ECCV'16
EUROPEAN CONFERENCE ON COMPUTER VISION

Iterative Reference Driven Metric Learning for Signer Independent Isolated Sign Language Recognition

Fang Yin, Xiujuan Chai, Xilin Chen

Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190, China

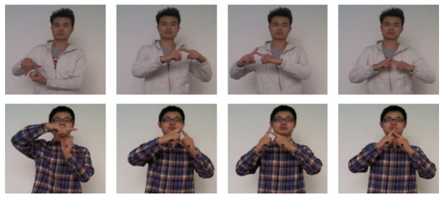
{fang.yin, xiujuan.chai, xilin.chen}@vipl.ict.ac.cn

Motivation

Sign Language Recognition(SLR) aims to build the bridge between the hard of hearing and the hearing, and debilitate the social isolation for the hard of hearing.

Motivation

- There exists big variations among different signers, due to the different heights of the signers or the habits during signing.
- A realistic SLR system must overcome the challenges brought by the inter-signer variations.



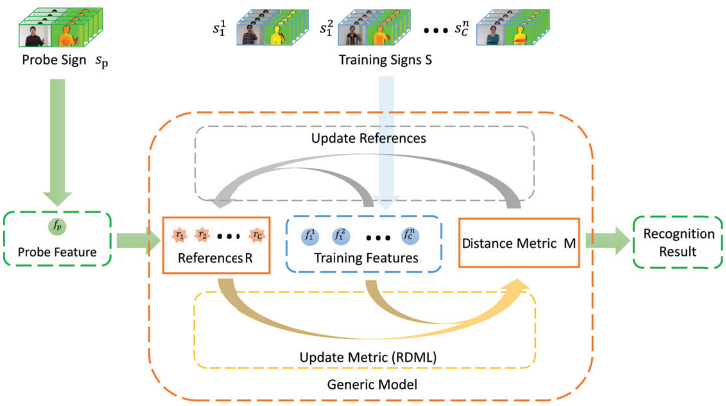
Examples of the sign Citizen performed by two different signers.

The Proposed Method

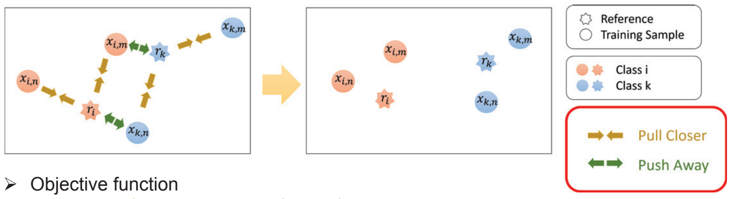
Basic Idea

- Use a signer invariant reference to represent each sign, and learn the metric accordingly.
- Optimize the reference and the metric alternately.

Framework of iRDML



Reference Driven Metric Learning (RDML)



- Objective function

$$\min_M f(M) = \sum_{i=1}^{C}\sum_{j=1}^{n_i} d(x_{i,j}, r_i) + \sum_{i=1}^{C}\sum_{j=1}^{n_i}\sum_{k=1}^{C}[1 + d(x_{i,j}, r_i) - d(x_{i,j}, r_k)]_+$$

$$d(x_{i,j}, r_k) = (x_{i,j} - r_k)^T M (x_{i,j} - r_k)$$

Iterative Reference Driven Metric Learning (iRDML)

- In RDML, the references are assumed to be known and can be initialized with the center of each class.
- The signer invariant references are learnt with the criterion similar to RDML.
- Alternately optimize ***M*** and ***R*** iteratively until converge.
  - A group of references ***R*** can be learnt given a metric ***M***.
  - A new metric ***M*** can be learnt given a group of references ***R***. (RDML)

Experimental Results

Evaluation on DEVISIGN Dataset

- Dataset

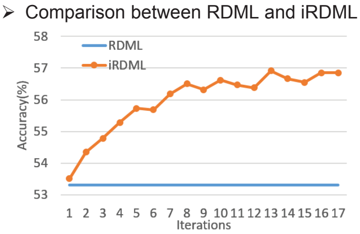
DEVISIGN[Chai,TR'14] dataset has a vocabulary size of 2000 from 8 different signers.

Evaluation protocol: 8 groups of data from 4 signers form the training set and the data from other 4 signers are the test data.

URL: vipl.ict.ac.cn/homepage/KSL/data.html

DEVISIGN Database

- Comparison between RDML and iRDML



- Comparison with existing methods

Method	Chai et al. [Chai, TR'14]				LMNN [Weinberger, NIPS'05]	Ours	
	HMM	DTW	ARMA	GCM		RDML	iRDML
Acc.(%)	34.44	38.35	39.03	51.81	51.49	53.30	**56.85**
Time(ms)	507.4	15778	1842	534.7	1.174	0.213	0.213

Evaluation on Human Motion Recognition

- Dataset

HDM05[Cho, VISAPP'14] consists of 2337 motion sequences from 65 actions.
Evaluation protocol: 10-fold cross validation

- Comparison with existing methods

Method	Cho et al.[Cho, VISAPP'14]				Ours
	ELM	SVM	MLP	Hybrid MLP	iRDML
Acc.(%)	91.57	94.95	95.20	95.59	**95.76**

Statistical Evaluation on SL Dataset

- Dataset

1000 signs × 7 signers × 1 repetition, for signer independent evaluation.

Evaluation protocol: leave-one-out cross validation.

- Comparison with existing methods

Method	G1	G2	G3	G4	G5	G6	G7	Ave.	sd.
HMM	57.4	57.1	58.7	55.9	55.9	61.0	47.4	56.2	4.3
DTW	61.7	60.5	66.6	60.6	33.5	49.2	16.2	49.8	18.5
ARMA	65.8	65.2	66.6	64.9	61.7	71.4	47.0	63.2	7.7
ITML	68	70.2	67.8	74.6	66.9	76.3	64.8	69.8	4.2
CSML	70.5	73.4	70.8	73.5	68.3	76.9	67.2	71.5	3.3
LMNN	70.0	72.1	69.0	75.6	69.1	77.7	66.2	71.3	4.0
iRDML	**75.0**	**78.5**	**74.4**	**78.9**	**74.1**	**81.4**	**72.1**	**76.3**	3.3

- The p-values given by the Student's distribution

Baseline/iRDML	HMM/iRDML	DTW/iRDML	ARMA/iRDML	LMNN/iRDML
p-value	0.000003	0.006209	0.000982	0.000046

Conclusion

- RDML is proposed to learn the distance between specific references and the training samples.
- iRDML is designed based on RDML to further explore more appropriate references and the corresponding distance metric.
- The effectiveness and efficiency of the proposed method is evaluated extensively on several public databases for both sign language recognition and human motion recognition tasks.