|  |  |
| --- | --- |
| **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 №4**

**«По известным датасетам для RGBD и собранному датасету»**

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

**(Шифр ПО «DeepFR»)**

**Н. Новгород**

**2019**

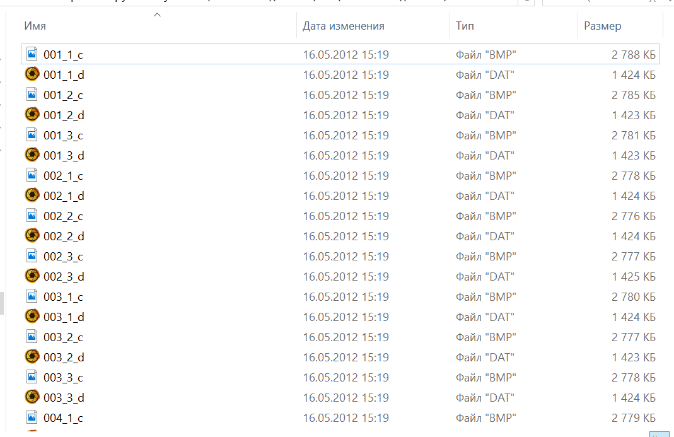
**Datasets overview**

**RGB-D Face database** - [http://vap.aau.dk/rgb-d-face-database/](http://www.vap.aau.dk/rgb-d-face-database/)

«This paper purposes an RGB-D database containing 1581 images (and their depth counterparts) taken from 31 persons in 17 different poses and facial expressions using a Kinect device. The faces in the images are not extracted neither in the RGB images nor in the depth hereof, therefore they can be used for both detection and recognition. The proposed database has been used in a face detection algorithm which is based on the depth information of the images. The challenges and merits of the database have been highlighted through experimental results»

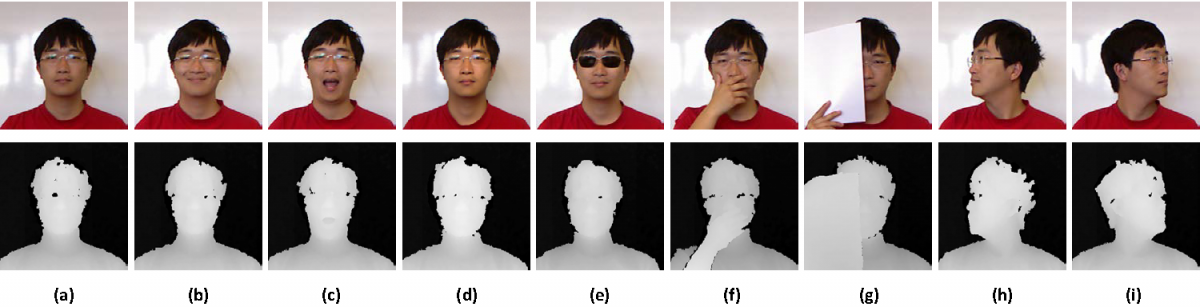
**Можно свободно скачать**

Пример: rgb - .bmp file, depth - .dat file (не открыла)



**Kinect Face Dataset** - <http://rgb-d.eurecom.fr/>

«The Dataset consists of the multimodal facial images of 52 people (14 females, 38 males) obtained by Kinect. The data is captured in two sessions happened at different time period (about half month). In each session, the dataset provides the facial images of each person in 9 states of different facial expressions, different lighting and occlusion conditions: neutral, smile, open mouth, left profile, right profile, occlusion eyes, occlusion mouth, occlusion paper and light on **[Figure 1]**. All the images are provided in three sources of information: the RGB color image, the depth map (provided in two forms of the bitmap depth image and the text file containing the original depth levels sensed by Kinect) as well as 3D. In addition, the dataset comes with the manual landmarks of 6 positions in the face: left eye, right eye, the tip of nose, left side of mouth, right side of mouth and the chin **[Figure 2]**. Other information of the person such as gender, year of birth, glasses (this person wears the glasses or not), capture time of each session are also available»



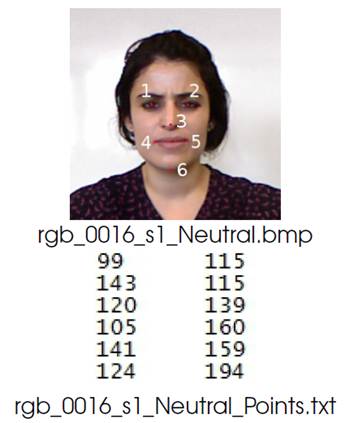
Figure 1: Illustration of different facial variations acquired in our database: (a) neutral face; (b) smiling; (c) mouth open; (d) strong illumination; (e) occlusion by sunglasses; (f) occlusion by hand; (g) occlusion by paper; (h) right profile face and (i) left profile face. (Upper: the RGB images. Lower: the depth maps aligned with above RGB images.).

