Reduced Analytical Dependency Modeling for Classifier Fusion

Andy J Ma and Pong C Yuen

Department of Computer Science, Hong Kong Baptist University, Hong Kong

Overview

Introduction

- Existing probability based classifier fusion methods rely on certain assumptions
  - Features are *conditionally independent* [1].
  - *Normal* distribution assumption [2].
  - *Product formulation* with assumption that posteriors will not deviate very much from the priors [3].
- These assumptions may not be true in practice.
- Optimal weighting methods [4] [5] do not fully make use of probabilistic properties.
- Classifier fusion by modeling dependency based on probabilistic properties.

Contributions

- Prove equivalent condition to independent assumption
- Develop a *novel* framework for dependency modeling by analytical function.
- Propose Reduced Analytical Dependency Modeling (RADM) for classifier fusion.
- Advantages of the proposed method
  - Distribution-free.
  - Without product formulation assumption.

Reduced Analytical Dependency Modeling

Motivation

- Given scores $s_{lm} = \Pr(\omega_l|\vec{x}_m)$
  - Independent Fusion [1]

$$\Pr(\omega_l|\vec{x}_1,\dots,\vec{x}_M) = P_0\left(L^{M-1}\prod_{m=1}^{M} s_{lm}\right) = P_0 h_{\text{Product}}(s_{l1},\dots,s_{Ml})$$

  - Linear Classifier Dependency Model [3]

$$\Pr(\omega_l|\vec{x}_1,\dots\vec{x}_M) = P_0\left(\sum_{m=1}^{M} a_m s_{lm} + \frac{1-M}{L}\right) = P_0 h_{\text{LCDM}}(s_{l1},\dots,s_{lM})$$

Dependency can be modeled by choosing a suitable function h

Analytical Dependency Modeling

- Consider h as analytical function

$$\Pr(\omega_l|\vec{x}_1,\dots\vec{x}_M) = P_0 \sum_{|n_1+\cdots+n_M|=0}^{\infty} \alpha_{l(n_1,\dots,n_M)} \prod_{m=1}^{M} s_{lm}^{n_m}$$

- By Bayes' rule and properties of marginal distributions,

$$s_{lm} = \sum_{r=0}^{\infty} G_{lmr}(\vec{\alpha}_{lmr}) s_{lm}^{r}$$

where G_{lmr} is a function on $\vec{\alpha}_{lmr}$.

- Independent assumption is equivalent to the solution to the following equation system is trivial.

$$G_{lm1} = 1, G_{lm0} = 0, G_{lm2} = 0, \dots, G_{lmr} = 0, \dots$$

- Model dependency by setting non-trivial solution.

Model Learning

- Approximate the model by

$$h_l(\vec{s}_l,\vec{\alpha}_l) = \sum_{|\theta|=0}^{K} \alpha_{l\theta}\vec{s}_l^{\,\theta}, 0 \le n_M \le R$$

- Consider a special case

$$\Pr(\omega_l|\vec{x}_{j1},\dots,\vec{x}_{jM}) = \delta_{jl} = \begin{cases}1, \omega_l = y_j\\ 0, \omega_l \ne y_j\end{cases}$$

- Learn the model by minimizing the regularized least square errors

$$\min_{\vec{\alpha},\vec{p}} \sum \left(h_l(\vec{s}_l,\vec{\alpha}_l) - \frac{\delta_{jl}}{p_j}\right)^2 + b\|\vec{\alpha}_l\|^2$$

