THIS DOCUMENT DESCRIBES THE DATASET AND HOW IT WAS COLLECTED

Summary

A dataset consisting of manually-measured heights and left and right MUAC as well as photos of 31 participants (11 female, 20 male; mean age 26.2 years ± 4.8 SD), was created for evaluation. Data collection was performed with approval from the UNSW Sydney Human Research Ethics Committee and all participants provided informed written consent. Manual measurement and photo capture of every participant were performed in the same session and completed within an hour avoiding the effects of daily body variations.

Manual measurement

The height was measured with the subject standing with feet slightly apart with the back of their head, their buttocks and their heels against a wall with the head positioned so that the horizontal line connecting the upper ear opening and the lower edge of the eye socket runs parallel to the floor; a laser range finder (Bosch GLM 7000; Class II laser, accuracy ± 1.5 mm) was used to measure the distance from the top of the subjects head to the ceiling, which was subtracted from the distance between the floor and the ceiling (measured with the same device). Participants were free to remove or remain in their footwear during height measurement for comfort and convenience.

The mid-upper arm is the midpoint between the acromion process (bony process on the shoulder blade) and the elbow (when the elbow is bent at approximately a 90° angle). Following identification of the mid-upper arm, the subject then relaxed their arm by their side and an un-stretchable tape measure (Ibis Medical; Infant head circumference measuring tape) was wrapped around the arm passing over the midpoint, touching the skin without compressing the tissues and skin; the mid-upper arm circumference (MUAC) is recorded to the nearest mm. In order to indicate the mid-upper arm in photographs, lines were drawn on both arms of the participants when their manual measurements were collected.

Photos

Photos of the participants were collected indoors with a tripod mounted Canon IXUS 110 IS digital camera with a resolution of 4000×3000 pixels.. The camera was set to auto focus, auto exposure time and ISO-200, and no zoom or image stabilization was applied. Participants' arms were bare from the shoulder down, and participants remained barefoot or in their footwear (as they were during the manual measurement of height) for their comfort and convenience.

During photo capture, participants were instructed to stand up straight, and look straight ahead with their arms by their side during photo capture with their medial line in the same plane as a checkerboard (with 7×10 squares, each square of which is 90×90 mm) which was perpendicular to the floor next to the participant.

Photos of each participant were taken from five views by changing the orientation of the participant relative to the camera lens - for the different views, the camera and checkerboard positions were fixed and the subject would turn and reposition their feet according to markings on the ground, whilst keeping their medial line in the same plane as the checkerboard: 0° corresponds to the subject facing the camera, 45° corresponds to the subject having turned clockwise by 45°, 90° corresponds to the subject having turned clockwise by another 45° (the side-view of the subject with the left arm closest to the camera) and so on.

A total of 744 photos were taken at three distances between the participants and the camera lens (i.e. three distances per participant: mean 3.39 m ± 0.30 m SD, mean 3.25 m ± 0.33 m SD and 3.12 m ± 0.32 m SD, respectively, with the difference between the furthest and the middle distance, and between the middle and the closest distance approximately equal to 150 mm). The camera was approximately 1.125 m above the floor. Variation in the camera-participant distances and the height of the camera were due to the equipment being set-up/packed-up in different locations (rooms of varying sizes) before/after every participant.

Manual labels

Each photo was manually annotated with a number of markers:

- Reference object: The paper on which the calibration checkerboard was also used as a reference object within the image. The corners of the paper were manually marked. The dimensions of the paper were 840 × 1186 mm.
- Floor line: The two corners at the bottom of the calibration checkerboard were marked as an indication of the location of the floor in the plane of the checkerboard (and hence the participant).
- Top of the head: the top of the participants head was marked
- Bounding box: The bounding box enclosing the silhouette of the participant was marked with the top edge passing through the top of the head, and the bottom edge passing through the lowest point of the feet.
- Edges of the left and right mid-upper arm: The edges of the left and right mid-upper arm were marked where the MUAC had been measured manually (indicated by a line drawn on the skin).

File structure and contents

Folders 0003 – 0033 contain the photos (in a subfolder XXXX_Images_BlurFace) and photo-related meta data (in a subfolder XXXX_Meta) for 31 study participants with numerical identifiers 0003 to 0033.

Each subfolder **XXXX_Images_BlurFace** contains 24 photos in .JPG format.

Each subfolder **XXXX_Meta** contains the following files:

- checkerboard_info.mat - This file contains two variables: checkerboard_info and checkerboard_info_names
  - checkerboard_info_names is a 1x5 cell array containing the names of the columns in the parameter checkerboard_info: 'Filename','Checkerboard Detected','# Squares Detected','Checkerboard Points' and 'Checkerboard Square Size (W x H mm)'
  - checkerboard_info is a 24x5 cell array containing the relevant values corresponding to checkerboard_info_names for each photo in the XXXX_Images_BlurFace folder for this participant
    - 'Filename' – the name of the corresponding .JPG image file
    - 'Checkerboard Detected' – a Boolean indicating whether the checkerboard was detected or not
    - '# Squares Detected' – a two element vector indicating the number of checkerboard squares detected
    - 'Checkerboard Points' – the image coordinates of the checkerboard points
    - 'Checkerboard Square Size (W x H mm)' – a two element vector containing the width and height (in mm) of an individual checkerboard square
- img_info.mat – This file contains two variables: img_info and img_info_names
  - img_info_names is a 1x5 cell array containing the names of the columns in the parameter img_info: 'Filename', 'View Angle', 'Distance', 'Height' and 'Rep'

