Int J Comput Vis (2013) 101:270–287
DOI 10.1007/s11263-012-0567-y

A Two-Layer Framework for Piecewise Linear Manifold-Based Head Pose Estimation

Jacob Foytik · Vijayan K. Asari

Received: 2 August 2011 / Accepted: 30 August 2012 / Published online: 19 September 2012

Abstract Fine-grain head pose estimation from imagery is an essential operation for many human-centered systems, including pose independent face recognition and human-computer interaction (HCI) systems. It is only recently that estimation systems have evolved past coarse level classification of pose and concentrated on fine-grain estimation. In particular, the state of the art of such systems consists of nonlinear manifold embedding techniques that capture the intrinsic relationship of a pose varying face dataset. The success of these solutions can be attributed to the acknowledgment that image variation corresponding to pose change is nonlinear in nature. Yet, the algorithms are limited by the complexity of embedding functions that describe the relationship. We present a pose estimation framework that seeks to describe the global nonlinear relationship in terms of localized linear functions. A two layer system (coarse/fine) is formulated on the assumptions that coarse pose estimation can be performed adequately using supervised linear methods, and fine pose estimation can be achieved using linear regressive functions if the scope of the pose manifold is limited. A pose estimation system is implemented utilizing simple linear subspace methods and oriented Gabor and phase congruency features. The framework is tested using widely accepted pose-varying face databases (FacePix(30) and Pointing'04) and shown to perform fine head pose estimation with competitive accuracy when compared with state of the art nonlinear manifold methods.

J. Foytik (✉) · V.K. Asari
University of Dayton, 300 College Park, Dayton, OH, 45469, USA
e-mail: jfoytik1@notes.udayton.edu

V.K. Asari
e-mail: vijayan.asari@notes.udayton.edu

Keywords Head pose estimation · Piecewise linear manifold · Coarse to fine · Phase congruency · Gabor filter

1 Introduction

The advantages of an accurate head pose estimation system span several research areas including gaze estimation, human-computer interaction (HCI), 3D face modeling, face expression recognition, and face recognition. In all of these examples, knowledge of the head orientation is a crucial task for analysis of the face. The need for noise invariant face recognition systems, capable of identifying individuals regardless of face pose angle, creates a reliance on accurate face pose estimation systems. In terms of face recognition, pose variation is deemed a noise factor which can be easily removed if the orientation of the face is known. Pose estimation systems allow faces of similar orientation to be compared, producing higher recognition accuracy. Additionally, head pose estimation is a necessary component in the analysis of a person's gaze direction, where a person's focus of attention can be obtained through the analysis of both head orientation and eye direction.

Recently, much attention has been given to the manifold class of techniques for pose estimation. These methods are based on the foundation that the high dimensional input images should theoretically lie in a compact subspace that purely defines head pose changes. The given input patterns are considered to be ill-suited for directly assessing pose, but can be transferred to a subspace that contains a higher density of valuable features. Since the head can only move with three degrees of freedom, including *yaw*, *pitch*, and *roll*, the observed high-dimensional image should theoretically lie in a low-dimension constrained by the allowable pose variation (Murphy-Chutorian and Trivedi 2008). Furthermore the