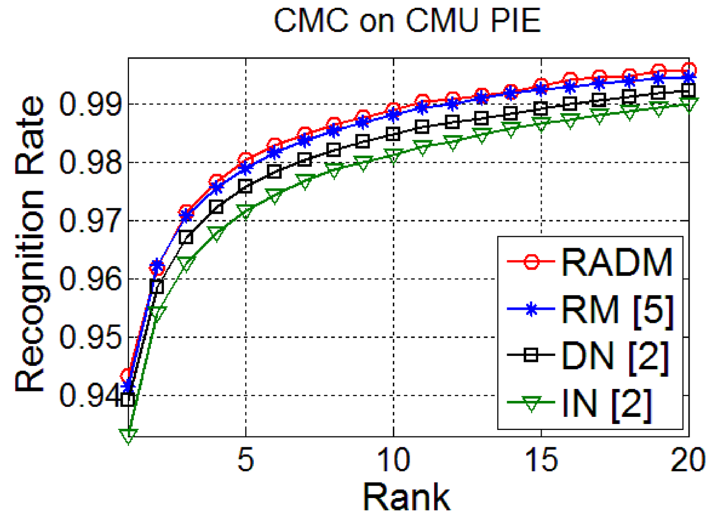
where $p_j = \prod_{m=1}^{M}\Pr(\vec{x}_{jm}) / \Pr(\vec{x}_{j1},\dots\vec{x}_{jM})$.

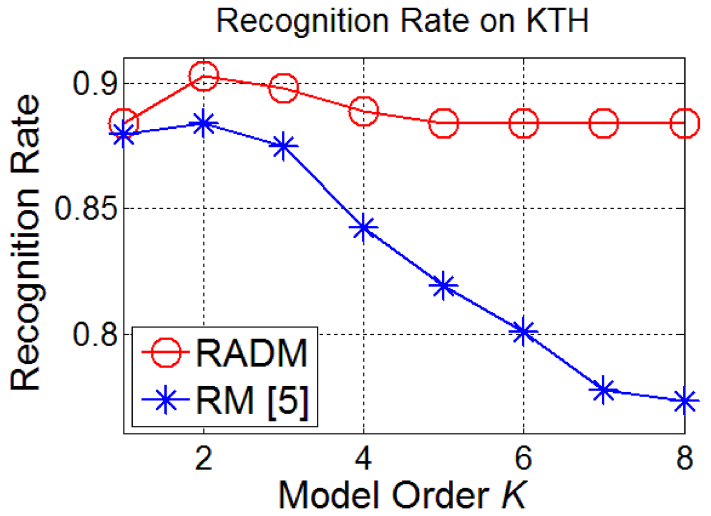
Experiments

Recognition accuracies (%) of fusion methods on different databases.

Test	Digit	Flower	CMU PIE	FERET	Weizmann	KTH
BestFea	94.77	70.39	88.87	83.33	82.22	78.70
Sum [1]	96.23	85.39	91.75	86.11	84.44	84.72
IN [2]	95.63	85.49	93.32	88.19	**85.56**	84.26
DN [2]	94.93	84.22	93.91	87.73	84.44	83.80
LCDM [3]	96.79	86.27	93.01	88.81	**85.56**	85.19
LP-B [4]	96.57	85.49	92.00	87.65	84.44	85.19
RM [5]	96.71	85.39	94.14	90.05	84.44	88.43
RADM	**96.84**	**88.04**	**94.34**	**90.97**	**85.56**	**90.28**

- RADM outperforms other classifier fusion methods.





- Compare the best four fusion algorithms on CMU PIE
- Compare RADM and RM [5] with different parameters.
  - Our method is robust to model order changes.

Reference

1. J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas. On combining classifiers. *TPAMI*, 20(3):226–239, 1998.
2. O. R. Terrades, E. Valveny, and S. Tabbone. Optimal classifier fusion in a non-bayesian probabilistic framework. *TPAMI*, 31(9):1630–1644, 2009.
3. A. J. Ma and P. C. Yuen. Linear dependency modeling for classifier fusion and feature combination. *TPAMI*, 2012, in press.
4. P. Gehler and S. Nowozin. On feature combination for multiclass object classification. *ICCV*, pages 221–228, 2009.
5. K. A. Toh, W. Y. Yau, X. Jiang, A reduced multivariate polynomial model for multimodal biometrics and classifiers fusion. TCSVT, 14, 224–233, 2004.