Figure 2: The marker point positions in the neutral image and the corresponding .txt file.

**Свободно скачать нельзя:**

«Please fill online this form to request for the Database. You need to be a representative for your organization (students are not accepted) and use your official email address in the organization in order to request for the Database. After filling the form, an email will be automatically sent to you with the Usage Agreement between your organization and EURECOM, in addition with the instruction to get the Database»

**Pandora dataset** - <http://imagelab.ing.unimore.it/pandora/>

Pandora has been specifically created for head center localization, head pose and shoulder pose estimation and is inspired by the automotive context. Pandora contains more than 250k full resolution RGB (1920x1080 pixels) and depth images (512x424) with the corresponding annotation; 110 annotated sequences using 10 male and 12 female actors. Each subject has been recorded five times. A Microsoft Kinect One device is used to acquire depth data, with a better quality than other datasets created with the first Kinect version, as reported in the paper.

Data: each frame of the dataset is composed of the RGB appearance image, the corresponding depth map, the 3D coordinates of the skeleton joints corresponding to the upper body part, including the head center and the shoulders positions. For convenience's sake, the 2D coordinates of the joints on both color and depth frames are provided as well as the head and shoulder pose angles with respect to the camera reference frame. Shoulder angles are obtained through the conversion to Euler angles of a corresponding rotation matrix.

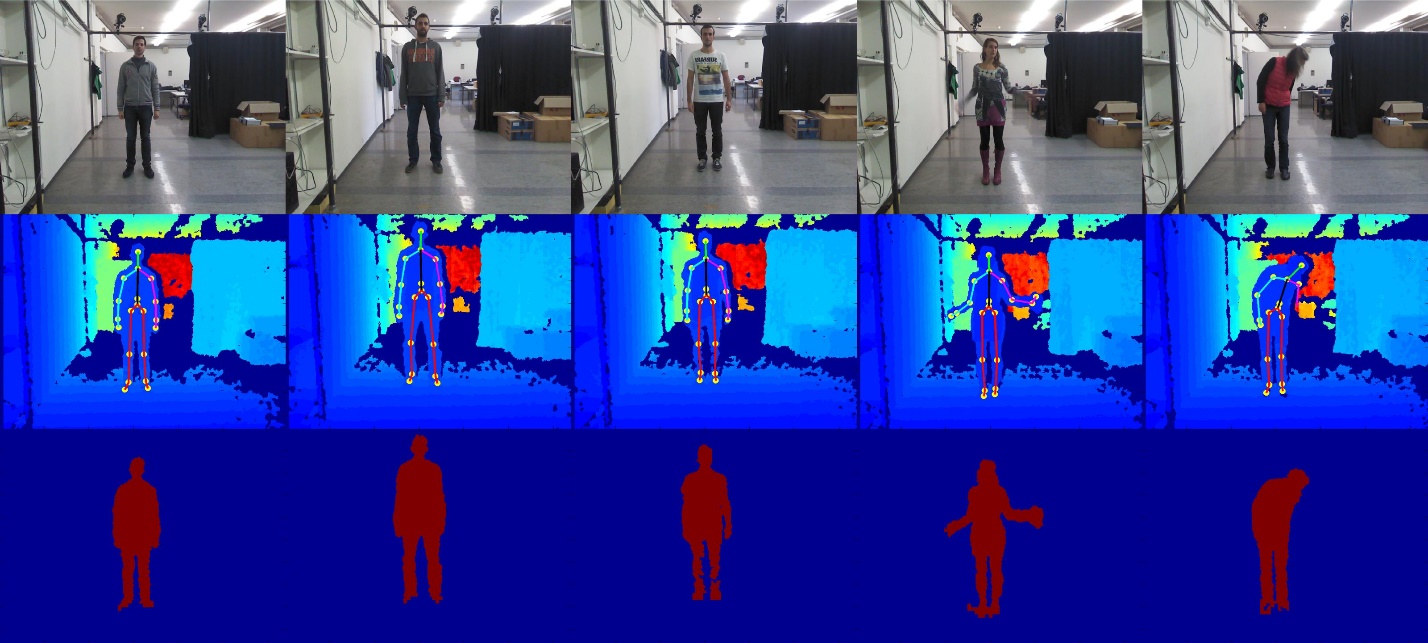


**Свободно скачать нельзя:**

«Human face is very personal, we only send the data to approved researchers. You need to be a representative for your organization (students are not accepted) and use your official email address in the organization in order to request for the Pandora dataset»

