NONLINEAR SUBSPACE CLUSTERING

Wencheng Zhu[1,3], *Jiwen Lu*[1,2,3] and *Jie Zhou*[1,2,3]

[1]Department of Automation, Tsinghua University, Beijing, 100084, China.

[2]State Key Lab of Intelligent Technologies and Systems, Beijing, 100084, China

[3]Tsinghua National Laboratory for Information Science and Technology (TNList), Beijing, 100084, China.

INTRODUCTION

Subspace clustering has been one important visual analysis task and has many potential applications such as image and motion segmentation, face clustering and so on. The objective of subspace clustering is to partition samples into different subspaces and seek the multi-cluster structure of data.

METHODOLOGY

Given the data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{d \times n}$, the output of m^{th} layer is denoted as

$$\mathbf{h}_i^{(m)} = g\left(\mathbf{W}^m \mathbf{h}_i^{(m-1)} + \mathbf{b}^{(m)}\right) \in \mathbb{R}^{d_m}. \quad (1)$$

the output $\mathbf{H}^{(M)}$ of the top layer in the neural network is defined as:

$$\mathbf{H}^{(M)} = [\mathbf{h}_1^{(M)}, \mathbf{h}_2^{(M)}, \dots, \mathbf{h}_n^{(M)}]. \quad (2)$$

NSC transforms the data matrix $\mathbf{X}$ into a nonlinear space and then conducts the subspace clustering iteratively. The objective function $\mathbf{J}$ of NSC can be formulated as:

$$\min_{\{\mathbf{W}^{(m)}, \mathbf{b}^{(m)}\}_{m=1}^{M}, \mathbf{C}} \quad \mathbf{J} = \mathbf{J}_1 + \alpha \mathbf{J}_2 + \beta \mathbf{J}_3 \quad (3)$$

$\mathbf{J}_1$ is the loss term and guarantees the rebuilding ability of the self-representation matrix in the nonlinear space, which is defined as:

$$\mathbf{J}_1 = \frac{1}{2} \sum_{i=1}^{n} \left\| \mathbf{h}_i^{(M)} - \mathbf{H}^{(M)} \mathbf{c}_i \right\|_F^2 \quad (4)$$

$\mathbf{J}_2$ utilizes the grouping effect and the effectiveness of the grouping effect, which is formulated as:

$$\mathbf{J}_2 = \frac{1}{2} \mathrm{tr}(\mathbf{C}\mathbf{L}\mathbf{C}^T) \quad (5)$$

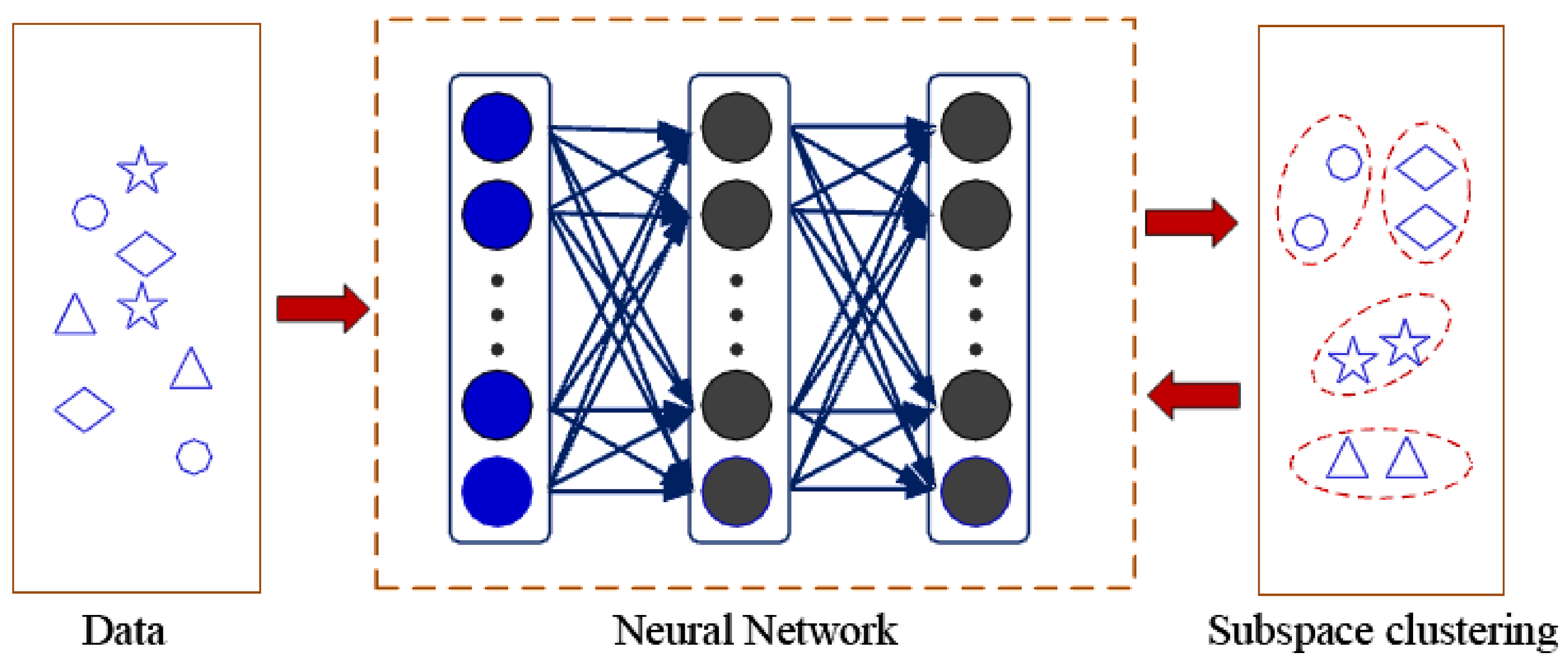
$\mathbf{J}_3$ is the regularization term and aims to avoid the model over-fitting, which is designed as:

