Fixation and Saccade based Face Recognition from Single Image per Person with Various Occlusions and Expressions

Xingjie Wei and Chang-Tsun Li

Department of Computer Science, University of Warwick, UK

x.wei@warwick.ac.uk

http://warwick.ac.uk/xwei

THE UNIVERSITY OF WARWICK

Unconstrained face recognition

- Partial occlusions and expression changes
- Intra-class variations > inter-class variations

→→ poor recognition performance !

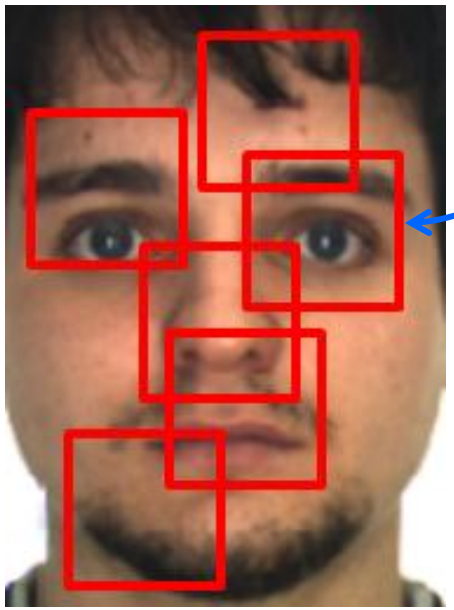
Single sample per person problem

- Real-world scenarios
  - law enforcement, driver license or passport card identification
- Learning-based method
  - will suffer because the training samples are very limited

Our method

- Does not require a training phase
- Matches faces by simulating the mechanism of fixations and saccades in human visual perception
- Is robust to to local deformations of faces

