Hong Shao, Shuang Chen, Jie-yi Zhao, Wen-cheng Cui, Tian-shu Yu, 2015. Face recognition based on subset selection via metric learning on manifold. *Frontiers of Information Technology & Electronic Engineering*, **16**(12):1046-1058. [doi:10.1631/FITEE.1500085]

Face recognition based on subset selection via metric learning on manifold

Key words: Face recognition, Sparse representation, Manifold structure, Metric learning, Subset selection

Contact: Shuang Chen
E-mail: chenshuang19891129@gmail.com
ORCID: http://orcid.org/0000-0001-7441-4749

Introduction

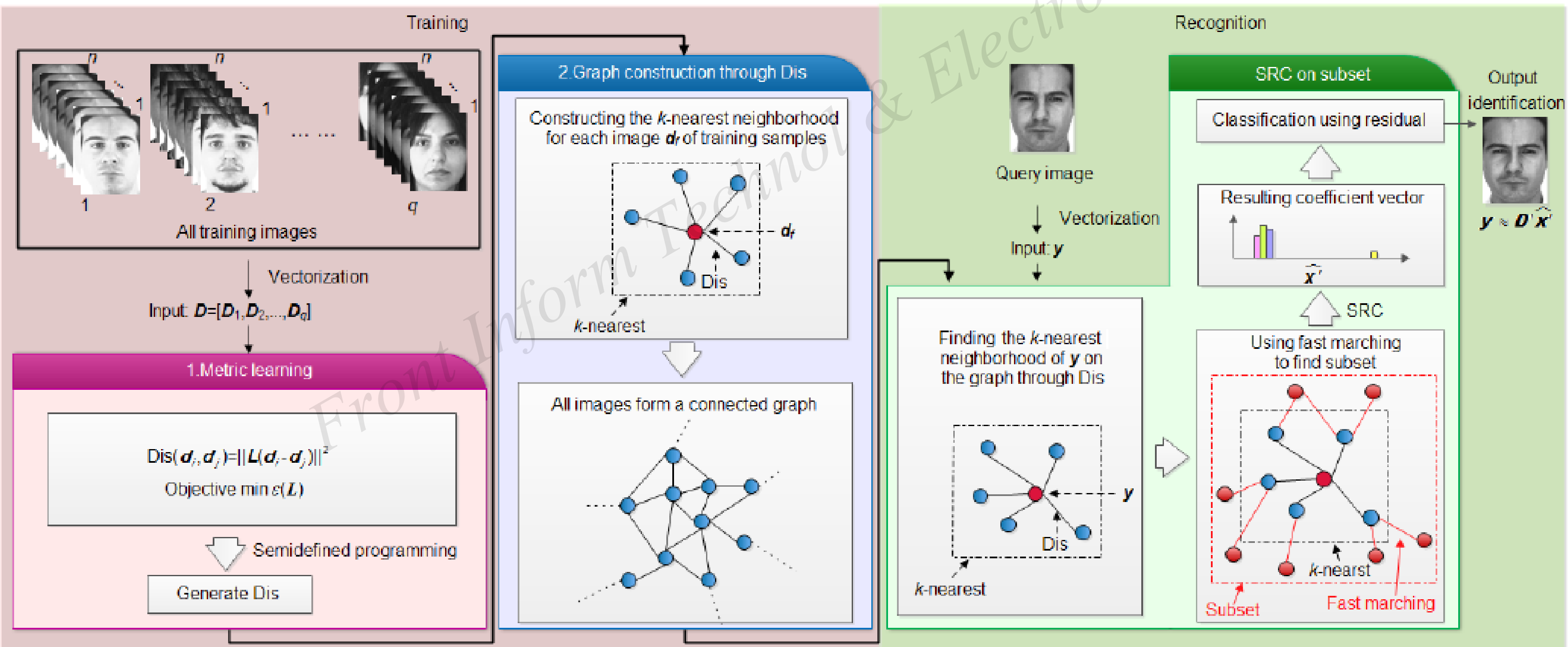
- In the face recognition based on SRC, rather than extracting complex features from the image, the image itself can be directly used as a vectorized form, under the assumption that the vectorized face image lies in a subspace sparsely spanned by other images from the same class.