$$\mathbf{J}_3 = \frac{1}{2} \sum_{m=1}^{M} \left(\|\mathbf{W}^{(m)}\|_F^2 + \|\mathbf{b}^{(m)}\|_F^2 \right) \quad (6)$$

The neural network can be updated by the following paradigm:

$$\begin{cases} \mathbf{W}^{(m)} = \mathbf{W}^{(m)} - \tau \dfrac{\partial \mathbf{J}}{\partial \mathbf{W}^{(m)}} \\ \mathbf{b}^{(m)} = \mathbf{b}^{(m)} - \tau \dfrac{\partial \mathbf{J}}{\partial \mathbf{b}^{(m)}} \end{cases} \quad (7)$$

The corresponding α and β are the positive parameters.



DATA & EVALUATION

The Extended Yale Face B

The USPS dataset

We employed two evaluation criteria including the clustering error (**CE**) and normalized mutual information (**NMI**) to evaluate the performances of different subspace clustering methods.

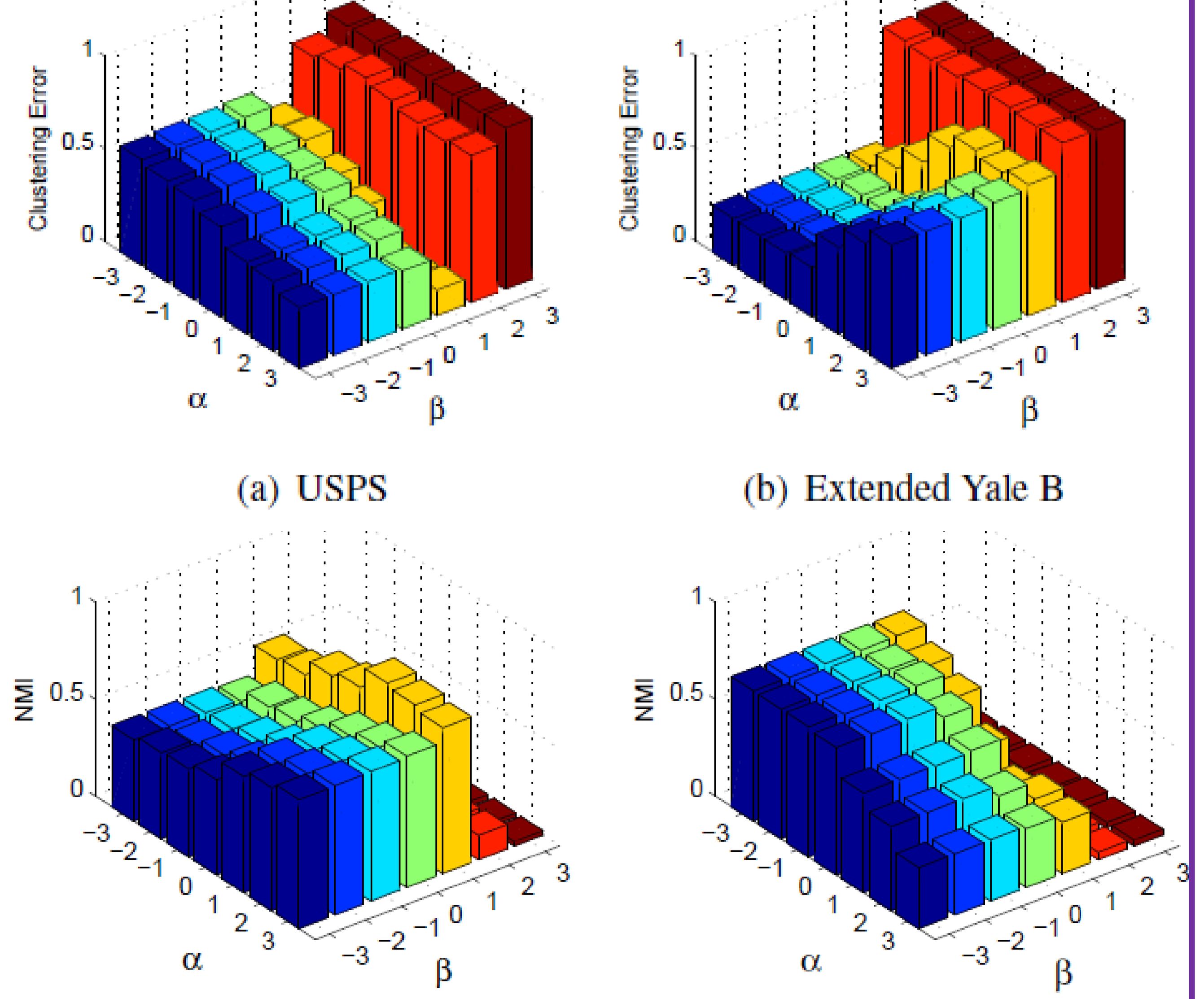
