ISSN 2395-1621

Evaluation of Dense 3D Reconstruction from 2D Face Images in the wild : An Overview

Reena Joseph, Preeti Pimpale, Rujuta Hirve, Pallavi Gosavi, Prof. S.R.Khade

UG student, Modern Education Society's College Of Engineering
Professor, Modern Education Society's College Of Engineering

ABSTRACT

As a leading technique in machine learning, Deep Neural Networks (DNNs) have developed themselves. DNNs were top performers on a wide range of tasks, including classification of images, recognition of voice, and recognition of face. In the real-time dataset, Convolutional neural networks (CNNs) were used in almost all the best performing methods on the Labeled Faces. Deep learning based object detection techniques have shown an efficacy to learn the object features directly from the data. The paper mainly focuses on providing a survey on various state-of the- art deep learning based object detection techniques.In general, a visual image is rich in details. Visualization of objects in 3D process is one of the essential ways of extracting information from 2D image slices. The purpose of this research is to visualize 3D objects using 2D images. Due to unknown 3D depth information, the 2D representation lacks a variety of views, realistic effects and information loss. This is the first challenge in 3D face reconstruction with real, accurate and higher 3D ground truth from single 2D in - the-wild images.

Keywords: **2D and 3D images, Deep Learning, DCNN, image processing, Feature Extraction, Face images.**

ARTICLE INFO

Article History

Received: 8th March 2020

Received in revised form : 8th March 2020

Accepted: 10th March 2020

Published online : 11th March 2020

I. INTRODUCTION

That being said, this problem is also a really difficult one, and it has not been until recent years that quality results are being obtained. In fact, this problem is usually split into different sub-problems to make it easier to work with, mainly face detection in an image, followed by the face recognition itself. However, the one that is currently mostly used, and providing the best results, consists in using Deep Learning (DL), especially the Convolutional Neural Networks (CNN). These methods are currently obtaining high quality results, so, after reviewing the current state of art, we decided to focus this project on them. A good example the transition has begun is the increasing progress in the development of 3D object material for cinema applications. While 3D content gives the viewer more concrete meaning, the transmission of 3D objects is much less than 2D broadcasting. Nevertheless, for many applications such as television, movies, and video games, 3D model can be called the next revolution.

The three dimensional face reconstructions from the 2D face image have motivated and attracted a lot of recent attention due to its wide range of applications in biometrics etc. Among those, the outstanding applications are face recognition and expression analysis. Most of the literature surveys on face recognition deals with 2D intensity images rather than 3D object face models. However, the 2D image of a human face changes for many reasons including the expressions, head pose, illumination, aging, makeup etc. Due to 3D nature of face, the pixel intensities in face images changes a lot more with changes in head pose and illumination than with a change in identity of the person. The advantage of using a reconstructed 3D object face model for biometrics is that such a model can be reconstructed offline only once, and can then be used under varying conditions subsequently.Two dimensional is a concept that describes anything that composes of length and width. In two dimensional everything in the image is presented at the same distance from the viewer. But users continuously demand richer, more immersive and closer to reality viewing experiences. After the introduction of color displays and high definition images, 3D object promises to be the next revolution in visual technology. The recent momentum in the production of 3D object content for cinema applications is a good example that the revolution has started. Nevertheless, although many stereoscopic 3D movies have been produced recently, there is a lack of 3D video content, especially for the 3DTV industry. Moreover, 3D object reproduction of conventional well-know 2D movies or TV programs is appealing for both users and content producers. Similarly, there are large no of 2D video available that exist in different compressed format. Though 3D content provide more realistic sense to viewer, the 3D object broadcasting is very less as compared to 2D broadcasting. However, 3D object can be regarded as the next revolution for many applications such as television, movies, and video games. The smash hit movie "Avatar" has demonstrated great success in the use of 3D object and announced the approach of the 3D era. Therefore, there is an urgent need for efficient and robust 2D image to 3D object conversion algorithms. The method which do not require human operator is called automatic methods and the method which require human operator is called as semi automatic method.

