Pattern Recognition Letters 50 (2014) 82–90

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Automatic alignment and reconstruction of facial depth images ☆

Giancarlo Taveira *, Leandro A.F. Fernandes

Instituto de Computação, Universidade Federal Fluminense (UFF), CEP 24210-240 Niterói, RJ, Brazil

ARTICLE INFO

Article history:
Available online 12 December 2013

Keywords:
Depth image
Alignment
Interpolation
Resampling
Face image

ABSTRACT

Face, gender, ethnic and age group classification systems often work through an alignment, feature extraction, and identification pipeline. The quality of the alignment process is thus central to the performance of the identification process. Furthermore, missing portions of depth information can greatly affect results. Appropriate image reconstruction is therefore crucial for the correct operation of those systems. This paper presents a simple and effective approach for the automatic alignment and reconstruction of damaged facial depth images. By using only four facial landmarks and the raw depth data, our approach converts a given damaged depth image into a smooth depth function, performs the 3D alignment of the underlying face with the face of an average person, and produces an aligned depth image having arbitrary resolution. Our experiments show that the proposed approach outperforms commonly used methods. For instance, we show that it improves the quality of a state-of-art gender classification technique.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The ability to retrieve information from facial depth images has many practical applications including face recognition, age group estimation, gender and ethnic group classification. Unfortunately, depth data is often damaged due to limitations intrinsic to off-the-shelf depth-image capturing systems. Examples include, but are not limited to, depth shadowing and the influence of reflective, refractive and infrared absorbing materials in the scene (Zhu et al., 2008). Also, the amount of pixels covering the imaged face and face's orientation often vary from image to image, making difficult or even impossible the use of captured images without the proper alignment and reconstruction of depth data (Szeliski, 2010).

Virtually every computer vision researcher that needs to perform alignment and reconstruction of facial depth data usually presents its own solution to the problem. A well-known technique is to identify some facial features by curvature, and compute the alignment based on them (Moreno et al., 2005). Solutions based on *principal component analysis* (PCA) have also been proposed (Stormer and Rigoll, 2008). However, most of the attempts do not make proper use of depth information while performing the alignment, restricting the solution to the 2D image plane. Also, linear interpolation is commonly used to fill the holes (Wu et al., 2010), leading to unnatural flat artifacts on the facial surface.

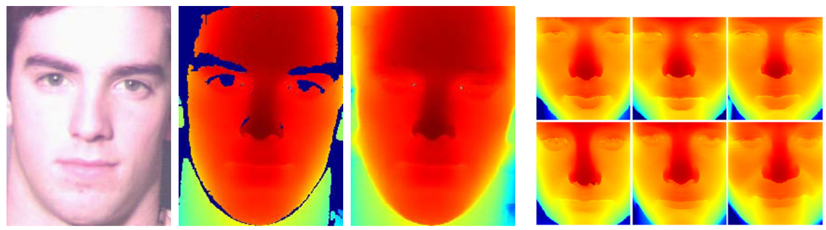
It is remarkable that the solutions commonly applied in the literature contradict the common wisdom that appropriate alignment and resampling techniques must be employed in order to produce corrected depth images from the original ones. For instance, it is important to make use of the depth information intrinsic to this kind of data in order to alignment the structures of interest (i.e., the faces) in the 3D space, not just on the image plane. Furthermore, by considering the nature of the surface of interest, it is clear that it is necessary to apply smooth non-linear interpolation techniques capable of reconstructing the damaged portions of the original image and also of producing depth values with sub pixel precision. With such care, the expectation is that the performance of depth-based classification techniques may be improved.

This paper presents a simple and effective method for aligning and reconstructing facial depth images from damaged depth data in a completely automatic way (Section 3). The approach uses information extracted from valid pixels to adjust a smooth *thin-plate spline* (TPS) interpolating function that naturally reconstructs the depth information of missing pixels (see Fig. 1) and computes smooth transitions among existing ones. The approach also explores facial landmarks in order to determine the actual position and orientation of the imaged face in the 3D space. The relation between the set of landmarks in the actual face and a set of canonical landmarks is used to map the shape of the imaged face to a standard space where the resulting aligned image is generated by ray casting the reconstructed surface. The developed ray casting scheme is derived from the *relief mapping* (RM) technique proposed by Policarpo et al. (2005) for real-time rendering of surface details mapped on coarse triangular meshes. Our approach easily fits into popular processing pipelines, and can be extended to produce correct color and normal map images to be used with the resulting depth images.

☆ This paper has been recommended for acceptance by Dmitry Goldgof.

* Corresponding author. Tel.: +55 21 2629 5665; fax: +55 21 2629 5669.
E-mail addresses: gtaveira@ic.uff.br (G. Taveira), laffernandes@ic.uff.br (L.A.F. Fernandes).