Face recognition by humans



- Fixation

- Saccade: quickly jump between fixations

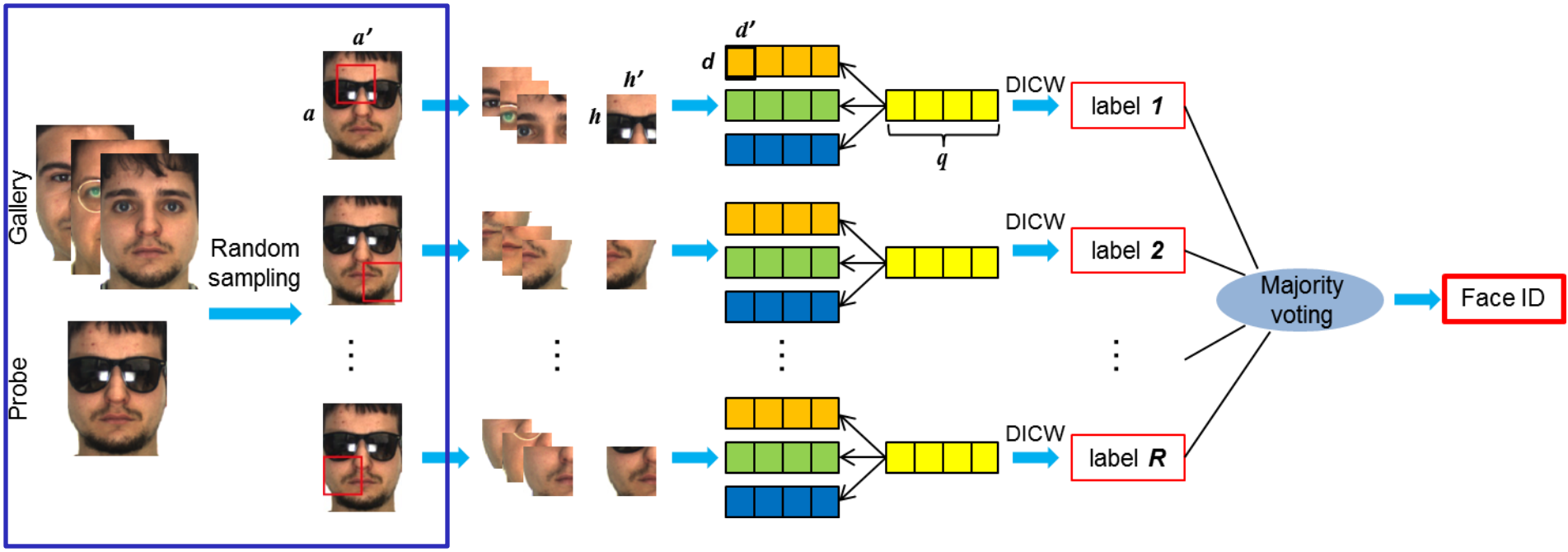
Three key observations

- Humans scan a series of fixations instead of the whole face

  – not all local areas are helpful due to occlusion and expression variations

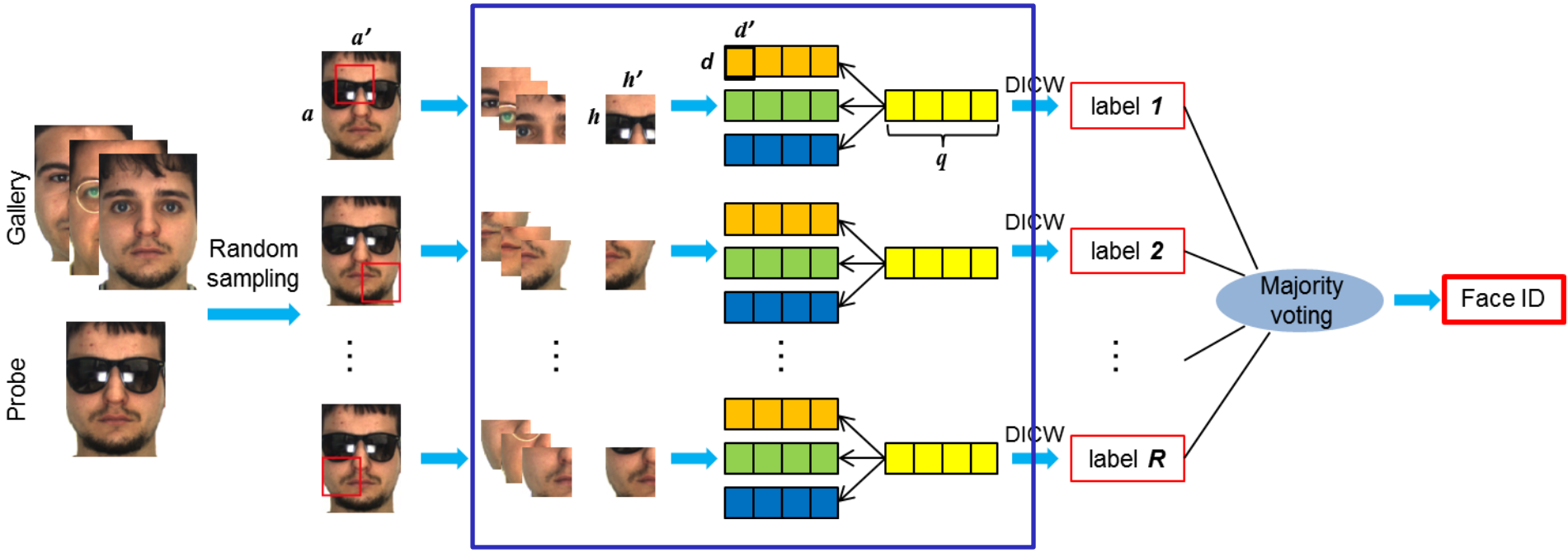
- Law of large numbers
  – one fixation → not sufficient
  – a large field of random selections: √
- The spatial relationship between facial features is very important

