

SHREC 2017 (<http://liris.cnrs.fr/eg3dor2017/#shrec>)

3D Hand Gesture Recognition Using a Depth and Skeletal Dataset

In conjunction with Eurographics 3DOR'2017 Workshop (<http://liris.cnrs.fr/eg3dor2017/>), April 23-24, 2017

Online DHG

The SHREC 2017 track on 3D Hand Gesture Recognition was dealing with the pre-segmented gesture from the Online-DHG dataset with is realized in an online scenario.

The online dataset provides 280 sequences of 10 unsegmented gestures occurring sequentially.

The unsegmented sequences for online recognition scenario are available here ([OnlineDHG.zip](#)).

Task description

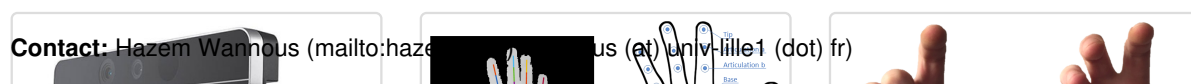
Recent virtual or augmented reality devices offer new 3D environment in which we need new and precise manner to interact. Hands can offer an intuitive and effective tool in these applications. Some gestures, as swipes, are more defined by the hand motion (called here **coarse** gesture) while others are defined by the hand shape through the gesture (called **fine** gesture). This difference between useful gestures have to be taken into account in a hand gesture recognition algorithm.

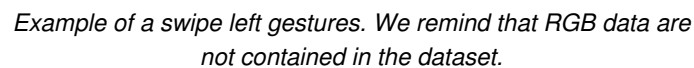
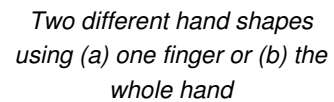
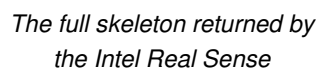
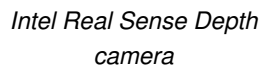
Effective and inexpensive depth sensors, like the Microsoft Kinect, are increasingly used in the domain of computer vision. By adding the third dimension into the game, depth images give new opportunities to many research fields, one of which is the **hand gesture recognition** area. Very recently, new devices as the Intel RealSense or the Leap Motion Controller also provide precise skeletal data of the hand and fingers in the form of a full 3D skeleton corresponding to 22 joints. Hand skeletal data could handle a precise information of the hand shape that HCI applications need in order to use the hand as a manipulation tool.

In this track, we present a new 3D dynamic hand gesture dataset which provides sequences of hand skeletal data in addition to the depth images. Such a dataset will facilitate the analysis of hand gestures and open new scientific axes to consider. This track aims to bring together researchers from the computer vision and machine learning communities in order to challenge their recognition algorithm of dynamic hand gesture using depth images **and / or** hand skeletal data.

Dataset

The dataset contains sequences of **14 hand gestures** performed in two ways: using one finger and the whole hand. Each gesture is performed between 1 and 10 times by 28 participants in 2 ways - described above - , resulting in **2800** sequences. All participants are right handed. Sequences are labelled following their gesture, the number of fingers used, the performer and the trial. Each frame of sequences contains a **depth image**, the coordinates of 22 joints both in the 2D depth image space and in the 3D world space forming a **full hand skeleton**. The Intel RealSense short range depth camera is used to collect our dataset. The depth images and hand skeletons were captured at 30 frames per second, with a resolution of the depth image of 640x480. The length of sample gestures ranges from 20 to 50 frames.





- » TAR.GZ (http://www-rech.telecom-lille.fr/shrec2017-hand/HandGestureDataset_SHREC2017.tar.gz)
- » RAR (http://www-rech.telecom-lille.fr/shrec2017-hand/HandGestureDataset_SHREC2017.rar)

The files of the dataset are structured as below:

...

```

\---gesture_14
train_gestures.txt
test_gestures.txt
display_sequence.m
display_sequence.py

```

For a sequence of size N:

- » **depth_n.png** contains the depth image of the n^{th} frame of the sequence.
 - » **general_informations.txt** contains a matrix of size $N \times 5$ (one line by frame). The format is as follows:
Timestamp in 10^{-7} seconds and hand region of interest in the depth image (x, y, width, height).
 - » **skeletons_image.txt** contains a matrix of size $N \times 44$. Each line contains the 2D hand joints coordinates in the **depth image** space. The format is as follows: $x_1 \ y_1 \ z_1 - x_2 \ y_2 \ z_2 - \dots - x_{22} \ y_{22} \ z_{22}$.
 - » **skeletons_world.txt** contains a matrix of size $N \times 66$. Each line contains the 3D hand joints coordinates in the **world** space. The format is as follows: $x_1 \ y_1 \ z_1 - x_2 \ y_2 \ z_2 - \dots - x_{22} \ y_{22} \ z_{22}$.
- The order of the joints in the line is: 1.Wrist, 2.Palm, 3.thumb_base, 4.thumb_first_joint, 5.thumb_second_joint, 6.thumb_tip, 7.index_base, 8.index_first_joint, 9.index_second_joint, 10.index_tip, 11.middle_base, 12.middle_first_joint, 13.middle_second_joint, 14.middle_tip, 15.ring_base, 16.ring_first_joint, 17.ring_second_joint, 18.ring_tip, 19.pinky_base, 20.pinky_first_joint, 21.pinky_second_joint, 22.pinky_tip.
- » **train_gestures.txt** and **test_gestures.txt** contains information about the train and the test sequences. These files contains respectively 1960 (70%) and 840 (30%) lines. Each line follow the following pattern:
id_gesture id_finger id_subject id_essai 14_labels 28_labels size_sequence
 - » **display_sequence.m** and **display_sequence.py** is respectively a matlab and a python script which charge and display a sequence. Dependencies for the python script: Scipy, Numpy and Matplotlib.

