Head Yaw Estimation via Symmetry of Regions
Bingpeng Ma, Annan Li, Xiujuan Chai, Shiguang Shan
byma@ucas.ac.cn
University of Chinese Academy of Sciences
Key Lab of Intelligent Information Processing of Chinese Academy of Sciences (CAS),
Institute of Computing Technology, CAS

Abstract
This paper proposes a novel method to estimate the head yaw rotations using the symmetry of regions. We argue and reveal that the symmetry between the two regions in the same horizontal row is closely relevant to the yaw rotation of head, while at the same time insensitive to the identity of the face. We first extract the Gabor features of a face image, then the covariance descriptors are used to compute the symmetries of Gabor features. Since the covariance matrix can eliminate the influence which is caused by rotations and illuminations, the proposed method is robust to these variations. In addition, the proposed method can be further improved by combining it with supervised learning method.

Motivations
In [1], Ma et al. argued that the asymmetry of the intensities in each row of the face image is closely relevant to the yaw rotation of head. However, the asymmetry based on the 1D signal is easily effected by other factors, for example, the misalignment.

Fig. 1 The relationship between the symmetry plane of the head and the center lines of the images.

In this paper, 2D Gabor features are divided into many regions with the same size. Covariance descriptor is taken as feature extractor to extract the symmetry of the symmetrical regions of the Gabor features.

Fig. 2 The symmetry measures of the different poses. (a) The symmetry of the 1D signals. (b) The symmetry of 2D regions.

This figure shows that the symmetry of 2D regions is more related to the pose varies than that of 1D signals.

CovGa
STEP 1:
- For an image I(x, y), we compute its convolution with Gabor filters:
  \[ G(\mu, \nu) = I(x, y) * \psi_{\mu, \nu}(z) \]
  We then apply "MAX" pooling over two consecutive scales:
  \[ G_{ij} = \max(G(2i - 1, j), G(2i, j)) \]

STEP 2:
- For each pixel, a 7-dimensional vector is computed to capture intensity, texture and shape statistics:
  \[ f_{ij}(x, y) = [x, y, G_{ij}, G_{ij}, G_{ij}, G_{ij}, G_{ij}, G_{ij}, G_{ij}] \]
- We divide Gabor features into small regions of equal size and overlap. Each region is represented by a covariance descriptor:
  \[ C_{ij,r} = \frac{1}{n-1} \sum_{(x,y) \in r} (f_{ij}(x, y) - f_{ij})(f_{ij}(x, y) - f_{ij})^T \]
- For region \( r \) and its' symmetrical region \( r' \) in the row, we compute their similarity \( d_{ij,r} \) and take it as the symmetry of regions:
  \[ d_{ij,r} = d(C_{ij,r}, C_{ij,r'}) = \sqrt{\sum_{p=1}^{P} \ln^2 \lambda_p(C_{ij,r}, C_{ij,r'})} \]
- The symmetry metrics are concatenated to form the image representation.
- The distance between two images is obtained by computing the Euclidian distance between these representations.

sCovGa:
- CovGa can be combined with the supervised method, such as LDA, to improve the discriminate ability.

FlowChart

Experiments
Cas-Peel database:
- 21 poses (7 yaw poses, 3 pitch poses).
- a subset, IDs range from 401 through 600.

Our own database:
- 3, 030 images of 102 subjects.
- The yaw and pitch poses range within \([-50°, +50°]\), intervals of 1°.
- The sample number is 30 for each class (i.e. yaw pose).

Conclusion
CovGa can improve the performance of head pose estimation.

References