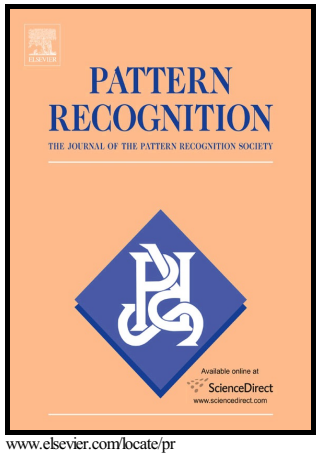

www.elsevier.com/locate/pr

Spectral Attribute Learning for Visual Regression

Ke Chen, Kui Jia, Zhaoxiang Zhang, Joni-Kristian Kämäräinen

PII: S0031-3203(17)30010-9
DOI: http://dx.doi.org/10.1016/j.patcog.2017.01.009
Reference: PR6011

To appear in: *Pattern Recognition*

Received date: 15 July 2016
Revised date: 10 December 2016
Accepted date: 6 January 2017

Cite this article as: Ke Chen, Kui Jia, Zhaoxiang Zhang and Joni-Kristia Kämäräinen, Spectral Attribute Learning for Visual Regression, *Patter Recognition,* http://dx.doi.org/10.1016/j.patcog.2017.01.009

Spectral Attribute Learning for Visual Regression

Ke Chen[a], Kui Jia[b,*], Zhaoxiang Zhang[c], Joni-Kristian Kämäräinen[a]

[a]*Department of Signal Processing, Tampere University of Technology, Tampere, Finland*
[b]*School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China*
[c]*Institute of Automation, Chinese Academy of Sciences, Beijing, China*

Abstract

A number of computer vision problems such as facial age estimation, crowd counting and pose estimation can be solved by learning regression mapping on low-level imagery features. We show that visual regression can be substantially improved by two-stage regression where imagery features are first mapped to an attribute space which explicitly models latent correlations across continuously-changing output. We propose an approach to automatically discover "spectral attributes" which avoids manual work required for defining hand-crafted attribute representations. Visual attribute regression outperforms direct visual regression and our spectral attribute visual regression achieves state-of-the-art accuracy in multiple applications.

Keywords: Facial age estimation, Crowd counting, Head Pose Estimation, Spectral learning, Attributes, Regression

1. Introduction

Visual regression maps imagery features to a continuous output space and is therefore a suitable tool for a number of computer vision applications such as facial age estimation, pedestrian counting and object pose estimation. These vision problems can also be formulated as multi-class classification problems (*e.g.* [1]), but recently visual regression based methods that naturally exploit continuous scalar-valued output label spaces have achieved superior results in important applications (*e.g.* age estimation [2, 3, 4], pedestrian density estimation [5, 4] and human body/face pose estimation [6, 7, 8]).

*Tel/fax: +86-20-87111435.
Email address: kuijia@scut.edu.cn (Kui Jia)