Head pose estimation

Objective

The objective of this project is to estimate the head pose. Given a 2D image containing a head in an unknown orientation, the required output is the orientation of the head. Supplementary, the second output is an image with a pose normalized face.

Motivation

People use the orientation of their heads to convey a lot of non verbal information. First of all, a person indicates who is the target of a conversation by looking at him. During the conversation, the head movements can be a form of gesturing, indicating dissent, confusion, consideration, agreement, surprise,…

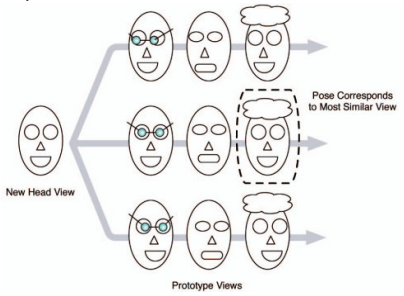
If two people have their visual attention on each other, referred as mutual gaze, this is often a sign that the two people are engaged in a discussion. Mutual gaze also occurs when people notice each other, e.g. a pedestrian and a car driver. Finally the head orientation could give information about the environment, e.g. when several persons are watching into the same direction.

Head pose estimation is also an important issue in 2D face recognition. Without taking into account head orientation, 2D recognition performance is usually poor.

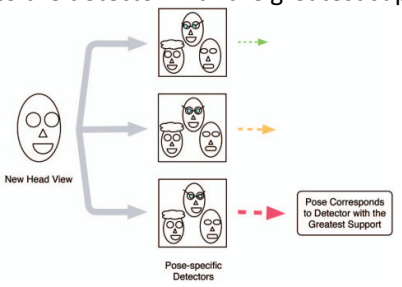
Methods

According to [1] eight fundamental approaches for head pose estimation can be distinguished:

- **Appearance template methods** compare a new image of a head to a set of example images (each labeled with a discrete pose) in order to find the most similar view.

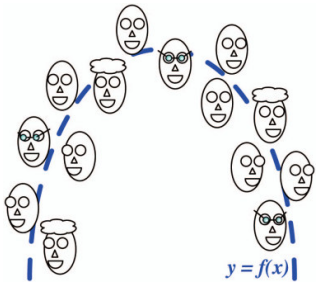


- **Detector array methods** train a series of head detectors each corresponding to a specific pose and assign a discrete pose to the detector with the greatest support.

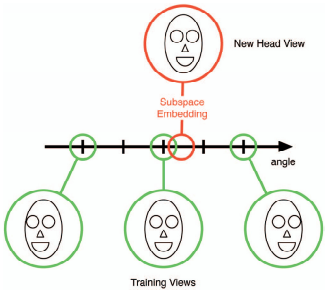


Detector arrays are similar to appearance templates, but for each pose, only one detector is defined based on a training with a supervised learning algorithm.

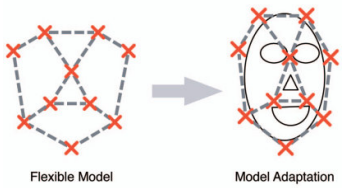
- **Nonlinear regression methods** use non-linear regression tools to develop a functional mapping from the image (or feature data) to a head pose measurement.



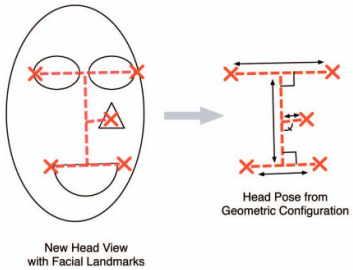
- Manifold embedding methods



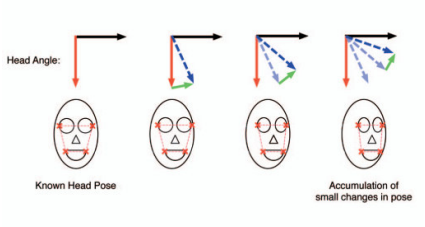
- **Flexible models** fit a non-rigid model to facial structure in the image of each individual. Head pose is estimated from feature-level comparisons or from the instantiation of the model parameters.



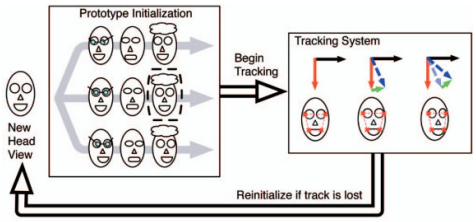
- **Geometric models** use the location of features such as eyes, mouth, and nose tip to determine the pose from their relative configuration.



- **Tracking methods** recover the global pose change of the head form the observed movement between video frames.



- **Hybrid methods** one or more of the aforementioned methods to overcome limitations of using a single approach.



Papers

From the various approaches that have been proposed for head pose estimation, I selected some interesting methods:

Overview:

[1] E. Murphy-Chutorian, M. M. Trivedi, "Head Pose Estimation in Computer Vision: A Survey", TPAMI vol. 31, no. 4, april 2009.

- **Appearance templates**

  [2] S. Niyogi and W. Freeman, "Example-Based Head Tracking," Proc. FG, pp. 374-378, 1996.

- **Non linear regression**

  [3] Y. Li, S. Gong, J. Sherrah, and H. Liddell, "Support Vector Machine Based Multi-View Face Detection and Recognition," Image and Vision Computing, vol. 22, no. 5, p. 2004, 2004.

  [4] H. Moon and M. Miller, "Estimating Facial Pose from a Sparse Representation," Proc. IEEE Int'l Conf. Image Processing, pp. 75-78, 2004.

  [5] E. Murphy-Chutorian and M. Trivedi, "Head Pose Estimation for Driver Assistance Systems: A Robust Algorithm and Experimental Evaluation," Proc. CITS , pp. 709-714, 2007.

- **Manifold embedding**

  [6] Y. Wei, L. Fradet, and T. Tan, "Head Pose Estimation Using Gabor Eigenspace Modeling" .

  [7] S. Srinivasan and K. Boyer, "Head Pose Estimation Using View Based Eigenspaces," Proc. ICPR, pp. 302-305, 2002.

  [8] J. Wu and M. Trivedi, "A Two-Stage Head Pose Estimation Framework and Evaluation," Pattern Recognition, vol. 41, no. 3, pp. 1138-1158, 2008.

- **Flexible models**

  [9] T. Cootes, K. Walker, and C. Taylor, "View-Based Active Appearance Models," Proc. FG, pp. 227-232, 2000.

Material

For this project, the Bosporus database is available. It contains 4666 facial images with known pose. Also manual indicated landmarks are available.

The database also contains 3D images, acquired with the "Inspeck Mega Capturor II 3D" scanner leading to 3D point clouds of approximately 35 000 points. Beside pose variations, expression variations and occlusions are present in the database.

Task

1. Read overview paper.
2. Summarize in a few sentences the head pose estimation strategies and comment the advantages and disadvantages of the different strategies.
3. Select a strategy and choose one (or more) papers using that strategy for implementation in Matlab (preferably). Motivate the selection. Discuss the implementation and the results.
4. Optionally, you can implement a face recognition algorithm following your pose estimation algorithm.
5. Write a small (about 10 pages) report on all this to be submitted together with the code by the end of may and prepare a Powerpoint presentation for the examination.
6. Make an appointment (as soon as possible) with me (Dirk.Vandermeulen@esat.kuleuven.be) for your examination to be held during the examination period (or before if you want).
7. Make the project in groups of maximum 2 people. Let me know by email which groups you form.

Questions need to be addressed to my assistant (Dirk.Smeets@uz.kuleuven.ac.be).