II. LITERATURE REVIEW

To measure the correlation between a color area and skin color, a stochastic model is adapted. Both hair-like features and Local Binary Pattern(LBP) features have been reused to create a cascaded classification. The improved classifier is implemented with the focus on skin color to identify the face area from a color image. A skin color-based face detection system using a boosting algorithm that emphasizes skin color information while finding non-skin color information to be phasizing. Since it has a range of applications abroad, human face detection is one of the most important topics in sin biometric study[1].

In the chromatic and pure color space YCrCb, color pictures of skin color separating elements of luminescence and chrominance. The maximum likelihood criterion is used to estimate a Gaussian probability density from skin samples, obtained from different ethnic groups. Adaptive holding to segment the faces within the skin regions detected. For this approach, the main motivation is to use excess entropy as a measure of an image's structural details. The experimental results showed good sensitivity to the implementation of good behavior[2].

The main difference from previously reported facial detection methods based on components is the use of online learning, suitable for highly repetitive activities. This results in faster and more precise face detection as the performance of the system improves with continuous use. In addition, the estimation of the standard deviation of face components and their relationships adds complexity. A component-based uncertain approach provides versatility to define an entity in appearance and geometry in terms of variability. To handle the face pose variation problem under different conditions, part-based face detection methods have been proposed [3].

A deep multi-view hazing network, known as deep multi-view hazing, where each layer of hidden nodes consists of view-specific and shared hidden nodes to learn hidden spaces from multiple views of data. Numerical tests on image data sets show the useful behavior of our deep multi-view hashing (DMVH) compared to the newly proposed multi-modal deep network as well as current shallow hazing models. Although the proposed part-based face detection methods have attained good results, their performance will decrease in case of occlusion or large pose variation, since those models only summed the scores of part-based detectors without exploring the correlations among the visibilities of different face parts[4].

Limited (often minimal) receptive fields of Convolutionary winner-take-all neurons produce large network size, resulting in approximately as many sparsely connected neural layers as found in mammals between retina and visual cortex. Only the neurons of the winners are trained. Many deep neural columns are experts in different ways on inputs preprocessed; their predictions are averaged. Graphics cards allow you to practice quickly. Our approach is the first to achieve near-human performance on the highly competitive MNIST handwriting benchmark. Deep learning originates from artificial neural networks and consists of multi-layer perceptron (MLP) of multi-hidden layers which is a deep learning structure. [5].

The influence on the efficiency of the recognition pipeline of landmark identification, CNN layer selection and pose model selection. In both verification and recognition (i.e. search) tasks, our novel representation produces better results than the state-of - the-art on IARPA's CS2 and NIST's IJB-A. Use a qualified CNN model to represent a face, as well as additional Bayesian metric learning to test the similarities between two face representations.[6]

III. PROPOSED SYSTEM ARCHITECTURE

This system basically focuses image construction from 2D image to 3D images. System used machine deep technique for feature extraction, training as well as testing respectively. 2D image dataset has been used for evaluating the entire system which already existsin 3D image dataset,used for evaluation as ground truth image. DCNN algorithm works for extracting the feature during train module execution while similar approach is performed during testing. Converted 3D image has evaluated with groundtruth image and measure the confusion matrix for classifying the accuracy of system.

Shape descriptors: To draw inferences about 3D objects in both computer vision and graphics literature, a wide corpus of shape descriptors has been created. Shape descriptors can be divided into two broad categories: 3D shape descriptors that work directly on objects ' native 3D representations, such as polygon meshes, voxel-based discretization's, point clouds, or implied surfaces, and view-based descriptors that define the shape of a 3D object by "how it appears" in a 2D projection array.

