LEARNING DISCRIMINANT GRASSMANN KERNELS FOR IMAGE-SET CLASSIFICATION

IEEE Signal Processing Society

Lei Zhang[1,2], Wenhui Liu[2], Xuezhi Xiang[2] , Xiantong Zhen[3]

1. Guangdong University of Petrochemical Technology 2. Harbin Engineering University, China. 3. The University of Texas at Arlington.

INTRODUCTION

- *Image-set classification has been widely researched in computer vision due to its widespread applications*
- *Image-set classification can preferably handle the conditions with multi-view cameras or larger within-class divergence tasks.*
- *Learning discriminant Grassmann kernels (DGK) for image-set classification is proposed*
- *Based on partial kernels of principal angels, the global kernels between image sets aggregate the partial kernels, which are learned by kernel alignment in a supervised learning framework*

GRASSMANN MANIFOLDS

Linear subspaces on Grassmann manifolds

Defination 1 *[10] The Grassmann manifold $\mathcal{G}(m, D)$ is the set of m-dimensional linear subspaces of the $\mathbb{R}^D$.*

✓*The matrix representation is constructed based on orthogonalization and normalization, which guarantees its uniqueness.*

Principal Angles

Defination 2 *[10] Principal angles $0 \leq \theta_1 \leq \ldots \leq \theta_m \leq \pi/2$ between two subspaces M_1 and M_2 in Grassmann manifold are defined recursively by*

$$\cos\theta_k = \max_{\mathbf{u}_k} \max_{\mathbf{v}_k} \mathbf{u}_k^\top \mathbf{v}_k$$

✓ *In practice, for the matrix representations M1 and M2 , the principal angles can be computed by SVD*

✓ $M_1^\top M_2 = U(\cos\Theta)V^\top$

where $\cos\Theta = \mathrm{diag}(\cos\theta_1, \ldots, \cos\theta_m)$.

GRASSMANN KERNEL LEARNING

Partial Grassmann kernels

✓ *Two points: **zi** and **zj***

✓ *Defining the partial Grassmann kernel* $K^p \in \mathbb{R}^{N\times N}$ *with respect to the p-th principal angle as follows:*

$$[K^p]_{ij} = k^p(\mathbf{z}_i, \mathbf{z}_j)$$

Where

$$k^p(z_i, z_j) = \cos^2\theta_p$$

✓ *Since* K^p *only reflects partial relations between two points based on the p-th principal angles, we refer it as the **partial Grassmann kernel***

Discriminant learning by kernel alignment

✓ *Equal weight combination*

$$[K]_{ij} = \sum_{p=1}^{P} k^p$$

✓*Discriminant learning by kernel alignment*

$$[K_w]_{ij} = \sum_{p=1}^{P} w_p k^p$$

•*Objective function*

$$\max_{\mathbf{w}} \rho(\bar{K}_w, \mathbf{y}\mathbf{y}^\top) = \max_{\mathbf{w}} \frac{<\bar{K}_w, \mathbf{y}\mathbf{y}^\top>_F}{\|\bar{K}_w\|}$$

$$[\bar{K}_w]_{ij} = [K_w]_{ij} - \frac{1}{m}\sum_{i=1}^{m}[K_w]_{ij} - \frac{1}{m}\sum_{j=1}^{m}[K_w]_{ij} + \frac{1}{m^2}\sum_{i,j=1}^{m}[K_w]_{ij}$$

$$\Rightarrow \min_{\mathbf{v}\geq 0} \mathbf{v}^\top M\mathbf{v} - 2\mathbf{v}^\top\mathbf{a} \Rightarrow \mathbf{w}^* = \mathbf{v}/\|\mathbf{v}\|.$$

$$\begin{cases} [M]_{kl} = \langle \bar{K}_k, \bar{K}_l\rangle_F. \\ \mathbf{a} = (\langle \bar{K}_1, \mathbf{y}\mathbf{y}^\top\rangle_F, \ldots, \langle \bar{K}_m, \mathbf{y}\mathbf{y}^\top\rangle_F)^\top \end{cases}$$