**BIWI RGBD-ID dataset** - <http://robotics.dei.unipd.it/reid/index.php/downloads>

The **BIWI RGBD-ID Dataset** is a RGB-D dataset of people targeted to long-term people re-identification from RGB-D cameras. It contains 50 training and 56 testing sequences of **50 different people**. The dataset includes synchronized RGB images (captured at the highest resolution possible with a Microsoft Kinect for Windows, i.e. 1280x960 pixels), depth images, persons' segmentation maps and skeletal data (as provided by Microsoft Kinect SDK), in addition to the ground plane coordinates. These videos have been acquired at about 10fps.



**Можно свободно скачать**

**IAS-Lab RGB-D Face dataset -** <http://robotics.dei.unipd.it/reid/index.php/downloads>

<http://robotics.dei.unipd.it/reid/index.php/8-dataset/9-overview-face>

The **IAS-Lab RGB-D Face Dataset** has been created to measure accuracy and precision of 2D and 3D face recognition algorithms based on data coming from consumer RGB-D sensors. The Kinect v2 (or Kinect One) has been used to acquire this dataset. The **training** dataset consists of **26 subjects** captured in 13 different conditions (with **pose**, **light** and **expression** variations), standing 1 or 2 meters from the sensor.

In order to represent a typical **service robotics** scenario, where few people have to be recognized and many others have to be classified as unknown, the **testing** dataset contains **19 subjects** and just four of them were also present in the training dataset. The other testing subjects can thus be considered as unknown.

The IAS-Lab RGB-D Face Dataset provides two different files for every frame:

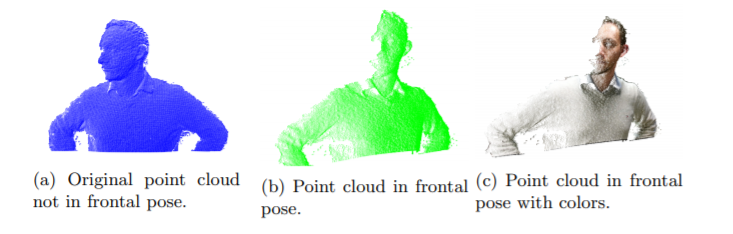
* RGB image (1920x1080 resolution)
* XYZRGB point cloud (960x540 resolution)

The point cloud is registered to the RGB image and its resolution is downsampled by two.

**Можно свободно скачать**

.png –rgb file, .pcd- point cloud file (не открывала)



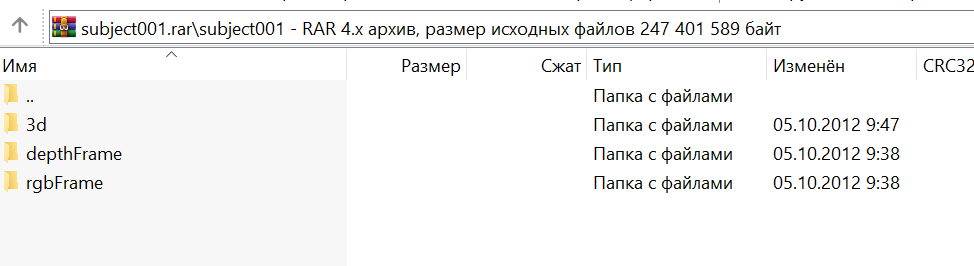
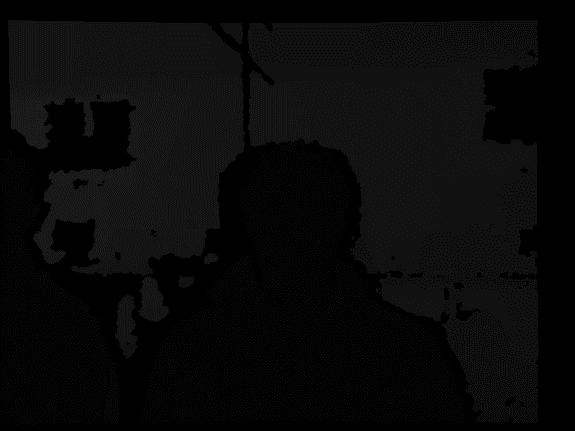


