Image and Vision Computing 40 (2015) 49–64

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Editor's Choice Article

Hallucination of facial details from degraded images using 3D face models☆

Matthaeus Schumacher [1], Marcel Piotraschke [1], Volker Blanz *

Institute for Vision and Graphics, University of Siegen, Germany

ARTICLE INFO

Article history:
Received 24 December 2013
Received in revised form 23 October 2014
Accepted 13 June 2015
Available online 24 June 2015

Keywords:
Face hallucination
3D models
Model-based deblurring
Occlusions
Faces

ABSTRACT

The goals of this paper are: (1) to enhance the quality of images of faces, (2) to enable 3D Morphable Models (3DMMs) to cope with severely degraded images, and (3) to reconstruct textured 3D faces with details that are not in the input images. Details that are lost in the input images due to blur, low resolution or occlusions, are filled in by the 3DMM and an additional texture enhancement algorithm that adds high-resolution details from example faces. By leveraging class-specific knowledge, this restoration process goes beyond what general image operations such as deblurring or inpainting can achieve. The benefit of the 3DMM for image restoration is that it can be applied to any pose and illumination, unlike image-based methods. However, it is only with the new fitting algorithm that 3DMMs can produce realistic faces from severely degraded images. The new method includes the blurring or downsampling operator explicitly into the analysis-by-synthesis algorithm.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Difficult imaging conditions due to blur, low resolution, noise, partial occlusions or non-uniform lighting are frequently encountered in many real-world applications, for example in law enforcement if a suspect has to be recognized in low quality image material. A number of image processing algorithms recover and enhance information that is present in the image, yet mostly invisible to the human eye. For example, deconvolution strives to invert the effect of blurring. Often the blur kernel (or point spread function, PSF) is unknown and has to be estimated from the image. A survey of this so-called *Deblurring by Blind Deconvolution* problem can be found in [1], and some more recent publications in this field include [2] and [3]. For low resolution video, deconvolution can be combined with methods to merge data from all frames [4,5]. On the other hand, structures that are degraded due to occlusion can be partly recovered with image inpainting methods [6].

These methods can be applied to any image material, because they make only very general assumptions about the image content. However, if it is known that part of an image shows a human face, it is possible to add new information that was not in the image to begin with. If the lower half of a face is occluded, we can still safely assume that a mouth and a chin have to be added, and we can estimate their pose angle and lighting from the upper half of the face. The same is true for blurred regions: even if the eye and eyebrow are only dark spots in the input image, we can fill in eyeball, iris, eyelashes and all other details of human eyes. Fig. 3 shows how this can be done with our proposed algorithm.

Given a mathematical model of the expected content of the image, such a model-based image enhancement can be achieved. In this paper, we rely on the *3D Morphable Model* (3DMM) [7] as a statistical description of the natural shapes and textures of faces. It is important to stress that the added image detail cannot be more than an educated guess, based on the prior information about faces on the one hand, and all the remaining information that is in the image on the other hand.

One solution would be to fill in the details from the average face, or any other random face. Our algorithm goes one step further by exploiting correlations in the set of human faces: After fitting the 3DMM to the degraded image, we obtain a best fit, then from this best fit we take the details and render these into the image, both in the case of blurred and partly occluded faces (Fig. 2). This idea is along the lines of 3D shape reconstruction from single images using 3DMMs [7], where the model is fitted to colors of pixels, and gives an estimate of depth. Recently, it has been shown that this inference is consistent with human expectation: when viewers see a frontal view of a face and then a choice of profiles that are all geometrically consistent with the front view, they tend to prefer those profiles that were calculated by the 3DMM [8], even if the choice includes the ground truth profiles of the face.

With this caveat, model-based inference of missing information may be a useful tool to obtain high quality images or 3D face models from degraded input data.

☆ Editor's Choice Articles are invited and handled by a select rotating 12 member Editorial Board committee. This paper has been recommended for acceptance by Shishir Shah.

* Corresponding author. Tel.: +49 271 740 2035.
E-mail addresses: schumacher@informatik.uni-siegen.de (M. Schumacher), piotraschke@nt.uni-siegen.de (M. Piotraschke), blanz@informatik.uni-siegen.de (V. Blanz).
[1] Tel.: +49 271 740 2036.

http://dx.doi.org/10.1016/j.imavis.2015.06.004
0262-8856/© 2015 Elsevier B.V. All rights reserved.