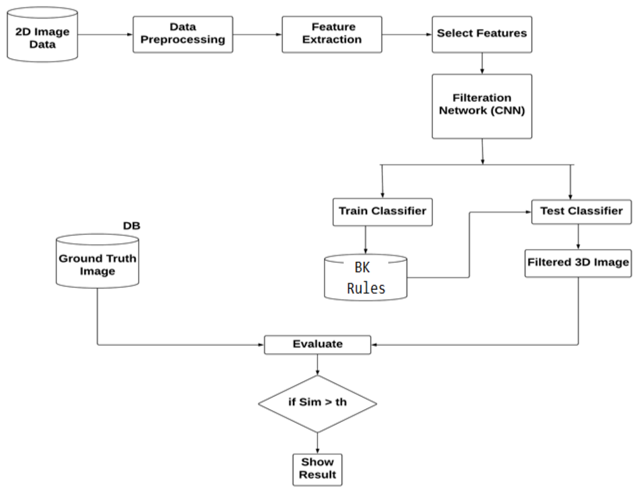

Figure 1 : System block workflow

Convolutional neural networks: Our research is also linked to recent progress in the identification of photos using CNNs. In particular, CNNs trained on large datasets such as Image Net have been shown to learn general-purpose image descriptors for a variety of vision tasks such as object detection, scene recognition, texture recognition and fine-grained classification. To train our network, use both image and form datasets

• System, a fully automatic 2D-to-3D conversion algorithm that takes 2D images or video frames as input and outputs 3D image pairs.

• The stereo images can be evaluate with given ground truth image.

• System is trained directly on 2D features from a dataset of 2D videos to minimize the pixel-wise reconstruction error of the right view when given the left view.

• Internally, the system network estimates a probabilistic disparity map that is used by a differentiable depth image-based rendering layer to produce the right view.

IV. CONCLUSION

We introduced a probabilistic system for matching two sets of features in "3D Face Modeling," extracted automatically from images, taking into account the global structure of the feature sets. The method of normalization is used to generate a doubly stochastic matrix denoting the match probabilities for the two feature sets. While the world is full of 3D forms, we understand that the world is mostly through 2D pictures, at least as human beings. Although even the use of such multiple 2D projections yields impressive performance in discrimination, we can achieve compactness, efficiency, and better accuracy by building descriptors that are aggregations of information from multiple perspectives.

V. FUTURE SCOPE

By quantitatively assessing the problem of 3D face reconstruction, we feel that a number of measures, including can further improve the performance of the system. Because most of the shape variance is in the middle of the nose, using segmented models. Improve the quality of the initial 3D models and the dense 3D correspondence performance. Finding a cost function that is more appropriate for the estimation of shape.

• Neural network learning may be embedded in the proposed system to reduce the semantic gap further.

• Now we are studying about reducing transfer of time and more accuracy to provide good performance.

• The extension of this work for the web based 2D image to 3D object will be used for commercial applications.

REFERENCES

[1] Ban, Yuseok, et al. "Face detection based on skin color likelihood."Pattern Recognition 47.4 (2014): 1573-1585.

[2] Chandrappa, D. N., M. Ravishankar, and D. R. RameshBabu. "Face detection in color images using skin color model algorithm based on skin color information."2011 3rd International Conference on Electronics Computer Technology. Vol. 1. IEEE, 2011.

[3] Lee, Kyoung-Mi. "Component-based face detection and verification."Pattern Recognition Letters 29.3 (2008): 200-214.

[4] Kang, Yoonseop, Saehoon Kim, and Seungjin Choi. "Deep learning to hash with multiple representations."2012 IEEE 12th International Conference on Data Mining. IEEE, 2012

[5] Cireşan, Dan, Ueli Meier, and Jürgen Schmidhuber. "Multi-column deep neural networks for image classification."arXiv preprint arXiv:1202.2745 (2012).

[6] AbdAlmageed, Wael, et al. "Face recognition using deep multi-pose representations."2016 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2016.

[7] Zhang, Qi, Roy Eagleson, and Terry M. Peters. "Rapid voxel classification methodology for interactive 3D medical image visualization."International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Berlin, Heidelberg, 2007.

[8] Su, Hang, et al. "Multi-view convolutional neural networks for 3d shape recognition."Proceedings of the IEEE international conference on computer vision. 2015.

[9] Bakic, Predrag R., Frederic JP Richard, and Andrew DA Maidment. "2D-to-3D x-ray breast image registration."International Workshop on Biomedical Image Registration. Springer, Berlin, Heidelberg, 2006.

[10] Qi, Yujuan, Yanjiang Wang, and Yuchi Liu. "Object Tracking Based on Deep CNN Feature and Color Feature."2018 14th IEEE International Conference on Signal Processing (ICSP). IEEE, 2018.

