

Revealing Smooth Structure of Visual Data by Permutation on Manifolds

Yi-Lei Chen

<https://sites.google.com/site/fallcolor/>

Chiou-Ting Hsu

<http://www.cs.nthu.edu.tw/~cthsu/candy.html>

Department of Computer Science,

National Tsing Hua University, Taiwan

Motivation

The ever-increasing scale of ubiquitous visual data inevitably raises new challenges in data organization. Although the sources of data collection are unpredictable, a collection of visual data usually shares similar topics and most of them are created for some special events. Some pioneer work, called photo sequencing, has been proposed to discover a set of still images for estimating the temporal order of a dynamic scene [1, 2] or the spatial order of an instantaneous event [3]. However, these methods rely on different hypotheses when capturing images, such as similar view points or ring-like camera arrangement, and therefore limit their applicability to general datasets.

In this paper, we address the issue of visual data organization by recovering an intrinsic order from an unorganized dataset. The proposed method exploits the inherent nature of manifold. This new perspective, posing no hypothesis on local topology of observed data, is simply built on the smoothness prior of manifold geometry. Under the observation that strong relation exists among visual content, we assume a visual dataset lies on a manifold and thus changes smoothly from point to point. By exploiting the linearity within nearby data points, our goal becomes to visit all of the data points along a manifold-guided order and to characterize the specific manifold's shape.

Methodology

The main idea of our method is to *construct a template of smooth manifold* and then *discover the optimal permutation between data points to fit this smooth shape*. As shown in Fig. 1, we use a pseudo graph to build the template and define the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathbf{W}$ by

$$w_{ij} = \exp\left(-\frac{(i-j)^2}{\eta_g^2}\right) \text{ and } d_{ij} = \begin{cases} \sum_k w_{ik}, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \quad (1)$$

The edge weight w_{ij} is simply measured according to the adjacency between indices to ensure the well-ordered smoothness, and η_g roughly controls the size of effective neighbourhood. Given N high-dimensional data points $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N]$, we attempt to uncover the intrinsic data order $i \mapsto \phi(i)$ under a guidance of smooth manifold geometry:

$$\begin{aligned} \min \mathbf{f}(\mathbf{P}) &= \min \sum_{1 \leq i, j \leq N} \|\mathbf{x}_{\phi(i)} - \mathbf{x}_{\phi(j)}\|_2^2 w_{\phi(i)\phi(j)} \\ &= \min \text{tr}(\mathbf{XPLP}^T \mathbf{X}^T), \end{aligned} \quad (2)$$

where \mathbf{L} is the Laplacian matrix derived from a pseudo graph, and $\mathbf{P} \in \Pi$ is a permutation matrix satisfying

$$\Pi = \{\mathbf{P} | p_{ij} \in \{0,1\}, \sum_{j=1}^N p_{ij} = 1, \forall i, \sum_{i=1}^N p_{ij} = 1, \forall j\}. \quad (3)$$

Equation (2) is an NP-hard combinatorial optimization problem and remains very challenging without off-the-shelf solvers. We thus propose a novel algorithm named *Permutation on Manifolds* (PoM) to estimate \mathbf{P} in polynomial time. Specifically, we use variable splitting technique to derive

$$\min \text{tr}(\mathbf{X}_p \mathbf{L} \mathbf{X}_p^T) \quad \text{s.t. } \mathbf{X}_p = \mathbf{X} \mathbf{P}, \quad (4)$$

Then we follow augmented Lagrange multiplier (ALM) method to solve the constrained optimization problem. The core of our algorithm boils down to one of the sub-problems with the following form:

$$\mathbf{P}^{(t+1)} = \arg \min_{\mathbf{P}} \|\mathbf{X} \mathbf{P} - \mathbf{Z}\|_F^2 \quad \text{s.t. } \mathbf{P} \in \Pi, \quad (5)$$

After we rephrase Equation (5) by a sequence of additive costs, our method can approach local optima efficiently with three guidelines:

- Relax binary integer constraints into nonnegative constraints;
- Allow each data point \mathbf{x}_i a limited number of moves to its k nearest data points among $\mathbf{z}_1, \dots, \mathbf{z}_N$; and
- Determine the best k , as small as possible, by binary search so that every \mathbf{z}_i can be accessed by $\mathbf{x}_1, \dots, \mathbf{x}_N$ at least once.

Note that we further study the potential of extending PoM algorithm to different applications rather than visual data organization. For example, we can seamlessly embed PoM into linear dimension reduction

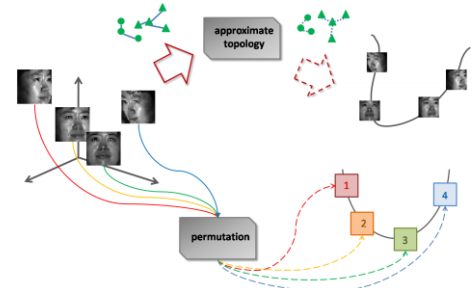


Fig. 1. Illustration of our proposed idea. (a) Data points in the original high-dimensional space; (b) the low-dimensional manifold approximated by classic methods; and (c) the smooth manifold approximated by permutation.

methods when the data affinity matrix is unavailable. In addition, including PoM into unsupervised manifold priors as a regularization term is also studied in our paper.

Results

We evaluate the proposed PoM algorithm with three experiments:

- We sample 300 data points from the Swiss roll dataset and justify our hypothesis by linking the data points along our estimated order (see Fig. 2).
- We embed PoM into linear dimension reduction and one of existing tensor completion method [4] to show the ability of PoM for regularization as manifold priors.
- We use two image sets “Basketball” and “Slide” released from [1] to verify that PoM is able to approximate a perceptually reasonable order from a given image set, which somewhat reveals the spatial order of cameras.

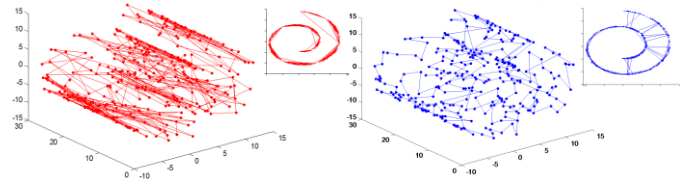


Fig. 2. The data-linking results of the 300 data points sampled from a Swiss roll obtained by ground truth labels (left) and PoM (right).

Conclusion

We summarize our contributions from three aspects: 1) to the best of our knowledge, this is the first work that leverages smooth manifold geometry to organize visual data; 2) the proposed pseudo graph generalizes the use of spectral graph theory and bypasses the measurement of data affinity in traditional methods; and 3) from our experiments, the proposed PoM algorithm shows its potential to serve as a core technique of numerous applications.

- [1] T. Basha, Y. Moses, and S. Avidan. Photo Sequencing. In *Proc. ECCV*, 2012.
- [2] T. Basha, Y. Moses, and S. Avidan. Space-Time Tradeoffs in Photo Sequencing. In *Proc. ICCV*, 2013.
- [3] H. Averbuch-Elor and D. Cohen-Or. RingIt: Ring-ordering Casual Photos of a Temporal Event. In *Proc. SIGGRAPH*, 2015.
- [4] Y. L. Chen, C. T. Hsu, and H. Y. Mark Laio. Simultaneous Tensor Decomposition and Completion Using Factor Priors. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 36(3): 577-591, 2014.