|  |  |
| --- | --- |
| **СОГЛАСОВАНО**  Доцент кафедры  ИАНИ ННГУ, к.ф.-м.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Д.А. Яшунин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. | **УТВЕРЖДАЮ**  Профессор кафедры  ИАНИ ННГУ, д.т.н.  \_\_\_\_\_\_\_\_\_\_\_\_\_ Н.В. Старостин  «\_\_\_\_»\_\_\_\_\_\_\_\_\_\_\_\_\_\_2019 г. |

**Пояснительная записка №2**

**«По методам детекции лиц по RGBD»**

**ОКР «Разработка и реализация программного обеспечения для обнаружения и распознавания лиц с использованием RGBD камеры»**

**(Шифр ПО «DeepFR»)**

**Н. Новгород**

**201**

**Обзор подходов детекции лиц по RGBD данным**

«RGB-D Face Recognition with Texture and Attribute Features»

Gaurav Goswami, Student Member, IEEE, Mayank Vatsa, Senior Member, IEEE, and Richa Singh, Senior Member, IEEE

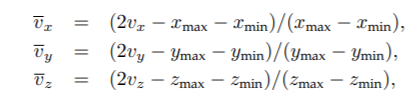
[**http://iab-rubric.org/papers/RGBD-Face.pdf**](http://iab-rubric.org/papers/RGBD-Face.pdf)

«An automatic face detector (Viola-Jones face detector) is applied on the RGB image to obtain the face region. The corresponding region is also extracted from the depth map to crop the face region in depth space. While texture feature descriptor does not require image size normalization, the images are resized to 100×100 to compute depth features. Depth map is then preprocessed to remove noise (holes and spikes). Depth map of a face is divided into 25×25 blocks and each block is examined for existence of holes and spikes. Depth values identified as the hole/spike are rectified using linear interpolation, i.e. assigned the average value of their 3×3 neighborhood»

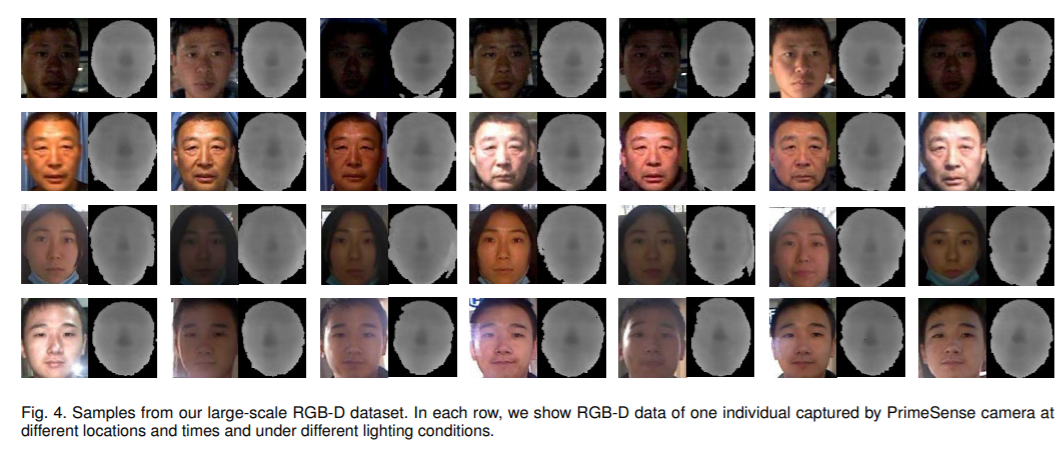
«Robust RGB-D Face Recognition Using Attribute-Aware Loss»

Luo Jiang, Juyong Zhang† , Member, IEEE, and Bailin Deng, Member, IEEE

[**https://arxiv.org/pdf/1811.09847.pdf**](https://arxiv.org/pdf/1811.09847.pdf)

