# Expression Robust 3D Face Recognition by Matching Multi-component Local Shape Descriptors on the Nasal and Adjoining Cheek Regions

Jiangning Gao, Adrian N Evans

Department of Electronic and Electrical Engineering
University of Bath, Bath, UK

# Outline

- Introduction and background
  - Why expression robust?
  - Why the nasal region and its environs?
  - Why surface normals?
- Preprocessing and landmarks localization
- Multi-component feature extraction
  - Local shape descriptor
  - Local patches evaluation
  - Square local shape descriptor
- Conclusions and future work

# Introduction and background

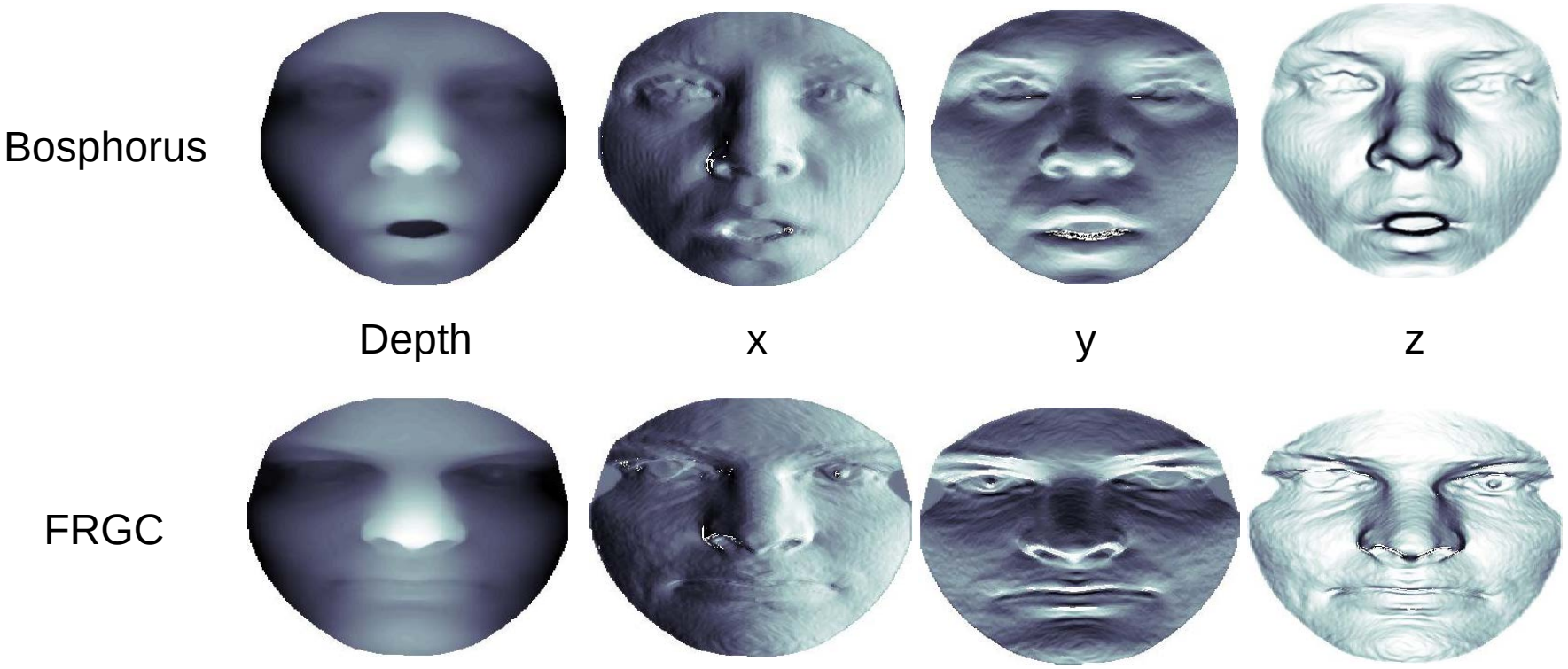
- **Why expression robust?**
- **Why the nasal region?**
  - Contains discriminative features
  - Easy to detect
  - Robustness to hair occlusions
  - Difficult to conceal and easy to align
  - Relatively expression invariant .... but not all parts!
- **Why its environs?**
  - The adjoining cheek may also contain additional discriminative features

[7]. J. Gao, M. Emambakhsh, and A. N. Evans, “A Low Dimensionality Expression Robust Rejector for 3D Face Recognition,” in Proc. IEEE International Conference on Pattern Recognition, Sweden, pp. 506-511, 2014.

# Introduction and background

## Why use the surface normals?

In addition to depth, the surface normal of each point determines the orientation of the surface and contains information on local shape variations.



