Cross-Age Face Verification by Coordinating with Cross-Face Age Verification

Liang Du, Haibin Ling
Department of Computer and Information Sciences, Temple University.

We present a novel framework for *cross-age face verification* (FV) by seeking help from its "competitor" named *cross-face age verification* (AV), i.e., deciding whether two face photos are taken at similar ages. While FV and AV share some common features, FV pursues *age insensitivity* and AV seeks *age sensitivity*. Such correlation suggests that AV may be used to guide feature selection in FV, i.e., by reducing the chance of choosing age sensitive features.

Driven by this intuition, we propose to learn a solution for cross-age face verification by coordinating with a solution for age verification.

The baseline algorithm is the greedy coordinate descent strategy used for function approximation [1]. The goal is to learn the model coefficients $\mathbf{p}$ to minimize a loss function of the form:

$$\mathcal{L}(\mathcal{D},\mathbf{p}) = \sum_{i=1}^{N} \ell(y_i, f(\mathbf{x}_i;\mathbf{p})), \tag{1}$$

where $\ell(\cdot,\cdot)$ is the loss function for an individual sample.

Specifically, a joint additive model is devised to simultaneously handling both tasks, while encoding feature coordination by a competition regularization term, as follows,

$$\begin{aligned}\mathcal{J}(\mathcal{D},\mathcal{A},\mathbf{p},\mathbf{a}) = &\sum_{i=1}^{N} \exp(-y_i f(\mathbf{x}_i;\mathbf{p})) \\ &+\sum_{i=1}^{N_a} \exp(-l_i f_a(\mathbf{z}_i;\mathbf{a})) + \lambda \mathbf{p}^\top M_c^\top M_a \mathbf{a}\ ,\end{aligned} \tag{2}$$

where $\mathbf{p}$ and $\mathbf{a}$ are the model parameters of FV and AV, respective. Then, an *alternating greedy coordinate descent* (AGCD) algorithm is developed to solve this joint model.

When fixing $\mathbf{a}$, the partial derivative of $\mathcal{J}(\mathbf{p},\mathbf{a})$, w.r.t. p_j is calculated by

$$\begin{aligned}-\frac{\partial \mathcal{J}(\mathbf{p},\mathbf{a})}{\partial p_j} = &\sum_{i=1}^{N}\Big(y_i h_j(\mathbf{x}_i)\exp\big(-y_i\sum_{k=1}^{N_h} p_k h_k(\mathbf{x}_i)\big)\Big) \\ &-\lambda\frac{\partial(\mathbf{p}^\top M_p^\top M_a \mathbf{a})}{\partial p_j}\ .\end{aligned} \tag{3}$$

Given the binary property of M_p, we have

$$\frac{\partial(\mathbf{p}^\top M_p^\top M_a \mathbf{a})}{\partial p_j} = (M_p(:,j))^\top \boldsymbol{\phi}_a = \boldsymbol{\phi}_a(s_p(j))\ , \tag{4}$$

where $M_p(:,j)$ is the j-th column of M_p, $\boldsymbol{\phi}_a(s_p(j))$ the $s_p(j)$-th element of $\boldsymbol{\phi}_a$, and $s_p(j)$ the feature selection of the j-th hypothesis in f. Finally, we have

$$\begin{aligned}-\frac{\partial \mathcal{J}(\mathbf{p},\mathbf{a})}{\partial p_j} = &\sum_{i=1}^{N}\Big(y_i h_j(\mathbf{x}_i)\exp\big(-y_i\sum_{k=1}^{N_h} p_k h_k(\mathbf{x}_i)\big)\Big) \\ &-\lambda\boldsymbol{\phi}_a(s_p(j))\ .\end{aligned} \tag{5}$$

Following the GCD strategy, we choose p_j that maximizes this term. This is similar to the base hypothesis selection under the current weight distribution in the classical AdaBoost algorithm, except that in our case there is an additional term $-\lambda\boldsymbol{\phi}_a(s_p(j))$. Therefore, the base hypothesis selection is not only to fit the current distribution of data, but also to discourage choosing features favored by the auxiliary task.

A similar alternating step is conducted for the auxiliary task learning, taking the partial derivative of $\mathcal{J}(\mathbf{p},\mathbf{a})$ w.r.t. a_j as

$$\begin{aligned}-\frac{\partial \mathcal{J}(\mathbf{p},\mathbf{a})}{\partial a_j} = &\sum_{i=1}^{N_a}\Big(l_i h_j(\mathbf{z}_i)\exp\big(-l_i\sum_{k=1}^{N_h} a_k h_k(\mathbf{z}_i)\big)\Big) \\ &-\lambda\boldsymbol{\phi}_p(s_a(j))\ ,\end{aligned} \tag{6}$$

where $s_a(j)$ is the feature selection of the j-th hypothesis in f_a.

To evaluate the proposed algorithm, we conduct cross-age face verification experiments using two benchmark cross-age face datasets, FG-Net and MORPH. Table 1 and Table 2 show the results with comparison of state-of-the-arts.

Table 1: Comparison with state-of-the-arts on FG-net. Results are quoted from the corresponding references except for "proposed".

Method	Year	EER (%)
Graph Matching [5]	2010	25.4
GOP [2]	2010	24.1
Landmark [7]	2010	23.6
Growth Model [8]	2012	22.3
AGCD (GOP)	proposed	21.7
AGCD	proposed	**19.4**

Table 2: Comparison with state-of-the-arts on MORPH. Results are quoted from [4] except for "proposed".

Method	Year	EER (%)
Bayesian Eigenface [6]	2005	9.7
GOP [2]	2010	10.5
Bagging LDA [3]	2011	10.2
NRML [4]	2014	8.6
MNRML [4]	2014	7.5
AGCD (GOP)	proposed	9.2
AGCD	proposed	**5.5**

The main contributions of our work:

- To the best of our knowledge, our study is the first one that treats age-sensitive information as a blessing rather than a curse for cross-age face verification.
- A novel framework is proposed to harness the task conflict in a principled way, and a new algorithm is developed in this framework. The algorithm is general and can be extended to other scenarios involving similar task competition constraints.
- Extensive experiments on benchmark datasets are conducted and new results are registered.

[1] J. Friedman. Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5):1189–1232, 2001.

[2] H. Ling, S. Soatto, N. Ramanathan, and D. Jacobs. Face verification across-age progression using discriminative methods. *IEEE T-IFS*, 5(1):82–91, 2010.

[3] Z. Li, U. Park, and A. Jain. A discriminative model for age invariant face recognition. *IEEE T-IFS*, 6(3):1028–1037, 2011.

[4] J. Lu, X. Zhou, Y. Tan, Y. Shang, and J. Zhou. Neighborhood repulsed metric learning for kinship verification. *PAMI*, 36(2):331–345, 2014.

[5] G. Mahalingam and C. Kambhamettu. Age invariant face recognition using graph matching. In *BTAS*, 2010.

[6] N. Ramanathan and R. Chellappa. Face verification across-age progression. In *CVPR*, 2005.

[7] T. Wu, P. Turaga, and R. Chellappa. Age estimation and face verification across aging using landmarks. *IEEE T-IFS*, 7(6):1780-1788, 2012.

[8] T. Wu and R. Chellappa. Age invariant face verification with relative craniofacial growth model. In *ECCV*, 2012.

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.