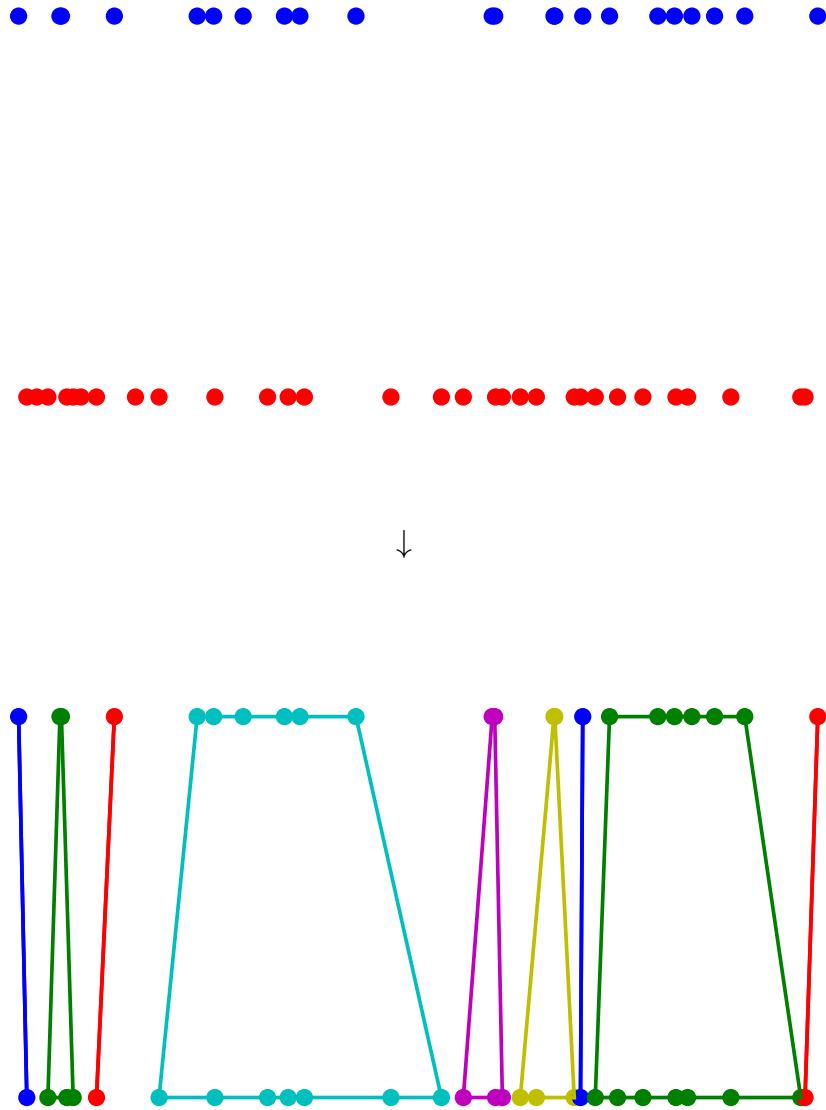


Figure 5: Assignment problem decomposition

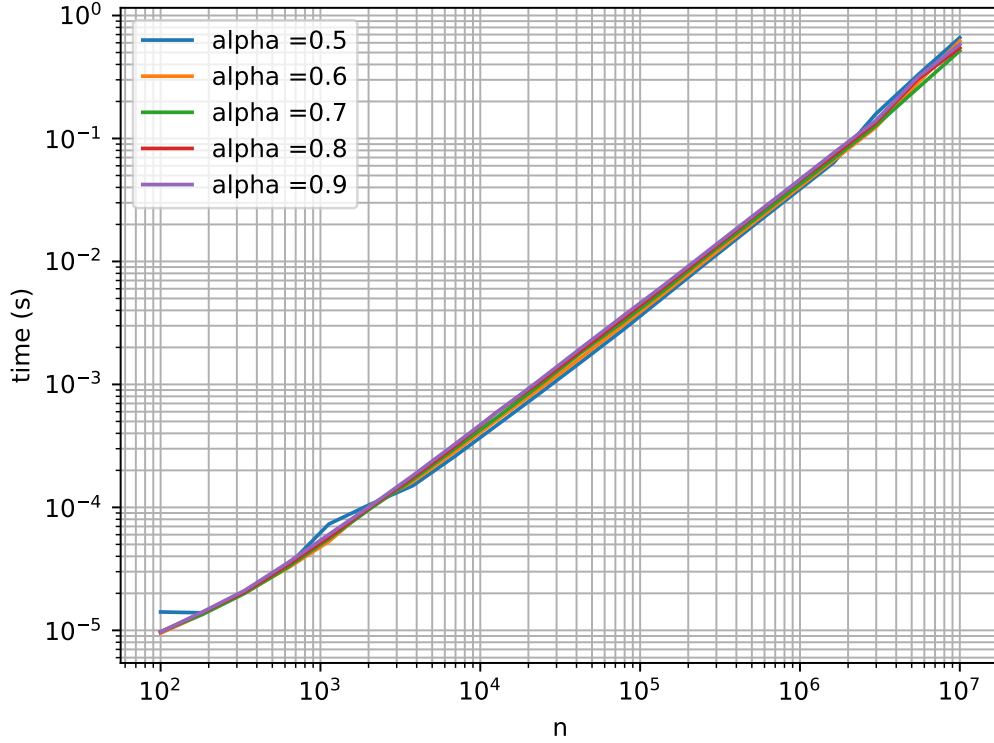


Figure 6: Mean time over 100 simulations for the decomposition of the assignment problem with respect to n , for different values of α , logscale

METTRE IMAGE TAILLE DECOMPOSITION

2 Sliced Partial Transport

While 1-d optimal transport has limited interest per se, it is the main ingredient of sliced optimal transport. Sliced optimal transport is a very fast algorithm that shares similar properties with optimal transport and works in n dimensions.

The partial 1-d optimal transport solution provided in the last section, while of quadratic complexity in the worst case, is fast (see Fig. 1, 4, ??) and is thus amenable to the computation of sliced optimal transport.

In the sliced framework, instead of minimizing directly the n -dimensional Wasserstein distance, we minimize the mean over all one-dimensional subspaces of the one-dimensional Wasserstein distance between the one-dimensional projections of the point sets :

