WMCA Database

The Wide Multi Channel Presentation Attack (WMCA) database consists of 1679 short video recordings of both bonafide and presentation attacks from 72 different identities. The data is recorded from several channels including color, depth, infra-red, and thermal.

Here we provide the preprocessed images of four channels for the data used in the reference publication.

The WMCA database is produced at Idiap within the framework of "IARPA BATL" and "H2020 TESLA" projects and it is intended for investigation of presentation attack detection (PAD) methods for face recognition systems.

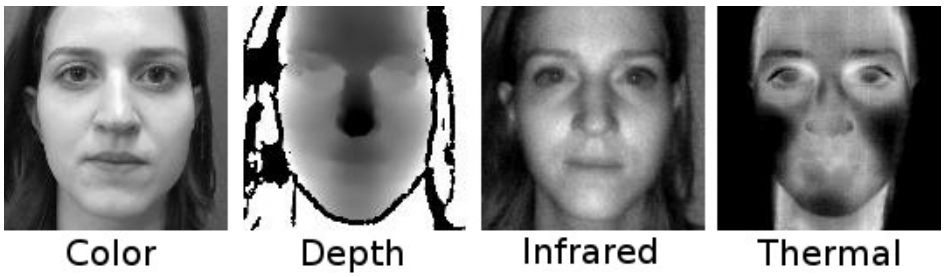

Figure 1.1: Example of preprocessed data from four channels for a bonafide presentation.

1.1 Acknowledgments to reference publication:

If you use this database, please cite the following publication:

@article{George_TIFS_2019,

author = {George, Anjith and Mostaani, Zohreh and Geissenbuhler, David and Nikisins, Olegs and Anjos, André and Marcel, Sébastien},

title = {Biometric Face Presentation Attack Detection with Multi-Channel Convolutional Neural Network},

journal = {IEEE Transactions on Information Forensics and Security},

year = {2019},

publisher = {IEEE}

}

1.2 Database Description:

The image data from different channels were captured with the following sensors.

1. "Intel RealSense SR300" which captures the color, depth, and infra-red channels.
2. "Seek Thermal Compact PRO" which captures data from thermal channel.

The two mentioned sensors captures data synchronously during 10 seconds. The specifications of the sensors used for image data can be found in Table 1.1.

The data was captured during several sessions with different environmental conditions. Session 4 was dedicated to presentation attacks only. The information about the sessions can be found in Table 1.2.

1.3 Presentations:

The presentations in the database can be grouped into different categories:

Table 1.1: Sensor description for WMCA data collection

Sensor	Channel	Resolution	Frame rate (approx.)
Intel RealSense SR300	Color	1920×1080	30 fps
Intel RealSense SR300	Depth	640×480	30 fps
Intel RealSense SR300	Infra-red	640×480	30 fps
Seek Thermal Compact PRO	Thermal	320×240	15 fps

Table 1.2: Session description for WMCA data collection

Session	Background	Illumination
1	uniform	ceiling office light
2	uniform	day-light illumination
3	complex	day-light illumination
4	uniform	ceiling office light
5	uniform	ceiling office light
6	uniform	side illumination with LED lamps
7	complex	ceiling office light

1. **Bonafide:** Includes bonafide data from participants with and without medical glasses. Some participants were asked additionally to wear a pair of unisex glasses with black rim that had no correction (Fig. 1.2).

2. **Glasses:** Different models of disguise glasses with fake eyes (funny eyes glasses) and paper glasses. These attacks constitute partial attacks (Fig. 1.3).

3. **Fake head:** Several models of mannequin heads were used, some of the mannequins were heated with a blower prior to capture (Fig. 1.4).

Figure 1.2: The unisex glasses with no correction used in some of the bonafide recordings.

Figure 1.3: Sample images of glasses used as presentation attacks.

4. **Print:** Printed face images on A4 matte and glossy papers using professional quality Ink-Jet printer (Epson XP-860) and typical office laser printer (CX c224e). The images were captured by the rear camera of an iPhone S6 and re-sized so that the size of the printed face is human like.

