UNIVERSITY OF MARYLAND

UMIACS

CVPR Workshop in Biometrics 2012

Data Insufficiency in Sketch Versus Photo Face Recognition

17 June 2012

Jonghyun Choi
Abhishek Sharma, David W. Jacobs, Larry S. Davis

Institute of Advanced Computer Studies
University of Maryland, College Park

Sketch-Photo Face Recognition

- Why is it important?
  - Automated criminal search by forensic sketch can reduce the time of crime investigation

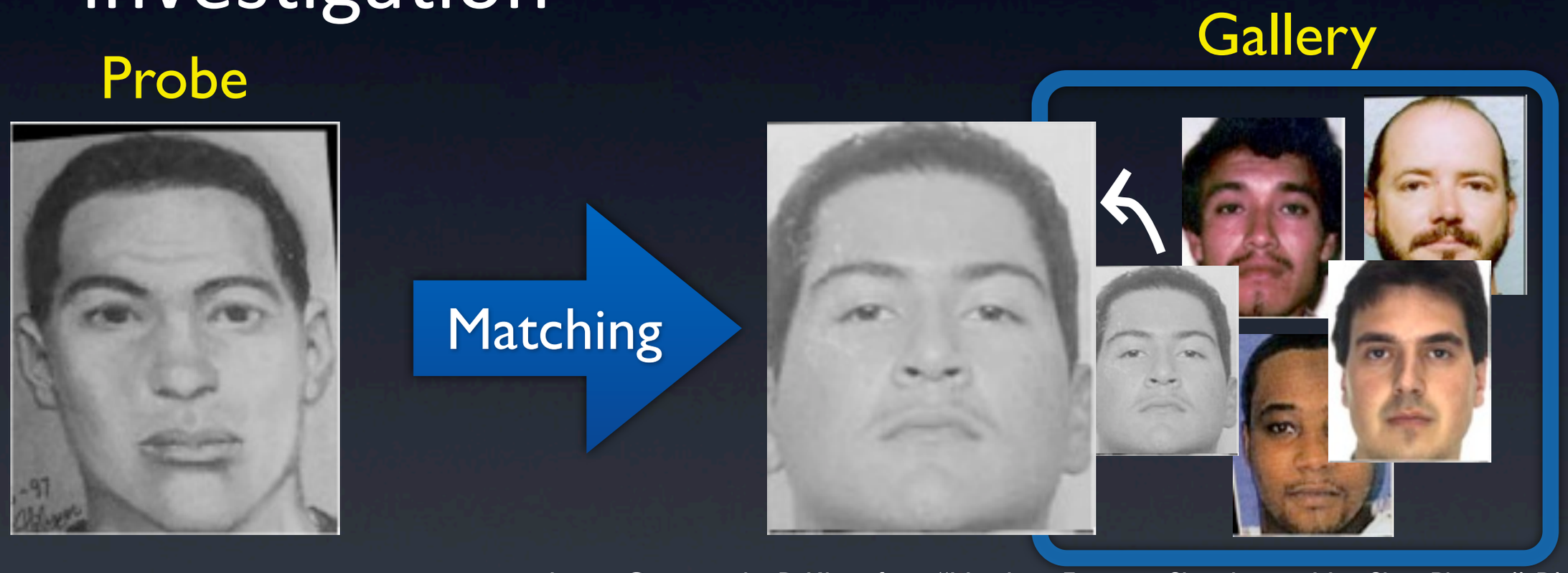
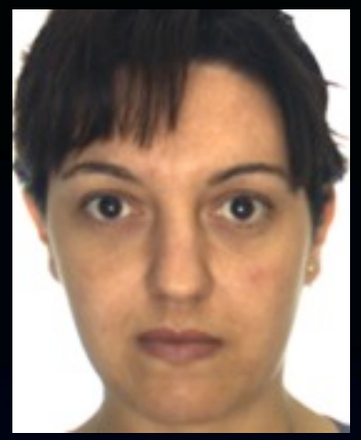
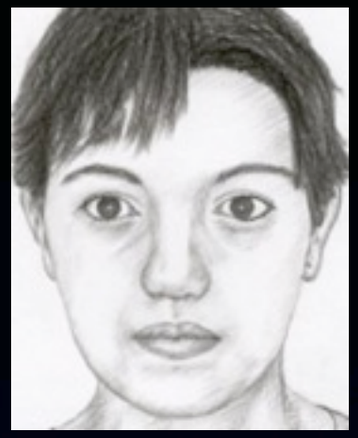


Image Courtesy by B. Klare from "Matching Forensic Sketches to Mug Shot Photos", PAMI 2011