# Databases

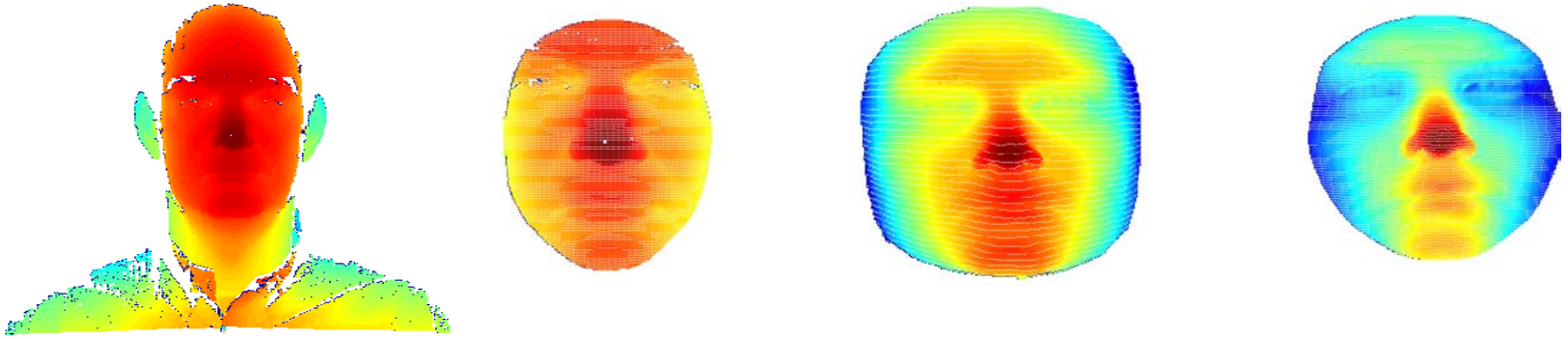
UNIVERSITY OF BATH

- FRGC
  - 557 unique subjects, 4950 captures
  - Spring2003 (Gallery)

    Fall2003 and Spring2004 (Probe)

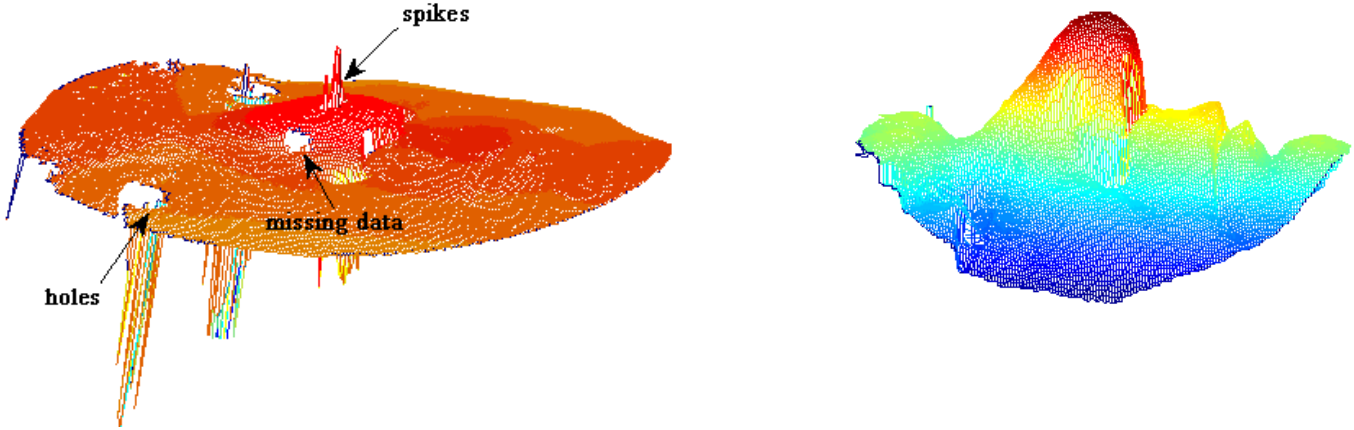
- Bosphorus
  - 105 unique subjects, ~3000 captures
  - Various expressions: neutral, happy, surprise, fear, sadness, anger and disgust

# Preprocessing



FRGC

Bosphorus




[2]. M. Emambakhsh, A. Evans, and M. Smith, “Using nasal curves matching for expression robust 3D nose recognition,” in Proc. Biometrics: Theory, Applications and Systems, 2013.

[3]. A. S. Mian, M. Bennamoun, and R. Owens, “An Efficient Multimodal 2D-3D Hybrid Approach to Automatic Face Recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 29, pp. 1927-1943, 2007.

# Landmark localization

24 landmarks on the nasal and adjoining region are automatically detected

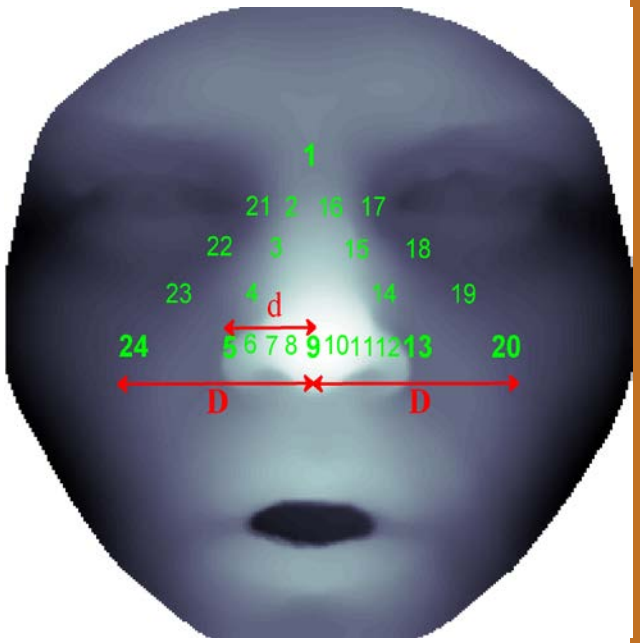
- To crop the nasal region (L1 to L16)
  - Nose tip (L9)
  - Saddle (L1)
  - Alar (L5, L13)

