Supplemental Material for "Can Visual Recognition Benefit from Auxiliary Information in Training?"

Qilin Zhang[1], Gang Hua[1], Wei Liu[2], Zicheng Liu[3], Zhengyou Zhang[3]

[1]Stevens Institute of Technolog, Hoboken, NJ
[2]IBM Thomas J. Watson Research Center, Yorktown Heights, NY
[3]Microsoft Research, Redmond, WA

In this supplemental material, we present the proof of Proposition 1 in Section 1 and additional experimental results with the the RGBD Object dataset [1] in Section 2.

1 Proof of Proposition 1

In this section, we present the proof of the Proposition 1 for Paper 498. First, the existence of the upper bound is proven in Section 1.1, then the proof that the sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is monotonic is presented in Section 1.2. With the Bolzano-Weierstrass theorem and the conclusions of Section 1.1 and Section 1.2, the Proposition 1 is proven.

1.1 Proof of the existence of the upper bound C_u

From the constraint in Eq. (6) of the paper:

$$\mathbf{W}_j^T \left[(1-\tau_j)\frac{1}{n}\boldsymbol{R}_j\boldsymbol{R}_j^T + \tau_j \mathbf{I}_n \right] \mathbf{W}_j = \mathbf{I}, \tag{1}$$

where τ_j denotes the pre-specified regularization parameter, $0 < \tau_j < 1$ ($j = 1, 2, \cdots, J$), we have

$$(1-\tau_j)\mathrm{tr}\left(\frac{1}{n}\mathbf{W}_j^T\boldsymbol{R}_j\boldsymbol{R}_j^T\mathbf{W}_j\right) + \tau_j \mathrm{tr}\left(\mathbf{W}_j^T\mathbf{W}_j\right) = n, \tag{2}$$

and therefore

$$\mathrm{tr}\left(\mathbf{W}_j^T\boldsymbol{R}_j\boldsymbol{R}_j^T\mathbf{W}_j\right) \leq \frac{1}{1-\tau_j},\ 0 < \tau_j < 1,\ \forall j = 1, 2, \cdots, J. \tag{3}$$

In addition, we have

$$\mathrm{tr}\left(\mathbf{W}_j^T\boldsymbol{R}_j\boldsymbol{R}_k^T\mathbf{W}_k\right) \leq \frac{1}{2}\left[\mathrm{tr}\left(\mathbf{W}_j^T\boldsymbol{R}_j\boldsymbol{R}_j^T\mathbf{W}_j\right) + \mathrm{tr}\left(\mathbf{W}_k^T\boldsymbol{R}_k\boldsymbol{R}_k^T\mathbf{W}_k\right)\right] \tag{4}$$

$$\leq \frac{1}{2}\left[\frac{1}{1-\tau_j} + \frac{1}{1-\tau_k}\right] < \infty, \tag{5}$$

where the inequality (4) follows the property

$$\mathrm{tr}(\mathbf{B}^T\mathbf{A}) \le \frac{1}{2}\left[\mathrm{tr}(\mathbf{A}^T\mathbf{A}) + \mathrm{tr}(\mathbf{B}^T\mathbf{B})\right], \tag{6}$$

which comes from the fact that $\mathrm{tr}\left((\mathbf{A}-\mathbf{B})^T(\mathbf{A}-\mathbf{B})\right) \ge 0$, because the matrix $(\mathbf{A}-\mathbf{B})^T(\mathbf{A}-\mathbf{B}) \succeq 0$. Since $\mathrm{tr}\left(\mathbf{W}_j^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k\right)$ is bounded, , $g(x) = x$ or x^2, we have

$$g\left(\mathrm{tr}\left(\mathbf{W}_j^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k\right)\right) \le \frac{1}{4}\left[\frac{1}{1-\tau_j} + \frac{1}{1-\tau_k}\right]^2 < \infty. \tag{7}$$

Considering $c_{jk} = 0$ or 1, we have

$$\sum_{j\neq k}^{J} c_{jk} g\left(\mathrm{tr}\left(\frac{1}{n}\mathbf{W}_j^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k\right)\right) \le \frac{1}{4n}\sum_{j\neq k}^{J}\left[\frac{1}{1-\tau_j} + \frac{1}{1-\tau_k}\right]^2 < \infty, \tag{8}$$

which shows that the sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is upper bounded by

$$C_u = \frac{1}{4n}\sum_{j\neq k}^{J}\left[\frac{1}{1-\tau_j} + \frac{1}{1-\tau_k}\right]^2 < \infty. \tag{9}$$

□

1.2 Proof that the sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is monotonically increasing

In this section, the monotonic property of the sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is presented. Following [2–4], we first present a Lemma, and then prove that

$$f(\mathcal{W}(s)) \le f(\mathcal{W}(s+1)),\ s = 1, 2, \cdots, \tag{10}$$

where s is the iteration index, $s = 1, 2, \cdots$.

