Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT®

PERGAMON

Pattern Recognition 38 (2005) 813–824

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Steerable pyramid-based face hallucination

Congyong Su[a, b, *], Yueting Zhuang[a, b], Li Huang[a, b], Fei Wu[a, b]

[a]College of Computer Science, Zhejiang University, 310027 Hangzhou, China
[b]Microsoft Visual Perception Laboratory of Zhejiang University, 310027 Hangzhou, China

Received 9 August 2004

Abstract

In this paper we propose a robust learning-based face hallucination algorithm, which predicts a high-resolution face image from an input low-resolution image. It can be utilized for many computer vision tasks, such as face recognition and face tracking. With the help of a database of other high-resolution face images, we use a steerable pyramid to extract multi-orientation and multi-scale information of local low-level facial features both from the input low-resolution face image and other high-resolution ones, and use a pyramid-like parent structure and local best match approach to estimate the best prior; then, this prior is incorporated into a Bayesian maximum a posterior (MAP) framework, and finally the high-resolution version is optimized by a steepest decent algorithm. The experimental results show that we can enhance a 24×32 face image into a 96×128 one while the visual effect is relatively good.

Keywords: Face hallucination; Super-resolution; Steerable pyramid; Local best match; Bayesian maximum a posterior (MAP)

1. Introduction

Super-resolution predicts an image's high-resolution part from its low-resolution one. Available super-resolution techniques can be divided into two classes: learning-based [1–7] and reconstruction-based [8–11]. In this paper we are concerned with learning-based techniques' application in faces, which can be called face hallucination [1] that predicts high-resolution face images from low-resolution ones. Face hallucination can be used in many fields, such as face recognition, face tracking, and facial expression analysis.

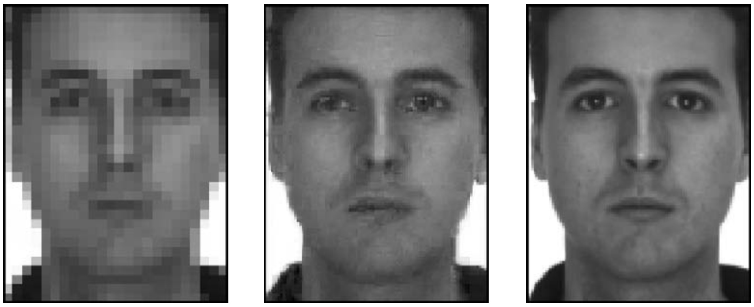
Face hallucination is very challenging because people are so familiar with faces. In this article, we propose an algorithm that predicts the high-resolution face images from the low-resolution ones, through learning from a training set which consists of a large number of other high-resolution face images.

Our considerations for face hallucination are as follows: (1) Markov random field (MRF) hypothesis: a pixel in a face image only relies on the pixels in its neighborhood. (2) Local low-level facial features: the multi-orientation and multi-scale information of local low-level facial features should be taken into account when calculating feature vectors. (3) Semi-exact correspondence: a pixel p in an input face image may correspond to p_t or any pixel of p_t's local area in the training set, rather than exactly corresponding to the pixel p_t with the same coordinates.

Based on the above analysis, we propose an algorithm, in which the steerable pyramid is used to capture multi-orientation and multi-scale information in face images, and a pyramid-like parent structure is applied to record the feature vectors of the pixel's neighborhood in various levels of the pyramid. For the third consideration, we employ the local best match algorithm with Gaussian kernel to find the best feature vector in the training set. Finally, the best prior

* Corresponding author. College of Computer Science, Zhejiang University, 310027 Hangzhou, China. Tel.: +86 130 6780 4358.
E-mail address: su@cs.zju.edu.cn (C. Su).

0031-3203/$30.00
doi:10.1016/j.patcog.2004.11.007



(a) Input 24 x 32 (b) Hallucinated (c) Original 96 x 128

Fig. 1. An experimental result. (a) The input low-resolution face image. (b) The hallucinated result of our algorithm. (c) The original high-resolution face image.

is incorporated into the Bayesian MAP framework and the optimal high-resolution face image is obtained by a steepest descent algorithm. An experimental result can be seen from Fig. 1, where the hallucinated face is very close to the original one, and oriented facial features are recovered well.

The remainder of this paper is organized as follows. The related work is surveyed in the next section. In Section 3, details of our algorithm are presented. Some experimental results are shown in Section 4. Section 5 contains the conclusion and proposal for future work.

2. Related work

There are some learning-based super-resolution algorithms. Freeman et al. [3,4] propose an example-based approach, and they train Markov networks to learn the fine details that correspond to different image regions seen at a low resolution and then use these learned relationships to predict fine details in other images. Image analogies [5] are based on a multi-scale auto-regression, and infer the high-resolution image from a low-resolution source one. Sun et al. [7] construct primal sketch priors to enhance the quality of the hallucinated high-resolution image. The above three algorithms are only applied to generic images, rather than face images. Baker and Kanade [1,2] propose a face hallucination algorithm to learn a prior on the spatial distribution of the image gradient for frontal face images, and use Bayesian MAP theory to hallucinate the high-resolution face image from the low-resolution one. The results are better than Freeman's, Hertzmann's and Sun's, but appear noisy at places. Liu et al. [6] develop a two-step statistical modeling approach that integrates both a global parametric model and a local nonparametric model. The resulting face images are smoother than Baker's and can retain global features well, such as explicit contours, coherent illumination and symmetry; however, because Liu et al. calculate the global model by PCA, the hallucinated faces appear a little different from the original faces.

To make the hallucinated face appear less noisy and keep it consistent with the original one, we choose a steerable pyramid to calculate our feature space rather than Laplacian pyramid and derivatives of Gaussian pyramid used in Ref. [1]. The steerable pyramid is generated by oriented steerable filters, which can preserve multi-orientation information better in our experimental results. Baker and Kanade use a line-like parent structure and exact pixel correspondence searching algorithm to find the best feature vector in the training set. Thus the hallucinated faces appear noisy at places. Accordingly, we propose a pyramid-like parent structure and local best-match approach to reduce noise and make the result smoother.

3. Theory and algorithms

3.1. Markov random field

Markov random field (MRF) theory provides a convenient and consistent way to model image pixels [12]. We assume that a face image is an MRF. By the definition of MRF, for every pixel p in a face image I, we have the following equation, where $N(p)$ is the neighborhood of p.

$$P(p|(I-p)) = P(p|N(p)). \tag{1}$$

Eq. (1) means a pixel in a face image only depends on its neighbor pixels.

MRF theory is often used together with statistical decision and estimation theories, and Bayesian maximum a posteriori (MAP) is a popular choice.

3.2. A Bayesian estimation to face hallucination

The Bayesian estimation problem comprises prior model and measurement (or likelihood) model. In face hallucination, let I_{hi} and I_{lo} denote the high- and low-resolution face images, respectively. The prior model is $P(I_{hi})$, and the