Emambakhsh *et al*., BTAS, 2013

- To explore the adjoining cheeks (L17 to L24)
- L20 and L24 localization
  - Proportional distance
  - Constant distance

Gao *et al*., ICPR, 2014

- 3 horizontal curves
  - L1822
  - L1923
  - L2024

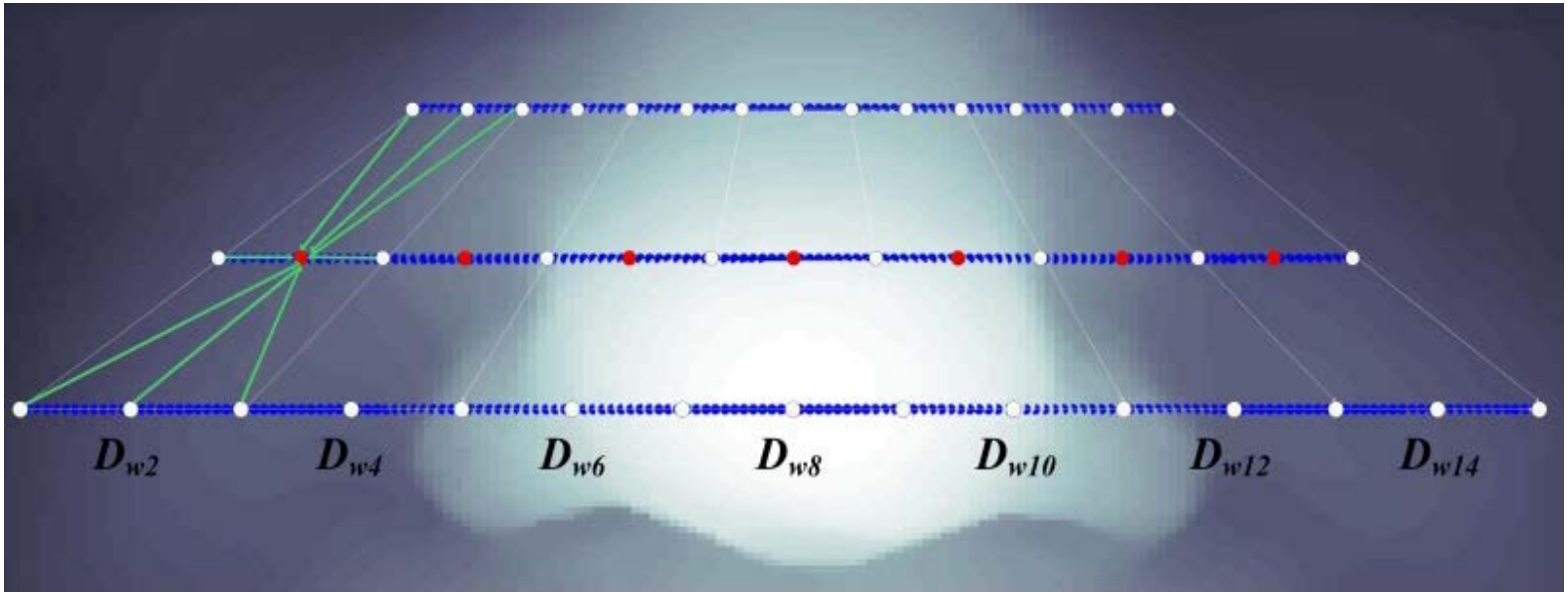



[2]. M. Emambakhsh, A. Evans, and M. Smith, “Using nasal curves matching for expression robust 3D nose recognition,” in Proc. Biometrics: Theory, Applications and Systems, 2013.

[7]. J. Gao, M. Emambakhsh, and A. N. Evans, “A Low Dimensionality Expression Robust Rejector for 3D Face Recognition,” in Proc. IEEE International Conference on Pattern Recognition, Sweden, pp. 506-511, 2014.