List of the 14 gestures

#	Name of the gesture	Type of the gesture
1	Grab	Fine
2	Tap	Coarse
3	Expand	Fine
4	Pinch	Fine
5	Rotation Clockwise	Fine
6	Rotation Counter Clockwise	Fine
7	Swipe Right	Coarse
8	Swipe Left	Coarse
9	Swipe Up	Coarse
10	Swipe Down	Coarse
11	Swipe X	Coarse
12	Swipe +	Coarse

#	Name of the gesture	Type of the gesture
13	Swipe V	Coarse
14	Shake	Coarse

Evaluation

We emphasized our main challenges compared to existing hand gesture datasets: (1) Study the dynamic hand gesture recognition using depth and full hand skeleton; (2) Evaluate the effectiveness of recognition process in terms of coverage of the hand shape that depend on the number of fingers used. The same movement is performed with one or more fingers, and the sequence can be labelled according to 14 or 28 label classes, depending on the gesture represented and the number of fingers used.

Indeed, labelling the sequences using the 14 gesture labels during the recognition process allows to judge if an algorithm can face the high coverage of hand shape while performing the same gesture with different number of fingers.

In the other hand, we can use 28 labels by grouping sequences following their type and the number of finger used to perform the gesture. In this manner, we will be able to evaluate the different methods on the task of fine-grained hand gesture recognition task.

The evaluation of proposed methods to this track will be performed as follow: the participants will have to return two files (following 14 or 28 gestures labels as described above) containing the label predicted by their algorithm for each hand gesture sequence in the **test** dataset (30% of the whole dataset). Please, if you can, add the time in millisecond for each sequence labeled. Each file will contain **840** lines formatted as follow:

id_gesture id_finger id_subject id_essai label_found time_millisecond

On the submitted results, we will compute the recognition accuracy compared to the provided ground truth. Please add a readme file in which you will precise the computer environment of your experiment (ex: 8GB RAM, Intel Core i5-3210m @ 2.5 GHz – 2 cores) and the data you are using (skeleton, depth or both).

Instructions to participants

Here are the different steps to follow by the participants:

- » Register by sending an email to quentin (dot) desmedt (a) telecom-lille (dot) fr (see contacts below). The registration should include entry name (name of the team/method) and contact information (name, affiliation, contact address and email of the participant(s));
- » Download the dataset (see dataset above):

Download the full dataset (~ 6 GB) (.rar)

or

Download the full dataset (~ 6 GB) (.tar.gz)
- » Run your algorithm(s) on the dataset;
- » Submit your results as a ZIP file containing the 2 result files (see evaluation above); Up to 4 ZIP files per group may be submitted, resulting from different parameters or methods;
- » Send a 1-page summary paper (PDF) to the organizers.

The track is now over, please feel free to send us your results in order to add them on the leaderboard.

Results

Method	Researchers	Accuracy 14 gestures	Accuracy 28 gestures	Date added
--------	-------------	----------------------	----------------------	------------

Method	Researchers	Accuracy 14 gestures	Accuracy 28 gestures	Date added
Skeleton-based Dynamic hand gesture recognition [1]	Quentin De Smedt, Hazem Wannous and Jean-Philippe Vandeborre	88.24	81.90	07 / 2017
Key frames with convolutional neural network [2]	Joris Guerry, Bertrand Le Saux and David Filliat	82.90	71.90	07 / 2017

[1] De Smedt, Q., Wannous, H., & Vandeborre, J. P. (2016). Skeleton-based dynamic hand gesture recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 1-9).

[2] De Smedt, Q., Wannous, H., Vandeborre, J. P., Guerry, J., Le Saux, B., & Filliat, D. (2017, April). SHREC'17 Track: 3D Hand Gesture Recognition Using a Depth and Skeletal Dataset. In 10th Eurographics Workshop on 3D Object Retrieval.

More details in the paper in 10th Eurographics Workshop on 3D Object Retrieval:

Quentin De Smedt, Hazem Wannous, Jean-Philippe Vandeborre, Joris Guerry, Bertrand Le Saux, David Filliat, *SHREC'17 Track: 3D Hand Gesture Recognition Using a Depth and Skeletal Dataset*, 10th Eurographics Workshop on 3D Object Retrieval, Lyon, France, April 23-24, 2017. [DOI: 10.2312/3dor.20171049 (<http://dx.doi.org/10.2312/3dor.20171049>)] [PDF (<https://hal.archives-ouvertes.fr/hal-01563505/document>)]

Schedule

~~January 16, 2017 — call for participation;~~

~~February 16, 2017 — deadline for registration by email;~~

~~February 20, 2017 — deadline for result submission by email;~~

~~February 21, 2017 — deadline for 1 page summary paper.~~

The track is now over, please feel free to send us your results in order to add them on the leaderboard.

Contacts

The organizers of this track are:

- » Quentin De Smedt, IMT Lille Douai, CRISTAL (UMR CNRS 9189), France;
- » Hazem Wannous, University Lille 1 / IMT Lille Douai, CRISTAL (UMR CNRS 9189), France;
- » Jean-Phillipe Vandeborre, IMT Lille Douai, CRISTAL (UMR CNRS 9189), France.

To contact the organizers, please use the main contact email address: Hazem Wannous (<mailto:hazem.wannous@univ-lille1.fr>).