«Using the RGB part of a facial image, we first detect the face region and five landmarks (the eyes, the nose, and two corners of the mouth) using MTCNN [52]. The face is then cropped to 112 × 96 by similarity transformation, and each RGB color component is normalized from the range [0, 255] into [−1, 1]. Afterward, we extract a face region from the corresponding depth image by transferring the RGB face region. Similar to [16], [53], we find the nose tip and crop the point cloud in the face region within an empirically set radius of 90mm. Then we move the center of the cropped facial scan to (0, 0, zopt) and reproject it onto a 2D image plane to generate a new depth map of size 112 × 96. The value zopt is chosen to enlarge the projection of facial scans onto the image plane as much as possible. Following [9], we compute the depth of each pixel with bilinear interpolation. Using this depth map, we generate a new point cloud under the camera coordinate system. Each point (vx, vy, vz) is further normalized as:

where (xmin, ymin, zmin) and (xmax, ymax, zmax) are the minimum and maximum x-, y- and z-coordinate values among all points, respectively. Augmenting the RGB face region with its normalized point cloud, we obtain a six-channel image with values in [−1, 1]6 , which is fed into the deep neutral network.»



«Face recognition using depth and infared pictures»

Soon-kak Kwon

[**https://www.jstage.jst.go.jp/article/nolta/10/1/10\_2/\_pdf/-char/en**](https://www.jstage.jst.go.jp/article/nolta/10/1/10_2/_pdf/-char/en)

Находится точка носа на основе карты глубины. Затем находится область лица

«Real-Time 3D Face Identification from a Depth Camera»

Rui Min, Jongmoo Choi, Gérard Medioni, Jean-Luc Dugelay

[**http://www.eurecom.fr/en/publication/3764/download/mm-publi-3764.pdf**](http://www.eurecom.fr/en/publication/3764/download/mm-publi-3764.pdf)

«The output from the PrimeSensor includes a RGB image and a depth map at 640×480 resolution. Although the face detection could be achieved by the popular ViolaJones’ method [13] using RGB images, it cannot segment the face/head region exactly from the background/body part. In addition, RGB images are sensitive to illumination variations. Therefore, we focus on the depth information from the range camera (which is illuminationinvariant). Because pixels on the head surface have close depth values, given a pre-defined threshold, it is easy to segment the head region according to the depth discontinuity. Given the segmented visible surface of a head, we first subsample points at a fixed resolution (60×60 in our system) and then compute corresponding real-world 3D coordinates. Any face with a lower resolution is automatically rejected as an invalid face candidate.»

[**https://github.com/cating341/RGB-D-Face-Recognition**](https://github.com/cating341/RGB-D-Face-Recognition)

Simple RGB-D face recognition implementation.

Input: Combine RGB and Depth images into 4 channels input data

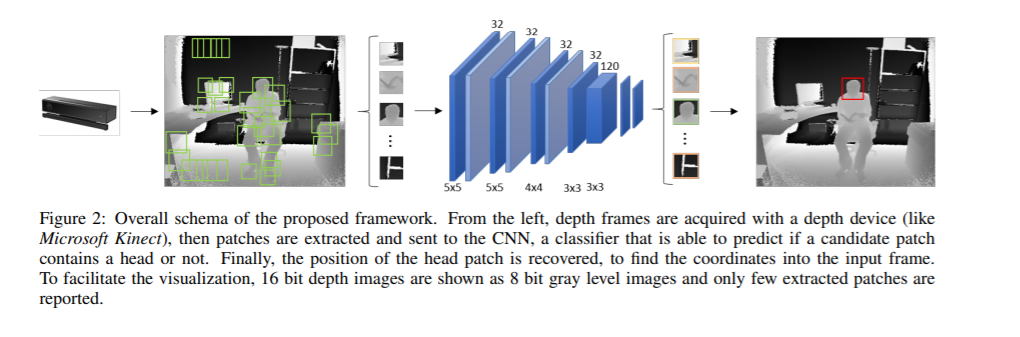
Network: ResNet

**Head detection**

«Head Detection with Depth Images in the Wild»