- However, when the dictionary is large and the representation is sparse, only a small proportion of the elements contributes to the l^1-minimization.

- This paper presents a face recognition method based on subset selection via metric learning on manifold.

Framework of our method

- The flowchart of the proposed framework



Design method (I)

- Metric learning for face recognition
  - This is the first stage of the proposed method. According to the formula, it is to learn a robust metric in terms of a linear projection so that the interclass and intraclass relationship, as well as the manifold structure, can be preserved.

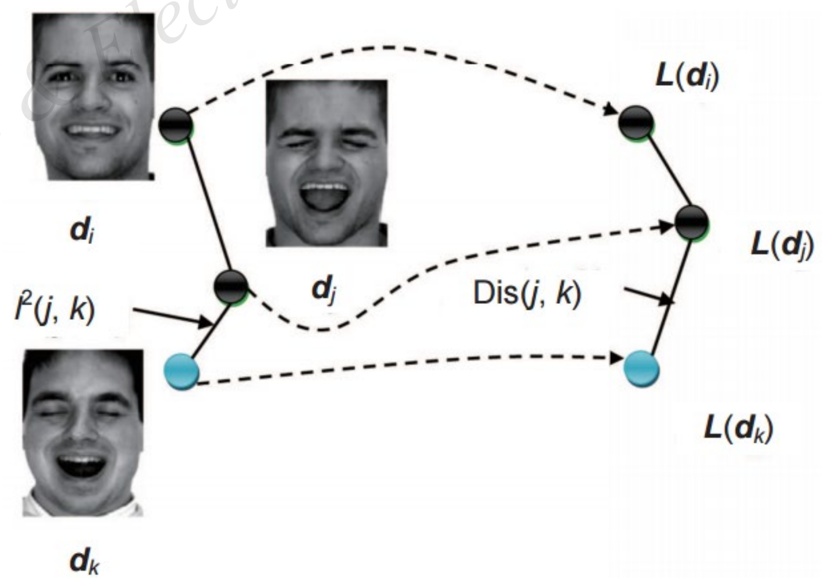
$$\mathrm{Dis}(\boldsymbol{d}_i, \boldsymbol{d}_j) = \|\boldsymbol{L}(\boldsymbol{d}_i - \boldsymbol{d}_j)\|^2$$

$$\min\ \varepsilon(\boldsymbol{L}) = \sum_{ij} \eta_{ij} \|\boldsymbol{L}(\boldsymbol{d}_i - \boldsymbol{d}_j)\|^2 + \lambda \sum_{ijl} \eta_{ij}(1 - c_{il}) \cdot [1 + \|\boldsymbol{L}(\boldsymbol{d}_i - \boldsymbol{d}_j)\|^2 - \|\boldsymbol{L}(\boldsymbol{d}_i - \boldsymbol{d}_l)\|^2]_+,$$

Design method (II)

- Graph construction using learned metric

➢ Using the metric Dis, the distance of face images from the same class can become closer. And we implement the *k*-nearest method to construct a weighted graph *G* = (*V*, *E*) on the whole training set ***D*,** and *k* is larger than the size of any training sample from the same class.



Design method (III)

