Face Recognition with Occlusion Using Dynamic *Image-to-Class* Warping (DICW)

Xingjie Wei, Chang-Tsun Li and Yongjian Hu

Department of Computer Science, University of Warwick, Coventry, CV4 7AL, UK

x.wei@warwick.ac.uk http://warwick.ac.uk/xwei

1. Introduction

Problem: Face recognition with occlusion

- Intra-class variations > inter-class variations
- Causes imprecise registration of faces

→→ Poor recognition performance!

Challenges: Why is it so difficult?

- No prior knowledge of occlusion
- Location,size,shape,texture -- unpredictable!

Our method:

- No occlusion detection
- No data-dependent training
- Works well with limited gallery images per person
- Efficient and appropriate for real applications

2. Method

Inspired by the **time series analysis** technique

Example: two similar time sequences

A = (3, 1, 10, 5, 6)

B = (3, 2, 1, 10, 5)

Bit-wise matching: distance = $\sqrt{0+1+81+25+1} \approx 10.39$

Warping:

A = (3, 1, 10, 5, 6)

B = (3, 2, 1, 10, 5)

Warping — distance is largely reduced

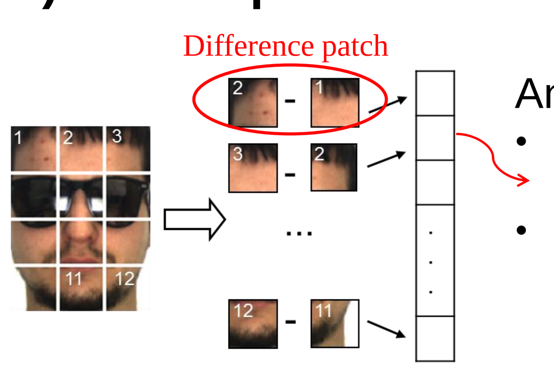
However, cross-matching is not allowed:

A = (3, 1, 10, 5, 6)

B = (3, 2, 1, 10, 5)

needs to maintain the *order* information

1) Face representation



An image → a patch sequence

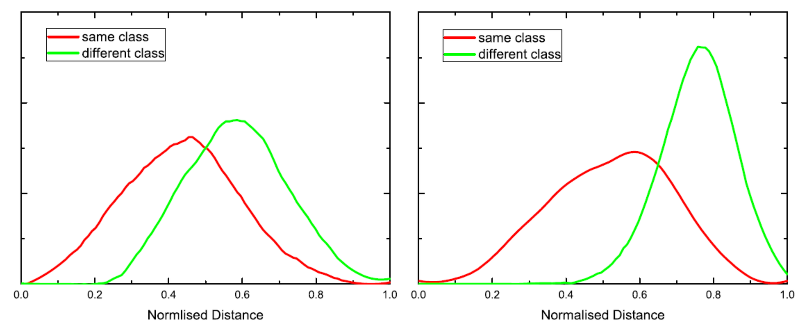
- Contains the *order* information of facial features
- Considers the inherent structure of the face

A face consists of forehead, eyes, nose, mouth and chin in a natural order

does not change despite occlusion or imprecise registration

Why use the difference patches?

- To enhance the detailed textured regions
- To enhance the *order* information of patches



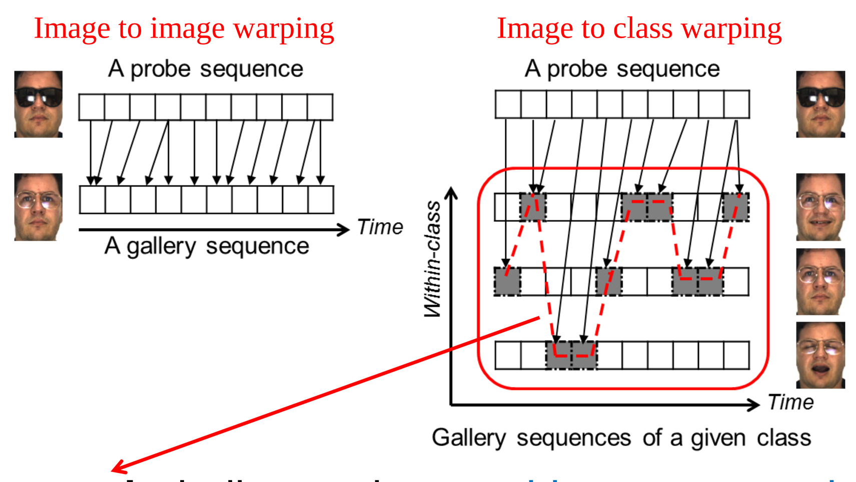
Using the original patches   Using the difference patches

2)Matching

Image to image warping →→ image to class warping

Compute the *Image-to-Class* distance

- From a probe sequence to all the gallery sequences of an enrolled class
- Each probe patch can be matched with patches from different gallery sequences