**Florence Superface dataset** - <https://www.micc.unifi.it/resources/datasets/florence-superface/>

The Florence Superface dataset comprises low-resolution and high-resolution 3D scans aiming to investigate innovative 3D face recognition solutions that use scans at different resolutions. Currently, 20 subjects are included in the dataset, but enrolling is still ongoing. For each subject, the dataset includes:

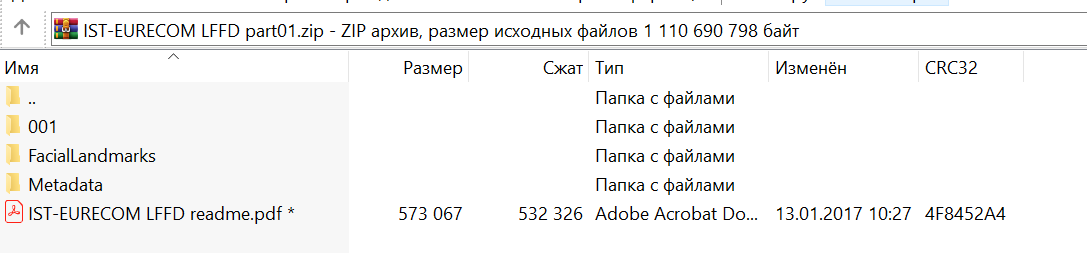
* A 2D/3D video sequence acquired with the Microsoft Kinect. During capture, subjects sit in front of the camera with the face at an approximate distance of 80cm from the sensor. Subjects are also asked to slightly rotate the head around the yaw axis up to an angle of about 60-70 degrees, so that both the left and right side of the face are visible to the sensor. This results in video sequences lasting approximately 10 to 15 sec. Videos are released as a sequence of depth (16 bits) and rgb (24 bits) frames in PNG format;
* A 3D high-resolution face scan acquired with the 3dMD scanner: 3D mesh with about 40,000 vertices and 80,000 facets; texture stereo image with a resolution of 3341 x 2027 pixels. The geometry of the mesh is highly accurate with an average RMS error of about 0.2mm or better (VRML format)