Define the function $r(\mathbf{Y}_j, \mathbf{Y}_k) \overset{\mathrm{def}}{=} \mathrm{tr}(\frac{1}{n}\mathbf{Y}_j^T\mathbf{Y}_k) = \mathrm{tr}(\frac{1}{n}\mathbf{W}_j^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k)$, therefore,

$$f(\mathbf{W}_1(s), \cdots, \mathbf{W}_J(s)) = \sum_{j,k=1,k\neq j}^{J} c_{jk} g[r(\mathbf{Y}_j(s), \mathbf{Y}_k(s))]. \tag{11}$$

Using this notation, we have the following Lemma:

Lemma 1. *Define*

$$f_j(\mathbf{W}_j) \overset{\mathrm{def}}{=} \sum_{k=1}^{j-1} c_{jk} g\left[r(\boldsymbol{R}_j^T\mathbf{W}_j, \mathbf{Y}_k(s+1))\right] + \sum_{k=j+1}^{J} c_{jk} g\left[r(\boldsymbol{R}_j^T\mathbf{W}_j, \mathbf{Y}_k(s))\right] \tag{12}$$

$$s.t.\ \mathbf{W}_j^T N_j \mathbf{W}_j = \mathbf{I},$$

then

$$f_j(\mathbf{W}_j(s)) \leq f_j(\mathbf{W}_j(s+1)),\ j = 1, \cdots, J. \tag{13}$$

Proof. We prove Lemma 1 in two cases, *i.e.*, $g(x) = x^2$ and $g(x) = x$.
Case 1: when $g(x) = x^2$, we have $f_j(\mathbf{W}_j)$ in the following form,

$$f_j(\mathbf{W}_j) = \sum_{k=1}^{j-1} c_{jk} \left(r(\boldsymbol{R}_j^T \mathbf{W}_j, \mathbf{Y}_k(s+1)) \right)^2 + \sum_{k=j+1}^{J} c_{jk} \left(r(\boldsymbol{R}_j^T \mathbf{W}_j, \mathbf{Y}_k(s)) \right)^2, \tag{14}$$

which can be written as

$$\begin{aligned} f_j(\mathbf{W}_j(s)) &= \frac{1}{n} \sum_{k=1}^{j-1} c_{jk} \theta_{jk}^{(s)} \mathrm{tr}(\mathbf{W}_j(s)^T \boldsymbol{R}_j \mathbf{Y}_k(s+1)) \\ &\quad + \frac{1}{n} \sum_{k=j+1}^{J} c_{jk} \theta_{jk}^{(s)} \mathrm{tr}(\mathbf{W}_j(s)^T \boldsymbol{R}_j \mathbf{Y}_k(s)) \qquad (15) \\ &= \frac{1}{n} \mathrm{tr} \left(\mathbf{W}_j(s)^T \boldsymbol{R}_j \left(\sum_{k=1}^{j-1} c_{jk} \theta_{jk}^{(s)} \mathbf{Y}_k(s+1) + \sum_{k=j+1}^{J} c_{jk} \theta_{jk}^{(s)} \mathbf{Y}_k(s) \right) \right), \qquad (16) \end{aligned}$$

where $\theta_{jk}^{(s)}$ are defined as

$$\theta_{jk}^{(s)} = \begin{cases} r\left(\mathbf{Y}_j(s), \mathbf{Y}_k(s+1)\right) & \text{if } \ k = 1, \cdots, j-1 \\ r\left(\mathbf{Y}_j(s), \mathbf{Y}_k(s)\right) & \text{if } \ k = j+1, \cdots, J \end{cases}. \tag{17}$$

Note that in Eq. (16), the term $\left(\sum_{k=1}^{j-1} c_{jk} \theta_{jk}^{(s)} \mathbf{Y}_k(s+1) + \sum_{k=j+1}^{J} c_{jk} \theta_{jk}^{(s)} \mathbf{Y}_k(s) \right)$ is equivalent to the definition of $\mathbf{Z}_j(s)$, hence $f_j(\mathbf{W}_j)$ can be simplified as $\frac{1}{n}\mathrm{tr}(\mathbf{W}_j^T \boldsymbol{R}_j \mathbf{Z}_j(s))$. Considering the following optimization problem:

