Dataset for Air-writing based on Ultrawideband Radar Sensor

This file explains the radar dataset collection setup, the included folders, the data format, and how to use the data for the air-writing recognition application. Additionally, a MATLAB script is provided to help getting started with the recognition process.

Data Collection

The ultrawideband radar used in this experiment is *Xethru* X4-M03 from *Novelda* [1]. Firstly, the radar is placed on a horizontal table where the user writes above the radar on the virtual writing space at a distance greater than 40cm from the radar. The nominal writing frame is 20cm × 20cm as shown in Figure 1. The radar detection range is set between 0.4m–1.55m to avoid missing any parts of air-written numbers. For more explanation on radar chipset and experimental setup, please refer to the original paper [2].

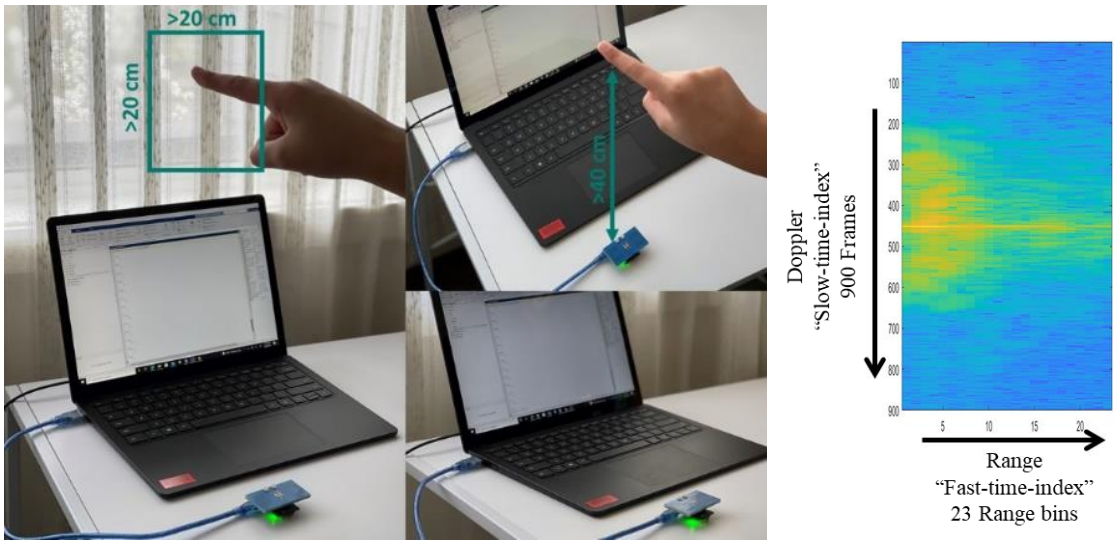

Figure 1 Experimental setup for data collection and a range-Doppler image sample [2].

The collected dataset includes air-written numbers from 0 to 9 using a uni-stroke writing technique. The acquired range-Doppler frames have a dimension of M × N = 900 × 23, where M=900 represents slow-time-index (number of frames) and N=23 represents the fast-time-index (range bins) as shown in Figure 1. A dataset of 180 samples is collected for each number.

The *Dataset* contains 4 sub-folders and a MATLAB script explained as shown in Figure 2 and as follows:

1. **2D-Range-Doppler Data CSV Format**
   This folder contains 10 sub-folders including the collected Range-Doppler samples for each digit in csv format, named as follows:
   "**Digit_x**" → sub-folder includes the samples for digit *x*.
   Inside each sub-folder the .csv file of the collected Range-Doppler data of a single sample for a single digit, named as follow:
   "**Digit_x_y.csv**" → sample number *y* of digit *x*.

2. **2D-Range-Doppler Data MAT Format**
   This folder contains 10 .mat files of the collected Range-Doppler samples for each digit in mat format, named as follows:
   "**Digit_x.mat**" → matrix of size (180x900x23) includes all the samples for digit *x*.

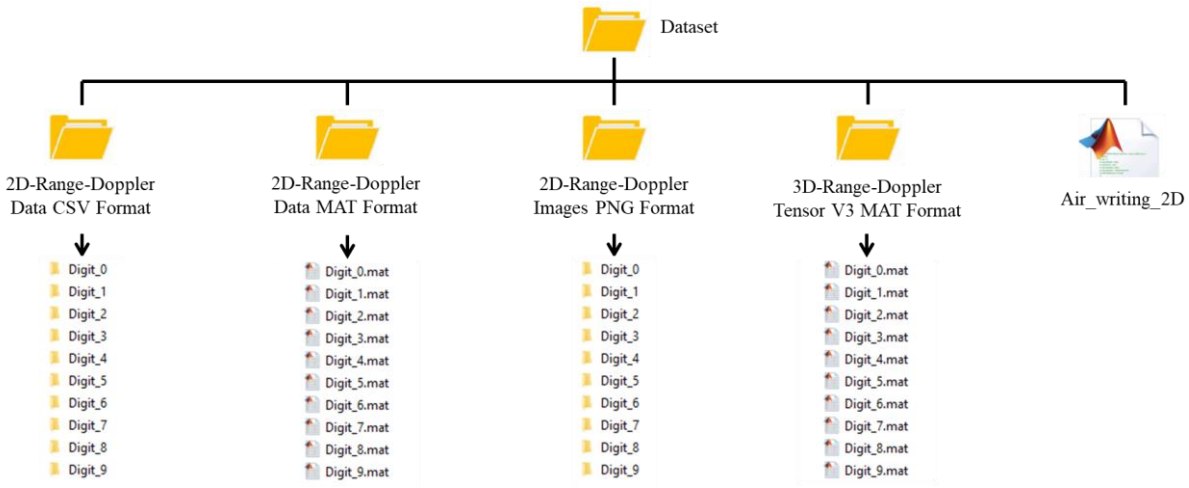

Figure 2 The structure of the dataset folder

3. **2D-Range-Doppler Images PNG Format**

   This folder contains 10 sub-folders including the Range-Doppler images for each digit in .png format, named as follow:

   "**Digit_x**" → sub-folder includes .png images for digit *x*.

   Inside each sub-folder there are 180 images of all samples for a single digit, named as follows:

   "**Digit_x_y.png**" → matrix of size (900x23) for sample number *y* of digit *x*.

4. **3D-Collected Data Tensor V3 MAT Format**

   The 3D data representation is developed from the collected raw data frame of size 900 × 23. The 3D tensor is created by converting the 2D frame into a series of L frames each of P blocks per frame. The resulting is a tensor of size L × P × N, where M=L × P.

   The developed 3D tensor arrangements, Variant 3 (V3) of size (15×60×23), is provided as the best reachable recognition accuracy arrangement. The folder contains 10 .mat file of the rearranged range-Doppler frames, named as follows:

   "**Digit_x.mat**" → matrix of size (180x15x60x23) of Range-Doppler data for digit *x*.

5. **MATLAB Script Labelled "Air_writing_2D"**

   A script for deep learning using 2D-CNN using the data from "**2D-Range-Doppler Images PNG Format**" folder.

License / Attribute

References

[1] A. Novelda. "Xethru X4M03 datasheet." http://laonuri.techyneeti.com/ wp-content/uploads/2019/02/X4M03 DATASHEET.pdf (accessed 15/06/2021)

[2] N. Hendy, H. M. Fayek and A. Al-Hourani, "Deep Learning Approaches for Air-Writing Using Single UWB Radar," in IEEE Sensors Journal, vol. 22, no. 12, pp. 11989-12001, 15 June 2022, doi: 10.1109/JSEN.2022.3172727.