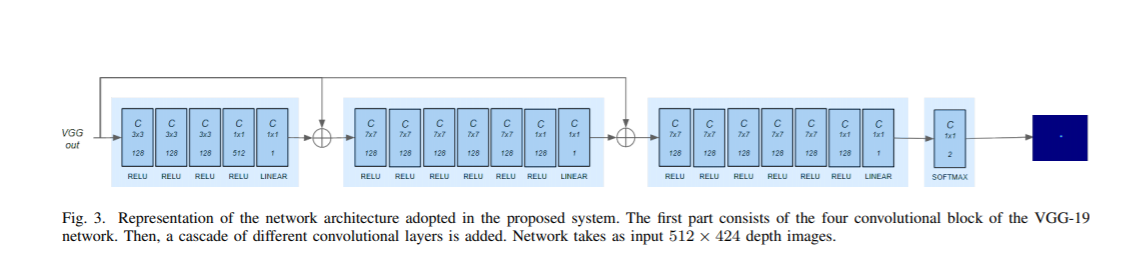
Diego Ballotta, Guido Borghi, Roberto Vezzani and Rita Cucchiara

[**https://arxiv.org/pdf/1707.06786.pdf**](https://arxiv.org/pdf/1707.06786.pdf)



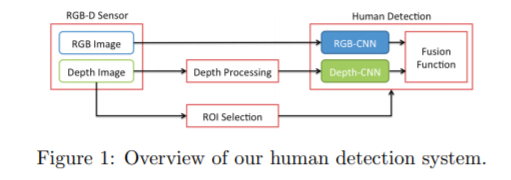
«Fully Convolutional Network for Head Detection with Depth Images»

Diego Ballotta, Guido Borghi, Roberto Vezzani and Rita Cucchiara

[**https://iris.unimore.it/retrieve/handle/11380/1159887/195233/icpr-2018-fully.pdf**](https://iris.unimore.it/retrieve/handle/11380/1159887/195233/icpr-2018-fully.pdf)

«Exploring RGB+Depth Fusion for Real-Time Object Detection»

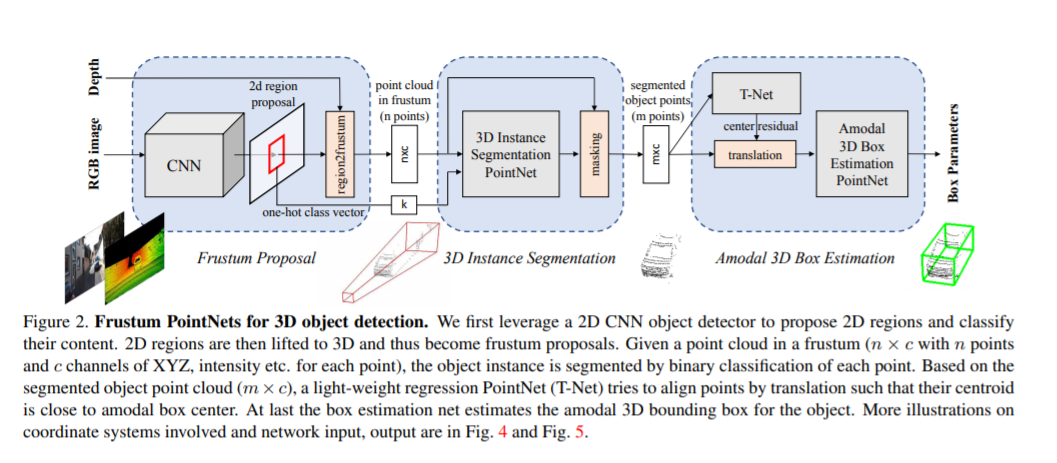
Tanguy Ophoff , Kristof Van Beeck and Toon Goedemé

[**https://www.mdpi.com/1424-8220/19/4/866/pdf-vor**](https://www.mdpi.com/1424-8220/19/4/866/pdf-vor)

«Frustum PointNets for 3D Object Detection from RGB-D Data»

Charles R. Qi, Wei Liu, Chenxia Wu, Hao Su, Leonidas J. Guibas

<http://zpascal.net/cvpr2018/Qi_Frustum_PointNets_for_CVPR_2018_paper.pdf>



**«rgb-d face recognition a comparative study of representative fusion schemes» - must see**