$$\max_{\mathbf{W}_j} \frac{1}{n} \mathrm{tr}(\mathbf{W}_j^T \boldsymbol{R}_j \mathbf{Z}_j(s)),\ \text{s.t. } \mathbf{W}_j^T N_j \mathbf{W}_j = \mathbf{I}, \tag{18}$$

whose solution is exactly

$$\mathbf{W}_j(s+1) = N_j^{-1} \boldsymbol{R}_j \mathbf{Z}_j(s) \left([\mathbf{Z}_j(s)^T \boldsymbol{R}_j^T N_j^{-1} \boldsymbol{R}_j \mathbf{Z}_j(s)]^{1/2} \right)^{\dagger}, \tag{19}$$

we have

$$\mathrm{tr}(\mathbf{W}_j(s)^T \boldsymbol{R}_j \mathbf{Z}_j(s)) \leq \mathrm{tr}(\mathbf{W}_j(s+1)^T \boldsymbol{R}_j \mathbf{Z}_j(s)). \tag{20}$$

Similarly, the following equations can be obtained:

$$\begin{aligned} f_j(\mathbf{W}_j(s)) &= \sum_{k=1}^{j-1} c_{jk}\theta_{jk}^{(s)} r(\mathbf{Y}_j(s), \mathbf{Y}_k(s+1)) + \sum_{k=j+1}^{J} c_{jk}\theta_{jk}^{(s)} r(\mathbf{Y}_j(s), \mathbf{Y}_k(s)) \\ &\leq \sum_{k=1}^{j-1} c_{jk}\theta_{jk}^{(s)} r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s+1)) \\ &\quad + \sum_{k=j+1}^{J} c_{jk}\theta_{jk}^{(s)} r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s)). \end{aligned} \tag{21}$$

Considering that c_{jk} is either 0 or 1, we hence have $c_{jk} = c_{jk}^2$. Applying the Cauchy-Schwartz inequality, we have

$$\begin{aligned} f_j(\mathbf{W}_j(s)) &\leq \sum_{k=1}^{j-1} c_{jk}^2\theta_{jk}^{(s)} r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s+1)) + \sum_{k=j+1}^{J} c_{jk}^2\theta_{jk}^{(s)} r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s)) \\ &\leq \sqrt{\sum_{k=1}^{j-1} c_{jk}\left(\theta_{jk}^{(s)}\right)^2 + \sum_{k=j+1}^{J} c_{jk}\left(\theta_{jk}^{(s)}\right)^2} \\ &\quad \cdot \sqrt{\sum_{k=1}^{j-1} c_{jk}\left(r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s+1))\right)^2 + \sum_{k=j+1}^{J} c_{jk}\left(r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s))\right)^2} \\ &= \sqrt{\sum_{k=1}^{j-1} c_{jk}\left(r(\mathbf{Y}_j(s), \mathbf{Y}_k(s+1))\right)^2 + \sum_{k=j+1}^{J} c_{jk}\left(r(\mathbf{Y}_j(s), \mathbf{Y}_k(s))\right)^2} \\ &\quad \cdot \sqrt{\sum_{k=1}^{j-1} c_{jk}\left(r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s+1))\right)^2 + \sum_{k=j+1}^{J} c_{jk}\left(r(\mathbf{Y}_j(s+1), \mathbf{Y}_k(s))\right)^2} \\ &= \sqrt{f_j(\mathbf{W}_j(s))} \cdot \sqrt{f_j(\mathbf{W}_j(s+1))}. \end{aligned} \tag{22}$$

We immediately have $f_j(\mathbf{W}_j(s)) \leq f_j(\mathbf{W}_j(s+1))$. This concludes the case 1 scenario.

Case 2: when $g(x) = x$, we have

$$\begin{aligned} f_j(\mathbf{W}_j) &= \sum_{k=1}^{j-1} c_{jk} r(\boldsymbol{R}_j^T \mathbf{W}_j, \mathbf{Y}_k(s+1)) \\ &\quad + \sum_{k=j+1}^{J} c_{jk} r(\boldsymbol{R}_j^T \mathbf{W}_j, \mathbf{Y}_k(s)). \end{aligned} \tag{23}$$

Therefore, we can have exactly the same equation as Eq. (16), except that $\theta_{jk}^{(s)} \equiv 1$ for all the cases. The same equation as in Eq. (20) can be obtained, which directly implies that $f_j(\mathbf{W}_j(s)) \leq f_j(\mathbf{W}_j(s+1))$. This concludes both the case 2 scenario and the entire proof of Lemma. □

