Further Results on the Subspace Distance

Xichen Sun, Liwei Wang, Jufu Feng
National Laboratory on Machine Perception, Center for information Science,
School of Electronics Engineering and Computer Science,
Peking University,
Beijing, P.R.China

Abstract
In previous papers [1, 2], we proposed a subspace distance. However, whether the subspace distance satisfies the triangle inequality was left open. In this note, we give positive answer to the open problem and prove our assertion.

Keywords: Subspace distance; Face recognition

1 Introduction
In previous papers [1, 2], we proposed a subspace distance. The subspace distance was used to analyze the so-called intrapersonal face subspaces and to design an adaptive Bayesian algorithm for human face recognition. If the subspace distance is a metric, it can also be applied to many other tasks such as the video based recognition.

However, whether the subspace distance satisfies the triangle inequality was left open. In [2, Section 5], we discussed two special cases in which the triangle inequality holds, and conjectured that it would be always true.

In this note, we give positive answer and prove our assertion.

2 The Triangle Inequality of Subspace Distance
For a $m$-dimensional subspace $U$ and a $n$-dimensional subspace $V$, their distance was defined as [2, Section 2, Definition 3]:

$$d(U,V) = \sqrt{\max(m,n) - \sum_{i=1}^{m} \sum_{j=1}^{n} (u_i^T v_j)^2}$$

where $u_1, u_2, \cdots, u_m$ and $v_1, v_2, \cdots, v_n$ are orthonormal bases of $U$ and $V$ respectively. Note that the distance has been shown to be independent on the choice of basis [2]. The main result of this note is given in the following theorem.

Theorem 1. Let $U$, $V$ and $W$ be arbitrary $m$, $n$ and $k$-dimensional subspaces of $\mathbb{R}^d$ ($k, m, n \leq d$). Then

$$d(U, V) \leq d(U, W) + d(W, V)$$

The equality holds if and only if $U = W$ or $V = W$. 

1
Proof. Let $u_1, u_2, \ldots, u_m$, $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_k$ be orthonormal bases of $U$, $V$ and $W$ respectively.

Consider a $d \times 2d$ rectangle matrix $M_U$ for subspace $U$. Denote the left half of $M_U$ as $M_U^l$, where $M_U^l$ is a $d \times d$ square matrix. Denote the right half of $M_U$ as $M_U^r$. $M_U^r$ is also a $d \times d$ square matrix. Formally, $M_U = [M_U^l, M_U^r]$.

Let $M_U^l = \sum_{i=1}^{m} u_i u_i^T$.

Let $M_U^r$ be a diagonal matrix, for which only the first $m$ diagonal elements be 1 and all the others be 0.

In like manner, we can define $M_V$ and $M_W$ using $v_1, v_2, \ldots, v_n$ and $w_1, w_2, \ldots, w_k$ respectively.

Now consider the Frobenius norm of the matrix $M_U - M_V$:

$$\|M_U - M_V\|_F = \sqrt{\|M_U^l - M_V^l\|_F^2 + \|M_U^r - M_V^r\|_F^2}$$

$$= \sqrt{\|M_U^l\|_F^2 + \|M_V^l\|_F^2 - 2 \cdot tr[(M_U^l)^T(M_V^l)] + |m - n|}$$

$$= \sqrt{m + n - 2 \sum_{i=1}^{m} \sum_{j=1}^{n} (u_i^T v_j)^2}$$

$$= \sqrt{2} \cdot \max(m, n) - \sum_{i=1}^{m} \sum_{j=1}^{n} (u_i^T v_j)^2$$

$$= \sqrt{2} \cdot d(U, V)$$

It is obvious that the Frobenius norm satisfies

$$\|M_U - M_V\|_F^2 = \|M_U - M_W\|_F^2 + \|M_W - M_V\|_F^2$$

and the equality holds if and only if $U = W$ or $V = W$. This completes the proof. \qed

3 Acknowledgements

This work was supported by Program for New Century Excellent Talents in University and NSFC (60575002) and NKBRPC (2004CB318000). Sun, X. thanks Daimler-Benz Foundation, Prof. Dr. K.-R. Mueller and other colleagues in Fraunhofer FIRST. IDA for their support and warm hospitality. Special thanks to Emili Zhang.

References