5. **Replay:** Electronic photos and videos. An iPad pro 12.9in was used for the presentations. The videos were captured in HD at 30 fps by the front camera of an iPhone S6 and in full-HD at 30 fps by the rear camera of the iPad pro 12.9in. Some of the videos were re-sized so that the size of the face presented on the display is human like.

6. **Rigid mask:** Custom made realistic rigid masks and several designs of decorative plastic masks (Fig. 1.5).

7. **Flexible mask:** Custom made realistic soft silicone masks (Fig. 1.6).

8. **Paper mask:** Custom made paper masks based on real identities. The masks were printed on the matte paper using both printers mentioned in the print category (Fig. 1.7).

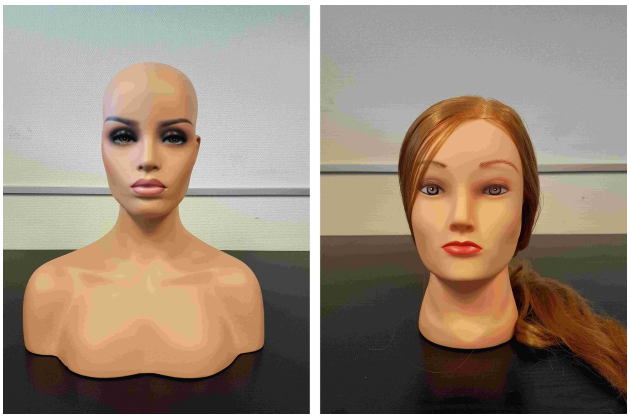

Figure 1.4: Sample images of mannequin heads.

Figure 1.5: Sample images of rigid masks.

The number of presentations in the database can be found in Table 1.3.

1.4 Evaluation:

All four channels of data obtained from "Intel RealSense SR300" and "Seek Thermal Compact PRO" are used in the reference publication. Since the consecutive frames are correlated, only 50 frames from each video were selected. The frames are uniformly sampled in the temporal domain. The total number of 1679 presentations including bonafide and presentation attacks were grouped into three subsets, train, dev, and eval. The data split is done ensuring almost equal distribution of PA categories and disjoint set of client

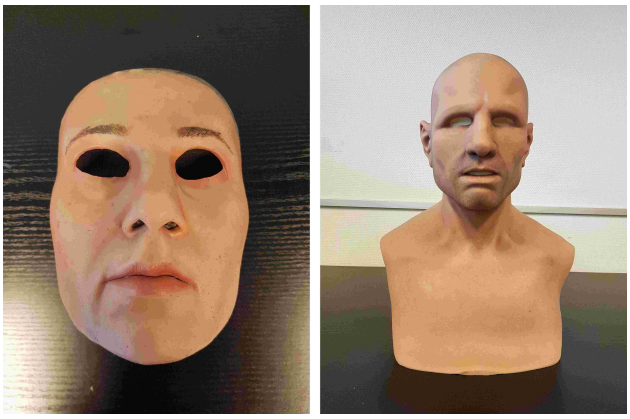

Figure 1.6: Sample images of flexible masks.

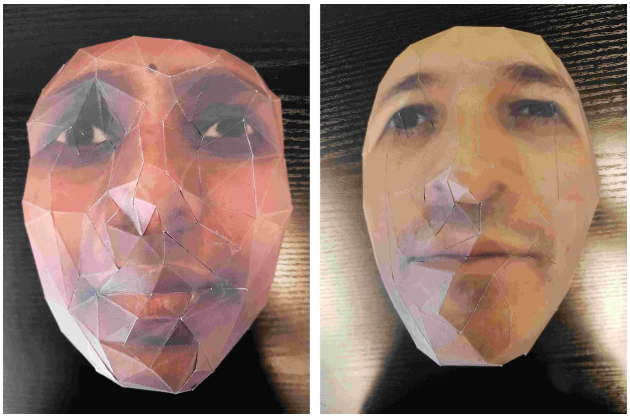

Figure 1.7: Sample images of paper masks.

identifiers in each set. Each of the PAIs had different client id. The split is done in such a way that a specific PA instrument will appear in only one set. A low level database interface is implemented that handles loading and spatial and temporal alignment of the data[1].