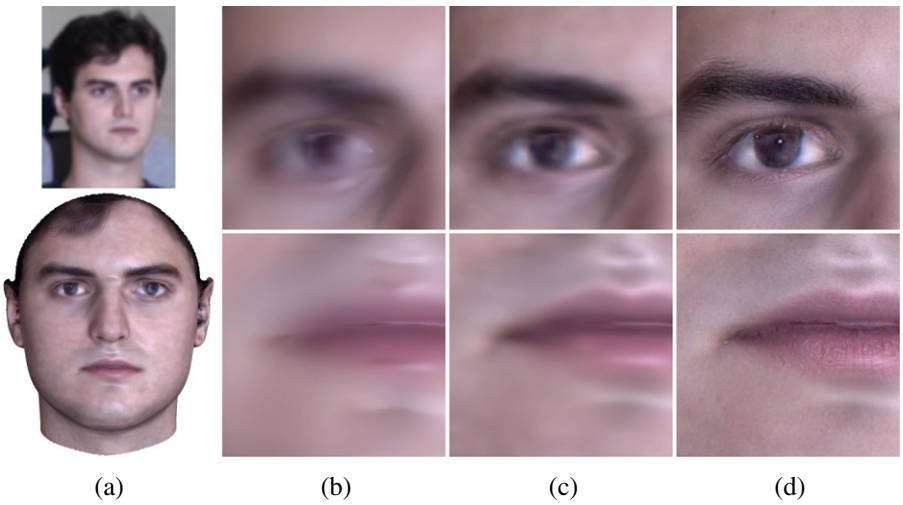

Fig. 1. a shows the blurred input image and the final result of the 3D reconstruction after applying our new approach. b to d are zoomed in views to the eye and mouth regions of the reconstructed 3D face model. While in 1b only the original texture was extracted from the photo, in 1c the deblurring described in Section 4.1 was applied and in 1d the high-resolution texture transfer of Section 5 was added to the results of column 1c.

For images of human faces, the model-based fill-in of facial details has become known as *Hallucinating Faces* [9]. For a recent survey, see [10]. For low resolution images, AAMs have been used to fill in missing details [11].

In contrast to this work, Baker et al. proposed a model that does not account for shape differences explicitly, and that uses a maximum a posteriori (MAP) estimator to estimate the high-resolution levels of a Gaussian Pyramid of registered images [9]. In a two-level approach based on global Eigenfaces and a local patch-based non-parametric Markov network, Liu et al. achieved very significant improvements in image resolution [12,13]. A separation of global face hallucination and local feature hallucination has been proposed in [14]. For face hallucination in video, Dedeoglu et al. use spatio-temporal consistencies and a domain-specific prior [15].

Soon after their initial development, Active Appearance Models (AAMs) were used to reconstruct missing structures in occluded regions [16]. Reconstruction of facial images both in case of partial occlusion and low resolution using a 2D Morphable Face Model, which bears similarities with AAMs, has been presented in [17,18].

Using separate reconstruction modules for 2D shape and texture that account for global structure and local detailed texture [19] can reconstruct occluded regions in images of faces. Another approach that is able to fill in occluded regions uses asymmetrical Principal Component Analysis (PCA) [20].

All of these algorithms use a statistical model of 2D faces restricted to poses that are close to frontal. Larger variations in pose have been handled by using a Gabor wavelet decomposition of faces and a set of linear mappings between wavelet features in different poses [21], or by exploiting large datasets, recent image matching techniques and MAP estimation [22].

Unlike these image-based methods, this paper proposes a 3D approach that is intrinsically invariant to changes in pose, size, illumination and other image parameters. Our strategy is to fit a 3DMM [7] to the input image with a novel fitting algorithm that is robust to the effects of blurring by explicitly simulating image blur in an analysis-by-synthesis. The algorithm works on any blur level, and estimates the appropriate level automatically. In addition to the facial details estimated by the 3DMM reconstruction (equivalent to an MAP estimate [23]), the algorithm adds details, such as eyelashes and pores, from other faces to obtain high resolution results.

A method for hallucinating 3D facial shapes from low resolution 3D scans using an radial basis function (RBF) regression that predicts curvatures and displacement images at high resolution from their low resolution version was presented by [24]. Unlike our algorithm, their input data are 3D, and no texture is used.

A major challenge in using a 3DMM for face hallucination or 3D reconstruction from low resolution images is to adapt the cost function to blurred input data. This challenge also occurs for generative face models in 2D, such as AAMs. An algorithm to make AAMs applicable to low resolution images was presented as *Resolution Aware Fitting* (RAF) [25]. Similar to our method, they include an explicit model of the downsampling or blurring in their cost function, and they compute the image difference in terms of pixels of the input image space and not in the shape-free texture space, as standard

Fig. 2. 3D reconstruction of non-frontal input image with occluded face regions. Parts of the input image are occluded due to glasses and facial hair. These have to be marked manually. b shows the 3D reconstruction with compensation of occluded areas. After reprojection and relighting of the reconstruction, a hallucination of occluded facial regions is possible (c).

Download English Version:

https://daneshyari.com/en/article/526793

Download Persian Version:

https://daneshyari.com/article/526793

Daneshyari.com