Evaluation metrics

For each submission, three scores corresponding to face (F), body (B) and hands (H) predictions will be computed. Participants will be ranked by the average rank position obtained on each of these three scores.

The scores (F, B and H) are the areas under the curves of the Cumulative Error Distribution for face, the Percentage of Correct Keypoints for body joints and the Success Rate for hands (computed for each hand, H^L and H^R, and then averaged):

$$F := AUC_{CED(0:0.25)}^{M_F} \qquad B := \frac{1}{P} \sum_{i=0}^{P} AUC_{PCK_i(0:0.5)}^{M_B}$$

$$H = \frac{H^L + H^R}{2} = \frac{(AUC_{SR_L(0:0.5)}^{M_L} + AUC_{SR_R(0:0.5)}^{M_R})}{2}$$

CED: Cumulative Error Distribution
PCK: Percentage of Correct Keypoints
SR: Success Rate
P: number of body joints considered
M_F, M_B, M_L, M_R: frames where the face, the body and the left/right hands are available and correct, respectively (checked by visual inspection).

To account for the error present in the body annotations, the area under the curve of the PCK plots is computed using only 10 bins, compared to the 100 bins used for face and hands.

The metrics are computed by normalizing the distances of each part by:
Face: <u>interpupil distance</u> (distance from the average of landmarks 38, 39, 41 and 42 to the average of 44, 45, 47 and 48)
Hands: <u>palm width</u> (distance from hand landmarks 5 to 17).
Body: <u>head size</u> (2.5 times the distance between landmarks 9 and 34 from the face).

Remarks:

- Faces, bodies and hands with inaccurate landmarks will not be taken into consideration for evaluation. These landmarks are manually flagged only in validation and testing sets.
- If the hand is not visible at a given frame, those hand landmarks and the wrist joint (from the body landmarks) will not be evaluated.