BP4D-Spontaneous

Binghamton-Pittsburgh 3D Dynamic (4D) Spontaneous Facial Expression Database [1] [2]

Contact:

Dr. Lijun Yin of Computer Science, State University of New York

Dr. Jeff Cohn of Psychology, University of Pittsburgh

USER GUIDE

Database Organization

BP4D-Spontaneous contains 3D dynamic spontaneous facial expression data of 41 subjects. Each subject went through 8 tasks. The elicitation methods of the 8 tasks and the corresponding emotion are listed in table 1.

TABLE 1: EIGHT TASKS FOR EMOTIONAL EXPRESSION ELICITATION

Task	Activity	Target Emotion
1	Talk to the experimenter and listen to a joke (Interview).	Happiness or Amusement
2	Watch and listen to a recorded documentary and discuss their reactions.	Sadness
3	Experience sudden, unexpected burst of sound.	Surprise or startle
4	Play a game in which they improvise a silly song.	Embarrassment
5	Anticipate and experience physical threat.	Fear or nervous
6	Submerge their hand in ice water for as long as possible.	Physical pain
7	Experience harsh insults from the experimenter.	Anger or upset
8	Experience an unpleasant smell.	Disgust

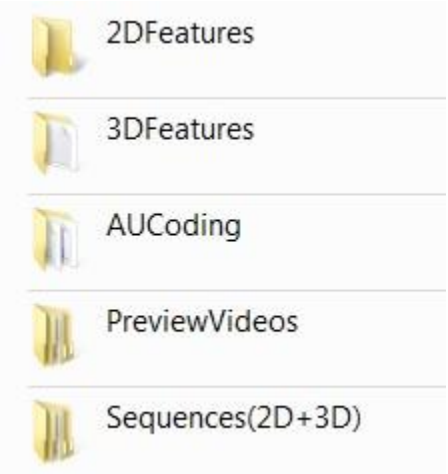

Figure 1: Structure of BP4D-Spontaneous

The database provides the user with 3D model sequences and 2D texture sequences, Metadata (including 2D tracking data, 3D tracking data, and Action Unit coding,) and preview 2D videos. Figure 1 shows the directories in the database. The contents and usage of the data is described in the following sections.

Model Data

Sequences(2D+3D) directory contains the model data of 41 subjects, indexed in FXXX or MXXX, which means female or male subjects. Each subject folder contains 8 tasks. Each task contains Wavefront .obj formatted 3D model sequence. Each frame associates with 3 files. For example, frame 0000 associates with 0000.jpg, 0000.mtl, and 0000.obj, among which, the material file (.mtl) records the material property of the model, and .jpg is the 2D texture image.

The database includes totally 368,036 frames.

Two Dimensional Tracking Data

Directory **2DFeatures** contains the 2D tracking data for each sequence in a MATLAB formatted binary file (.mat). The file is named as [Gender][Index]_T[TaskID].mat. [Gender] is either F or M. [Index] is the three digits ID of the subject. [TaskID] is one digit ID for 8 tasks.

Each file contains 2D tracking data for one task based on the 2D texture image sequence with image size of 1040x1392. Each file records one variable named 'fit', which should be a column structure array. Each element contains a structure with two members. 'pred' is a 49x2 matrix that reported 49 two dimensional tracking points for the frame. Their relation and indices can be found in Figure 2 below. 'pose' is another structure contains pose information. It provides two representations, 'rot' is the 3x3 rotation matrix and 'angle' is a vector reporting the rotation degree along the axes. If certain frame does not have 2D tracking information, 'pred' is empty and can be verified by MATLAB function isempty().

There are totally 328 files.

Three Dimensional Tracking Data

Directory **3DFeatures** contains the 3D tracking data which is arranged in the same directory pattern as the model data is. The 3D coordinate is defined in left-handed Cartesian coordinate system. Each frame is tracked with 83 feature points on 3D surface. Their relation and indices can be found in Figure 3 below. All the coordinates are recorded in the TBND file following the index order. Besides, the first line in the TBND file reports the pose information of the model in degrees. The three values are the angle of pitch, yaw and roll in range of [-180, 180].