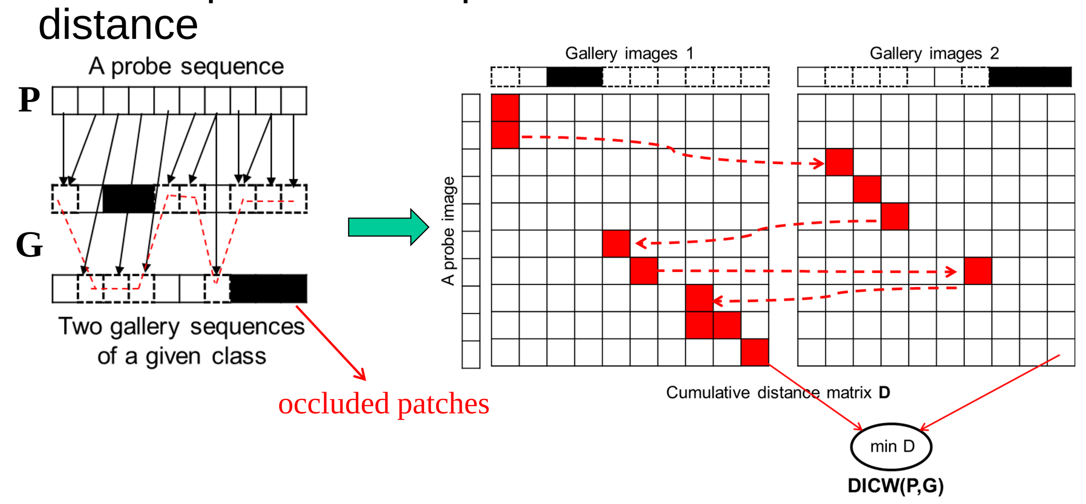
Warping path: indicates the matching correspondence of patches between a probe sequence to all gallery sequences of a given class

- Constraints: Boundary, Continuity, Monotonicity & Window →→ maintain the *order* information

How to compute the *Image-to-Class* distance?

Dynamic Programming (DP)

- Computes the cumulative distance for all paths
- Each step selects the predecessor with the minimal distance



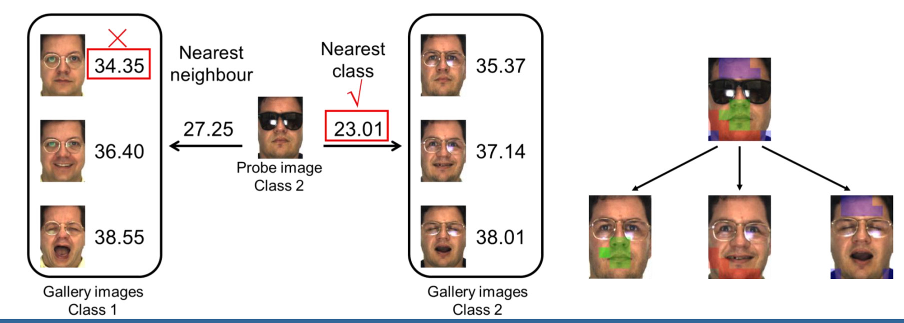
Why does it work?

- Tries every possible warping path and select the one with minimal overall distance