Popular Benchmark in Literature

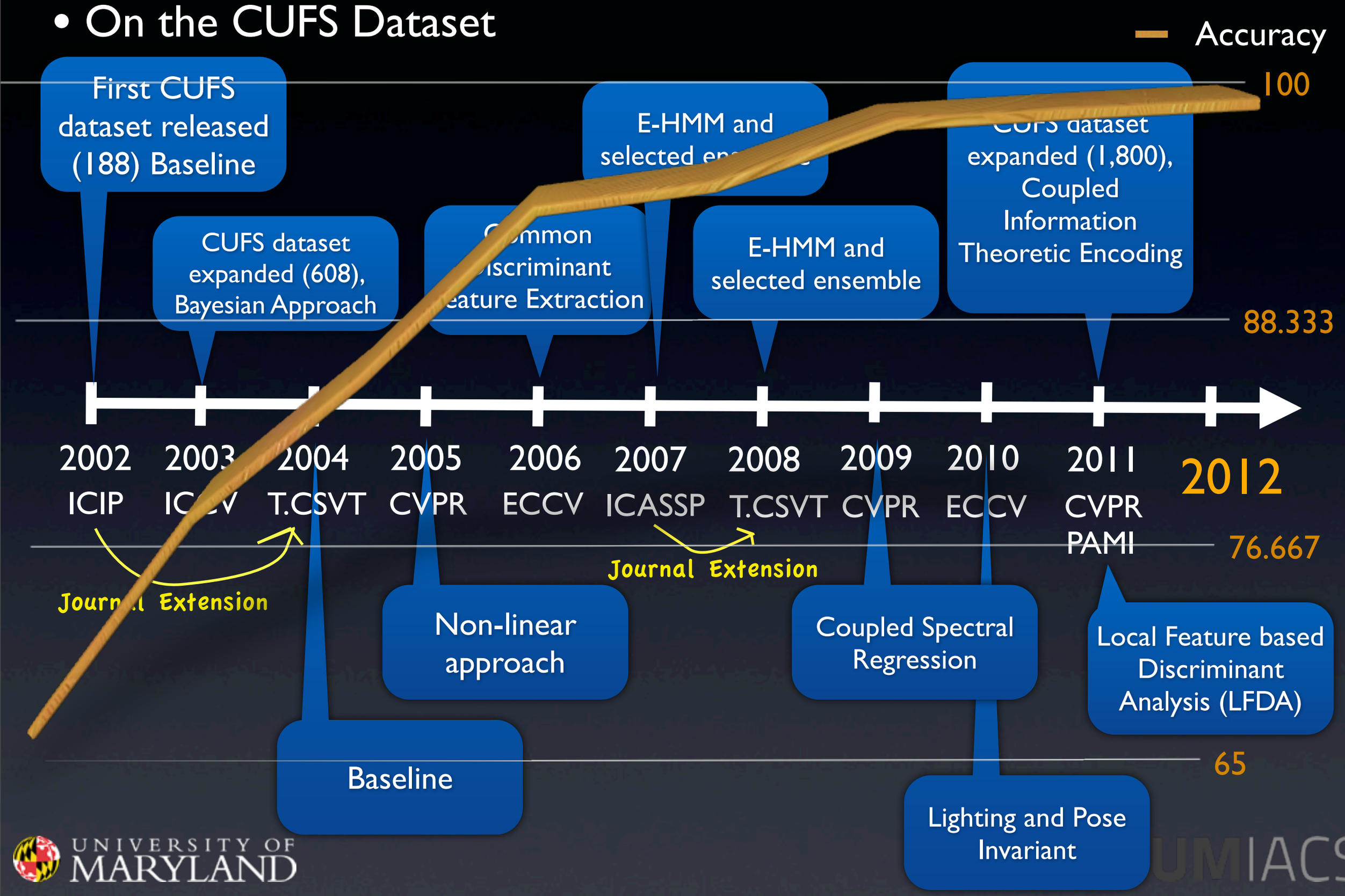



- CUFS dataset[1]
  - Public benchmark dataset
  - Promoting the initial research
    - A controlled dataset
      - ✓ Well lit photos, neutral expression, frontal poses
  - Many approaches evaluated so far

[1] Tang and Wang, "Face Photo Recognition using Sketch", ICIP 2002

Timeline of Research

- On the CUFS Dataset



Summary of Previous Results

Approach	#Train	#Test	Rate (%)
Sketch Synthesis			
Tang and Wang			
Tang and Wang			
Nonlinear			
E-HMM			
MS MRF+LDA	306	300	96.3
MS MRF+LDA	88	10	
MS MRF+W.PCA	88	10	
Modelling Modality G			
PLS-subspace	88	10	
Klare *et al.*	306	300	99.47
CITP	306	300	99.87

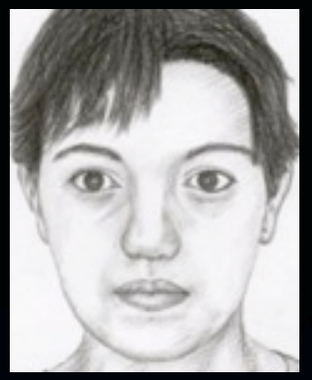
Is the problem solved?

Nearly perfect result!

Yes

Yes for CUFS Dataset

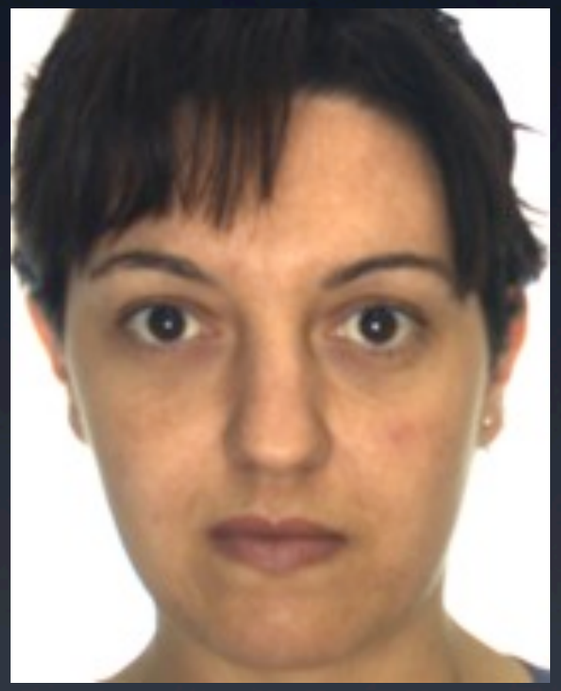
CUFS Dataset

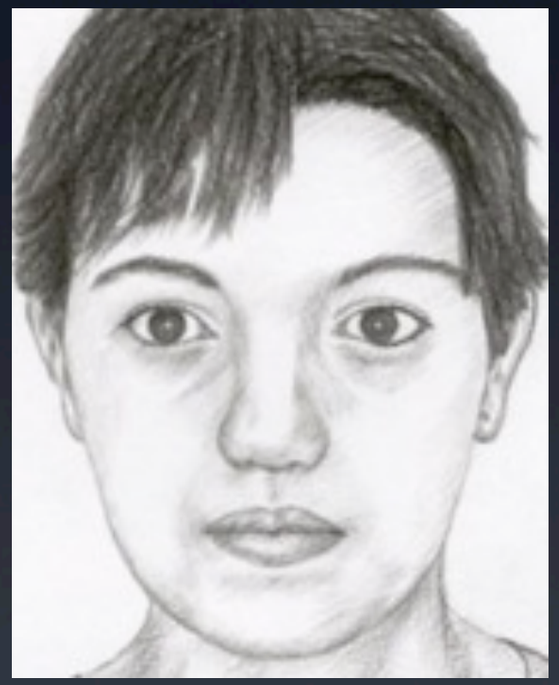



- 606 photo-sketch pairs
  - Random partitioning for training/testing
- Combined with CUFSF (2011) from FERET
  - Total 1,800 photo-sketch pairs
- *viewed sketch* dataset
  - Provides well-aligned photo-sketch pairs
    - Good for analysis of difference in sketch and photo domains without any other factors' interventions

CUFS Dataset (Cont'd)