Five hundred and forty-four frames do not contain 3D tracking information (.tbnd). Their location is reported in **3DFeatures\Loss.txt**. There are totally 367,492 TBND files. A few frames are partially un-trackable, indicated by (0,0,0) as certain vertices coordinates.

Action Units Coding Data

Directory **AUCoding** contains the action units (AU) coding of the sequences. The comma-separated values (.csv) file saves the coding information. There are two types of FACS annotation files: occurrence (**AU_OCC**) and intensity (**AU_INT**) files. For each task sequence, approximately 500 frames (20 seconds) are coded. The occurrence file is named as [Gender][Index]_T[TaskID].csv. [Gender] is either F or M. [Index] is the three digits ID of the subject. [TaskID] is one digit ID for 8 tasks. And the intensity file is named as [Gender][Index]_T[TaskID]_[AUID].csv, where [AUID] is the action unit which is coded for the corresponding task.

Each video file has a single occurrence annotation file which contains the annotations for all action units. The occurrence files are in the CSV format and are essentially numerical matrices in which each column corresponds to a single action unit and each row corresponds to a single video frame. The files also contain a header row with action unit labels and a header column with frame numbers. The proper way to parse these files is to drop the headers and then index for the video frame(s) and action unit(s) of interest. The following examples in MATLAB code may be helpful:

OCC = csvread('F001_T1.csv'); %import annotations for one video file

FRNO = OCC(2:end,1); %get all frame numbers

CODES = OCC(2:end,2:end); %get codes for all action units

AU12 = CODES(:,12); %index action unit 12 codes for all video frames

FR10 = CODES(FRNO==10,:); %index all action unit codes for video frame 10

AU12_FR10 = CODES(FRNO==10,12); %index code for action unit 12 for video frame 10

The occurrence codes themselves are either 0 for absent, 1 for present, or 9 for missing data (unknown). Codes will be missing when the video frame was not coded by the manual annotator (not all frames and AUs were). Occlusion was not coded in the occurrence files and a rule was used that if an action unit was present before an occlusion and was still present after it, the action unit was marked as occurring during the occlusion.

Each video file has multiple intensity annotation files: one per action unit. The action unit annotated in each file is indicated in its file name. The intensity files are also in the CSV format and are essentially numerical matrices in which each row corresponds to a single video frame. The first column lists the frame numbers and the second column lists the intensity codes. There is no header row in these files. The following examples in MATLAB code may be helpful:

INT = csvread('F001_T1_AU12.csv'); %import annotations for one video file

FRNO = INT(:,1); %get all frame numbers

AU12_FR10 = INT(FRNO==10,2); %index code for action unit 12 for video frame 10

The intensity codes themselves are either 0 for absent, 1 for present at the A level, 2 for present at the B level, 3 for present at the C level, 4 for present at the D level, 5 for present at the E level, or 9 for missing data (unknown). Codes will be missing when the video frame was not coded by the manual annotator (not all frames were) and when the face was occluded during the video frame.

There are totally 328 occurrence files. Since there are 5 AUs coded with intensity (AU06, AU10, AU12, AU14, and AU17), there are totally 1640 intensity files.

Preview Videos

Directory **PreviewVideos** saves the AVI videos generated from the scaled 2D textures of each model sequence. Frame rate is 25 fps. Each subject's videos are stored in his/her own directory. Each subject's directory contains 8 videos, which are named as [Gender][Index]_T[TaskID].avi. [Gender] is either F or M. [Index] is the three digits ID of the subject. [TaskID] is one digit ID for 8 tasks.

There are totally 328 files.

Notice about the Index

The frame index of model data starts from 0, while frame index of all AU coding and preview videos starts from 1. For example, for a coded model frame N, its corresponding AU coding frame index is N+1, and it should be the $(N+1)^{th}$ frame in the corresponding video.

Reference

[1] Zhang, X., Yin, L., Cohn, J. F., Canavan, S., Reale, M., Horowitz, A., & Liu, P., "A High-Resolution Spontaneous 3D Dynamic Facial Expression Database". In Proceedings of 10th IEEE International conference on automatic face and gesture recognition (FG'13), April, 2013.