$$\min_{\tilde{X}} \int_{S^{d-1}} W_S(\text{Proj}_{\omega}(\tilde{X}), \text{Proj}_{\omega}(Y)) d\omega \quad (1)$$

where S^{d-1} is the $(d-1)$ -dimensional sphere of directions.

To do so, we use a classical gradient descent with mini-batch size of one. We start by initializing $X_0 = X$. Then, at each iteration k , we choose a random direction ω and compute the gradient of $E_{\omega}(X_k) = W_S(\text{Proj}_{\omega}(X_k), \text{Proj}_{\omega}(Y))$ which is :

$$\nabla_X E_\omega(X_k) = Proj_\omega(X_k) - Proj_\omega(Y \circ a) \quad (2)$$

where a is the optimal partial assignment between $Proj_\omega(\tilde{X})$ and $Proj_\omega(Y)$.

We update X_k with the step:

$$X_{k+1} = X_k - \eta_k \nabla_X E_\omega(X_k)$$

With $\eta_k > 0$ the learning rate. Formally, in order to be sure of the algorithm convergence, η_k must be chosen such as $\sum_{k=0}^{+\infty} \eta_k = +\infty$ and $\sum_{k=0}^{+\infty} \eta_k^2 < +\infty$. In fact, this algorithm works well with a constant learning rate in some cases.

At the end of the algorithm, we get X^* an approximation of the optimal partial assignment of X to Y .

