FTRDBJ Semantic Indexing Systems for TRECVID 2010

Kun TAO
France Telecom (R&D) Orange Labs, Beijing

Nov. 15, 2010

Overview

- 2009 HLFE Systems
  - 7 CEGL features & 6 SIFT features
  - 3 late fusion runs & 3 early fusion runs
- 2010 SIN Systems
  - 7 CEGL features & 12 features based on local descriptor
  - 3 late fusion runs & 1 early fusion run
  - 30 concept “FT-30” corpus
  - A cross-domain run

Overview

- 4 runs

ID	TYPE	DESCRIPTION	MAP
1	F_A	classifier-level-combination of 19 low-level feature SVMs with equal weights	0.070
2	F_A	linear weighted combination of 19 feature SVMs through logistic regression	0.075
3	F_C	cross-domain fusion between the results of run_2 and the results of 05-09 TRECVID models	0.070
4	L_A	kernel-level-combination of 14 low-level features with equal weighted multiple kernel learning	0.063

Overview

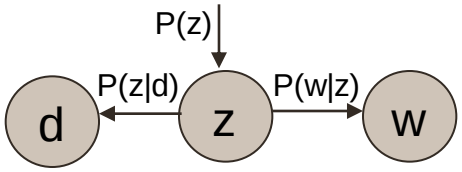
- FT-30
  - *Airplane_Flying*, Boat_Ship*, Bus*, Cityscape*, Classroom*, Demonstration_Or_Protest*, Hand*, Nighttime*, Singing*, Telephones**
  - *Animal+, Dark-skinned_People+, Flowers+, Running+, Sitting_Down+,*
  - *Anchorperson, Beach, Bicycles, Cats, Chair, Charts, Construction_Vehicles, Crowd, Female_Person, House_Of_Worship, Instrumental_Musician, Laboratory, Roadway_Junction, Shopping_Mall, Sports,*.

Features

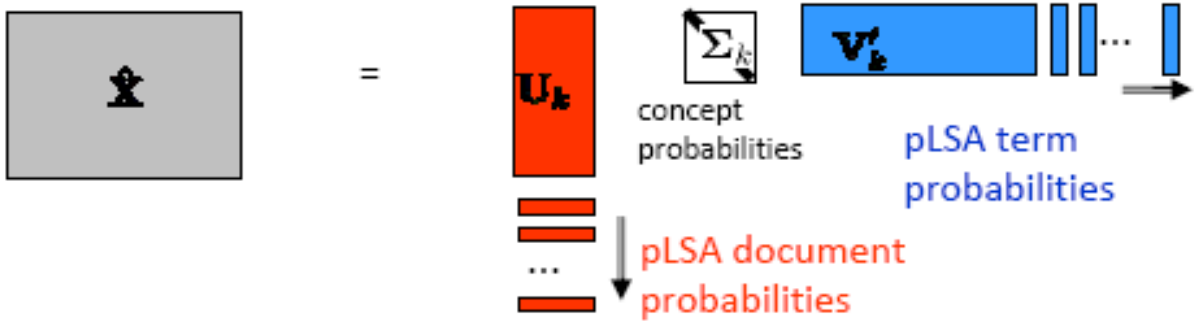
- 7 CEGL
  - Color Auto-Correlograms (CAC), Color Coherence Vector (CCV), Grid Color Moments (GCM), Edge Coherence Vector (ECV), Edge Direction Histogram (EDH), Gabor feature (Gabor) and Local Binary Patterns (LBP)
- 12 local descriptor features
  - SIFT, Dense-SIFT, SIFT-no_orientation
  - Pyramid HOW, PLSA
  - Soft -Assignment
  - HOG

Features

- PLSA



$$\hat{p}_{\mathrm{LSA}}(d,w) = \sum_z p(d|z)\, p(z)\, p(w|z)$$



Features

- Soft –Assignment

$$Weight_{ni} = \frac{1/(ni * Dist_{ni})}{\sum_{i=1}^{3}(1/(i * Dist_i))} \qquad ni = 1,2,3$$

- HOG

124-D Descriptors → **Pyramid Histograms**

"Object Detection using Histograms of Oriented Gradients". http://www.pascal-network.org/challenges/VOC/voc2006/slides /dalal.pdf.
Jianxiong Xiao et al. "SUN Database: Large-scale Scene Recognition from Abbey to Zoo",CVPR 2010

Features

- MAP of different features
  - (60% of dev. dataset for training SVM, 40% for evaluation)