[2] Zhang, Xing, Lijun Yin, Jeffrey F. Cohn, Shaun Canavan, Michael Reale, Andy Horowitz, Peng Liu, and Jeffrey M. Girard. "BP4D-Spontaneous: a high-resolution spontaneous 3D dynamic facial expression database." Image and Vision Computing 32, no. 10 (2014): 692-706 (Special issue of the Best of FG 13).

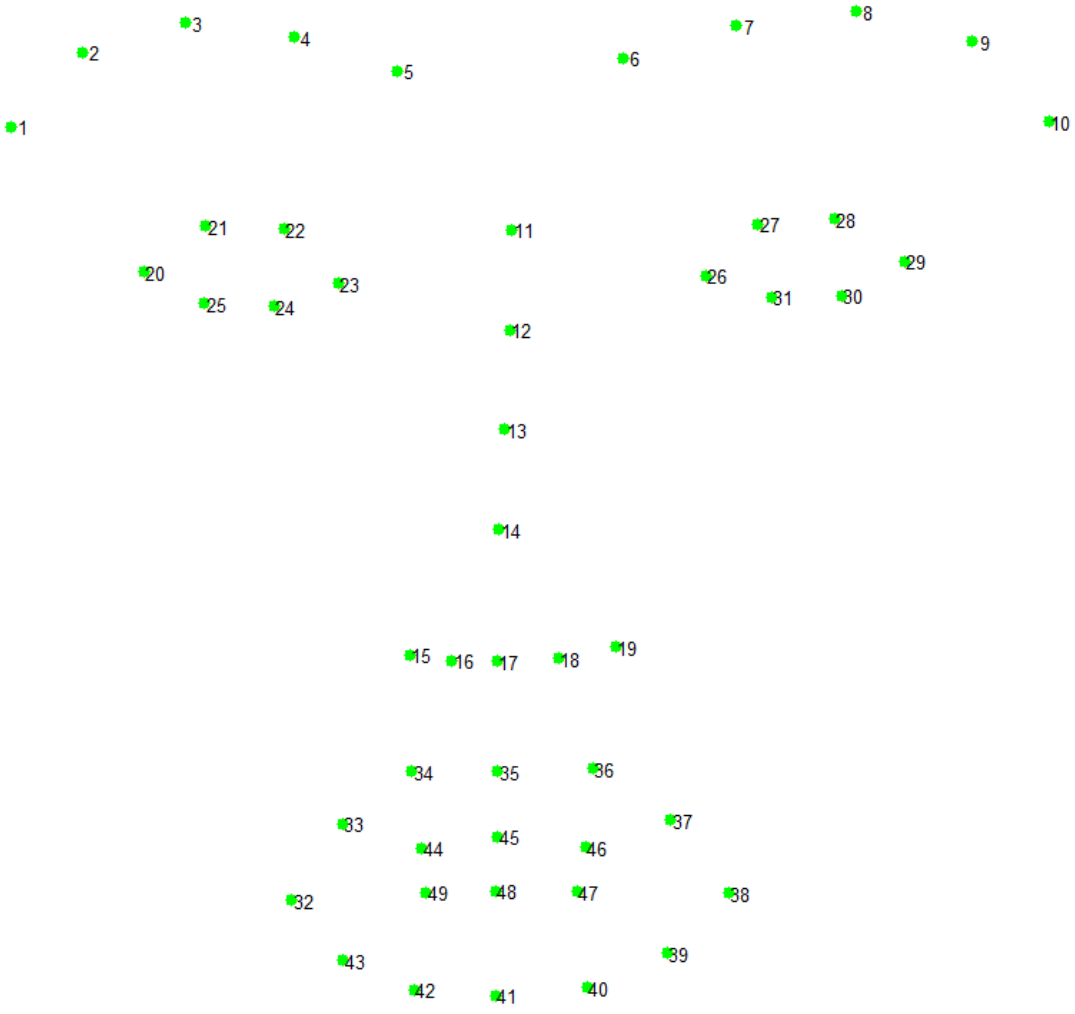

Figure 2: 2D Feature points' indices and location (Start from index one).