# Proposed local shape descriptor

- Three Curves
  - Top (L18L22), middle (L19L23) and bottom (L20L24)
  - Resampled to 15 points per curve
- 7 windows
- Depth differences between the central points (red points) on the middle curve and the surrounding points (white points) on the top, middle and bottom curves.
- Feature size per capture: $8 \times 7 = 56$



# Proposed local shape descriptor

## Difference outperforms direct depth

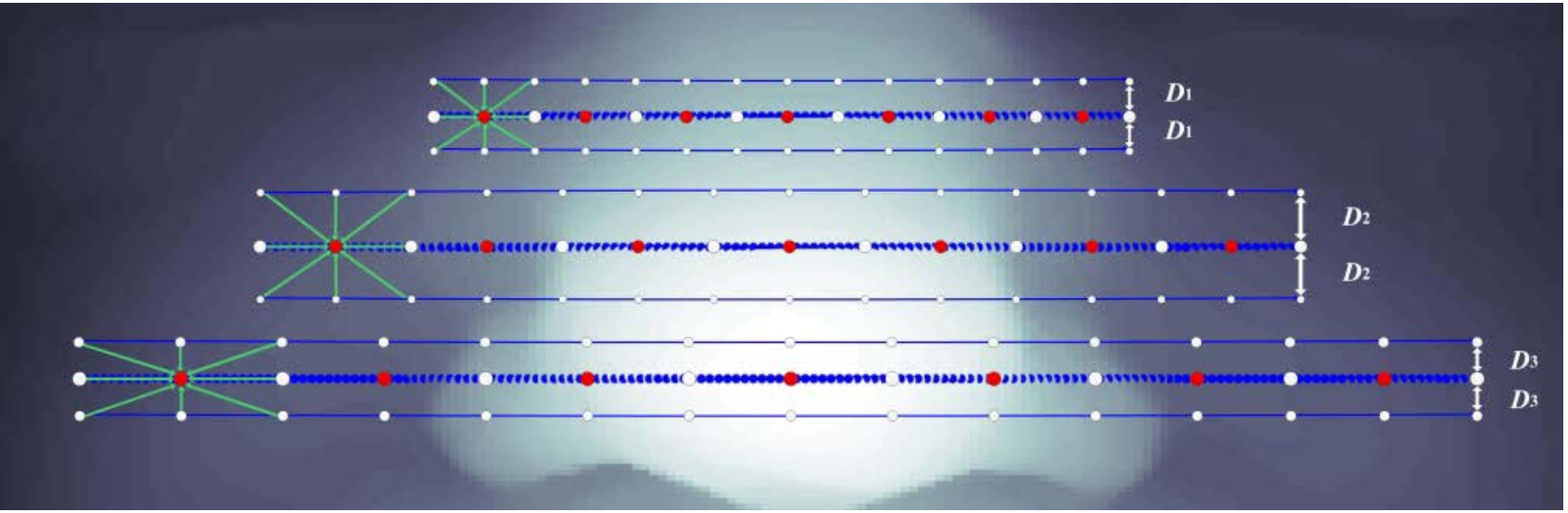
- Kernel Fisher's Analysis (KFA)
- Rank one recognition rate ($R_1RR$)
- Equal error rate (EER)

| | Scenarios | Depth Difference | | Direct Depth | |
|---|---|---|---|---|---|
| | | $R_1RR$ | EER | $R_1RR$ | EER |
| **Bosphorus** | All vs All | 88.98% | 4.93% | 85.82% | 5.10% |
| **Experiment 3 in FRGC** | All vs All | 83.53% | 5.34% | 79.32% | 5.32% |
| | All vs Neutral | 90.50% | 2.88% | 89.02% | 3.42% |
| | All vs Non-neutral | 70.91% | 7.02% | 61.85% | 8.56% |
| Features size of each capture | | $8 \times 7 = 56$ | | $15 \times 3 = 45$ | |

- Depth difference performs better for all scenarios, in particular under expressions
- More locally?

# Local Features Extraction

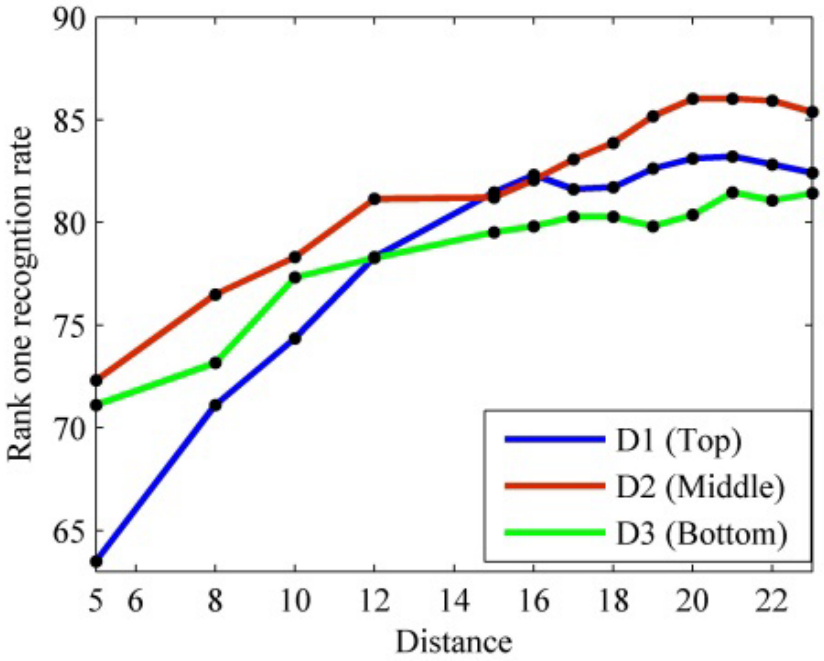
- Multi-component: Depth, Surface normals (*x*, *y* and *z*)
- 7 windows
- Feature size per capture: $8 \times 7 \times 3 = 168$
- $D_1$, $D_2$ and $D_3$ determines the size of the windows