Group Name	Feature Name	Dim.	MAP
S6	SIFT.HOW	512	0.117
	SIFT.2L-PHOW	2560	0.138
	SIFT. 3L-PHOW-PLSA	512	0.118
	DENSE-SIFT.HOW	512	0.166
	DENSE-SIFT.2L-PHOW	2560	0.169
	DENSE-SIFT. 3L-PHOW-PLSA	512	0.178
SS3	SIFT.HOW-SOFT	512	0.134
	SIFT-NO-ORIENTATION. HOW-SOFT	512	0.148
	DENSE-SIFT. HOW-SOFT	512	0.167

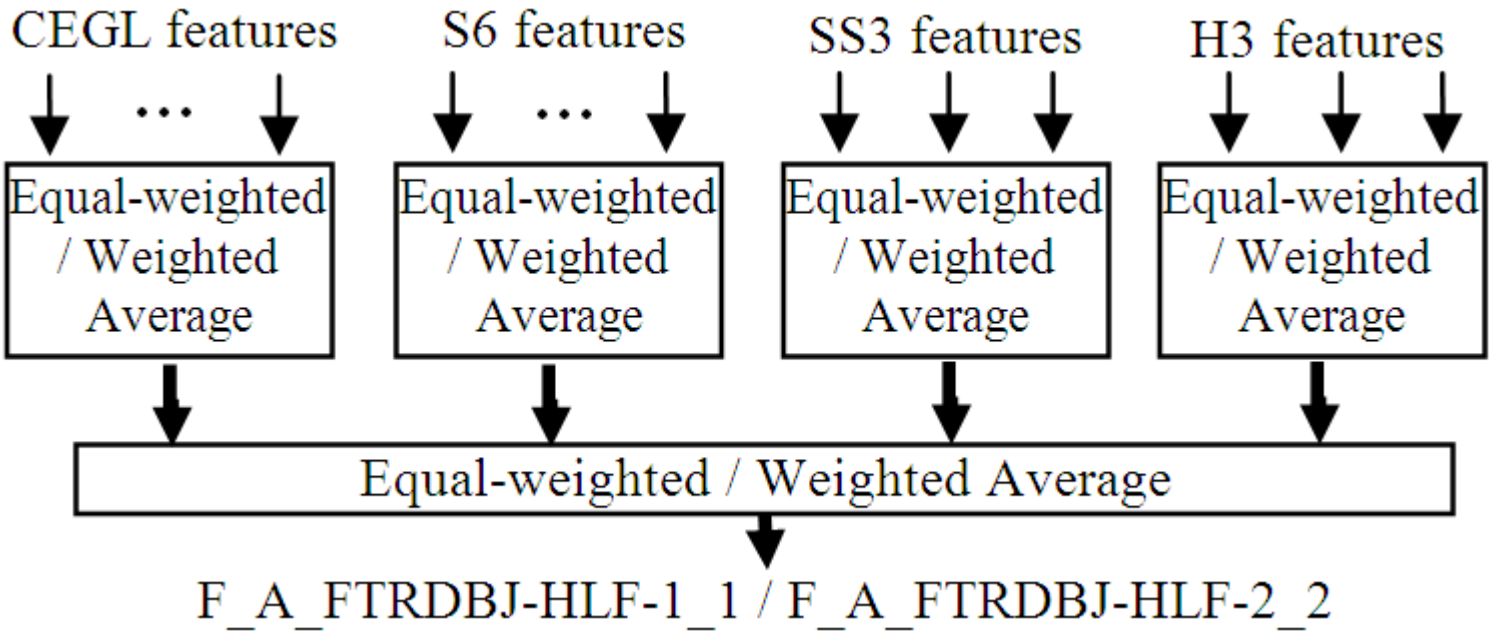
Features

- MAP of different features

Group Name	Feature Name	Dim.	MAP
CEGL	Color Auto-Correlograms (CAC)	256	0.051
	Color Coherence Vector (CCV)	360	0.083
	Grid Color Moments (GCM)	108	0.041
	Edge Coherence Vector (ECV)	320	0.035
	Edge Direction Histogram (EDH)	365	0.047
	Gabor feature (Gabor)	240	0.037
	Local Binary Patterns (LBP)	256	0.051
H3	HOG.HOW	512	0.127
	HOG.2L-PHOW	2560	0.133
	HOG. 3L-PHOW-PLSA	512	0.129

Basic Structure

- 2-Step Late Fusion



- Kernel-level early fusion

Unified Weights

- Motivation
  - Hard to evaluation all 130 concepts×19 features
  - Supported by internal evaluation
- LIBLINEAR were used in all modules of 2-step fusion

