

K3HI: Kinect-based 3D Human Interaction Dataset

A quality depth sensor, the Microsoft Kinect, is now in millions of homes. Yet there is no publicly accessible dataset for two-person interaction recognition based on the Kinect. We are collecting a massive Kinect-based 3D human interaction dataset.

Download the dataset

Version 1 of the dataset is available as a zip archive [here](http://www.lmars.whu.edu.cn/prof_web/zhuxinyan/DataSetPublish/data/Data0829.zip)

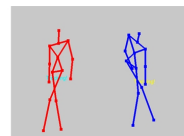
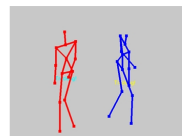
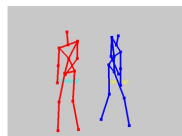
(http://www.lmars.whu.edu.cn/prof_web/zhuxinyan/DataSetPublish/data/Data0829.zip). Results that refer to this version are available in the following publication:

Tao Hu, Xinyan Zhu, Wei Guo, Kehua Su. Efficient Interactions Recognition through Positive Action based Representation Mathematical Problems in Engineering (**SCI, In Review**).

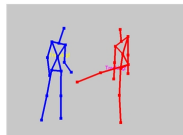
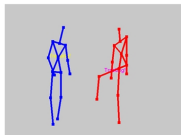
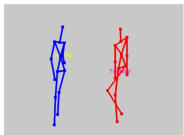
Tao Hu, Xinyan Zhu, Wei Guo. Two-person Interaction Recognition Based on Key Poses Journal of Computing and Information System (**EI, In Review**).



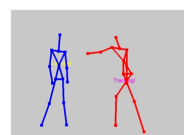
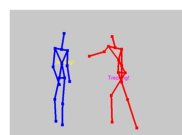
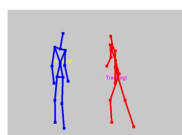
(a) Approaching



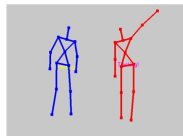
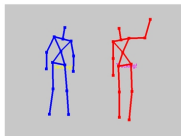
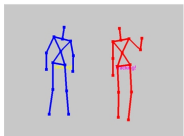
(b) Departing



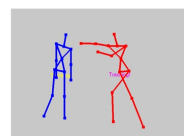
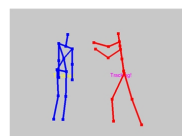
(c) Kicking



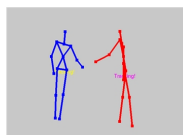
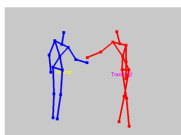
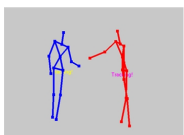
(d) Punching



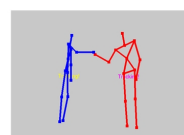
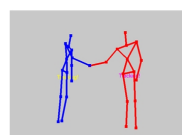
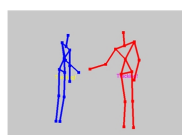
(e) Pointing



(f) Pushing



(g) Exchanging an object



(h) Shaking hands

Details

There are few publicly accessible test datasets to assess two-person interaction recognition approaches. The recent proliferation of a cheap but effective depth sensor, the Microsoft Kinect, has created more opportunities for quantitative analysis of complex human activities. Compared with the traditional video camera, the Kinect has the advantage of synchronous acquisition of color and depth images; with the use of depth maps, 3D information of a scene from a particular point of view is easily computed.

The most important data in our dataset is the spatial information (3D coordinates) of the two persons' skeletons. In order to ensure the integrity and continuity of target data, the original RGB and depth information were ignored when capturing data. An articulated skeleton for each person was extracted by OpenNI software with NITE (Natural Interaction Middleware) provided by PrimeSense.

A skeleton was represented by the 3D positions of 15 joints, including head, neck, left shoulder, right shoulder, left elbow, right elbow, left hand, right hand, torso, left hip, right hip, left knee, right knee, left foot and right foot

Thank you for your interest.

Made by [Tao Hu \(http://www.lmars.whu.edu.cn/prof_web/zhuxinyan/index.html#/\)](http://www.lmars.whu.edu.cn/prof_web/zhuxinyan/index.html#/).