Framework

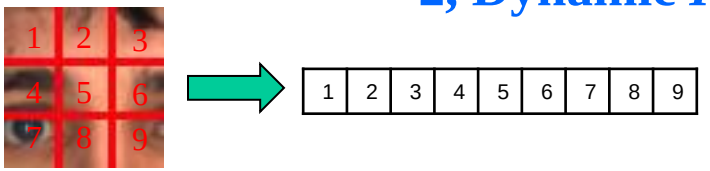


1, Simulating fixations and saccades by random sampling

Framework



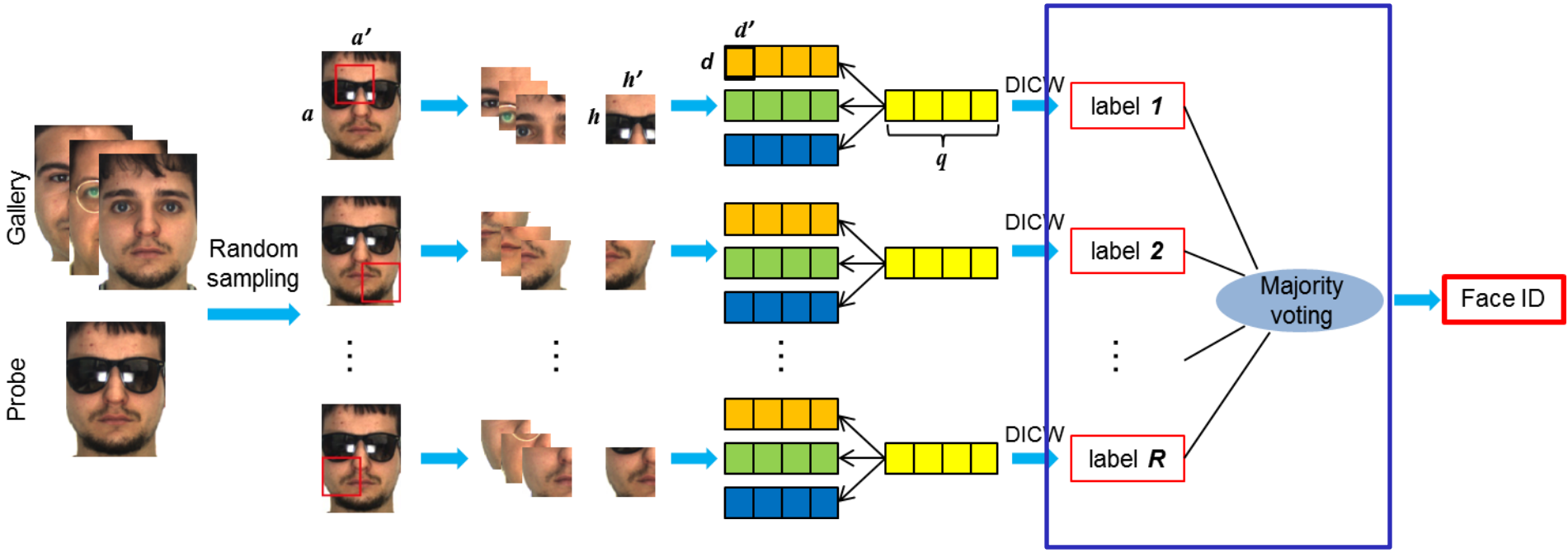
2, Dynamic *Image-to-Class* Warping



- Each fixation is represented as a sequence
- Consider the facial order information

X.Wei, C.-T. Li, and Y. Hu. *Face recognition with occlusion using dynamic image-to-class warping (DICW)*. IEEE International Conference on Automatic Face and Gesture Recognition (FG'13), 2013

Framework



3, Majority voting for fixations

Majority voting for fixations

- Each fixation has the possibilities to classify the face correctly or wrongly
- Correct recognition rate = probability of the consensus being correct
- Combined decision is wrong only if a majority of the fixations votes are wrong and they all make the *same* misclassification

Experimental results

- Randomly located occlusions
  - FRGC database, 100 subjects, 2 sessions
  - Single gallery image per person
  - Random occlusion: 0%~50%

Experimental results

- Randomly located occlusions

Table 1. Recog. rates (%) on the FRGC database

Occlusion	0%	10%	20%	30%	40%	50%
LSVM[4]	69.5	65.8	57.3	36.8	36.8	22.3
SRC-block[29]	65.8	55.8	47.8	39.8	32.5	22.8
DICW[27]	79.3	77.8	77.3	72.8	70.8	64.5
Ours	**84.2**	**82.4**	**80.3**	**78.1**	**73.2**	**69.6**

Experimental results

- Real disguise and expressions
- AR database, 100 subjects, 2 sessions
  - Gallery: neural expression face from session 1

Experimental results

Table 2. Rec. rates (%) on the AR database (occlusion)

Method	Session 1		Session 2		Avg.	M/H[1]
	Sung.	Scarf	Sung.	Scarf		
Stringface[5]	88.0	96.0	76.0	88.0	87.0	No
FARO[6]	90.0	85.0	-	-	87.5	No
SRC-block[29]	86.0	87.0	49.0	70.0	73.0	No
PD[22]	98.0	90.0	-	-	94.0	No
SOM[23]	97.0	95.0	60.0	52.0	76.0	No
CTSDP[8]	-	-	-	-	90.6	No
DICW[27]	**99.0**	97.0	93.0	81.0	92.5	No
$PWCM_{0.5}$[11]	97.0	94.0	72.0	71.0	83.5	Yes
CTSDP[8]	-	-	-	-	**98.5**	Yes
Ours	**99.0**	**98.7**	**93.7**	**94.9**	96.6	No

[1] Occlusion mask/threshold training required

Experimental results

Table 3. Rec. rate (%) on the AR database (expression)

Method	Session 1			Session 2			Avg.
	Sm.	An.	Sc.	Sm.	An.	Sc.	
Stringface[5]	87.5	87.5	25.9	-	-	-	67.0
FARO[6]	96.0	-	60.0	-	-	-	78.0
SRC[29]	98.0	89.0	55.0	79.0	78.0	31.0	71.7
PD[22]	**100.0**	97.0	93.0	88.0	86.0	63.0	87.8
SOM[23]	**100.0**	98.0	88.0	88.0	90.0	64.0	88.0
DMMA[16]	99.0	93.0	69.0	85.0	79.0	45.0	78.3
DICW[27]	**100.0**	99.0	84.0	91.0	92.0	44.0	85.0
CTSDP[8]	**100.0**	**100.0**	**95.5**	**98.2**	**99.1**	**86.4**	**96.5**
Ours	**100.0**	**100.0**	91.4	94.5	98.0	58.6	90.4

lower computational complexity than CTSDP

Thank you

- Questions ?

- Xingjie Wei
- x.wei@warwick.ac.uk
- http://warwick.ac.uk/xwei
- Department of Computer Science, University of Warwick, UK