With the conclusion in Lemma 1, we proceed with the proof that the sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is monotonically increasing. Consider the following subtraction

$$
\begin{aligned}
&\sum_{j=1}^{J}[f_j(\mathbf{W}_j(s+1)) - f_j(\mathbf{W}_j(s))] \\
&\quad = \sum_{j=1}^{J}\sum_{k=1}^{j-1} c_{jk} g \left[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s+1)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s+1)\right)\right] \\
&\quad + \sum_{j=1}^{J}\sum_{k=j+1}^{J} c_{jk} g \left[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s+1)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s)\right)\right] \\
&\quad - \sum_{j=1}^{J}\sum_{k=1}^{j-1} c_{jk} g \left[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s+1)\right)\right] \\
&\quad - \sum_{j=1}^{J}\sum_{k=1}^{j-1} c_{jk} g \left[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s+1)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s+1)\right)\right] \\
&\quad = \frac{1}{2}\Big[\sum_{j,k=1,k\neq j}^{J} c_{jk} g \big[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s+1)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s+1)\right)\big] \\
&\quad - \sum_{j,k=1,k\neq j}^{J} c_{jk} g \big[\operatorname{tr}\left(\frac{1}{n}\mathbf{W}_j(s)^T \boldsymbol{R}_j \boldsymbol{R}_k^T \mathbf{W}_k(s)\right)\big]\Big] \geq 0. \qquad (24)
\end{aligned}
$$

The last equation in Eq. (24) follows the Lemma 1. This implies that

$$f(\mathbf{W}_1(s), \cdots, \mathbf{W}_J(s)) \leq f(\mathbf{W}_1(s+1), \cdots, \mathbf{W}_J(s+1)) \qquad (25)$$

$$\text{i.e., } f(\mathcal{W}(s)) \leq f(\mathcal{W}(s+1)),\ s = 1, 2, \cdots. \qquad (26)$$

Using Eq. (8), Eq. (26), the bounded sequence $f(\mathcal{W}(s)), s = 1, 2, \cdots$ is monotonically increasing.

According to the Bolzano-Weierstrass theorem, the sequence will converge, *i.e.,* Proposition 1 is proven.

2 Additional Experimental Results

With the same RGBD Object dataset and experimental settings from [1] and [5], we have also conducted additional experiments with the state-of-the-art HMP-based features [5]. These additional results are summarized in Table 1–2 (too many entries to fit in a single page).

As is seen in Table 1–2, the overall recognition accuracy improves significantly across all methods, as compared to the EMK-based features [1]. In a large portion of the categories, perfect recognition is achieved even with the naive baseline

Table 1. Accuracy Table Part 1 for the Multi-View RGBD Object Instance recognition with HMP features, the highest and second highest values are colored red and blue, respectively. The remaining part is in Table 2.