- SRC on subset

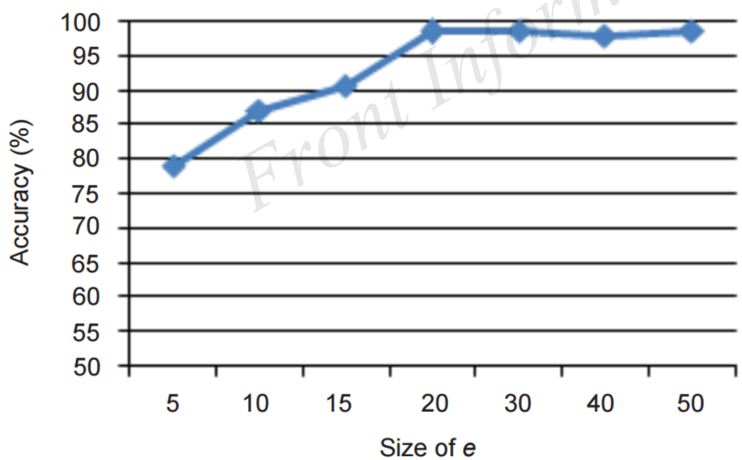
➢ Given the graph G, we seek to find the subset with size e along the coinciding manifolds, where e is the number of the images that are selected as the candidates to form the dictionary for the SRC.

➢ Along with the starting k images, the $e-k$ images with the shortest geodesic distances to the query image on G form the new dictionary $\boldsymbol{D}'$, assuming that the query image lies in the subspace spanned by the images from the new dictionary $\boldsymbol{D}'$.

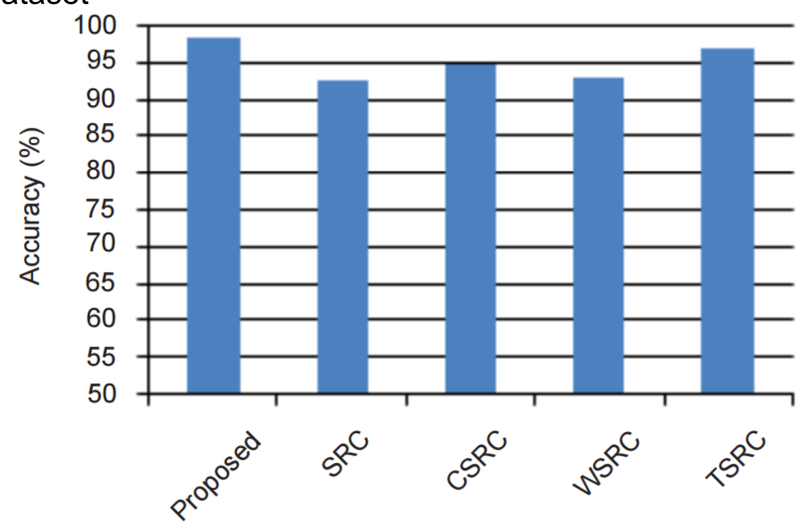
Major results

(1) Analysis of recognition performance on the AR database:

Fourteen face images with a variety of expressions and illuminations from a single person in the AR dataset



