A New Face Recognition Algorithm based on Dictionary Learning for a Single Training Sample per Person

Yang Liu yl504@cam.ac.uk Ian Wassell ijw24@cam.ac.uk

The number of the training samples per person has a significant impact on face recognition (FR) performance. For the single training sample per person (STSPP) problem, most traditional FR algorithms exhibit performance degradation owing to the limited information available to predict the variance of the query sample. This paper proposes a new method for the STSPP problem in FR, namely the Learn-Generate-Classify (LGC) method. The overall framework of the LGC method is presented in Fig.1.



Figure 1: The overall framework of LGC method

The first step is using the dictionary learning method to capture the intra-class variance from the generic set. We assume that the faces of the same subject in different views can be represented as the same one in the discriminant space. For a specific subject, the representation of its samples from different views with respect to the corresponding view-dictionary will be the same. Consequently, for the i^{th} subject, we have the following set of equations

$$\begin{cases} x_{i1} = D_1 \alpha_i + e_1 \\ \dots \\ x_{ip} = D_p \alpha_i + e_p \\ \dots \\ x_{iN} = D_N \alpha_i + e_N \end{cases}$$
(1)

where the sparse representation α_i is shared among the different view conditions of the subject *i*. The Dictionary D_p is the corresponding dictionary in the p^{th} view condition. There are a total of *N* different view conditions and e_p is the residual for the recovered image based on the dictionary D_p and sparse representation α_i . In order to make the learned dictionary both reconstructive and discriminat, the dictionary learning process should add another constraint to encourage the images from the same subject to have similar sparse coefficients and those from different subject to have dissimilar sparse coefficients representations. More specifically, we need to consider a new label consistency constraint, called the 'discriminant sparse-code error' and combine it with the reconstruction error to form a unified objective function, i.e.,

$$\arg\min_{D,W,\alpha} \|X - D\alpha\|_2 + \lambda_1 \|Q - W\alpha\|_2 + \lambda_2 \|\alpha\|_1, \qquad (2)$$

where λ_1 controls the relative contribution of the reconstructive term and the discriminant term. The consistently label Q is the ground truth for which dictionary columns should contribute to each of the training images. The matrix W is a linear transformation matrix and the linear transformation $g(W, \alpha) = W\alpha$ transforms the original sparse codes α to the most discriminate sparse feature domain. The objective function Eq.(2) is not jointly convex to D, W, α . Therefore, we solve this problem by breaking it into two sub problems, and alternately update the unknown variables. It involves a sparse coding stage using a pursuit algorithm, followed by an update of the dictionary.

The second step is generation of synthesized images. To generate multiple synthetic samples from the single training sample, two steps are Computer Laboratory, University of Cambridge

required. The first step is to recover the discriminant representation α_i over the generic training set, which aims to represent the query sample as a linear combination of the images from the generic set. The second step is to synthesize a series of virtual images using α_i and the corresponding view-basis dictionary, which represents an information flow from the discriminant representation back to the observation space.

The third step is using the Sparse Representation Classifier (SRC) framework for classification. The pseudo code of SRC algorithm can be found in [4]. The overall LGC algorithm is summarized in Algorithm 1.

Algorithm 1 Learn-Generate-Classify (LGC) Method					
Initialization: <i>D</i> and <i>Q</i>					
1: Variation Dictionary Learning					
2: while not converge do					
3: Update α in Eq.(2) with fixed Dictionary D and Q					
4: Update dictionary <i>D</i> atom by atom via K-SVD Algorithm [1]					
5: end while					
6: Generate Synthetic images					
7: Classify the test sample via SRC Algorithm [4]					
Output: Classification result					

Finally, we verified the effectiveness of the new LGC method on the CMU Multi-pie database, with different illumination, expression and pose variation factors. Table 1 and Table 2 show the recognition rates achieved by the different methods for illumination and pose variation respectively. According to the experimental results, it shows that the LGC method demonstrates a promising FR performance with only a STSPP.

Table	1:	Recognition	Rates	for	illumination	variation
-------	----	-------------	-------	-----	--------------	-----------

Session	S2	S3	S4
NN	43.64%	40.23%	38.99%
SVM	43.64%	40.53%	42.50%
SRC	44.66%	38.80%	43.21%
AGL [3]	84.37%	79.51%	79.33%
ESRC [2]	89.25%	84.06%	87.40%
LGC	90.72%	88.42%	89.68%

Table 2: Recognition Rates for pose variation								
Session	P05-0-S2	P04-1-S3	P04-1-S4					
NN	19.01%	8.99%	6.72%					
SVM	18.54%	8.73%	6.72%					
SRC	18.64%	9.21%	6.99%					
AGL	50.71%	23.58%	19.88%					
ESRC	53.95%	29.74%	22.58%					
LGC	56.77%	33.74%	29.25%					

- Michal Aharon, Michael Elad, and Alfred Bruckstein. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *Signal Processing, IEEE Transactions on*, 54(11):4311– 4322, 2006.
- [2] Weihong Deng, Jiani Hu, and Jun Guo. Extended SRC: Undersampled face recognition via intraclass variant dictionary. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(9):1864– 1870, 2012.
- [3] Yu Su, Shiguang Shan, Xilin Chen, and Wen Gao. Adaptive generic learning for face recognition from a single sample per person. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 2699–2706. IEEE, 2010.
- [4] John Wright, Allen Y Yang, Arvind Ganesh, Shankar S Sastry, and Yi Ma. Robust face recognition via sparse representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(2): 210–227, 2009.