Category	SVM	SVM2K	KCCA	KCCA +L	RGCCA	RGCCA +L	RGCCA +AL	DCCA
apple	87.62	**93.33**	**92.86**	77.14	92.38	87.62	**93.33**	**93.33**
ball	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
banana	73.74	77.78	**81.31**	**81.31**	**80.30**	75.25	**80.30**	**80.30**
bell pepper	82.07	**84.06**	78.88	77.29	73.31	80.88	73.31	**83.27**
binder	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
bowl	83.08	86.54	86.54	86.15	**87.31**	77.69	**87.69**	**87.69**
calculator	**100**	**100**	**100**	**99.44**	**100**	**100**	**100**	**100**
camera	**99.17**	**100**	**100**	**100**	**99.17**	**99.17**	**99.17**	**99.17**
cap	**100**	**100**	**100**	**100**	**100**	**97.08**	**100**	**100**
cellphone	**98.43**	95.29	**95.81**	94.76	**95.81**	**98.43**	**95.81**	**95.81**
cereal box	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
coffee mug	92.88	85.45	96.90	94.43	**99.69**	91.64	**99.69**	**100**
comb	96.67	**97.33**	**97.33**	96.67	**98.00**	95.33	**98.00**	**98.00**
dry battery	86.34	84.14	85.46	76.65	**87.67**	85.90	**88.55**	**88.55**
flashlight	97.34	97.34	**97.87**	95.21	**100**	97.34	**100**	**100**
food bag	**100**	**100**	**100**	**100**	**100**	**96.88**	**100**	**100**
food box	**100**	**99.84**	**100**	**99.84**	**100**	**100**	**100**	**100**
food can	98.08	93.02	94.39	81.67	**99.18**	98.08	**99.32**	**99.32**
food cup	96.32	**98.90**	97.06	96.32	97.43	97.06	97.43	**98.53**
food jar	**98.73**	**100**	**100**	91.46	**100**	97.15	**100**	**100**
garlic	90.91	**98.40**	**98.40**	**96.26**	95.99	91.98	95.99	95.99
glue stick	**100**	**100**	**100**	**100**	**100**	**99.68**	**100**	**100**
greens	76.76	89.73	89.19	85.41	85.41	81.08	**90.27**	**91.35**
hand towel	**99.63**	**100**	**100**	**100**	**100**	**99.63**	**100**	**100**
instant noodles	**100**	99.51	**99.76**	**99.76**	**100**	**100**	**100**	**100**
keyboard	94.55	95.05	95.05	94.55	**96.04**	94.55	**96.53**	**96.53**
kleenex	**98.87**	**97.36**	**97.36**	96.98	**98.87**	**98.87**	**98.87**	**98.87**
lemon	45.02	**47.41**	**47.41**	45.82	**47.81**	44.62	**47.81**	**47.81**
light bulb	87.67	**97.26**	**96.58**	95.21	91.78	89.04	92.47	92.47
lime	**67.78**	62.22	61.11	58.33	62.78	**68.33**	62.78	62.22
marker	**100**	93.93	98.26	96.75	99.57	**100**	**99.78**	99.57
mushroom	**99.42**	**100**	**100**	**100**	**100**	**99.42**	**100**	**100**
notebook	**100**	**100**	**100**	**100**	**99.62**	**100**	**100**	**100**
onion	**98.74**	94.64	95.90	95.90	96.53	**99.37**	**98.74**	**98.74**
orange	72.46	97.58	**99.03**	**98.07**	96.62	69.57	**99.03**	**99.03**
peach	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
pear	90.04	**93.24**	90.75	90.04	91.81	88.61	91.81	**93.95**
pitcher	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
plate	**99.66**	**100**	**100**	99.32	95.59	**100**	99.32	99.32
pliers	**92.58**	89.96	90.39	90.39	**93.45**	**92.58**	**93.45**	**93.45**
potato	63.71	**71.81**	**74.13**	65.64	65.25	67.57	71.04	71.43
rubber eraser	69.61	**75.49**	71.57	69.61	70.59	**74.02**	70.59	71.08
scissors	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
shampoo	**99.35**	**99.68**	**99.68**	98.39	**99.68**	94.19	**99.68**	**99.68**
soda can	**100**	**100**	**99.55**	99.10	**100**	**100**	**100**	**100**

Table 2. Continued from Table 1: Accuracy Table Part 2 for the Multi-View RGBD Object Instance recognition with HMP features, the highest and second highest values are colored red and blue, respectively.

Category	SVM	SVM2K	KCCA	KCCA+L	RGCCA	RGCCA+L	RGCCA+AL	DCCA
sponge	99.49	**100**	**100**	**99.83**	**99.83**	**99.83**	**99.83**	**99.83**
stapler	80.12	73.59	76.56	76.56	81.90	81.90	**82.20**	**82.49**
tomato	69.69	**81.59**	**83.85**	**83.85**	80.45	68.56	80.45	80.45
tooth brush	**99.48**	**100**	**99.48**	96.91	98.97	**99.48**	98.97	98.97
tooth paste	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**
water bottle	**97.32**	95.98	95.71	94.64	96.25	**96.78**	96.25	96.25
average	92.62	93.35	93.87	91.85	93.93	92.40	**94.35**	**94.62**

SVM algorithm, However, the advantage of the proposed method is revealed in some of the challenging categories.

Overall, with the HMP-based features, "KCCA+L" and the "RGCCA+L" algorithms cannot match the SVM baseline, and the "SVM2K" and "KCCA" algorithms are only marginally better than the baseline. The "RGCCA" and "RGCCA+AL" algorithms offer some improvements, while the proposed DCCA algorithm achieves the highest overall recognition accuracy. Considering there are more than 13,000 testing samples, the two percent performance improvement means correctly classifying an additional amount of more than 200 samples.

References

1. Lai, K., Bo, L., Ren, X., Fox, D.: A large-scale hierarchical multi-view rgb-d object dataset. In: Robotics and Automation (ICRA), 2011 IEEE International Conference on, IEEE (2011) 1817–1824
2. Tenenhaus, A., Tenenhaus, M.: Regularized generalized canonical correlation analysis. Psychometrika **76** (2011) 257–284
3. Tenenhaus, A.: Kernel generalized canonical correlation analysis. Actes des 42è journées de Statistique (2010)
4. Hanafi, M.: Pls path modelling: computation of latent variables with the estimation mode b. Computational Statistics **22** (2007) 275–292
5. Bo, L., Ren, X., Fox, D.: Unsupervised feature learning for rgb-d based object recognition. ISER, June (2012)