SELF-PACED LEAST SQUARE SEMI-COUPLED DICTIONARY LEARNING FOR PERSON RE-IDENTIFICATION

Wei Xu, Haoyuan Chi, Lei Zhou, Xiaolin Huang, Jie Yang
Institution of Image Processing and Pattern Recognition, Shanghai Jiao Tong University, China

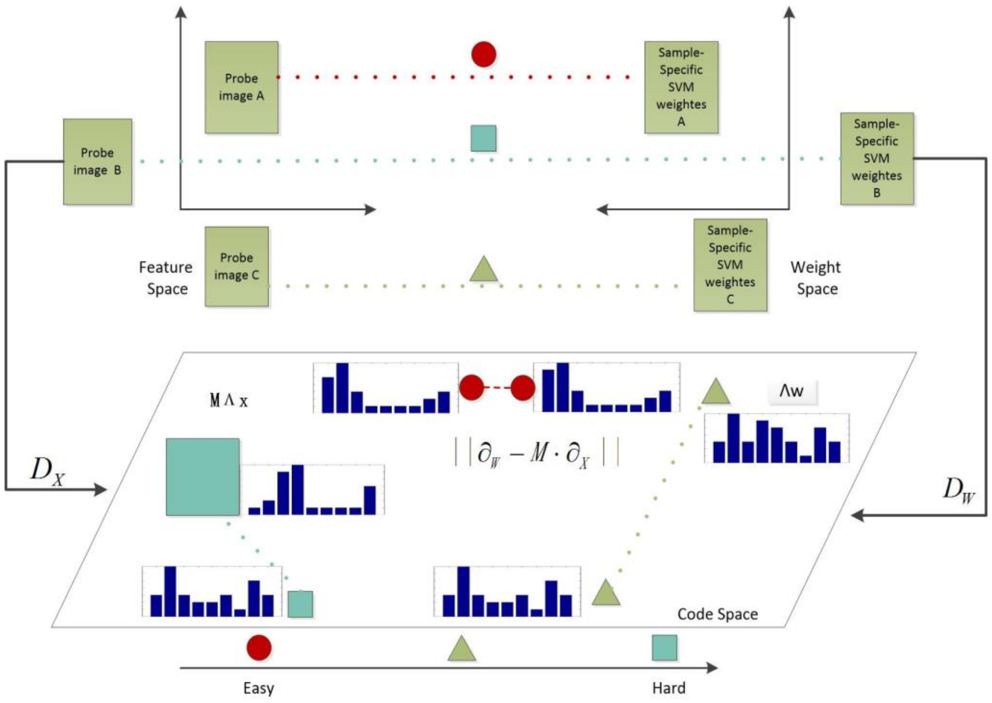
Overview

person re-identification: match people across disjoint camera views

Motivation: LSSCDL algorithm formulates the person re-identification problem into a binary classification problem and learn a classifier specifically for each pedestrian. It consists of 3 steps: Sample-Specific SVM Learning, LSSCDL and Pedestrian Matching.

Contribution:
1. Propose SLSSCDL algorithm to overcome the bad local minima problem of solving LSSCDL.
2. Introduce an undirected proximity graph in SLSSCDL to preserve the local similarities in feature space and weight space.

Application of SLSSCDL to person re-identification



The right side shape size represents sparse codes of dictionaries D_W for Sample-Specific SVM weights and the left side represents the sparse codes of dictionaries D_X for the probe images multiple with the mapping function M between the sparse code space of feature and weight. Intuitively, our method first selects the sample associated with the red circle, then with the yellow triangle and finally with the blue square.