The performance of our method under varying *e* values

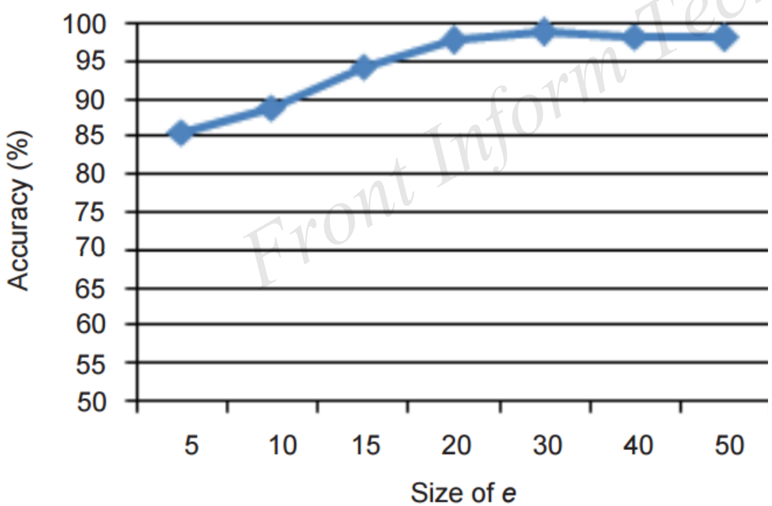


The comparison of the selected methods in accuracy

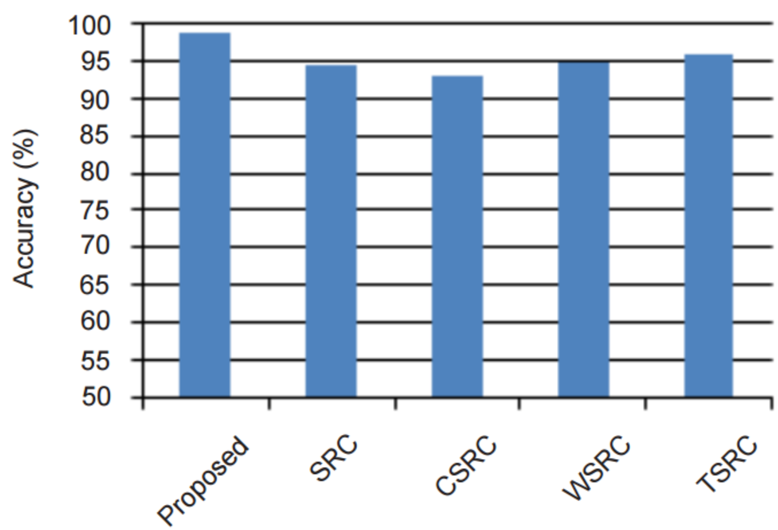
Major results (Con’d)

(2) Analysis of recognition performance on the extended Yale B database:

Face images from a single person in the extended Yale B database



The performance of our method under varying *e* values

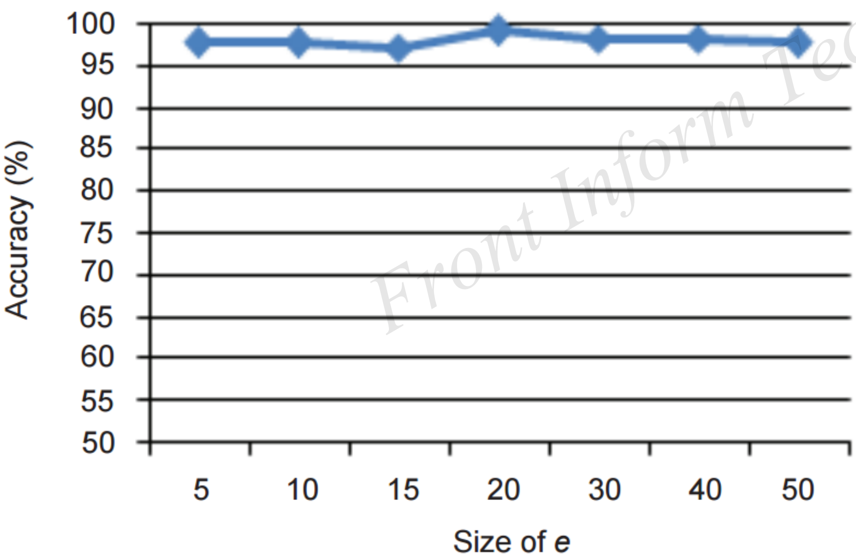


The comparison of the selected methods in accuracy

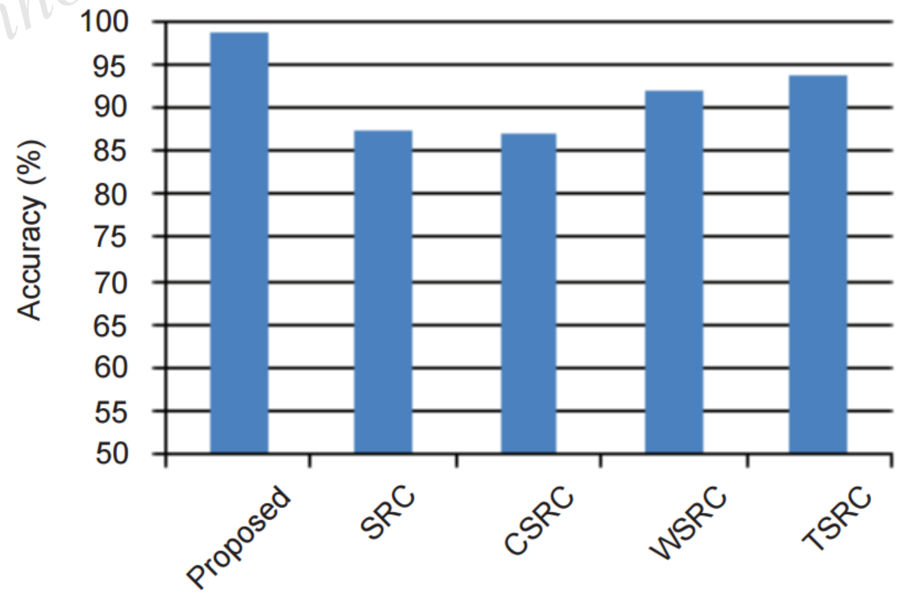
Major results (Con’d)

(3) Analysis of recognition performance on the ORL database:

Several face images of two classes from the ORL database



The performance of our method under varying *e* values

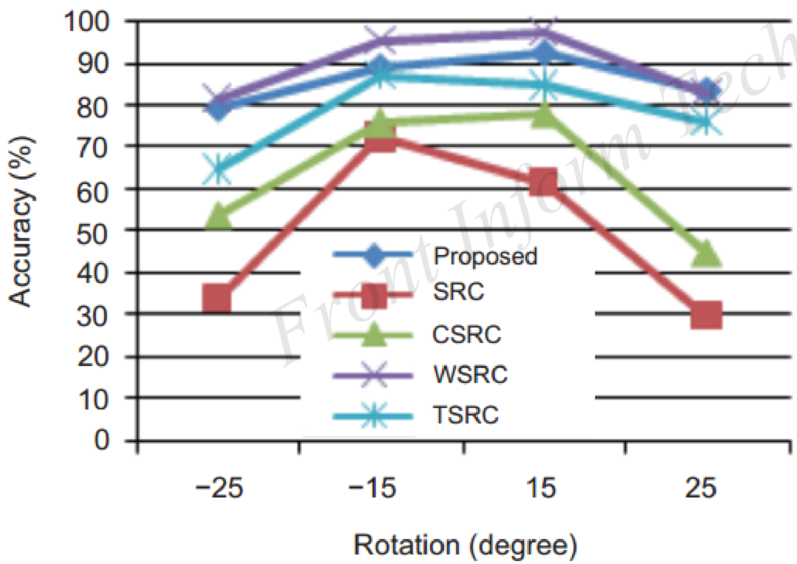


The comparison of the selected methods in accuracy

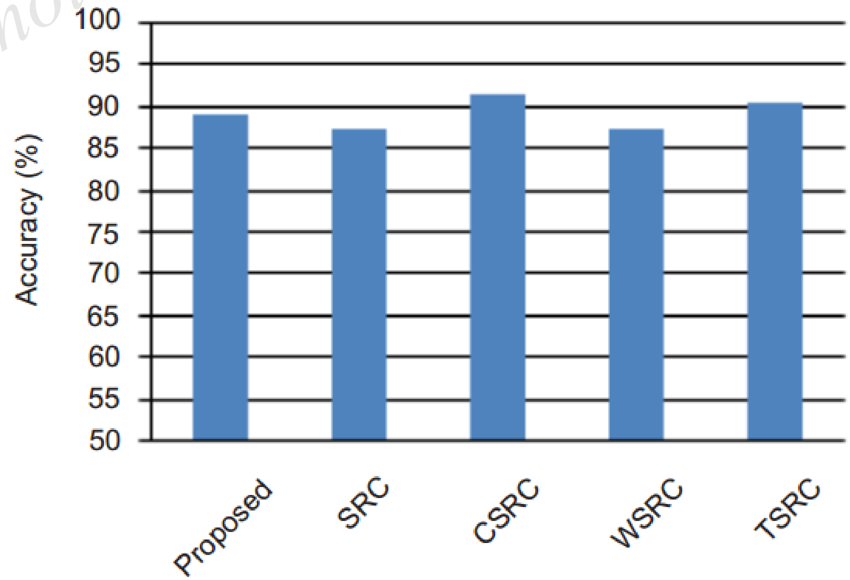
Major results (Con’d)