RESULT

Method	SSC	LRR	LSR1	LSR2	SMR	NSC
CE	33.1	38.6	30.3	26.9	26.1	**25.0**
NMI	58.4	54.5	57.8	65.3	66.1	**67.1**

The Extended Yale Face B

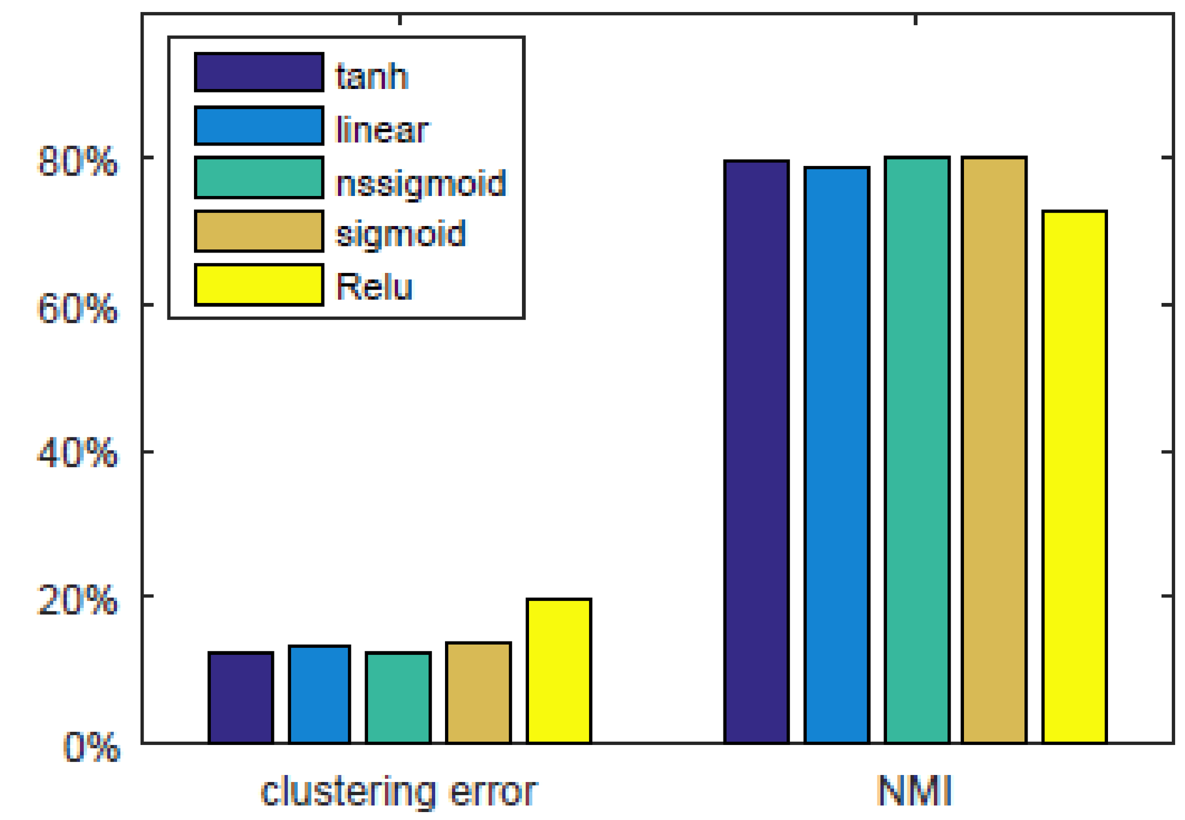
Method	SSC	LRR	LSR1	LSR2	SMR	NSC
CE	41.5	29.3	29.4	31.1	29.5	**12.4**
NMI	59.6	67.6	68.0	66.6	69.6	**79.0**

USPS



(a) USPS (b) Extended Yale B

(c) USPS (d) Extended Yale B



CONCLUSION

In this paper, we have proposed a nonlinear subspace clustering method (NSC) for image clustering. NSC simultaneously transforms the original feature space into a nonlinear space. Experimental results have clearly shown that our NSC achieve superior results than four state-of-the-art subspace clustering methods.

Email: **zwc17@mails.tsinghua.edu.cn**

lujiwen@tsinghua.edu.cn