All our experiments are conducted on Windows10 computer with 4.5GHz Intel Core I9 10850K CPU and 32-GB RAM, MATLAB 2021a.

1. Data sets

In the experiments, the data sets used in the experiments are all public data sets. The details of these data sets are shown in Table 1.

Table 1: Data sets details

index	data set	samples	features	clusters
1	COIL20	1440	1024	20
2	MINIST2k	4000	784	10
3	MSRA25	1799	256	12
4	Mnist05	3495	784	10
5	PIE	11560	1024	68
6	USPS	9298	256	10

Among the 6 data sets, COIL20(Nene et al., 1996), MINIST2k(Xu et al.) are all image data sets. PIE(Gross et al., 2010) is a typical traffic scene recorded using vehicle-mounted cameras containing more than 6 hours of footage to study pedestrian behavior in traffic. MSRA25(Li et al., 2009) data set is a face data set consisting of 1799 images from 12 individuals of different genders. The USPS [1] data set contains 7291 training images and 2007 testing images. The image is 16*16 grayscale pixels.

References

Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S., 2010. Multi-pie. Image and vision computing 28, 807–813.

Li, H., Wang, M., Hua, X.S., 2009. Msra-mm 2.0: A large-scale web multimedia dataset, in: 2009 IEEE International Conference on Data Mining Workshops, IEEE. pp. 164–169.

Nene, S., Nayar, S., Murase, H., 1996. Columbia image object library (coil-20). Technical Report CUCS006-96. .

Xu, J., Han, J., Xiong, K., Nie, F., . Robust and sparse fuzzy k-means clustering, in: IJCAI, pp. 2224–2230.

---

[1] https://www.kaggle.com