(4) Experimental results of five tests on the FERET database:

Seven face images of one class from the FERET database



The comparison of the selected methods in test 1

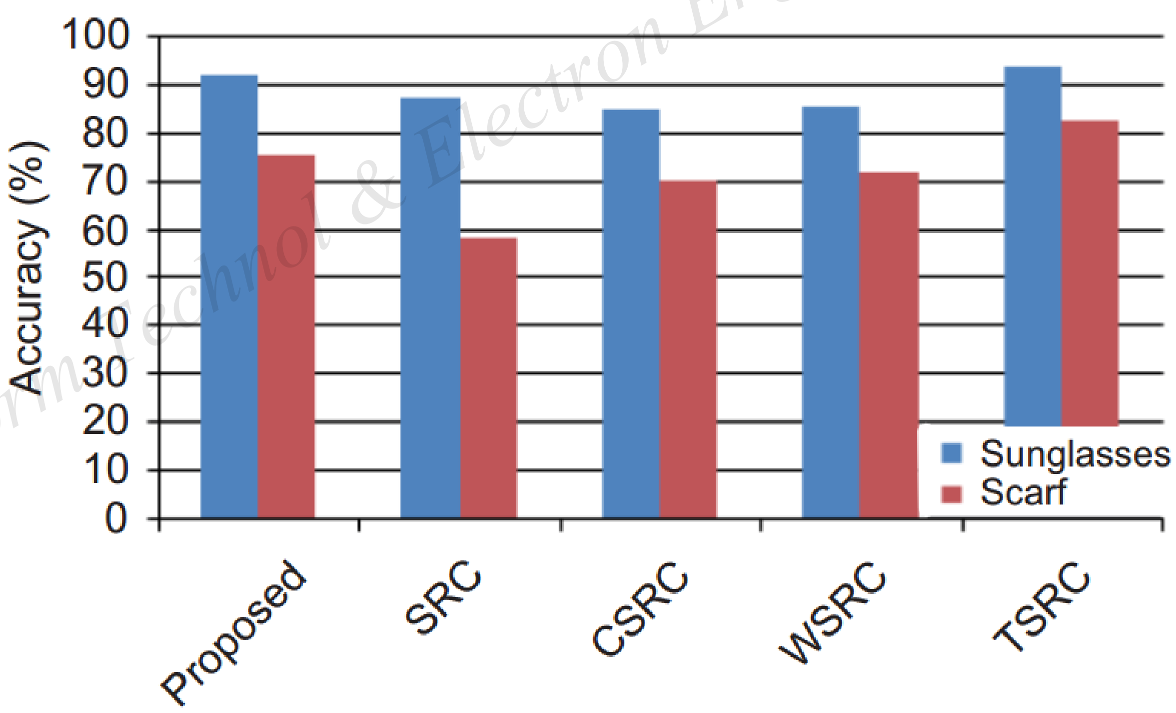


The performance of each selected method under various degrees of rotation (tests 2–5)

Major results (Con’d)

(5) Recognition under the sunglasses and scarf occlusion:

The images of sunglasses and scarf occlusion

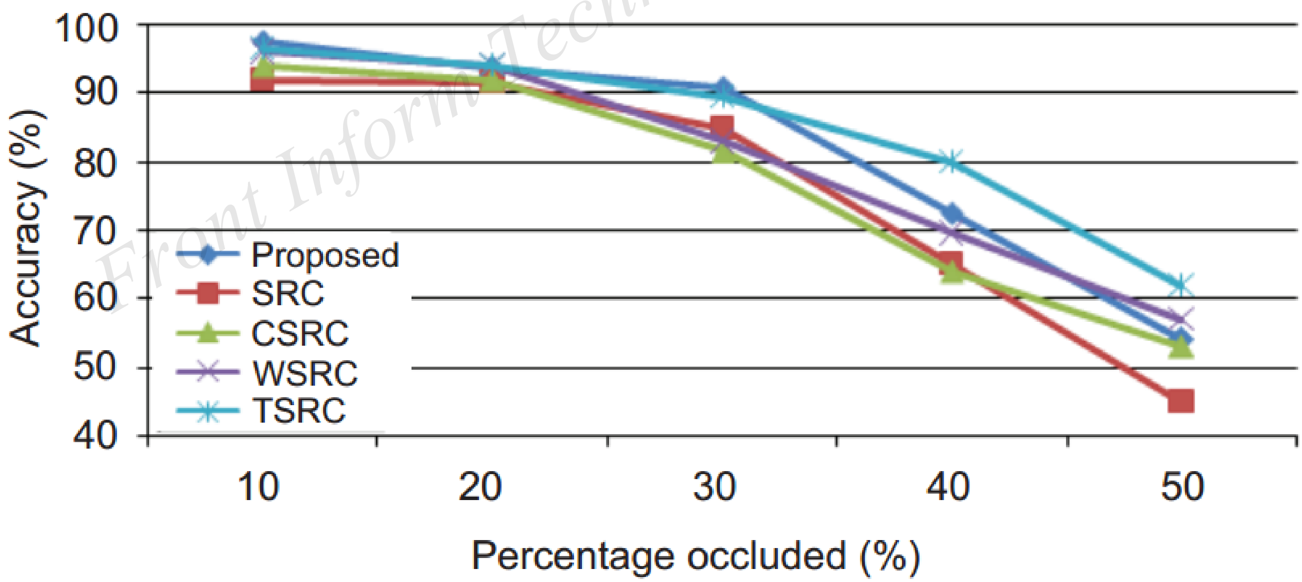


The comparison results for both cases

Major results (Con'd)

(6) Recognition under block occlusion:

The images of block occlusion



The performance of the proposed and the selected methods under various block occlusions

Major results (Con’d)

(7) Comparison of recognition speed:

Method	Recognition time (s)			
	AR	Yale B	ORL	FERET
Proposed				
Subset size: 5	0.207(+0.211)	0.134(+0.157)	0.073(+0.086)	0.110(+0.145)
Subset size: 10	0.219(+0.208)	0.138(+0.150)	0.082(+0.106)	0.117(+0.151)
Subset size: 20	0.220(+0.203)	0.145(+0.127)	0.102(+0.098)	0.118(+0.172)
Subset size: 30	0.227(+0.232)	0.141(+0.139)	0.108(+0.110)	0.129(+0.161)
Subset size: 50	0.239(+0.276)	0.158(+0.145)	0.110(+0.142)	0.134(+0.173)
Subset size: 100	0.349(+0.321)	0.192(+0.207)	0.171(+0.161)	0.187(+0.199)
Subset size: 150	0.401(+0.410)	0.219(+0.304)	0.203(+0.172)	0.220(+0.275)
SRC	78.405(+94.223)	41.089(+33.460)	11.900(+17.112)	28.142(+34.447)
CSRC	4.419(+7.889)	3.744(+3.412)	2.590(+3.077)	3.198(+4.106)
WSRC	4.018(+4.980)	3.146(+3.484)	2.183(+3.421)	2.871(+3.897)
TSRC	0.632(+0.703)	0.561(+0.475)	0.440(+0.344)	0.479(+0.543)

The number on each entry represents the average recognition time and the number in brackets is the extra time for maximum recognition cost

Conclusions

Rather than other face recognition methods based on SRC, this method is intuitive and easy to implement. According to the experimental results, because of integrating metric learning and subset selection, the proposed method achieves promising recognition performance with remarkable efficiency.