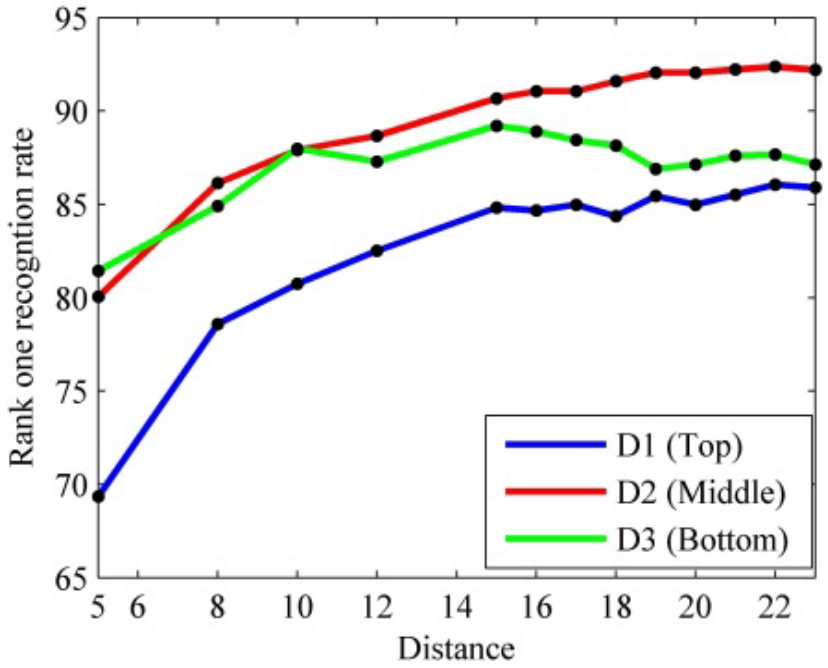
# Local Features Extraction

UNIVERSITY OF BATH

## Evaluation of $D_1$, $D_2$ and $D_3$ on depth




FRGC

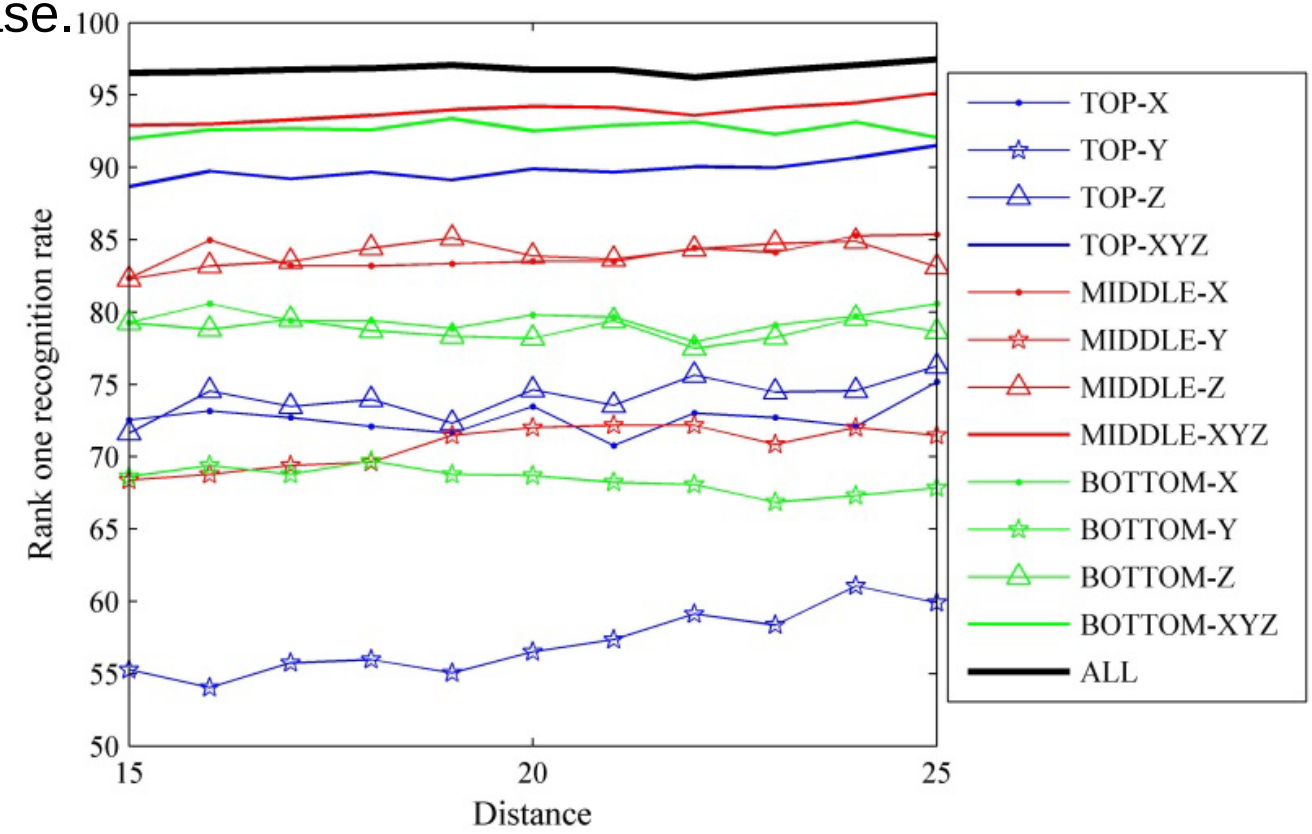



Bosphorus

# Local Features Extraction

## Evaluation of $D_1$, $D_2$ *and* $D_3$ on