3 Results

3.1 Color Transfer

3.2 Point Cloud Registration

Conclusion

Appendix

3.3 Proof of the optimality of algorithm 2

By mathematical induction:

Let $n \in \mathbb{N}$. Let t be the nearest neighbor assignment from X to Y . Let $P(m)$ be the property : algorithm 2 gives the optimal assignment for $card(X) = m$ and $card(Y) = n$.

Initial step: It is trivial for $m=1$

Inductive Step: Assume that $P(m)$ is true. We want to show that $P(m+1)$ is true.

Tanks to $P(m)$, we have $a_m : \llbracket 1; m \rrbracket \rightarrow \llbracket 1; n \rrbracket$ the optimal partial assignment from $\{x_i\}_{i \in \llbracket 1; m \rrbracket}$ to Y .

If $a_m[m] < t[m+1]$, then a_{m+1} such that $a_{m+1}(i) = a_m(i)$ if $i \leq m$ and $a_{m+1}(m+1) = t(m+1)$ is clearly optimal.

If $t[m+1] \leq a_m[m]$, then, x_{m+1} cannot be optimally assigned to any y_l with $l > a_m[m] + 1$. Indeed, as the nearest neighbor of x_{m+1} is (strictly) before $y_{a_m[m]+1}$, any other assignment of x_{m+1} would have a higher cost.

Moreover, an optimal 1-D assignment a has no intersections. That means that there is no $x_i < x_j$ such that $y_{a[j]} < y_{a[i]}$. Therefore x_{m+1} cannot be assigned to y_l with $l < a_m[m]$, and if x_{m+1} is assigned to $y_{a_m[m]}$, then $a_{m+1}[r : m] = a_m[r : m] - 1$.

As a consequence, is only two cases left, which are the case one and the case 2 of the algorithm. The optimal partial assignment a_{m+1} is that which has the smaller cost between case 1 and case 2. Therefore a_{m+1} as described by the algorithm is optimal.

Conclusion: $P(m)$ is true for $m \in \llbracket 1; n \rrbracket$. Therefore algorithm 2 is optimal for all $n \in \mathbb{N}$ and for all $m \in \llbracket 1; n \rrbracket$.

References

- [1] Nicolas Bonneel and David Coeurjolly. Spot: Sliced partial optimal transport. *ACM Transactions on Graphics (SIGGRAPH)*, 38(4), 2019.
- [2] Francois Pitie, Anil C Kokaram, and Rozenn Dahyot. N-dimensional probability density function transfer and its application to color transfer. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1*, volume 2, pages 1434–1439. IEEE, 2005.
- [3] Julien Rabin, Gabriel Peyré, Julie Delon, and Marc Bernot. Wasserstein barycenter and its application to texture mixing. In *International Conference on Scale Space and Variational Methods in Computer Vision*, pages 435–446. Springer, 2011.
- [4] Nicolas Bonneel, Julien Rabin, Gabriel Peyré, and Hanspeter Pfister. Sliced and radon wasserstein barycenters of measures. *Journal of Mathematical Imaging and Vision*, 51(1):22–45, 2015.
- [5] Jean-David Benamou. Numerical resolution of an “unbalanced” mass transport problem. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 37(5):851–868, 2003.
- [6] Alessio Figalli. The optimal partial transport problem. *Archive for rational mechanics and analysis*, 195(2):533–560, 2010.
- [7] Nicolas Bonneel, Kalyan Sunkavalli, Sylvain Paris, and Hanspeter Pfister. Example-based video color grading. *ACM Trans. Graph.*, 32(4):39–1, 2013.
- [8] François Pitié, Anil C Kokaram, and Rozenn Dahyot. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding*, 107(1-2):123–137, 2007.
- [9] Julien Rabin, Julie Delon, and Yann Gousseau. Regularization of transportation maps for color and contrast transfer. In *2010 IEEE International Conference on Image Processing*, pages 1933–1936. IEEE, 2010.