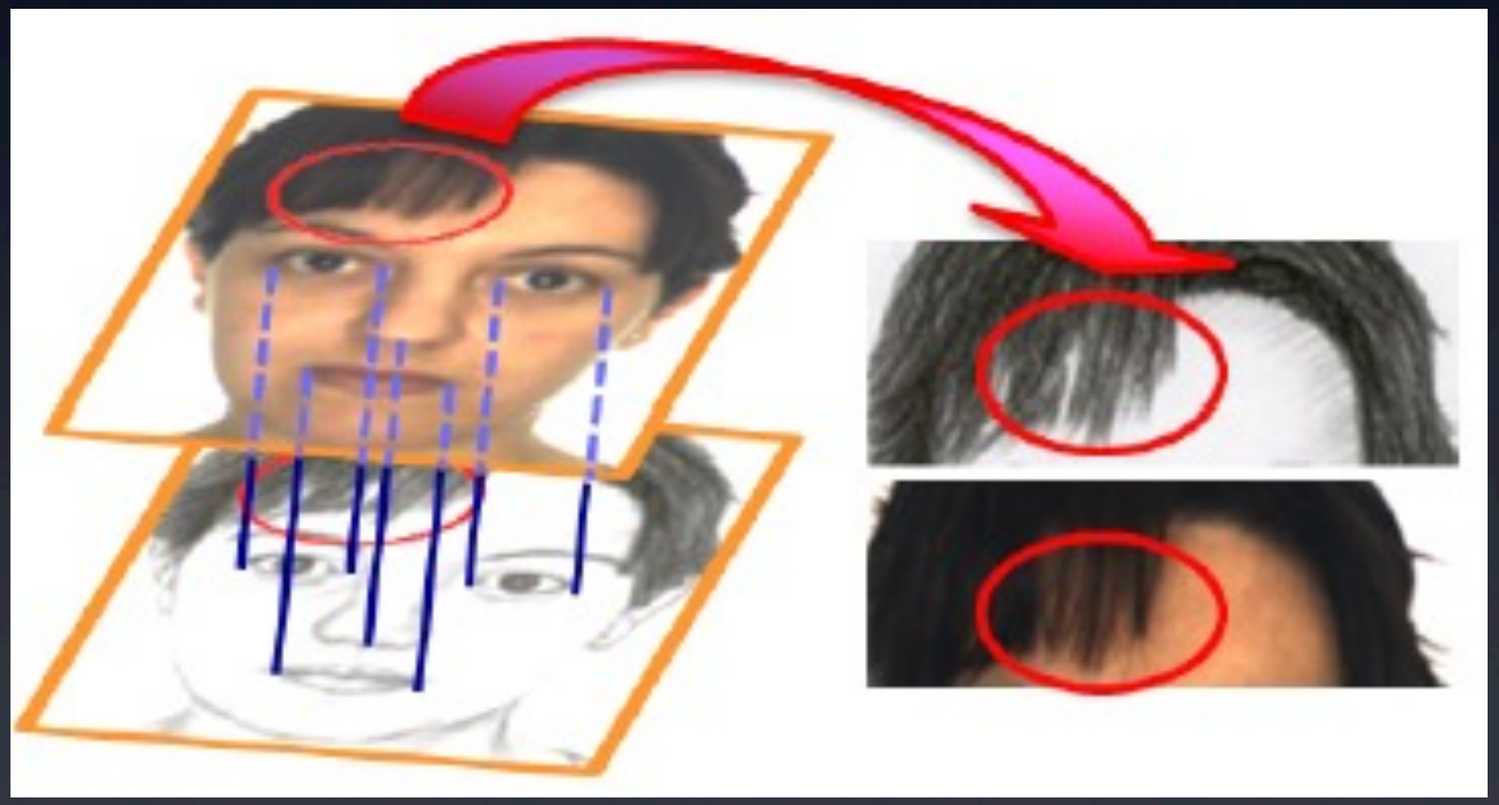
- *viewed sketch* dataset
  - Sketch is drawn by artists
    - Capturing subtle edge similarity (e.g. hair style)
  - Pre-processed to make them well aligned as well





CUFS Dataset (Cont'd)

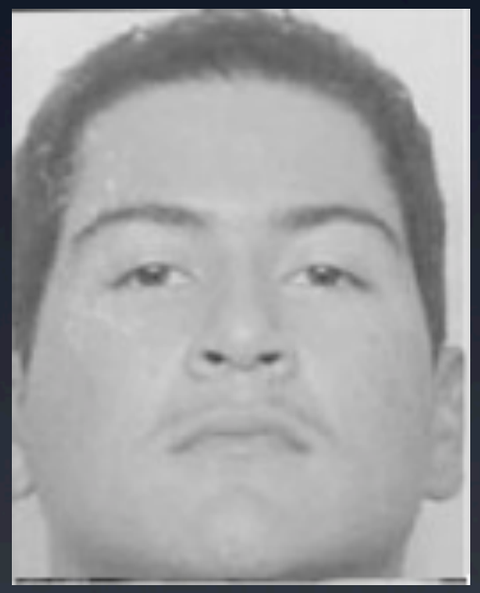
- *viewed sketch* dataset
  - Sketch is drawn by artists
    - Capturing subtle edge similarity (e.g. hair style)
  - Pre-processed to make them well aligned as well



Real-World Scenario for Sketch-Photo Face Recognition

1. Eye-witness describes the criminal's facial traits ***verbally***
2. Forensic artists draw the sketch according to the verbal description
   - *Forensic sketch*[1]

➡ Viewed sketch is not realistic



[1] B. Klare et al.,"Matching Forensic Sketches to Mug Shot Photos", PAMI 2011

*Images courtesy from B. Klare

Insufficiency of the CUFS Dataset

- Well-alignment of CUFS dataset
  - Good for initial research on domain difference (photo-sketch) w/o intervention of other factors
  - But simplifies the problem too much
    - Simple edge matching techniques might work
    - Ignores true variability of sketch-face recognition:
      - ✓ Mis-alignment of fiducial components
      - ✓ Semantic description of shapes of fiducial component
      - ✓ No precise description for subtle difference (*e.g.* hair style)

Today, We Show

- To obtain good result on CUFS dataset
  - Discriminative edge matching technique is enough
    - Outperforms state of the art in face identification setting
  - No effort in reducing modality gap is required
    - Thus no training set except gallery set is required
    - But even outperforms state of the art in bigger set

Discriminative Edge Matching

- Edge features
  - Gabor wavelet response
    - Blurred edge tendency: Macro edge
  - CCS-POP[1]
    - Micro-edgelet
- Discriminative weight on the feature
  - Build a one-vs-all PLS model

[1] Choi et al.,"A Complementary Local Feature for Face Identification", WACV 2012

Partial Least Sqaures

- A supervised dimension reduction technique by maximizing covariance of weighted independent variable (X) and weighted dependent variable (Y)

feature label

$$\hat{w} = \max_{|w|=1} \operatorname{cov}(Xw, Y)^2$$

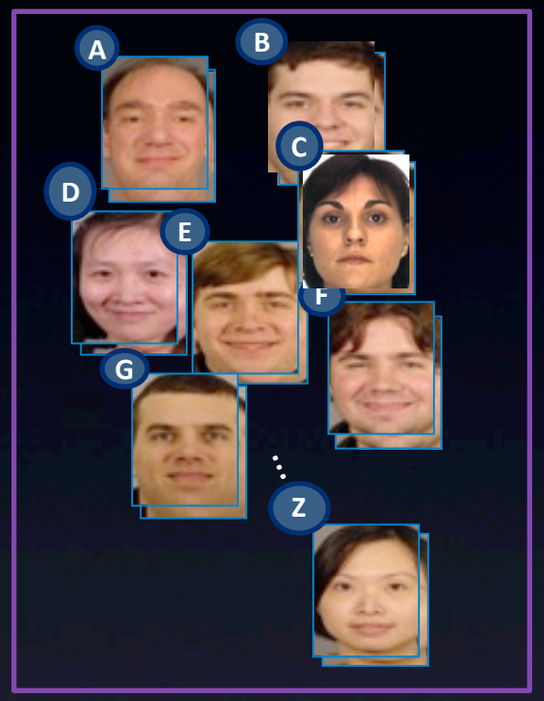
- Using NIPALS algorithm[1] to obtain the regression solution from X to Y