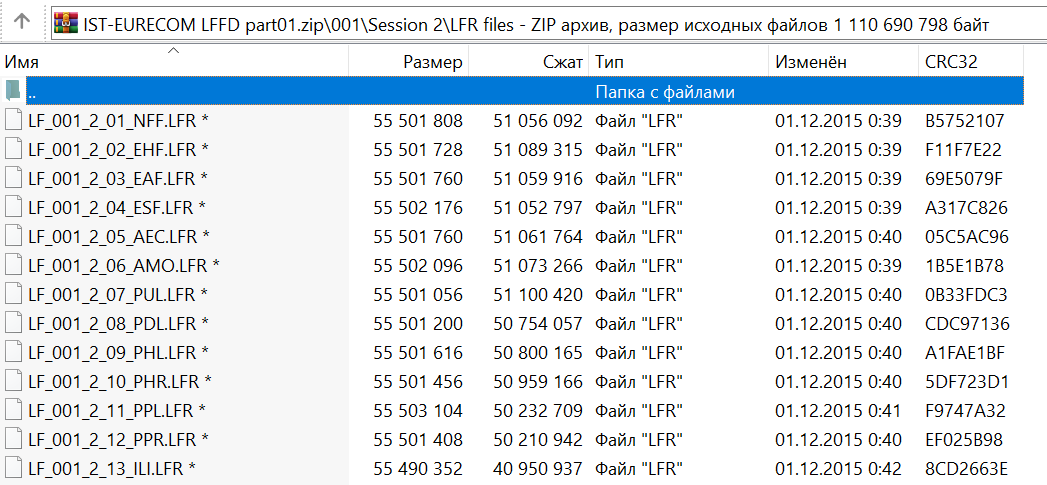
**Можно свободно скачать**



**IST-EURECOM Light Field Face Database** - <http://www.img.lx.it.pt/LFFD/>

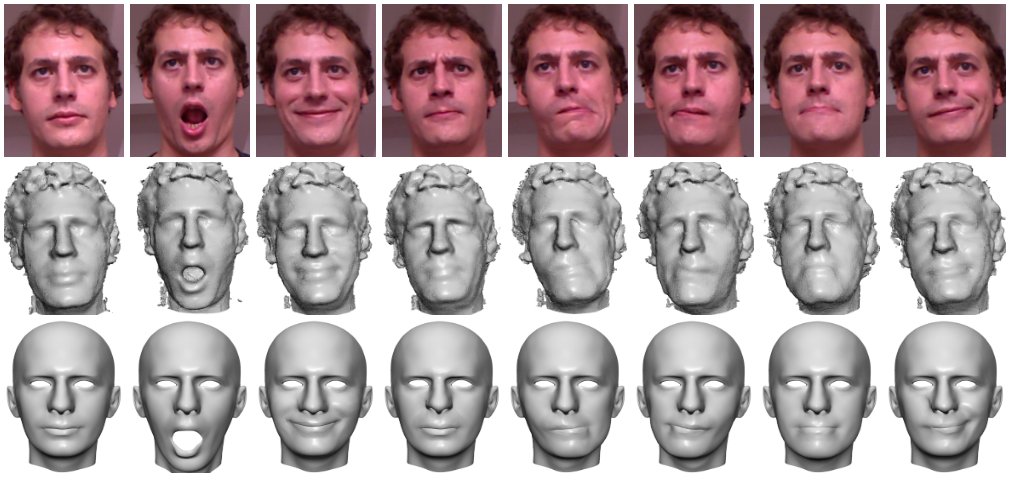
The database includes 20 image shots per person in each of the two sessions, with several facial variations including emotions, actions, poses, illuminations, and occlusions.  The IST EURECOM Light Field Face Database is the first database to include the raw light field images, sample 2D rendered images and the corresponding depth maps along with a rich collection of metadata, including the location of a set of facial landmarks.

**Можно скачать файлы, но необходимо заполнить какую-то форму, и авторы пришлют пароль для расшифровки данных**



**FaceWarehouse** - <http://kunzhou.net/zjugaps/facewarehouse/>

«FaceWarehouse is a database of 3D facial expressions for visual computing applications. Using a Kinect RGBD camera, we captured 150 individuals aged 7-80 from various ethnic backgrounds. For each person, we captured the RGBD data of her different expressions, including the neutral expression and 19 other expressions such as mouth-opening, smile, kiss, etc. For every RGBD raw data record, a set of facial feature points on the color image such as eye corners, mouth contour and the nose tip are automatically localized, and manually adjusted if better accuracy is required. We then deform a template facial mesh to fit the depth data as closely as possible while matching the feature points on the color image to their corresponding points on the mesh. From these fitted face meshes, we construct a set of individual-specific expression blendshapes for each person»

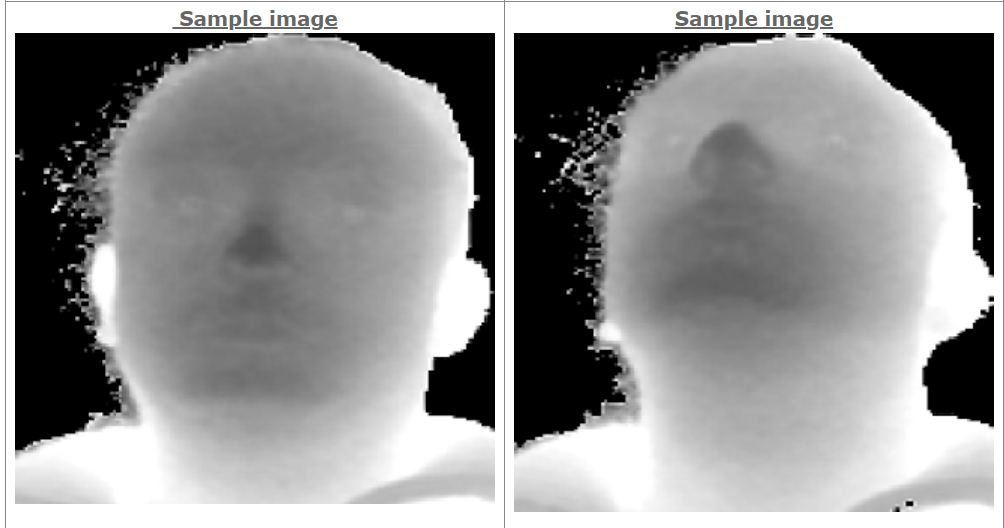


**Свободно скачать нельзя**:

«We make all RGBD scans, feature points and the fitted meshes available for academic research purposes. However, as human face is very personal, we only send the data to approved researchers. To obtain a copy, please send an email»