surface normals

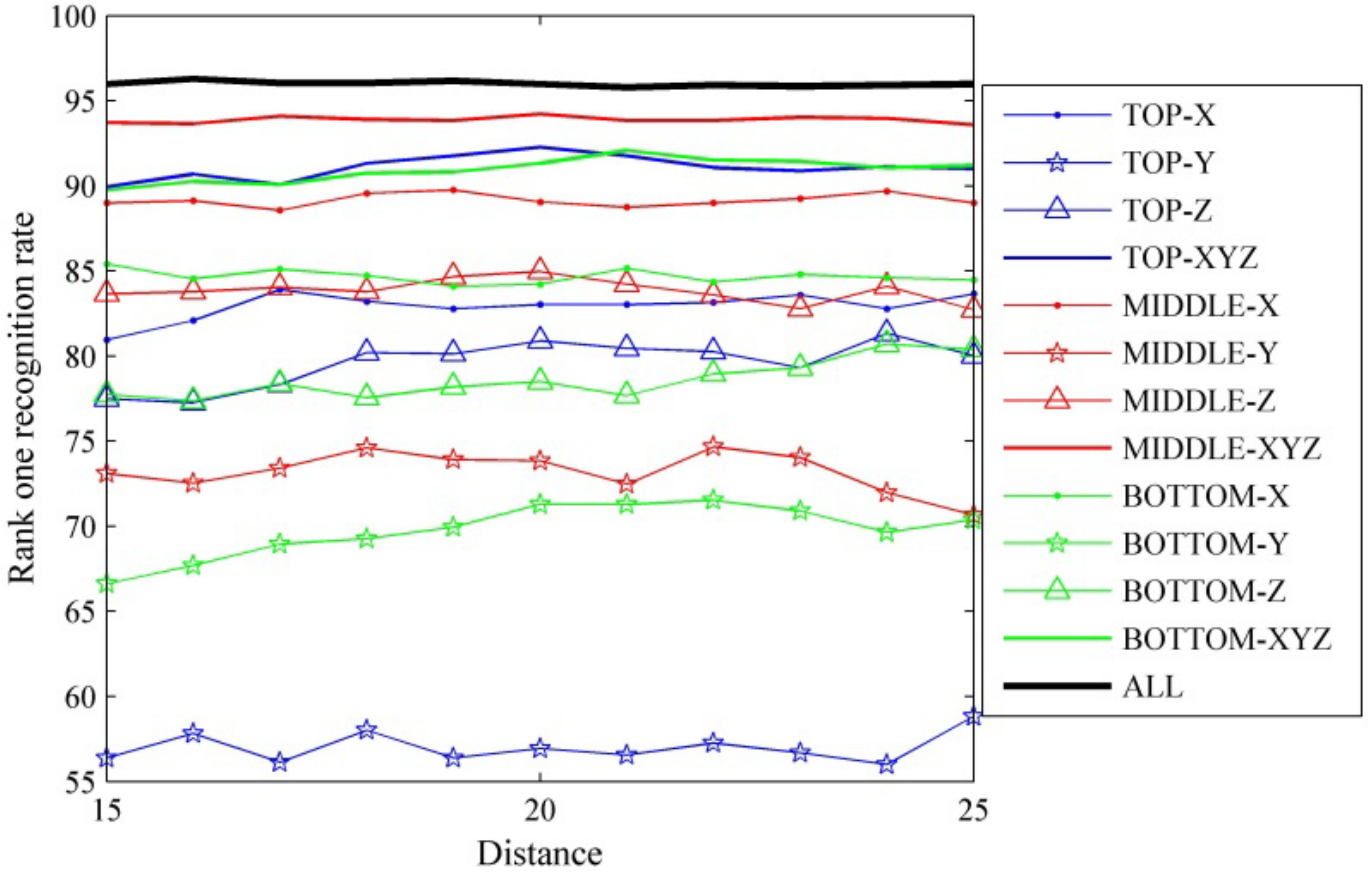
$R_1RR$ of features extracted from horizontal curves with different distances ($D_1$, $D_2$ and $D_3$) for Top, Middle and Bottom lines, respectively, using three channels ($x$, $y$ and $z$) on the Bosphorus database.



