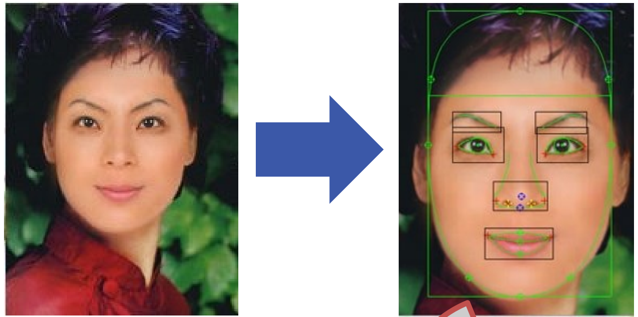
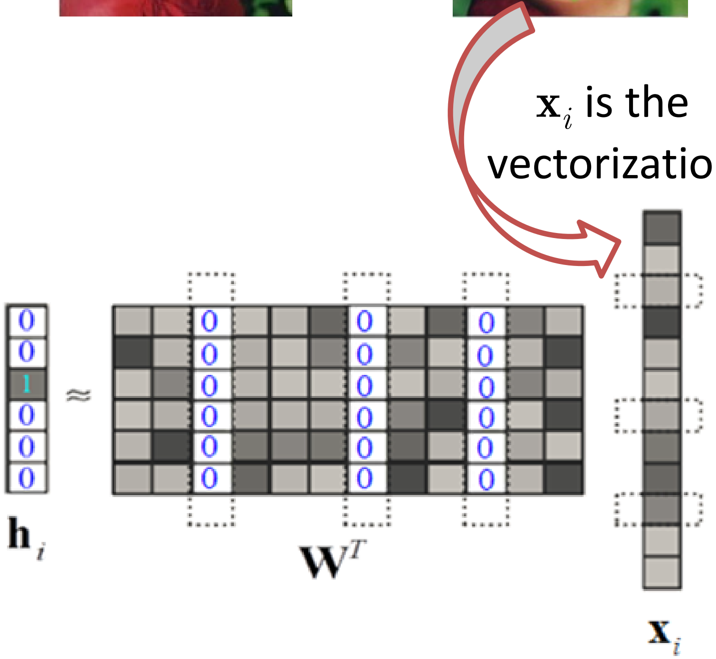


Which is more discriminative?



$$\min_{\mathbf{W}} \left\{ \begin{array}{c} \left\|\mathbf{W}^T\mathbf{X}-\mathbf{H}\right\|_F^2 \\ +\beta\left\|\mathbf{W}\right\|_{2,1} \\ +\alpha\Upsilon\left(\mathbf{W},\mathbf{X},\mathbf{H}\right) \end{array} \right\}$$

Regression based feature selection is frequently used to address this practical problem.

Preserving Ordinal Consensus: Towards Feature Selection for Unlabeled Data

Jun Guo, Heng Chang, and Wenwu Zhu

Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua University

- Our motivation:
  - Exploit one-to-one correspondence.

$$\mathbf{W}^T\mathbf{X} = (\mathbf{w}^1)^T\mathbf{x}^1 + \ldots + (\mathbf{w}^i)^T\mathbf{x}^i + \ldots + (\mathbf{w}^j)^T\mathbf{x}^j + \ldots$$

  - Exploit feature-level ordinal information.

Definition

Given distance function $dis(\cdot,\cdot)$ and projection function $\Phi(\cdot)$. A data point $\mathbf{z}_i$ and its neighbors $\mathbf{z}_u$ and $\mathbf{z}_v$ form a triplet. The projection of this triplet is defined as an ***Ordinal Consensus Preserving*** process when the following condition holds: If $dis(\mathbf{z}_i,\mathbf{z}_u)\leq dis(\mathbf{z}_i,\mathbf{z}_v)$, then $dis(\Phi(\mathbf{z}_i),\Phi(\mathbf{z}_u))\leq dis(\Phi(\mathbf{z}_i),\Phi(\mathbf{z}_v))$.

Preserving Ordinal Consensus: Towards Feature Selection for Unlabeled Data

Jun Guo, Heng Chang, and Wenwu Zhu
Tsinghua-Berkeley Shenzhen Institute (TBSI), Tsinghua University