[1] H. Wold, Partial Least Squares, 1985

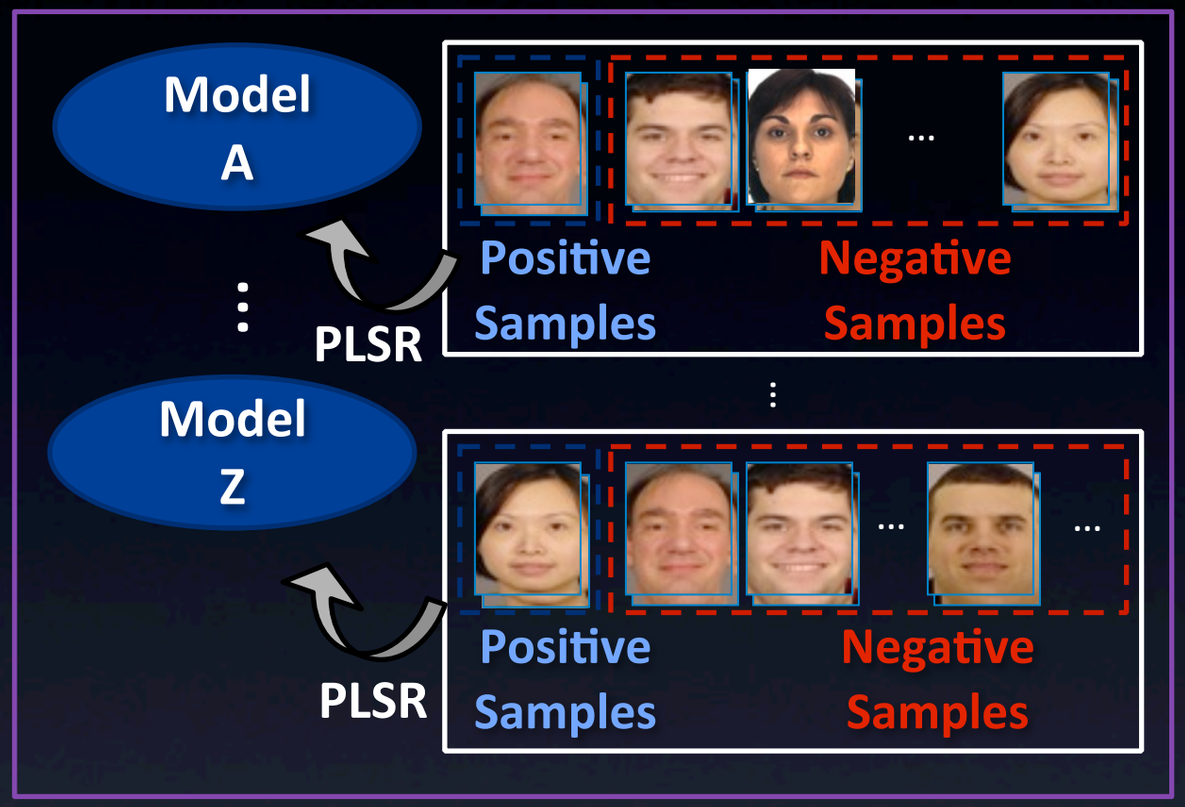
Overall System Diagram[1]

Model Building (Training)

Gallery

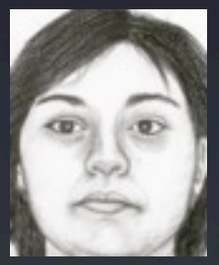


Build "One-vs-All" PLS regression models



Testing

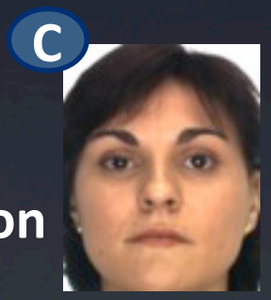
Probe



Regression

Regression responses

Identification Result



[1] Schwartz et al., "A Robust and Scalable Approach to Face Identification", ECCV 2010

Experimental Setup

- CUFS+CUFSF (CUFS) dataset
  - 1,800 pairs of sketch-face
  - No extra training set: Use all pairs for test
  - Photo to Sketch / Sketch to Photo experiments
- Comparison to previous work
- Various image cropping
  - Tight/Loose crop
  - Horizontal/Vertical strip crop
  - Fiducial component crop