- img_info is a 24x5 cell array containing the relevant values corresponding to img_info_names for each photo in the XXXX_Images_BlurFace folder for this participant
  - 'Filename' – the name of the corresponding .JPG image file
  - 'View Angle' – a string indicating the view angle of the image file (‘0’, ‘45’,…,’315’)
  - 'Distance' – an integer indicating the distance between the camera and the subject (1 for furthest distance, 2 for middle distance, and 3 for closest distance)
  - 'Height' – an integer equal to 1 (not used)
  - 'Rep' – an integer equal to 1 (not used)

- labels_info.mat - This file contains two variables: labels_info and label_info_names
  - labels_info_names is a 1x5 cell array containing the names of the columns in the parameter labels_info: 'Filename', 'Reference Object Pts', 'Floor Pts', 'Head Pts', 'Bounding Box Rect', 'Left MUA Pts', 'Right MUA Pts'
  - labels_info is a 24x5 cell array containing the relevant values corresponding to labels_info_names for each photo in the XXXX_Images_BlurFace folder for this participant
    - 'Filename' – the name of the corresponding .JPG image file
    - 'Reference Object Pts' – the image coordinates of the four corners of the reference object (top left corner, top right corner, bottom left corner, bottom right corner)
    - 'Floor Pts' – the image coordinates of the corners of the reference object in contact with the floor
    - 'Head Pts' – the image coordinates of the top of the subject’s head
    - 'Bounding Box Rect' – the rectangle defining the bounding box of the subject’s silhouette (minimum x-coordinate, minimum y-coordinate, width, height)
    - 'Left MUA Pts' – the image coordinates of the two points defining the edges of the left mid-upper arm
    - 'Right MUA Pts' – the image coordinates of the two points defining the edges of the right mid-upper arm

- msk_info.mat – This file contains two variables: msk_info and msk_info_names
  - msk_info_names is a 1x5 cell array containing the names of the columns in the parameter msk_info: 'Filename' and 'Mask Rect'
  - msk _info is a 24x5 cell array containing the relevant values corresponding to msk_info_names for each photo in the XXXX_Images_BlurFace folder for this participant
    - 'Filename' – the name of the corresponding .JPG image file
    - 'Mask Rect'– the rectangle defining position of the blurring mask for the subject’s face (minimum x-coordinate, minimum y-coordinate, width, height)

- refObj_info.mat – This file contains two variables: refObj_info and refObj_info_names
  - refObj_info_names is a 1x5 cell array containing the names of the columns in the parameter refObj_info: 'Filename', 'Width (mm)' and 'Height (mm)'
  - refObj_info is a 24x5 cell array containing the relevant values corresponding to refObj_info_names for each photo in the XXXX_Images_BlurFace folder for this participant
    - 'Filename' – the name of the corresponding .JPG image file
    - 'Width (mm)' – an integer defining the width in mm of the reference object
    - 'Height (mm)' – an integer defining the height in mm of the reference object

The folder **checkerboard_calibration_90x90** contains the photos used to perform camera calibration and a file ‘cam_par.mat’ which contains the resultant camera calibration parameters. Cam_par.mat contains two variables cameraParams and estimationErrors – these variables were generated by the MATLAB camera calibrator (Mathwork R2016a) using default settings (see https://au.mathworks.com/help/vision/ref/cameracalibrator-app.html,

https://au.mathworks.com/help/vision/ref/cameraparameters.html and https://au.mathworks.com/help/vision/ref/cameracalibrationerrors.html).

The folder **manual_measurements** contains a file manual_measurements.mat which contains two parameters defining the subject demographic information and manual measurements:

- manual_measurements_names is a 1x18 cell array containing the names of the columns in the parameter manual_measurements: 'Numerical Identifier', 'Sex', 'Age', 'Weight', 'Height', 'Height', 'Height', 'Head Circumference', 'Head Circumference', 'Head Circumference', 'Left Upper Arm Length', 'Left MUAC', 'Left MUAC', 'Left MUAC', 'Right Upper Arm Length', 'Right MUAC', 'Right MUAC', 'Right MUAC'
- manual_measurements is a 31x18 cell array containing the relevant values corresponding to manual_measurements_names for each subject
  - 'Numerical Identifier' – An integer value 3 - 33
  - 'Sex' – A string representing the sex of the subject (‘F’ or ‘M’)
  - 'Age' – An integer representing the age of the subject in years
  - 'Weight' – An double representing the self-reported weight of the subject
  - 'Height', 'Height', 'Height' – Three measurements of height in mm
  - 'Head Circumference', 'Head Circumference', 'Head Circumference' – Three measurements of head circumference in mm
  - 'Left Upper Arm Length' – The length of the left upper arm in mm
  - 'Left MUAC', 'Left MUAC', 'Left MUAC' – Three measurements of the left MUAC in mm
  - 'Right Upper Arm Length' – The length of the right upper arm in mm
  - 'Right MUAC', 'Right MUAC', 'Right MUAC' – Three measurements of the right MUAC in mm