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.12.007



(a) Original color (b) Original depth (c) Reconstruction (d) Resulting aligned images

Fig. 1. For the same subject: (a) the original color image, (b) the original damaged depth image, and (c) the image with reconstructed depth information produced by our technique. Six aligned and reconstructed depth images of different subjects are presented in (d). Images (b), (c) and (d) are presented in false-color, where dark red pixels denote the surface closest to the camera. Notice in (c) the smooth transition of depth values in the originally corrupted portions (navy blue pixels in (b)). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Our experiments (Section 4) show that the approximation errors produced by our method are smaller than those using linear interpolation for reconstruction with *iterative closest point* (ICP) for 3D alignment of depth data and with 2D alignment of the depth images (see Fig. 2). We also present a comparative study among four distinct interpolation methods (nearest-neighbor, linear, natural-neighbor and thin-plate spline) using the proposed alignment method in the 3D domain. Each interpolation method was applied as part of state-of-art gender classification processes proposed by Wu et al. (2010, 2011). Since the classification techniques receive surface normals computed from reconstructed facial depth images as input, the performance of these classification models as a function of the input images indicate the quality and the influence of each interpolation method on the result.

2. Related work

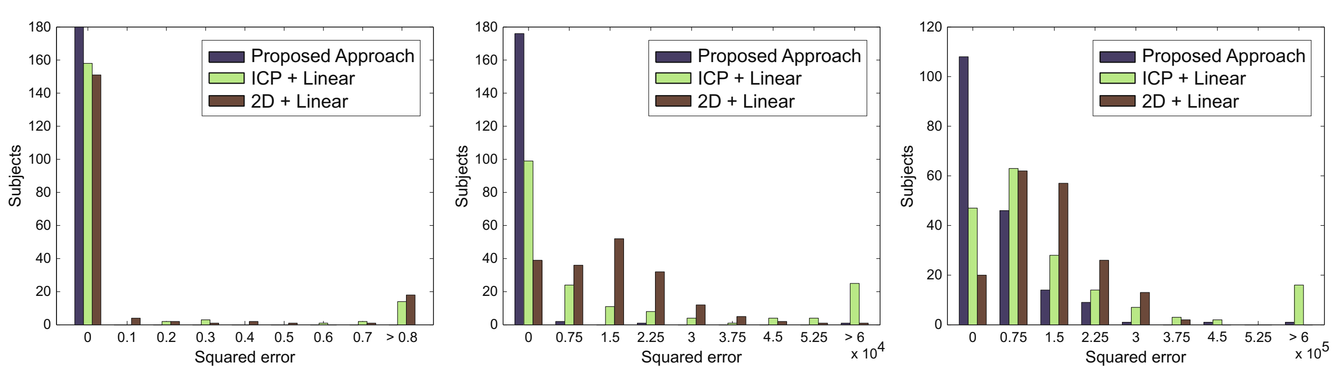
This section discusses the use of TPS on the interpolation of facial color images, the use of interpolation schemes to resample facial depth data, and alignment schemes for aligning human body surfaces.

Rosen (1996) developed the Java applet entitled "AlexWarp". Since its creation, the applet has gained popularity among internet users world-wide for its simple and fast method of facial image warping. When the user provides one pair of landmark points, the applet determines the region to warp, warps it, and then outputs the warped picture. One major drawback of the AlexWarp applet is that transformations can only be applied one at a time. The AlexWarp applet works on colored images in the 2D domain with a limited number of control points.

Whitbeck and Guo (2006) implemented an applet as an improvement over the AlexWarp program. In their implementation, TPS was used to allow more control points to be added instead of just one pair, as in the AlexWarp. The TPS was applied in a 2D domain in order to interpolate warped colored images. In our work, we use TPSs on depth images. We use all points from the original depth image and do not intend to apply warping to the data.

Guo et al. (2004) created an average morphable shape represented by a TPS to be used in face recognition applications. The average face was created using a database of 60 individuals (33 males and 27 females) containing only records of asian people which resulted in a model restricted to a particular ethnic group. The facial landmarks, a total of 7, were manually set. In order to reconstruct the face of a subject, they projected the colored 2D image over the average 3D model. A reduced number of control points and the usage of a single TPS are some of the limitations of their work. The authors reported that the results, although not great, showed an interesting potential. In our work, we use several TPS functions to build a different 3D model for each subject. Also, we developed an adaptive block scheme in order to allow the use of all depth values of the image pixels while performing the reconstruction of damaged depth information.

Moreno et al. (2005) developed a 3D face modeling system and used two face recognition methods to test their model, one based on PCA and another one based on *support vector machine* (SVM). Their system aims to work on face images with varying poses, in situations where there is no control over the depth data acquisition. They reported that median and Gaussian filters were applied in the pre-processing stage in order to remove noise and to smooth the curvature of the resulting surfaces. Instead, we propose the use



(a) Minimum (b) Mean (c) Maximum

Fig. 2. Histograms showing the distribution of (a) minimum, (b) mean and (c) maximum squared error values computed from the depth values of a reference image and images produced using the proposed alignment and reconstruction approach (blue), a common 3D ICP-based alignment method with linear reconstruction (green), and a common 2D alignment technique with linear reconstruction (red). Notice that the error values of the proposed approach are smaller than those of the common approaches. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Download English Version:

https://daneshyari.com/en/article/534213

Download Persian Version:

https://daneshyari.com/article/534213

Daneshyari.com