→→ the large distance causes by occluded patches won't be considered

- Exploits the information from different gallery images

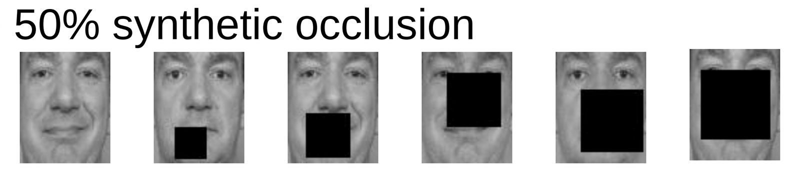
→→ image-to-class distance ↓



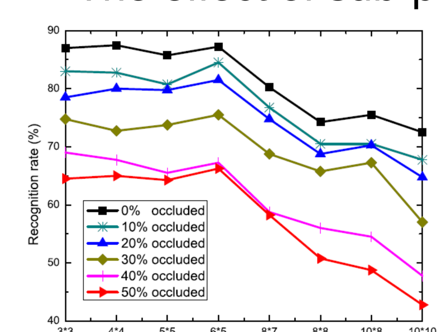
3. Results

Randomly located occlusions : FRGC

- 0% ~ 50% synthetic occlusion

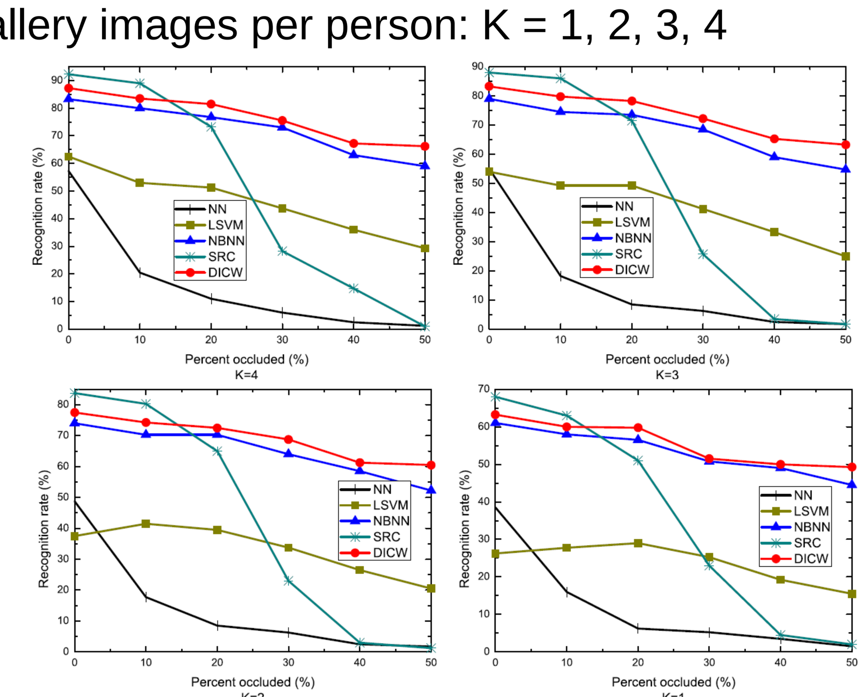


- The effect of sub-patch size



- Robust to different patch size in an appropriate range
- Smaller size, higher computational & memory cost
- Sizes from 4 × 4 to 6 × 5 pixels are recommended

- Gallery images per person: K = 1, 2, 3, 4

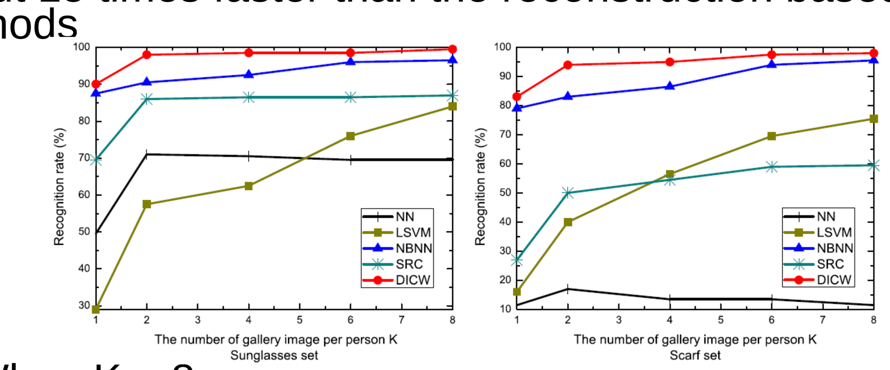


- SRC performs slightly better than DICW at the beginning then drops significantly
- Patch based NBNN and DICW perform better than others on the whole
- When K = 1, *Image-to-Class* distance degenerates into the *Image-to-Image* distance

Real disguise: AR

- Gallery images per person: K = 1, 2, 4, 6, 8

About 15 times faster than the reconstruction based methods

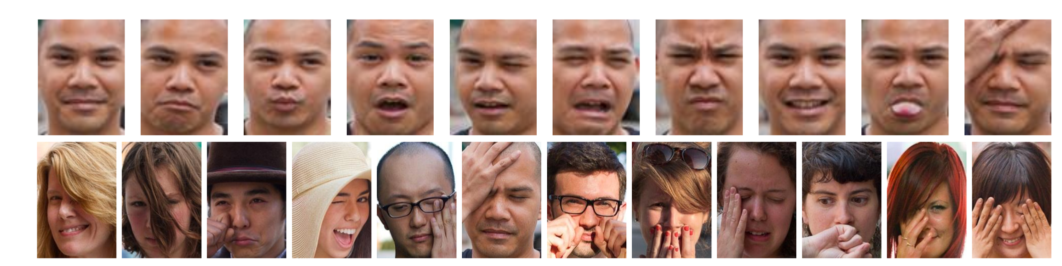


- When K = 8

	Sunglasses	Scarf	Average
SRC-partition	97.5	93.5	95.5
LRC	96.0	26.0	61.0
CRC-RLS	91.5	95.0	93.3
R-CRC	92.0	94.5	93.3
SEC-MRF	99.0~**100**	95.0~ 97.5	97~**98.8**
l_{struct}	99.5	87.5	93.5
Proposed DICW	99.5	**98.0**	**98.8**

The best recognition rate reported

Realistic images



The face we make (www.thefacewemake.org)

- Frontal view faces of strangers on the streets, captured under totally uncontrolled conditions in real environment
- Face images are not well registered
- Includes hand occlusion which is difficult to be detected by skin colour based models

Gallery size per person	1	3	5	8
NN	56.4	66.4	70.9	71.8
LSVM	19.1	30.0	41.8	48.2
NBNN	52.7	66.4	74.6	73.6
SRC	60.0	67.3	70.9	70.9
Proposed DICW	**61.8**	**76.4**	**77.3**	**81.8**

Conclusion

The proposed DICW:

- is robust to various types of occlusion
- performs consistently even when only single gallery image is available for each person
- can adopt other image descriptors such as LBP and Gabor