6. Supplemental

A. The detailed derivation of Eq.(13) is as the following: We denote the derivative of Eq.(12) regarding to j as H,

1) When $c^* > 0$:

a. If $j \geq c^*$: $H = \beta + \alpha + j - y = 0$, we can get $j = y - \alpha - \beta$, and with the assumption $j \geq c^*$, y has to satisfy $y \geq \alpha + \beta + c^*$ to get this solution;

b. If $0 < j < c^*$: $H = \beta - \alpha + j - y = 0$, which leads to $j = y + \alpha - \beta$, and similarly due to the prerequisite $0 < j < c^*$, we should have $0 < y + \alpha - \beta < c^*$;

c. If $j \leq 0$: $H = -\beta - \alpha + j - y = 0$, we can get $j = y + \alpha + \beta$, so $y + \alpha + \beta \leq 0$ as well;

2) When $c^* < 0$:

a. If $j \geq 0$: $H = \beta + \alpha + j - y = 0$, we can get $j = y - \alpha - \beta$, and with the assumption $j \geq 0$, only when $y \geq \alpha + \beta$ we can get this solution;

b. If $c^* < j < 0$: $H = -\beta + \alpha + j - y = 0$, then we have $j = y - \alpha + \beta$, also $c^* < y - \alpha + \beta < 0$;

c. If $j \leq c^*$: $H = -\beta - \alpha + j - y = 0$, we can get $j = y + \alpha + \beta$, so when $y + \alpha + \beta \leq c^*$ we can get this solution.

B. The derivative details for Eq.(21) is as the following:

We suppose the optimal solution is $c^* = a_i$, and it's obvious that the subgradient of $|a_i - c^*|$ is any element in the interval of [-1, 1].

1) When $c^* > 0$:

To make the derivative of Eq.(20) regarding to c^* equal to 0, we should have

$$-1 \leq 2\lambda(i-1) - 2\lambda(v-i) + \gamma q \leq 1, \tag{23}$$

which leads to

$$\frac{2v\lambda - \gamma q}{4\lambda} + \frac{1}{2} - \frac{1}{4\lambda} \leq i \leq \frac{2v\lambda - \gamma q}{4\lambda} + \frac{1}{2} + \frac{1}{4\lambda}, \tag{24}$$

i has to be an integer as an index, so $\left\lceil \frac{2v\lambda - \gamma q}{4\lambda} \right\rceil$ is an appropriate value for it. Also, an index i has to be larger than 0, we should have $2v\lambda > \gamma q$, and since the assumption is $c^* > 0$, only when $a_{\lceil \frac{2v\lambda - \gamma q}{4\lambda} \rceil} > 0$ we can get this solution;

2) When $c^* > 0$:

To make the derivative of Eq.(20) regarding to c^* equal to 0, we should have

$$-1 \leq 2\lambda(i-1) - 2\lambda(v-i) - \gamma q \leq 1, \tag{25}$$

which leads to

$$\frac{2v\lambda + \gamma q}{4\lambda} + \frac{1}{2} - \frac{1}{4\lambda} \leq i \leq \frac{2v\lambda + \gamma q}{4\lambda} + \frac{1}{2} + \frac{1}{4\lambda}, \tag{26}$$

Again, $\left\lceil \frac{2v\lambda + \gamma q}{4\lambda} \right\rceil$ is an appropriate value for i, and the index should not exceed v, thus we have

$$\frac{2v\lambda + \gamma q}{4\lambda} \leq v, \tag{27}$$

which leads to

$$2v\lambda \geq \gamma q, \tag{28}$$

similarly due to the assumption $c^* < 0$, only when $a_{\lceil \frac{2v\lambda + \gamma q}{4\lambda} \rceil} < 0$ we can get this solution.

When there is no such $a_{\lceil \frac{2v\lambda - \gamma q}{4\lambda} \rceil}$ or $a_{\lceil \frac{2v\lambda + \gamma q}{4\lambda} \rceil}$, it's easy to see $c^* = 0$ minimizes Eq.(20).

C. Below are some figures that are not included in the main body of the paper due to space limitation:

The residual plot of the ADMM solver is presented in Fig.8, it's converging to 0 as the iteration increases:

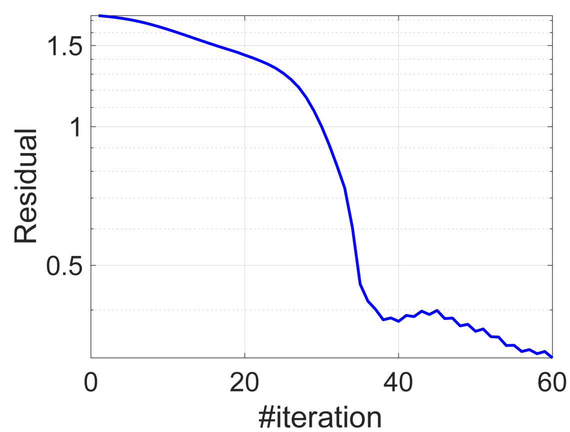

Figure 8. Residual with update.

To show the advantage of combining multi-view feature sets, in subspace clustering, we run our proposed method with increasing number of views on four datasets. Comparison of NMI is shown in Fig.9:

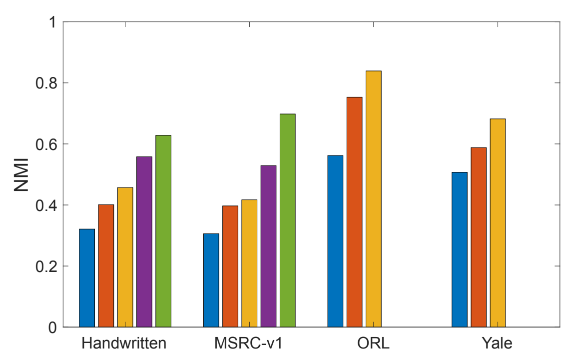

Figure 9. NMI comparison with increasing views, 1-5 views for Handwritten and MSRC-v1, 1-3 views for ORL and Yale.