# Local Features Extraction

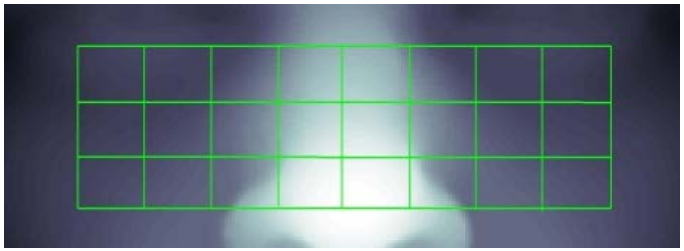
UNIVERSITY OF BATH

## FRGC

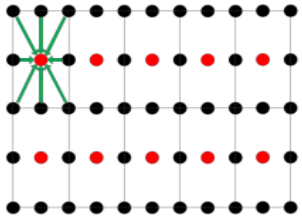


# Local patches evaluation

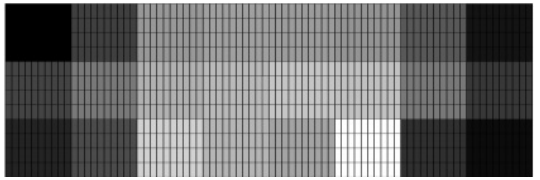
- 24 patches on the nasal and cheek region
- Each patch is resampled to 11 × 5 points, resulting in 10 windows
- 4 discriminative power maps (depth, *x*, *y* and *z*)
- Brighter values denote larger $R_1RR$



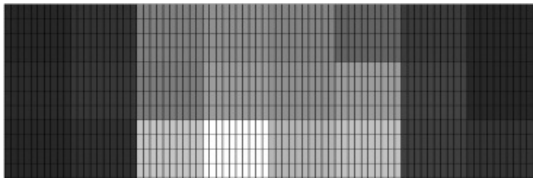
(a) 24 patches



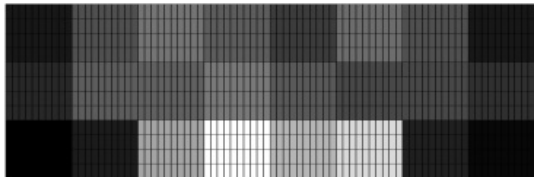
(b) Proposed Local shape descriptor



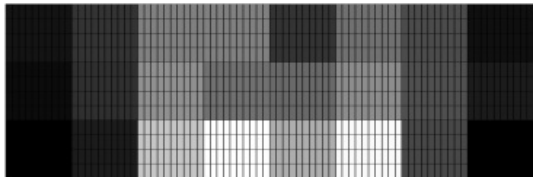
(c) Depth



(d) *x*



(e) *y*

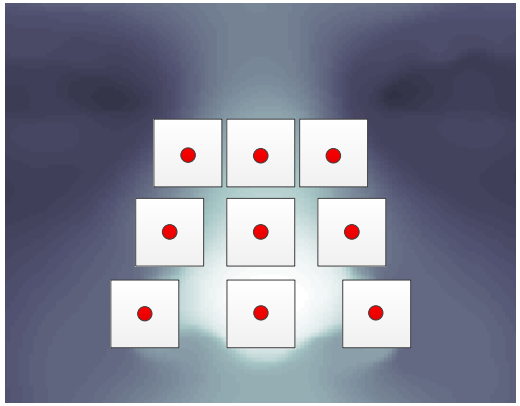


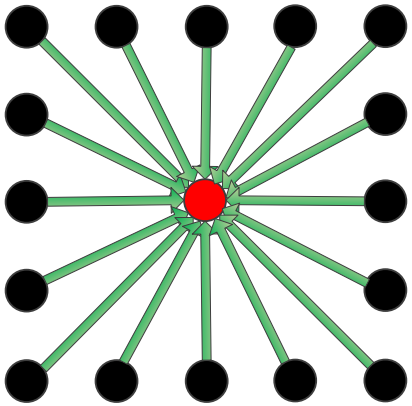
(f) *z*

Patches located on the boundary of the nasal and adjoining cheek regions perform better than those only wholly on the nasal region which shows the potential of the cheek regions immediately adjoining the nose.