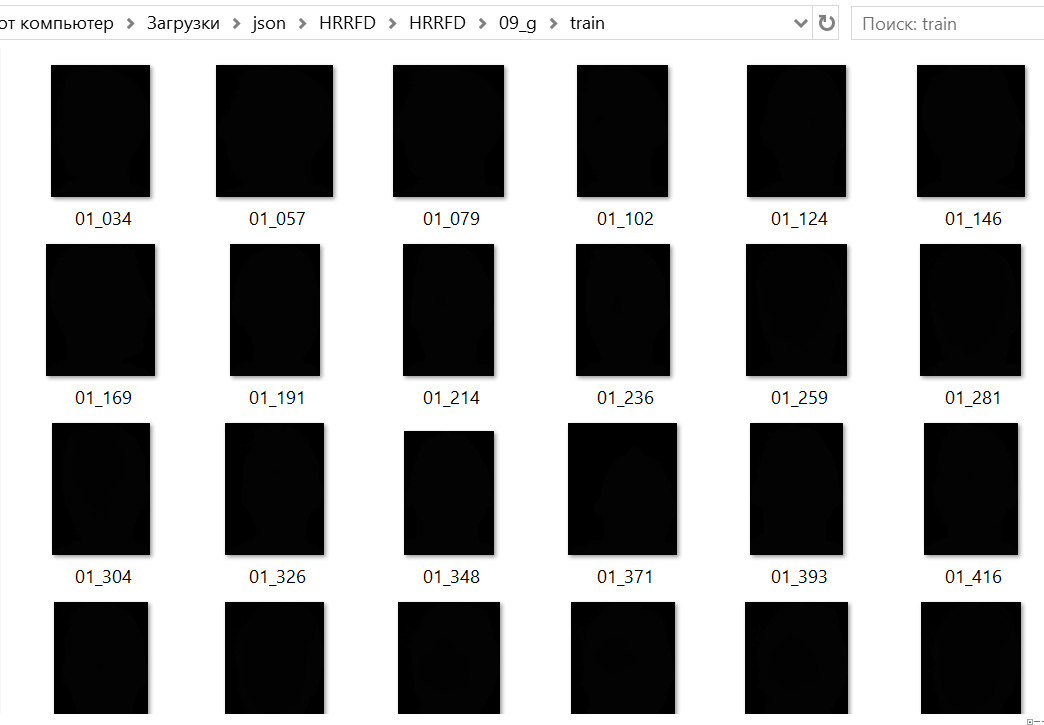
**High Resolution Range based Face Database (HRRFaceD)** -<https://sites.google.com/site/hrrfaced/>

A face database is presented composed by a set of high resolution range images acquired by the latest generation of range / depth cameras: the [Microsoft Kinect 2](http://www.microsoft.com/en-us/kinectforwindowsdev/newdevkit.aspx) (second generation). The database is composed by the faces of 18 people, acquired from different poses: frontal, lateral, etc. The faces of some of the people have been acquired with and without glasses.



**Можно свободно скачать**

Но скачалось только непонятно что:



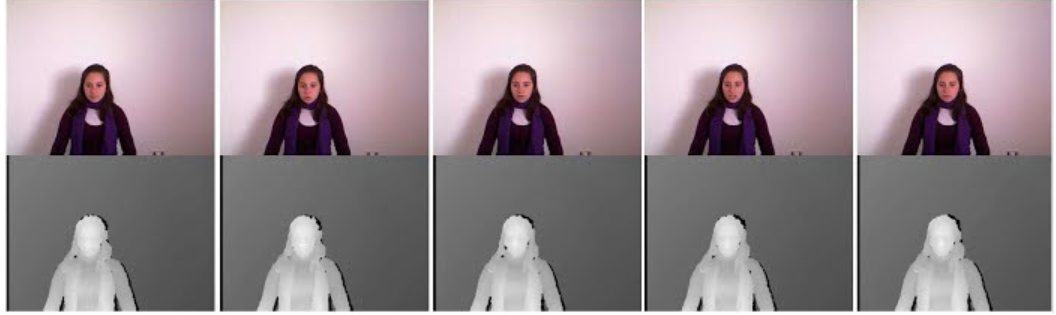
## **Biwi Kinect Head Pose Database -** <https://data.vision.ee.ethz.ch/cvl/gfanelli/head_pose/head_forest.html>

The dataset contains over 15K images of 20 people (6 females and 14 males - 4 people were recorded twice). For each frame, a depth image, the corresponding rgb image (both 640x480 pixels), and the annotation is provided. The head pose range covers about +-75 degrees yaw and +-60 degrees pitch. Ground truth is provided in the form of the 3D location of the head and its rotation. Even though our algorithms work on depth images alone, we provide the RGB images as well. Please note that this is a database acquired with frame-by-frame estimation in mind, not tracking. For this reason, some frames are missing.