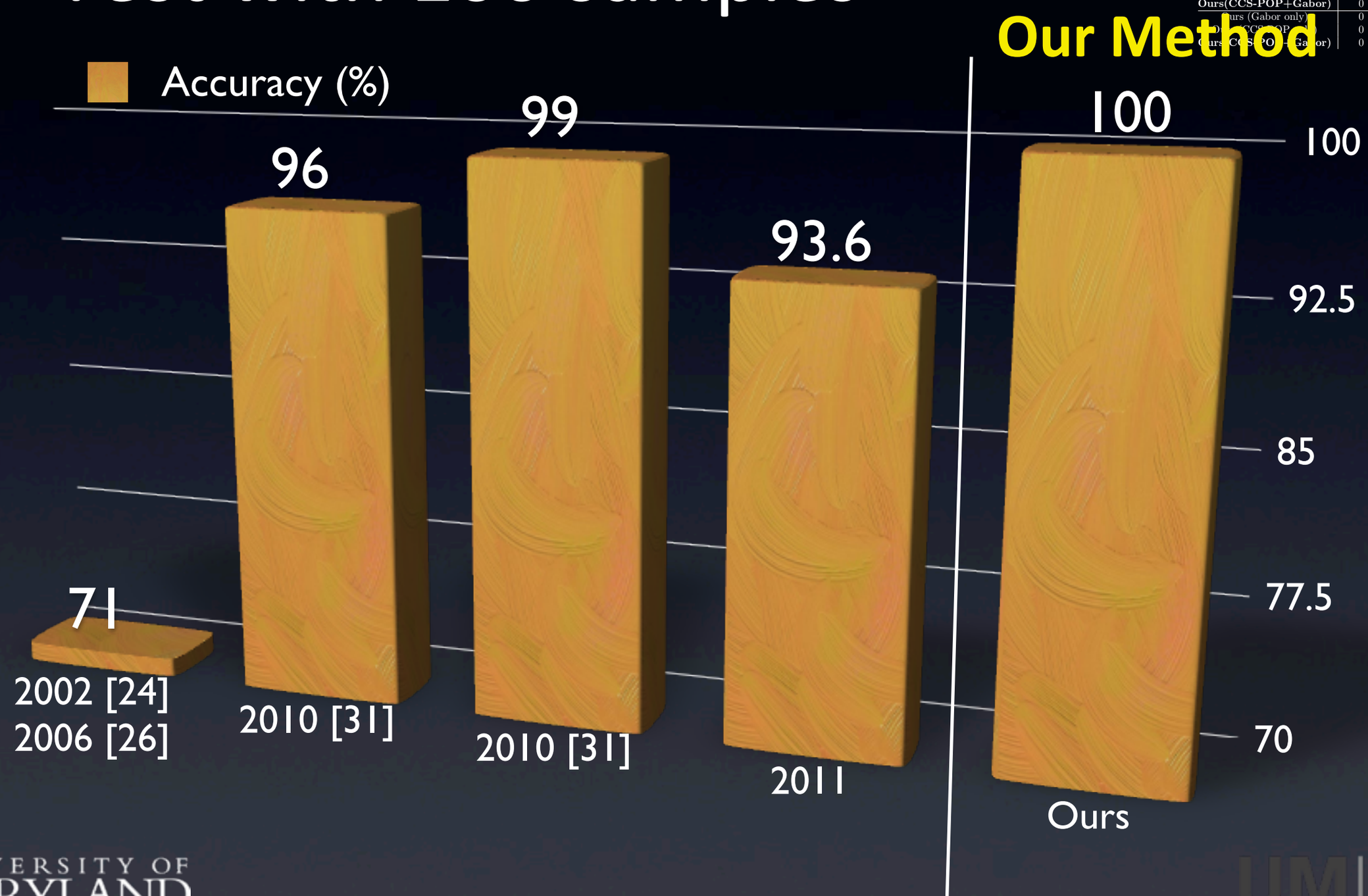
Experimental Results

Approach	#Train	#Test	Rate (%)
Sketch Synthesis			
Tang and Wang [24,26]	88	100	71
Tang and Wang [25]	306	300	81.3
Nonlinear [13]	306	300	87.67
E-HMM [4,35]	306	300	95.24
MS MRF+LDA [30]	306	300	96.3
MS MRF+LDA (from [31])	88	100	96
MS MRF+W.PCA [31]	88	100	99
Modelling Modality Gap			
PLS-subspace [21]	88	100	93.6
Klare*et al.* [9]	306	300	99.47
CITP [32]	306	300	99.87
Ours (Gabor only)	0	300	99.80 ± 0.44
Ours (CCS-POP only)	0	300	95.53 ± 0.90
Ours(CCS-POP+Gabor)	0	300	**100**
Ours (Gabor only)	0	1,800	99.50
Ours (CCS-POP only)	0	1,800	96.28
Ours(CCS-POP+Gabor)	0	1,800	**99.94**

Experimental Results

Approach	#Train	#Test	Rate (%)
Sketch Synthesis			
Tang and Wang [24,26]	88	100	71
Tang and Wang [25]	306	300	81.3
Nonlinear [13]	306	300	87.67
E-HMM [4,35]	306	300	95.24
MS MRF+LDA [30]	306	300	96.3
MS MRF+LDA (from [31])	88	100	96
MS MRF+W.PCA [31]	88	100	99
Modelling Modality Gap			
PLS-subspace [21]	88	100	93.6
Klare*et al.* [9]	306	300	99.47
CITP [32]	306	300	99.87
Ours (Gabor only)	0	300	99.80 ± 0.44
Ours (CCS-POP only)	0	300	95.53 ± 0.90
Ours(CCS-POP+Gabor)	0	300	**100**
Ours (Gabor only)	0	1,800	99.50
Ours (CCS-POP only)	0	1,800	96.28
Ours(CCS-POP+Gabor)	0	1,800	**99.94**

- Test with 100 samples

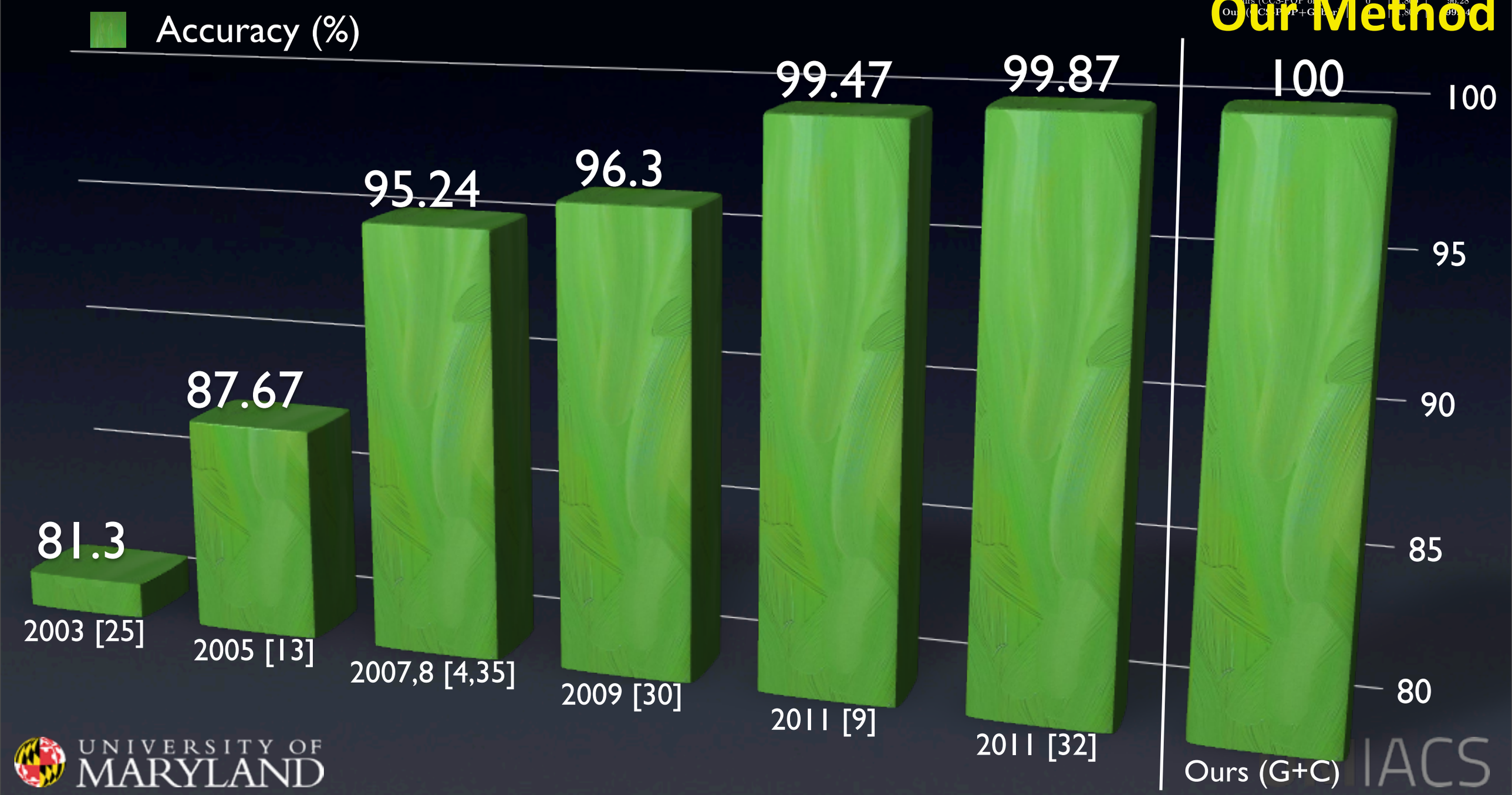


Experimental Results

- Test with 300 samples

Approach	#Train	#Test	Rate (%)
Sketch Synthesis			
Tang and Wang [24,26]	88	100	71
Tang and Wang [25]	306	300	81.3
Nonlinear [13]	306	300	87.67
E-HMM [4,35]	306	300	95.24
MS MRF+LDA [30]	306	300	96.3
MS MRF+LDA (from [31])	88	100	96
MS MRF+W.PCA [31]	88	100	99
Modelling Modality Gap			
PLS-subspace [21]	88	100	93.6
Klare*et al.* [9]	306	300	99.47
CITP [32]	306	300	99.87
Ours (Gabor only)	0	300	99.80 ± 0.44
Ours (CCS-POP only)	0	300	95.53 ± 0.90
Ours(CCS-POP+Gabor)	0	300	**100**
Ours (Gabor only)	0	1,800	99.50
Ours (CCS-POP only)	0	1,800	96.28
Ours (CCS-POP+Gabor)	0	1,800	**99.94**