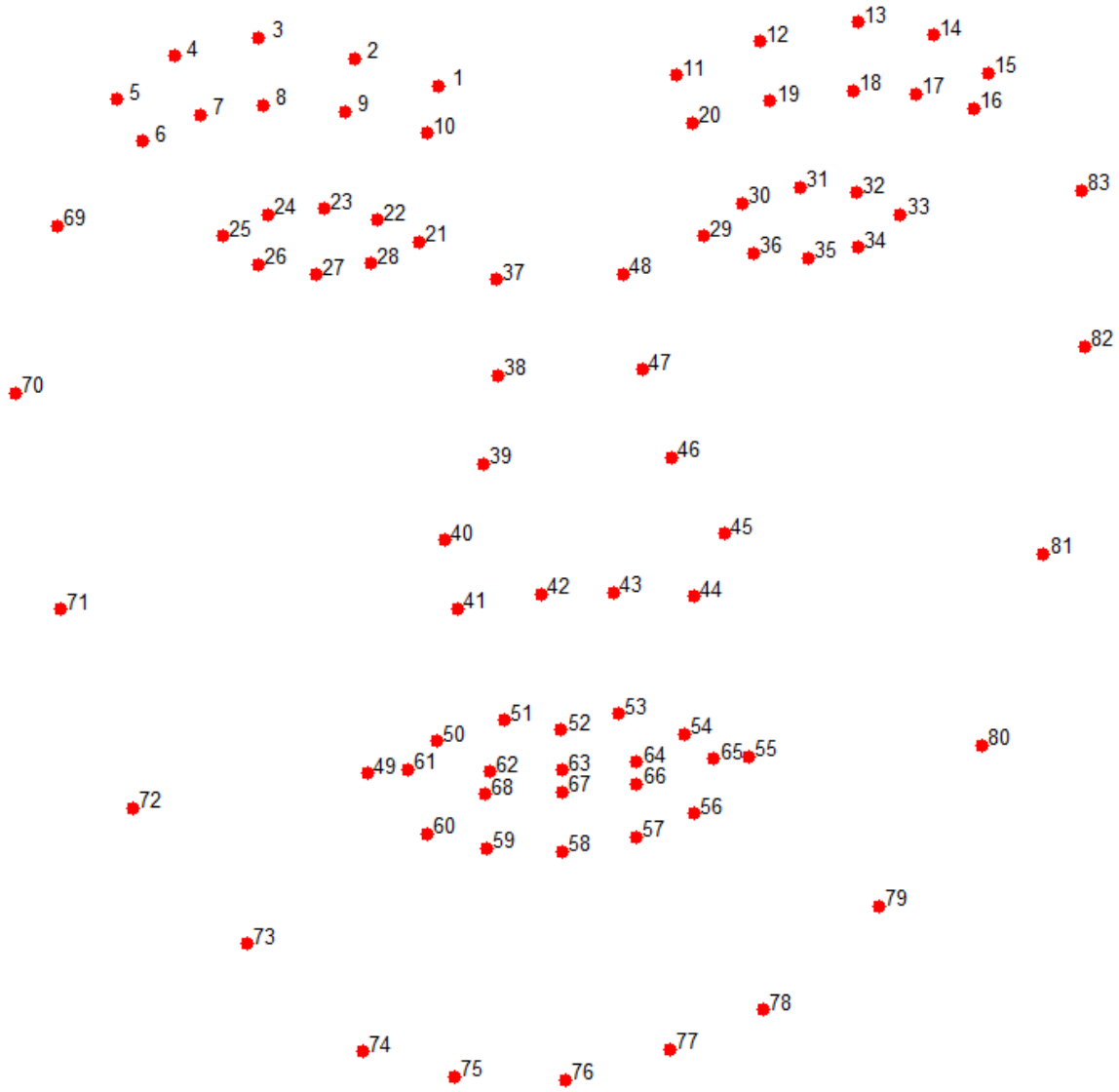

Figure 3: 3D Feature points' indices and location (Start from index one).