# Local patches evaluation

## Square Local Shape Descriptor

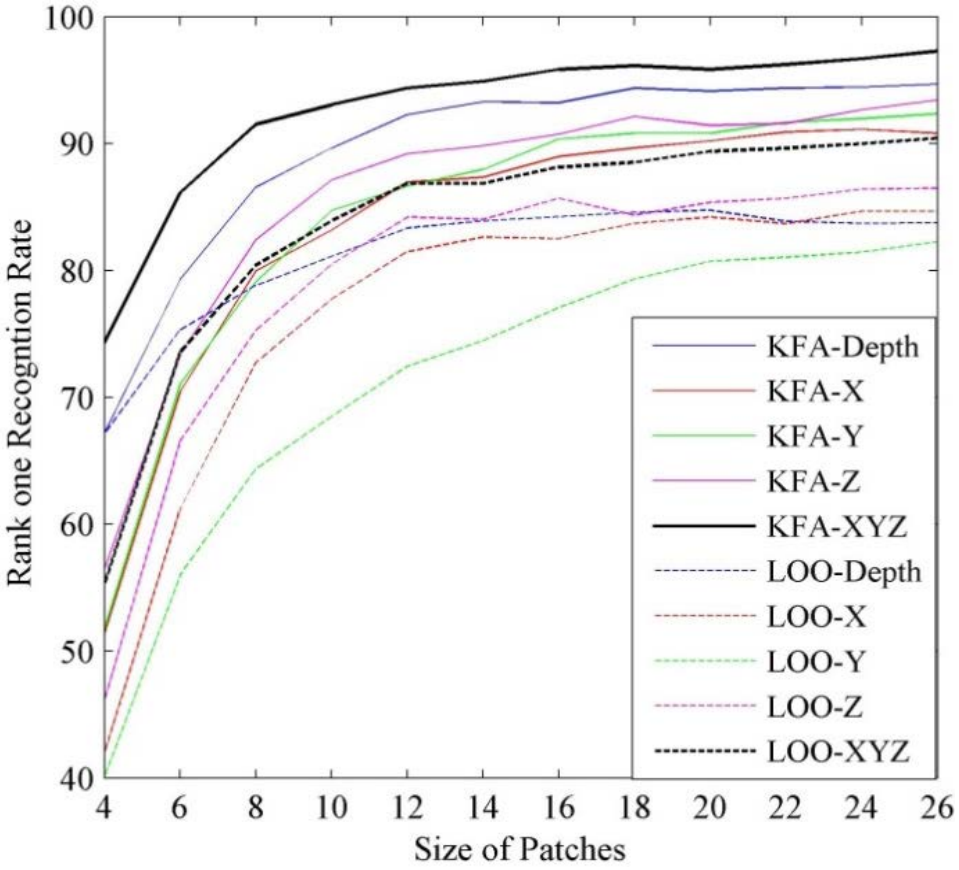




- 9 patches, placed on regions with higher discriminative power
- Feature sets size: $9 \times 16 = 144$
- Window size: 4 to 26

# Local patches evaluation

Rank one recognition rate calculated by both KFA and Leave One Out (LOO) classifier using different size of windows (patches) on the Bosphorus database.



# Comparison

Comparison of EERs tested on the FRGC database (Experiment 3) using nasal region and its environs

| Algorithm | Matching | Neutral | Non-neutral |
|---|---|---|---|
| **Chang et al. [1]** | ICP | 0.12 | 0.23 |
| **Emambakhsh et al. [2]** | KFA-Poly | 0.08 | 0.18 |
| **Gao et al. [7]** | KFA-Poly | 0.03 | 0.06 |
| **Proposed method** | KFA-Poly | 0.03 | 0.05 |

Comparison of $R_1$RRs tested in the FRGC (Experiment 3) and Bosphorus Databases using nasal region and its environs

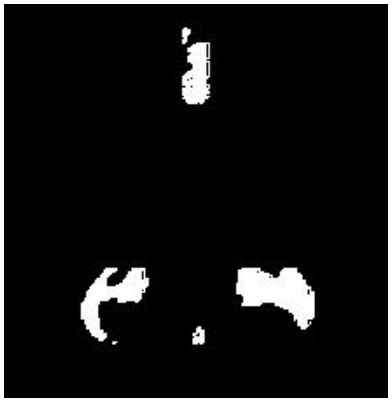
| Algorithm | FRGC | | Bosphorus |
|---|---|---|---|
| | Neutral | Non-neutral | |
| **Chang et al. [1]** | 96.6% | 82.7% | |
| **Emambakhsh et al. [2]** | 90.87% | 81.61% | 97.44% |
| **Gao et al. [7]** | 92.68% | 75.32% | 93.61% |
| **Proposed method** | 93.38% | 84.13% | 97.76% |

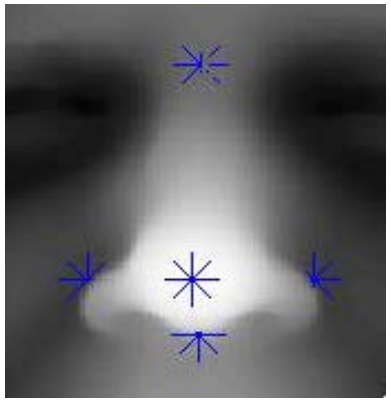
# Conclusions

- A novel local depth and surface normals descriptor is proposed to explore the discriminative features on the nasal region and adjoining parts of the cheek.
- A further analysis of the discriminatory power of patches shows that the adjoining regions have the potential to produce good recognition performance.
- On the basis of the proposed descriptor, a relatively small set of features extracted from the nasal and adjoining cheek regions produce a $R_1RR$ of 97.76% and an EER of 1.32%.

# Future Work

- Finding more effective patches for recognition
- Feature extraction using reconstructed captures
  - The Photoface device
  - Albedo and SN
- Landmark localization using surface normals





Thank you!