**Можно скачать свободно**

**MIRACL-VC1** - <https://sites.google.com/site/achrafbenhamadou/-datasets/miracl-vc1>

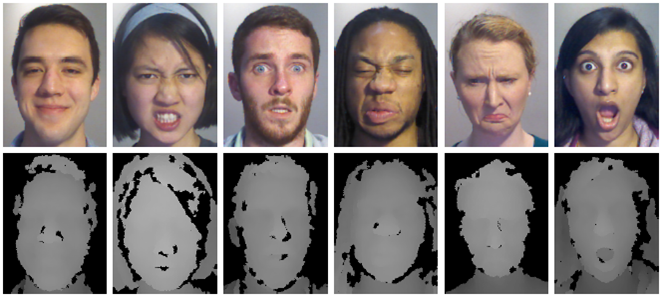
## «MIRACL-VC1 is a lip-reading dataset including both depth and color images. It can be used for diverse research fields like visual speach recognition, face detection, and biometrics. Fifteen speakers (five men and ten women) positioned in the frustum of a MS Kinect sensor and utter ten times a set of ten words and ten phrases (see the table below). Each instance of the dataset consists of a synchronized sequence of color and depth images (both of 640x480 pixels).  The MIRACL-VC1 dataset contains a total number of 3000 instances»



**Можно свободно скачать**

**VT-KFER dataset** - <https://computervisiononline.com/dataset/1105138604>

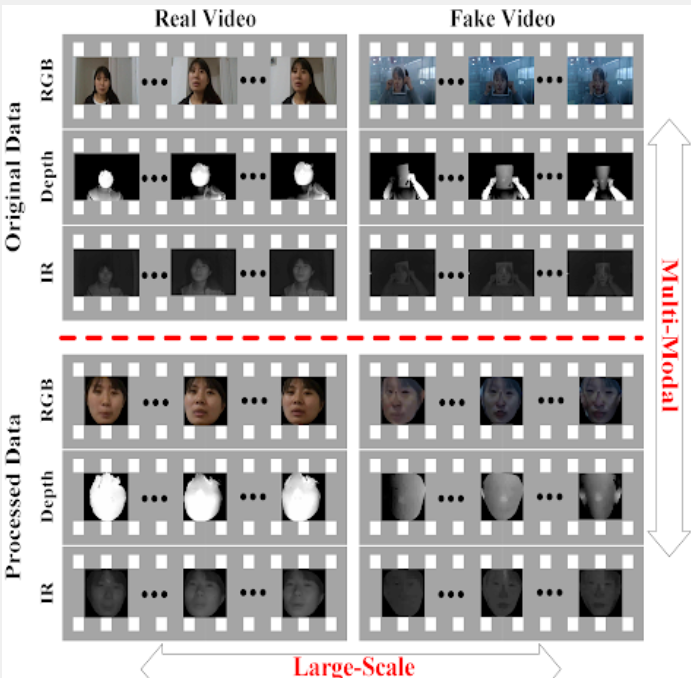
VT-KFER dataset is a RGBD+time facial expression recognition dataset collected using the Kinect 1.0 sensor in both scripted (acted) and unscripted (spontaneous) scenarios. This fully annotated dataset includes seven expressions (happiness, sadness, surprise, disgust, fear, anger, and neutral) for 32 subjects (males and females) aged from 10 to 30 and with different skin tones. Both human and machine evaluation were conducted. Each scripted expression was ranked quantitatively by two persons.