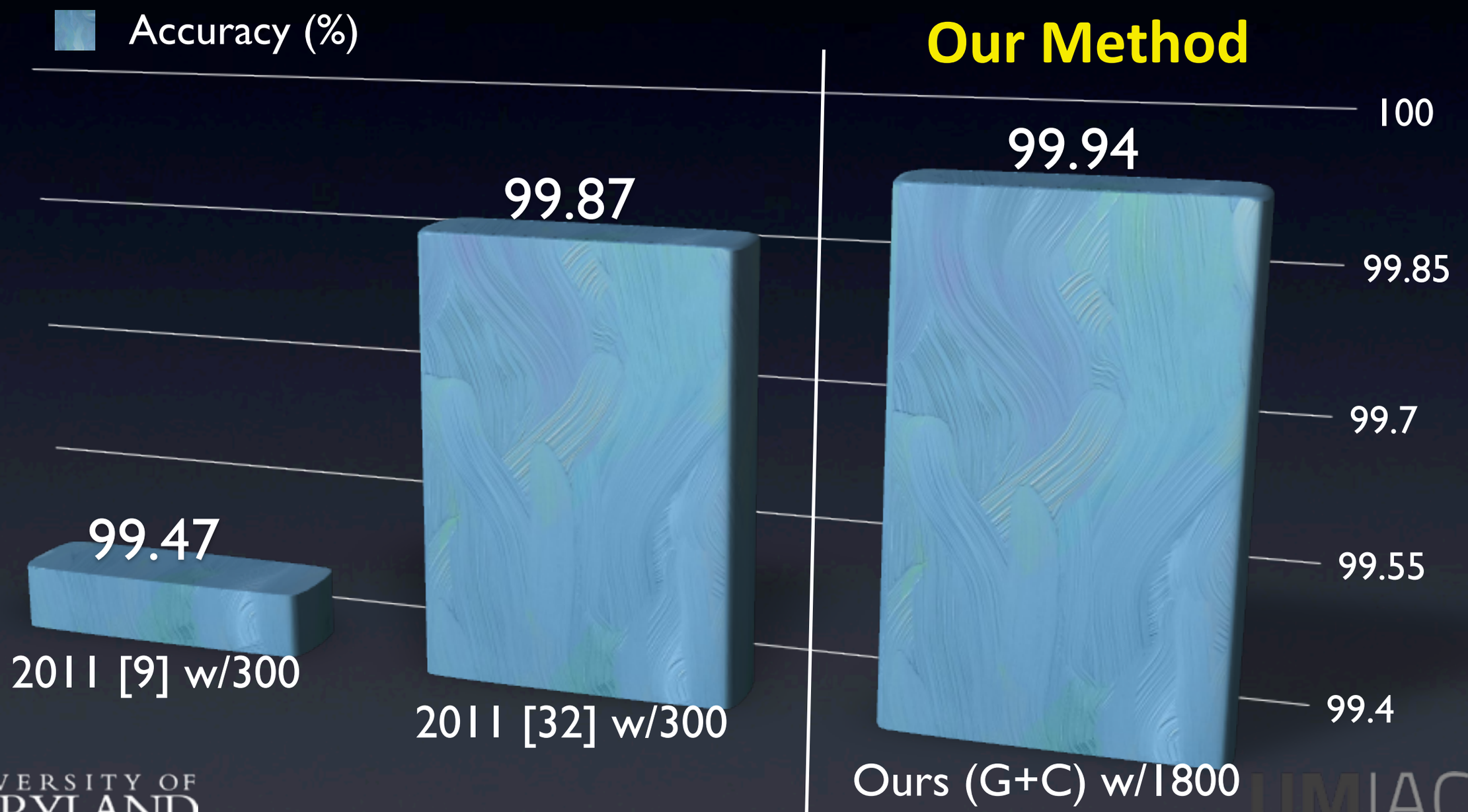
Our Method



Experimental Results

Approach	#Train	#Test	Rate (%)
Sketch Synthesis			
Tang and Wang [24,26]	88	100	71
Tang and Wang [25]	306	300	81.3
Nonlinear [13]	306	300	87.67
E-HMM [4,35]	306	300	95.24
MS MRF+LDA [30]	306	300	96.3
MS MRF+LDA (from [31])	88	100	96
MS MRF+W.PCA [31]	88	100	99
Modelling Modality Gap			
PLS-subspace [21]	88	100	93.6
Klare*et al.* [9]	306	300	99.47
CITP [32]	306	300	99.87
Ours (Gabor only)	0	300	99.80 ± 0.44
Ours (CCS-POP only)	0	300	95.53 ± 0.90
Ours(CCS-POP+Gabor)	0	300	**100**
Ours (Gabor only)	0	1,800	99.50
Ours (CCS-POP only)	0	1,800	96.28
Ours(CCS-POP+Gabor)	0	1,800	**99.94**

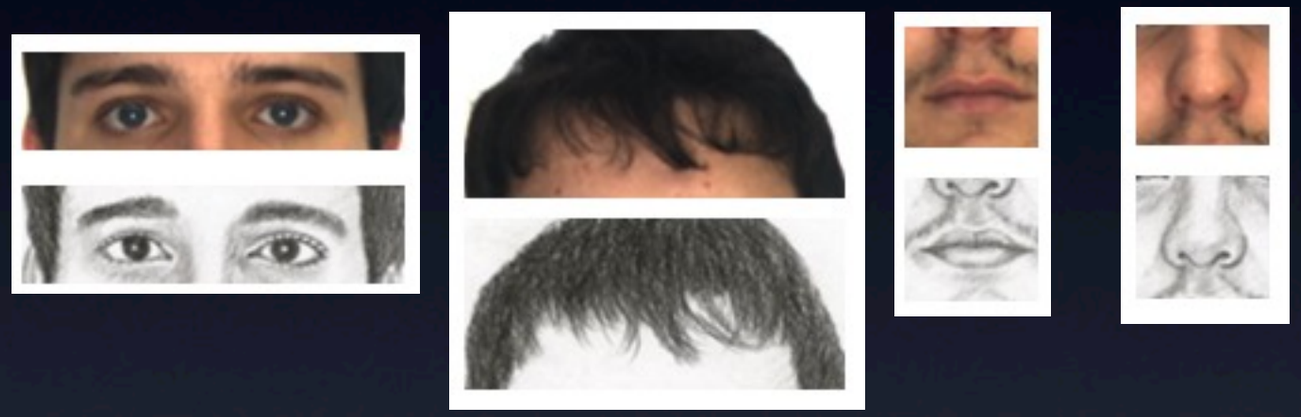
- Test with 1,800 samples
  - Compared to the results tested with 300 samples