The preprocessing for the color channel is done in several steps. First a face detection is performed. Once the face bounding box is obtained, face landmark detection is performed in the detected face bounding box. Then the images are aligned by transforming them such that the eye centers and mouth center are aligned to predefined coordinates. The aligned face images are converted to grayscale, and resized, to the resolution of 128×128 pixels. The preprocessing stage for non-RGB channels is done by reusing the facial landmarks

[1] `https://gitlab.idiap.ch/bob/bob.db.batl`

Table 1.3: Statistics for WMCA database

Category	Number of presentations
Bonafide	347
Glasses	75
Fake head	122
Print	200
Replay	348
Rigid mask	137
Flexible mask	379
Paper mask	71
TOTAL	**1679**

detected in the color channel and then a similar alignment procedure is performed. The images are then normalized to convert the range of the non-RGB images to 8-bit format.

The preprocessed method is described in details in the reference publication and the implementation is available publicly [2]. The number of presentations in train, dev, and eval subset for the grandtest protocol of the reference publication is mentioned in Table 1.4.

Table 1.4: Statistics for train, dev, and eval subset of grandtest protocol

Subset	Bonafide	Presentation Attack	Both
train	124	441	565
dev	108	449	557
eval	115	442	557
TOTAL	**347**	**1332**	**1679**

Some example images of attack presentations from all channels are illustrated in the following.

[2] https://gitlab.idiap.ch/bob/bob.paper.mccnn.tifs2018

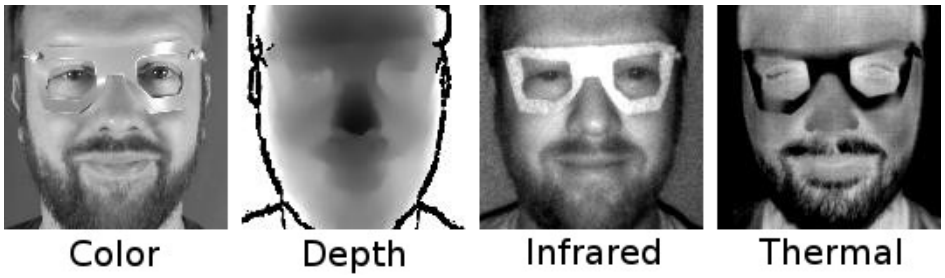

Figure 1.8: Example of preprocessed data from four channels for glasses attack.

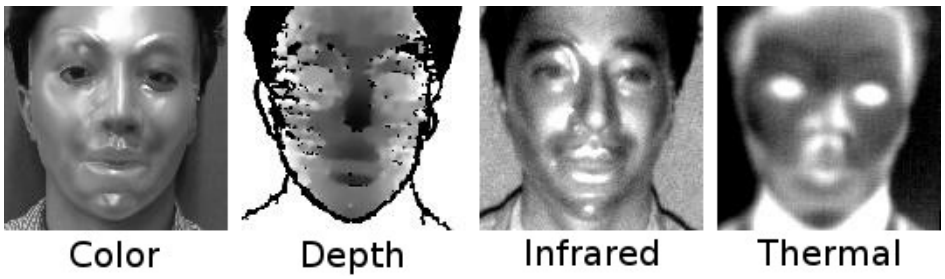

Figure 1.9: Example of preprocessed data from four channels for rigid mask attack.

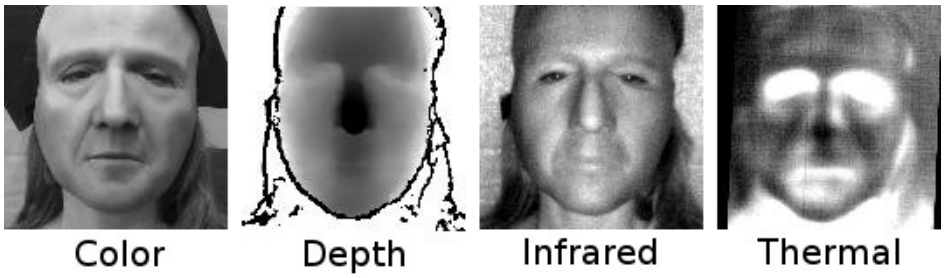

Figure 1.10: Example of preprocessed data from four channels for flexible mask attack.