Self-Paced Least Square Semi-Coupled Dictionary Learning

input: probe image matrix X_P, weight matrix W_P, parameters λ; λ_Λ; λ_M; λ_D; λ_S
Output: feature dictionary D_X, weight dictionary D_W, mapping matrix M

$$\min_{D_X,D_W,M,\Lambda_X,\Lambda_W,V} E(D_X,D_W,M,\Lambda_X,\Lambda_W,V)$$
$$= \left\|(X^P - D_X\Lambda_X)V\right\|_F^2 + \left\|(W^P - D_W\Lambda_W)V\right\|_F^2$$
$$+ \lambda\|(\Lambda_W - M\Lambda_X)V\|_F^2 + \lambda_\Lambda\|\Lambda_X\|_F^2 + \lambda_\Lambda\|\Lambda_w\|_F^2$$
$$+ \lambda_M\|M\|_F^2 + \lambda_D\|D_X\|_F^2 + \lambda_D\|D_W\|_F^2$$
$$+ \lambda_S(Tr(\Lambda_X^T L_X \Lambda_X) + Tr(\Lambda_W^T L_W \Lambda_W)) + f(V,k)$$

$Tr(\Lambda_X^T L_X\Lambda_X)$, $Tr(\Lambda_W^T L_W\Lambda_W)$ Are the local smoothness, λ_S is the regularization parameters of local smoothness, V is the diagonal sample weights *Matrix*, Λ, λ_Λ, λ_M, λ_D are regularization parameters to balance the terms in the objective function and Λ_X, Λ_W denote the coding coefficients.

Prediction

Given a test probe image x_{pt}, the corresponding weight vector w_{tp} can be derived with the learned dictionary pair D_w, D_x and mapping matrix M.

$$\min_{\{\alpha_x\}} \| x_t^p - D_x\alpha_x \|_F^2 + \lambda_\Lambda \| \alpha_x \|_F^2$$
$$\alpha_W = M\alpha_X \qquad w_t^p = D_W\alpha_W$$

With the corresponding weight vector wtp , the prediction is

$$F_i(x_i^p, x_j^g) = w_i^p * \phi(x_i^p, x_j^g) + b_i$$

$$F_i(x_i^p, x_j^g) = \begin{cases} \geq 0, & y_j^p = 1 \\ < 0, & y_j^p = -1. \end{cases}$$

where y_{jp} =+1 represents that (x_{pi} , x_{gj}) is a correct matching pair, while y_{jp} =-1 indicates the incorrect matches

Experimental and Discussion

1. **SPL**: Learning the easiness of the feature and weight pairs
2. **Smooth**: Preserving local similarities
3. **SLSSCDL(ours)**: Combining the easiness and local similarities

Table 1. Testing results on the VIPeR dataset (P=316). The CMC (%) at rank 1, 10, and 20 are listed

Method	rank=1	rank=10	rank=20
SLSSCDL (ours)	**42.77**	**85.37**	**93.76**
LSSCDL [8]	42.66	84.27	91.93
LOMO+XQDA [3]	40.00	80.51	91.08
KISSME [5]	19.60	62.20	77.00
SDALF [1]	19.87	49.37	65.73
PRDC [17]	15.66	53.86	70.09
LSSCDL[8]+smooth	42.70	84.87	92.78
LSSCDL[8]+SPL	42.73	84.96	93.42

Table 2. Testing results on the QMULGRID database (P=900). The CMC (%) at rank 1, 10, and 20 are listed

Method	rank=1	rank=10	rank=20
SLSSCDL (ours)	**22.60**	**52.40**	**63.44**
LSSCDL [8]	22.40	51.28	61.20
LOMO+XQDA [3]	16.56	41.84	52.40
PRDC [17]	15.66	53.86	70.09
RankSVM[18]	10.24	33.28	43.68
LSSCDL[8]+smooth	22.50	51.87	62.58
LSSCDL[8]+SPL	22.55	52.06	63.02

Convergence analysis:

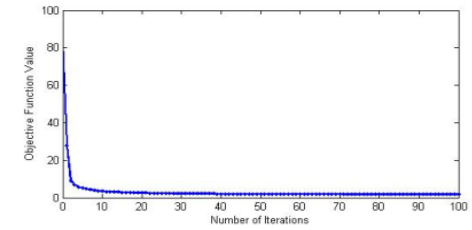


Fig. 2. The objective function value at varying number of iterations on QMUL GRID dataset.