**Не удалось найти доступ или скачать**

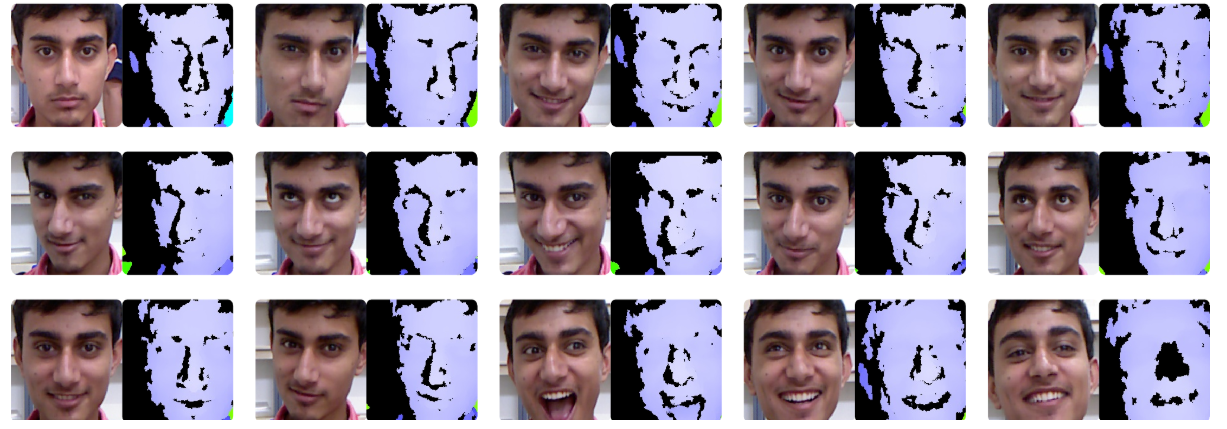
**CASIA-SURF Dataset** - <https://sites.google.com/qq.com/chalearnfacespoofingattackdete/dataset>

**Chalearn CASIA-SURF** is a face Anti-spoofing dataset which consists of 1, 000 subjects and 21, 000 video clips with 3 modalities (RGB, Depth, IR). In the proposed dataset, each sample includes 1 live video clip, and 6 fake video clips under different attack ways (one at tack way per fake video clip).  We use the Intel RealSense SR300 camera to capture the RGB, Depth and Infrared (IR) videos simultaneously.



**Не удалось найти доступ или скачать**

**IIIT-D Kinect RGB-D Face Database -** <http://www.iab-rubric.org/resources/rgbd.html>



**Можно свободно скачать, но необходимо писать письмо для запроса пароля на расшифровки файлов???**

**Generation depth data from 2d (RGB) data (???)**

<https://github.com/ialhashim/DenseDepth>

<https://github.com/shalomma/DEN> -

<https://github.com/nianticlabs/monodepth2> &

<https://github.com/fangchangma/sparse-to-dense.pytorch>

<https://github.com/juniorxsound/ReTouch>

<https://github.com/Rostifar/NYUDepthNet>

<https://github.com/imran3180/depth-map-prediction>

|  |  |  |  |
| --- | --- | --- | --- |
| Описание | Суммарное количество | Количество фотографий в датасете | Датасет |
| **Доступны, можно скачать** | ~84K | 1581 | [http://vap.aau.dk/rgb-d-face-database/](http://www.vap.aau.dk/rgb-d-face-database/) |
| ~40K | <http://robotics.dei.unipd.it/reid/index.php/downloads>  (там человек целиком, а не только лицо) |
| ~718 | <http://robotics.dei.unipd.it/reid/index.php/8-dataset/9-overview-face>  (rgb + point cloud) |
| ~7.2K | <https://www.micc.unifi.it/resources/datasets/florence-superface/> |
| ~17K (?) | <https://sites.google.com/site/hrrfaced/> (скачались просто черные изображения) |
| ~15K | <https://sites.google.com/site/hrrfaced/> |
| ~3K | <https://sites.google.com/site/achrafbenhamadou/-datasets/miracl-vc1> |
| **Можно скачать по запросы / можно скачать, но нужен запрос для расшифровки данных** | ~253K | ? | <http://rgb-d.eurecom.fr/> (?, more than 1500) |
| 250K | <http://imagelab.ing.unimore.it/pandora/> |
| ~500 | <http://www.img.lx.it.pt/LFFD/> |
| ~3K | <http://kunzhou.net/zjugaps/facewarehouse/> |
| ? | <http://www.iab-rubric.org/resources/rgbd.html> (немного) |
| **Неизвестно** | ? | ? | <https://computervisiononline.com/dataset/1105138604> |
| ? | <https://sites.google.com/qq.com/chalearnfacespoofingattackdete/dataset>  (антиспуф, видео?) |