Unified Weights

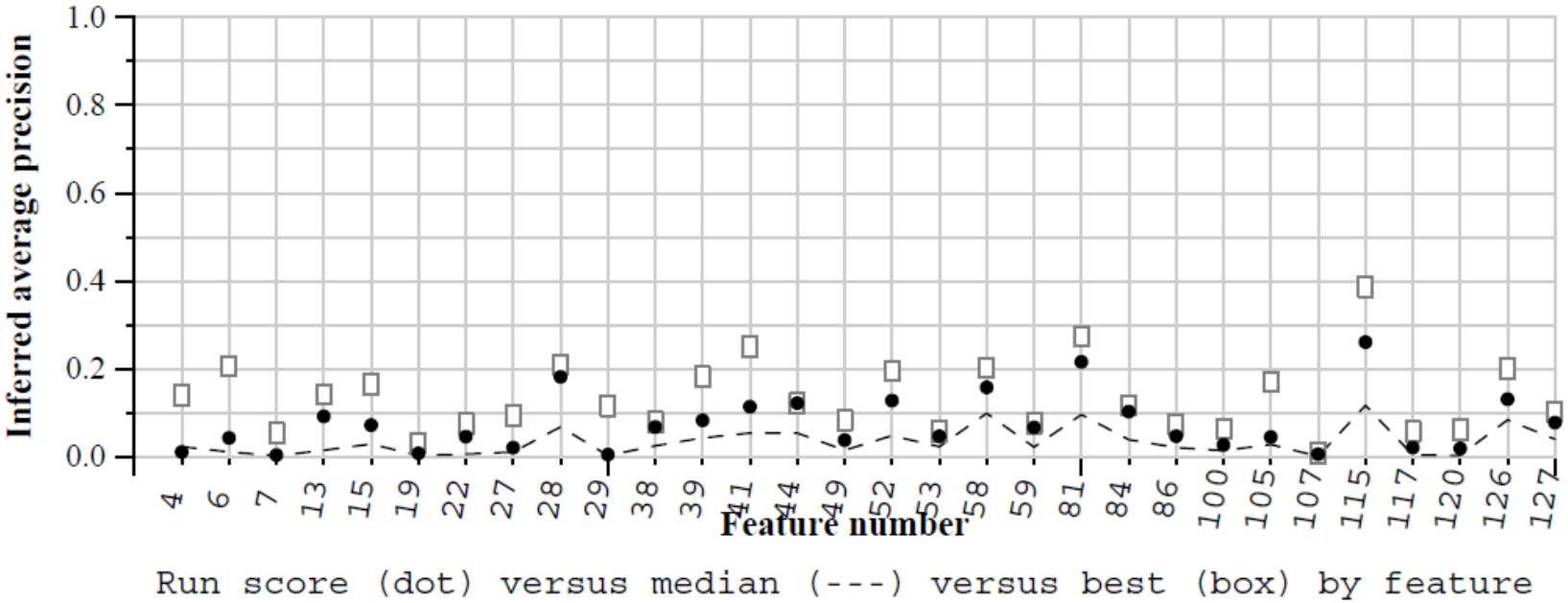
- Results
  - 60% for SVM, 20% for LR, 20% for evaluation

Two-step Fusion Results on FT-30 Corpus

Fusion Method / Group	Equal weighted (MAP)	Logistic regression with respective weights (MAP)	Logistic regression with unified weights (MAP)
CEGL	0.0634	0.0594	0.0640
H3	0.0513	0.0509	0.0513
S3	0.0635	0.0626	0.0636
S6	0.0648	0.0650	0.0640
2-Step Fusion	0.0704	0.0699	0.0694

Unified Weights

- Our best run



- Something more about generalization problem

Cross-domain

- Data level & classifier level
  - 60% for SVM, 20% for weights, 20% for evaluation

Cross-Domain Results based on SS3 and FT-30 Corpus

Model ID	Model Description	MAP
Model-1	Models trained on 2010 dataset	0.084
Model-2	Models trained on 05-09 dataset	0.027
Model-3	Models trained on 2010+05-09 dataset (data-level fusion)	0.075
Model-4	Equal weighted fusion of Model-1and Model-2 (classifier-level fusion)	0.063
Model-5	Weighted Average fusion of Model-1and Model-2 (classifier-level fusion)	0.086

Conclusion & Future works

- Using unified weights is a valuable choice
- The balance between feature numbers and computation cost
- Need further research on cross-domain

Thanks! Any questions?