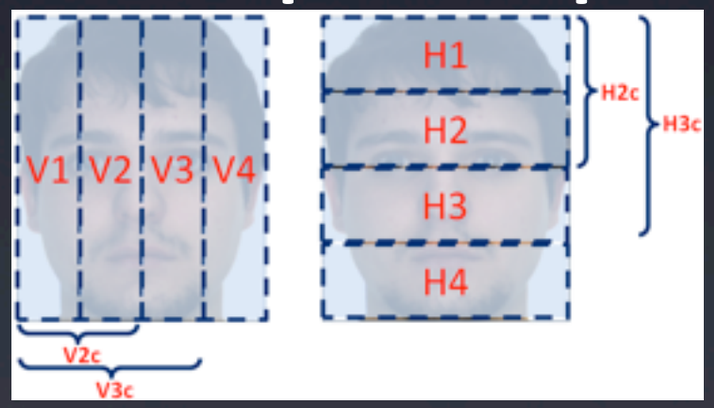
Different Cropping

- Tight/Loose Crop

- Fiducial Component Crop



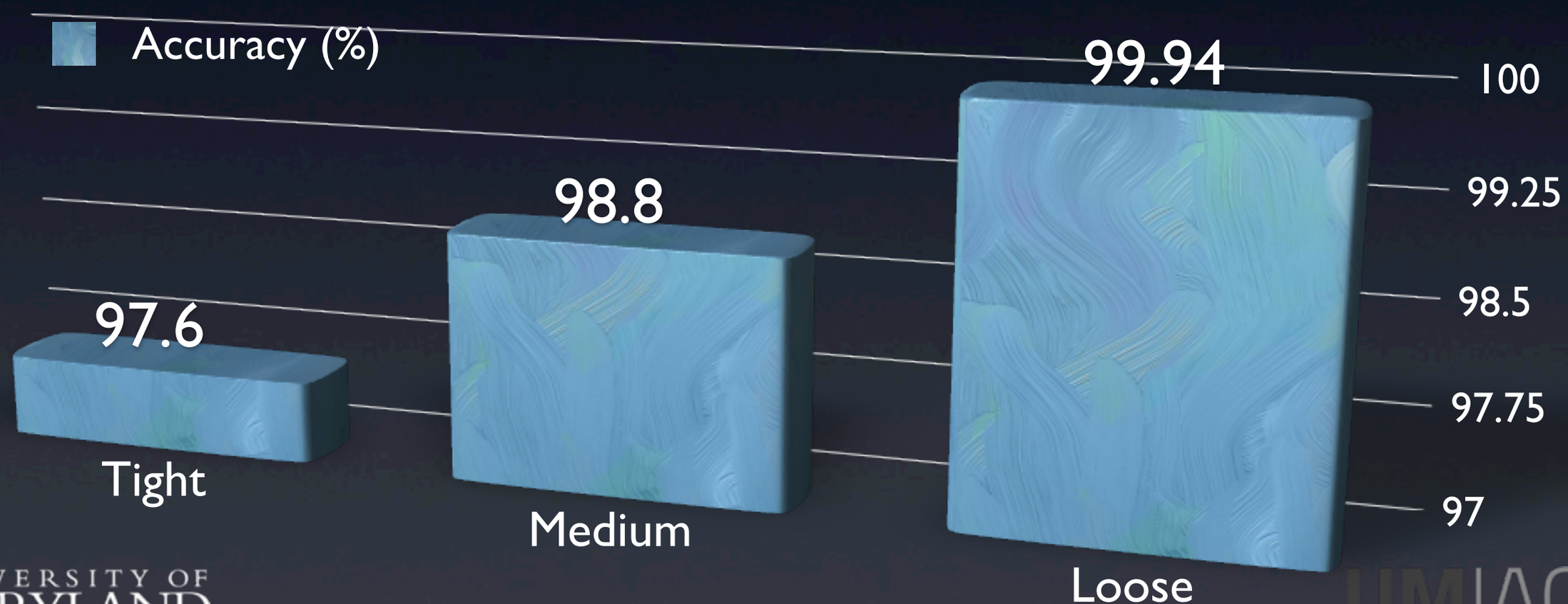
- Horizontal/Vertical Strip Crop



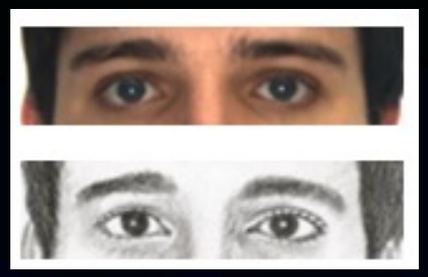
Results on Tight/Loose Cropping

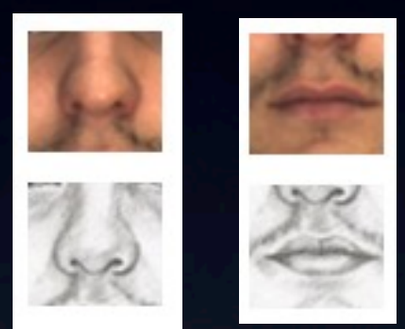
Table 1. Comparison of rank 1 face identification rate (%) according to components of a face. G+C denotes a combined feature of Gabor and CCS-POP. "S to P" denotes Sketch to Photo and "P to S" denotes Photo to Sketch.

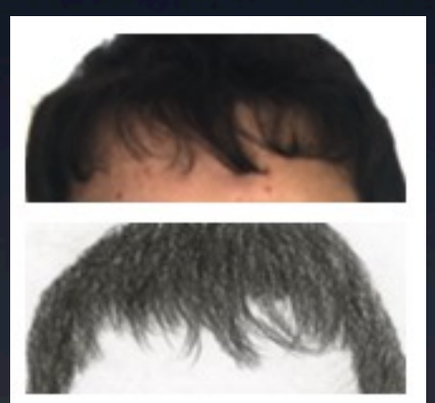
Feature	Gabor		G+C	
Regions	P to S	S to P	P to S	S to P
Tight	95.78	96.33	97.2	97.6
Medium	96.90	97.80	98.16	98.8
Loose	98.06	99.50	99.39	**99.94**



