

Annotation Manual for the MA-52 dataset

A Descriptions of Micro-Actions

ID	Categories	Descriptions
Body-level label: Body (A)		
A1	Shaking body	The movement of the upper body from side to side
A2	Turning around	The movement that causes changes in the facing angle
A3	Sitting straightly	The dynamic process of straightening the back, without involving the maintenance of an upright sitting
A4	Shrugging	Shrugging or moving the shoulders back and forth, indicating movements involving only the shoulders
A5	Rising up	The preparatory movements before standing up from a seated position
Body-level label: Head (B)		
B1	Nodding	The movement of continuous up-and-down head
B2	Shaking head	The movement of continuous turn of the head left and right
B3	Turning head	Turn the head backward or sideways, typically as a single motion
B4	Tilting head	Lower the head slowly or abruptly, typically as a single motion
B5	Bowing head	Tilt the head toward the shoulder, distinguishing between tilting and turning
B6	Head up	Raise the head slowly or abruptly, typically as a single motion
Body-level label: Upper limb (C)		
C1	Illustrative gestures	The illustrative gestures made with either the left arm or the right arm
C2	Other finger movements	Excluding finger movements as defined within the upper limb
C3	Hands touching fingers	The movement of changing from having the fingers of both hands uncrossed to interlaced
C4	Stretching arms	The movement of raising or extending the arm forward or sideways
C5	Waving	The movement of waving one's hand or arm, often accompanied by shaking one's head
C6	Scratching arms	The movement of scratching the other arm with one hand
C7	Spreading hands	The movement of holding one's hand open with the palm facing up
C8	Playing objects	The movement of playing with or adjusting objects in hand.
C9	Putting hands together	The movement of bringing both hands closer together until palms together
C10	Retracting arms	The movement of extending and retracting the arm
C11	Pointing oneself	The movement of the hand that indicates direction and self-reference
C12	Clenching fist	The movement of bending the fingers into a fist, including partial clenching of the fingers
C13	Rubbing hands	The movement of touching and rubbing the hands together
Body-level label: Lower limb (D)		
D1	Tiptoe	Place the foot sideways or rest the heel on a stool with the toe touching the ground
D2	Retracting feet	Retract the foot towards the body. If both feet retract into a tiptoe position, it is considered tiptoe
D3	Shaking legs	The movement of shaking the leg up and down, either continuously or briefly
D4	Stretching feet	The movement of extending the foot forward or sideways, often accompanied by leaning forward or turning
D5	Closing legs	The movement of moving the legs from apart to together
D6	Spread legs	The movement of increasing the distance between the knees
D7	Curling legs	The movement of sitting with one leg crossed over the other
D8	Crossing legs	The movement of legs gradually turning from crossing to crossing
Body-level label: Body-hand (E)		
E1	Scratching or touching chest	The movement of bringing the hand close to and touching the front of the upper body
E2	Scratching or touching neck	The movement of bringing the hand close to and touching the neck area
E3	Scratching or touching back	The movement of bringing the hand close to and touching the back of the upper body
E4	Arms akimbo	The movement of placing one or both hands on the hips
E5	Crossing arms	The movement of crossing one's arms over one's chest
E6	Scratching or touching shoulder	The movement of bringing the hand close to or touching the shoulder area
Body-level label: Head-hand (F)		
F1	Touching nose	The movement of bringing the hand close to and touching the nose area
F2	Scratching or touching face	The movement of touching the cheek area, excluding the nose, forehead, and chin
F3	Playing or tidying hair	Hand movements that adjust hair on the sides or top of the head, including women arranging their hanging hair
F4	Scratching or touching hindbrain	Touch the back of the head or the hair on the back of the head
F5	Pushing glasses	The movement of bringing the hand close to and touching the glasses
F6	Rubbing eyes	The movement of bringing the hand close to and touching the area around the eyes
F7	Scratching or touching forehead	The movement of scratching or touching the forehead
F8	Touching ears	The movement of bringing the hand close to and touching the ear, or touching and rubbing the ear
F9	Covering mouth	The movement of covering the mouth with the hand
F10	Covering face	The movement of covering the face region, excluding the mouth, with one or both hands
Body-level label: Leg-hand (F)		
G1	Touching legs	The movement of touching and rubbing the leg with one hand
G2	Patting legs	Move one hand close to and touch the leg in a dynamic process
G3	Scratching legs	Scratch the calf area with the hand, excluding the ankle
G4	Scratching feet	The movement of scratching the foot with the hand, excluding the legs

Table 1: The descriptions of each micro-action in the Micro-Action-52 dataset [1].

B Annotation Process

The Micro-Action-52 (MA-52) dataset [1] contains 52 micro-action categories along with 7 body part labels and encompasses a full array of realistic and natural micro-actions, accounting for 205 participants and 22,422 video instances collated from the psychological interviews.

The micro-actions are defined in Table 1 and categorized using a two-level labeling system: body-level categories (body, head, upper limb, lower limb, body-hand, head-hand, and leg-hand) and action-level categories (52 different micro-action categories). Each micro-action is described in Table 1.

For a complete interview video, we annotate the micro-actions in three steps. (1) The first step involves self-censorship by volunteer annotators, each responsible for independently editing and annotating one complete video. This step facilitates the identification and correction of potential errors. (2) The second step is cross-checking, which aims to reduce biases and inaccuracies that might arise from the volunteers' subjective interpretations of the annotations. When discrepancies arise, they are reconciled based on a majority consensus. (3) In the third step, a third-party team reviews the data annotations for ultimate verification following the completion of all microaction video annotations.

References

- [1] Dan Guo, Kun Li, Bin Hu, Yan Zhang, and Meng Wang. Benchmarking micro-action recognition: Dataset, methods, and applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.