EXPERIMENTS AND RESULTS

Corpus
•ETH dataset
•USCD traffic dataset

Set-based object recognition on ETH dataset

✓ *As can be seen, the performance has been largely improved from 94.0% (equal weight combination) to 95.5% by the proposed discriminant learning (DGK).*

✓ *The proposed DGK achieves state-of-the-art performance which is better than most of the compared methods in Table 1.*

Discriminant Grassmann kernels (DGK)	**95.5%**
Grassmann kernels (Eq. (5))	**94.0%**
Kernel Fisher Discriminant (KFD) [21]	81.1%
Marginal Fisher Analysis (MFA) [22]	80.1%
Manifold-Manifold Distance (MMD) [23]	85.0%
Mutual Subspace Method (MSM) [24]	83.3%
Manifold Discriminant Analysis (MDA) [25]	89.0%
Discriminant Canonical Correlations (DCC) [6]	91.7%
Log-Euclidean metric learning (LEML) [8]	94.8%
Graph embedding discriminant analysis (GEDA) [11]	92.3%
Localized multi-kernel metric learning (LMKML) [26]	94.5%

Table 1. Performance comparison on the ETH dataset

Video-based traffic congestion classification on USCD

✓ *The proposed DGK has largely improved the Grassmann kernels from 90.5% (equal weight combination) to 92.1% (DGK), which demonstrates the great effectiveness of the proposed supervised learning framework via kernel alignment for image-set classification.*

✓ *The proposed DGK achieves state-of-the-art performance which is better than most of the compared methods in Table 2.*

Discriminant Grassmann kernel (DGK)	**92.1%**
Grassmann kernels (Eq. (5))	**90.5%**
Linear dynamical system (LDS) [27]	87.5%
Compressive sensing LDS (CS-LDS) [27]	89.1%
Grassmann discriminant analysis (GDA) [10]	92.5%
Covariance discriminative learning (CDL)[7]	91.7%
NN classifier on Hellinger distance [5]	91.3%
NN classifier on J-divergence [5]	91.0%
Discriminant analytic stationary subspace analysis (DASSA) [28]	91.7%
Discriminant non-linear stationary subspace analysis (DNLSSA+RBF kernel) [28]	94.5%

Table 2. Performance comparison on the UCSD dataset.

Conclusion

- *Presented discriminant Grassmann kernel (DGK) learning for image-set classification, which learned the weight of partial Grassmann kernel by kernel alignment with target kernel constructed by labels*
- *Evaluated the DGK on the ETH dataset and traffic congestion classification on the UCSD dataset*

References

[5] "Beyond gauss: Image-set matching on the riemannian manifold of pdfs," in ICCV, 2015
[6] "Discriminative learning and recognition of image set classes using canonical correlations," T-PAMI, 2007.
[7] "Covariance discriminative learning: a natural and efficient approach to image set classification," in CVPR, 2012
[8] "Log-euclidean matric learning on symmetric positive define manifold with application to image set classification," in ICML, 2015
[10] "Grassmann discriminant analysis: a unifying view on subspace-based learning," in ICML, 2008
[11] "Graph embedding discriminant analysis on grassmannian manifolds for improved image set matching," in CVPR, 2010
[21] "A mathematical programming approach to the kernel fisher algorithm," in NIPS, 2000.
[22] "Graph embedding and extensions: a general framework for dimensionality reduction," T-PAMI, 2007.
[23] "Manifold-manifold distance with application based on image set," in CVPR, 2008
[24] "Face recognition using temporal image sequence," in FG, 1998.
[25] "Manifold discriminant analysis," in CVPR, 2009
[26] "Image set classification using holistic multiple order statistics features and localized multi-kernel metric learning," in ICCV, 2013
[27] "Compressive acquisition of dynamic scenes," in Springer Berlin Heidelberg, 2013
[28] "Discriminative non-linear stationary subspace analysis for video classification," T-PAMI, 2014.