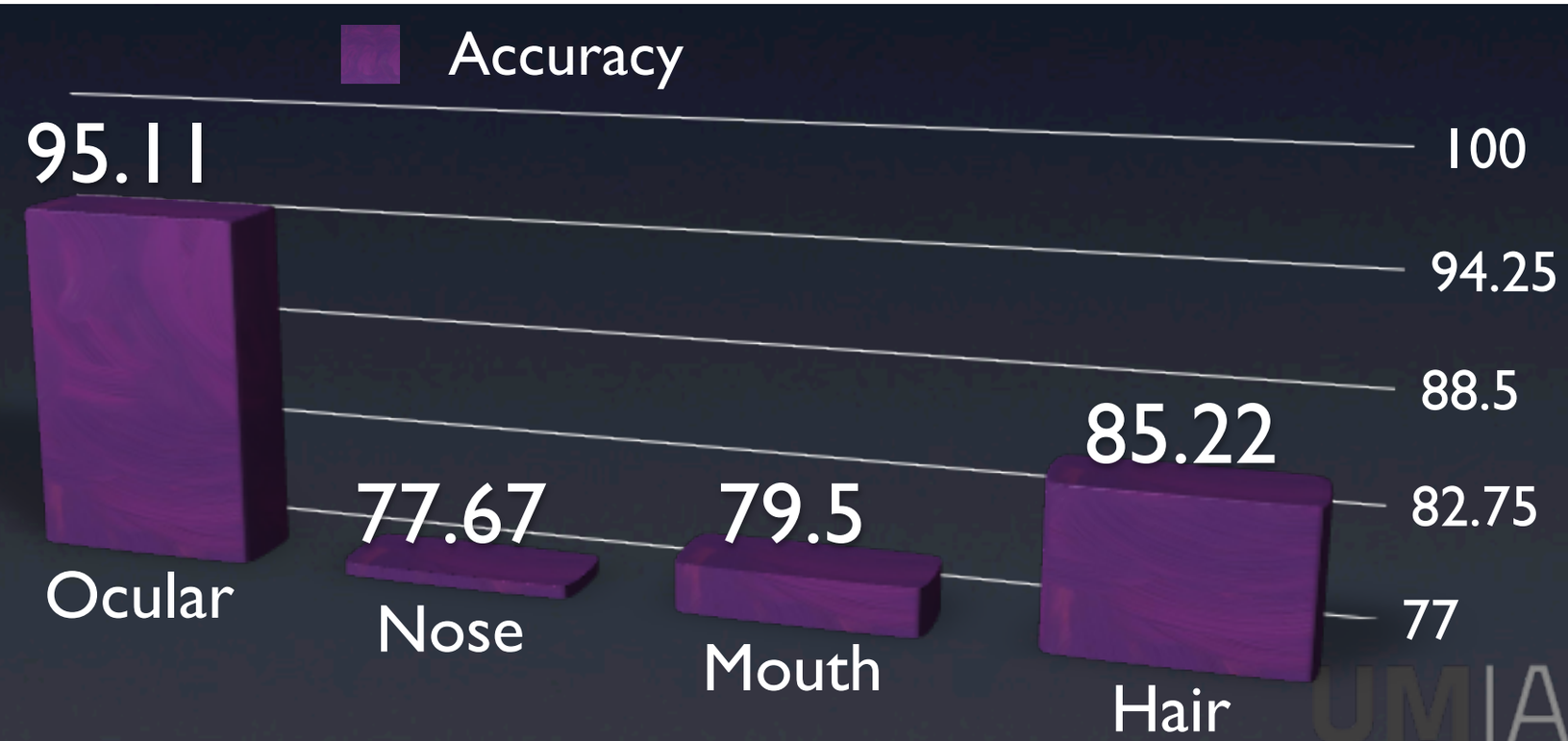
Results on Fiducial Component Cropping



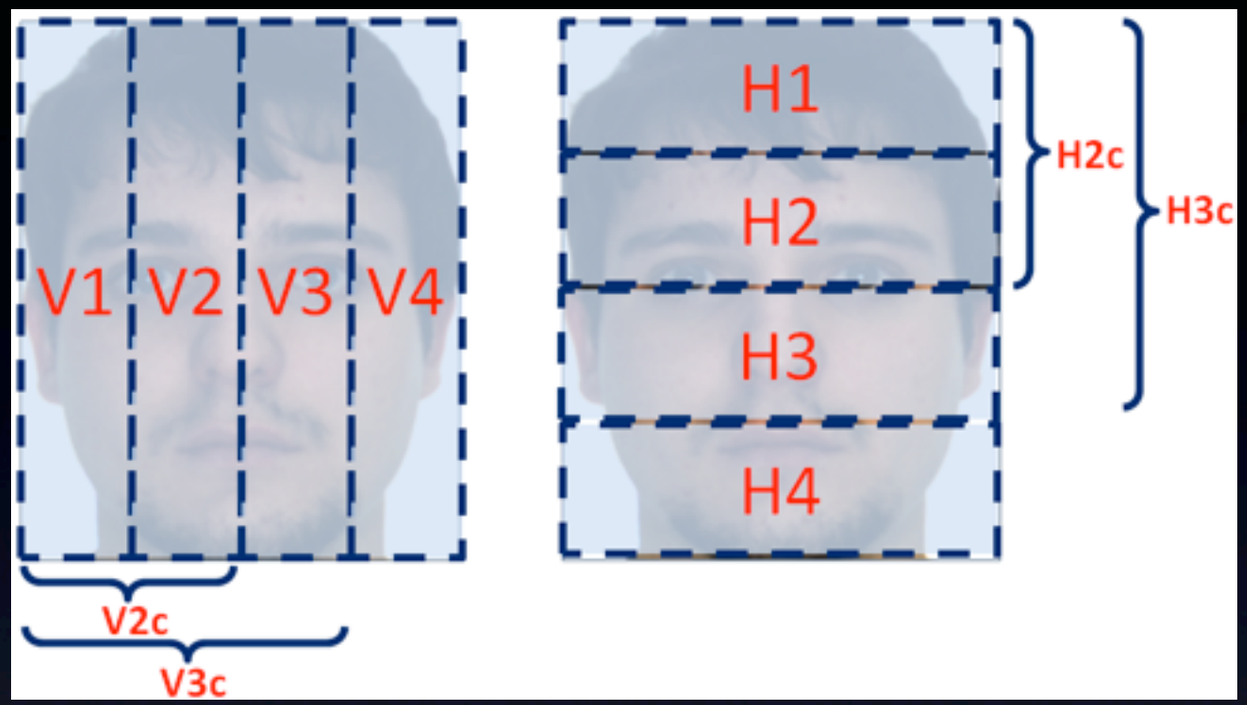




Feature	Gabor		G+C	
Regions	P to S	S to P	P to S	S to P
ocular	93.50	**93.83**	93.89	**95.11**
nose	74.11	75.00	76.06	77.67
mouth	77.11	77.39	76.94	79.50
hair	79.33	81.78	81.67	85.22



Results on Strip Cropping



Feature	Gabor		G+C	
Region ID	P to S	S to P	P to S	S to P
H1	76.28	78.00	75.28	78.83
H2	90.94	92.56	93.00	94.50
H3	90.06	92.06	91.89	94.39
H4	86.28	89.17	86.44	90.00
V1	89.22	91.00	89.83	93.94
V2	90.61	90.83	92.50	94.22
V3	89.89	91.33	92.06	94.28
V4	87.33	89.83	87.22	90.44
H1	76.28	78.00	75.28	78.83
H2c	97.5	98.67	99.06	99.72
H3c	97.61	99.22	99.50	99.89
full(H4c)	98.06	99.50	99.39	**99.94**
V1	89.22	91.00	89.83	93.94
V2c	98.33	99.11	99.27	99.67
V3c	98.22	99.17	99.38	99.72
full(V4c)	98.06	99.50	99.39	**99.94**

Discussion & Conclusion

- A simple discriminative edge analysis can perform well in overly-reduced problem of sketch-photo matching
- Now is the time to move on to more challenging dataset
- We suggest a guideline for new dataset (Please refer to our paper)

Thank you!

Q/A

Authors are supported by

MURI Grant N00014-08-10638 from Office of Naval Research