Gesture Annotation Key: Summer 2016, CSU CwC Project

Isaac Wang, Mohtadi Ben Fraj, Pradyumna Narayana, Dhruva Patil, Gururaj Mulay, Rahul Bangar, J. Ross Beveridge, Bruce A. Draper and Jaime Ruiz

Last Update: May 25, 2017

There are two tracks for labeling each video. The first track describes the physical properties of the subject's body, head, arms, and hands. The second track describes, when possible, the assumed intention of the subject. While perfect separation of description of gesture from intention of gesture is impossible, this distinction will be respected while annotating videos. Here first, the annotation key for gestures is laid out, followed by the annotation key for intention.

Labelling Order

```
body - head - arms - hands - left arm (LA) - left hand (LH) - right
arm (RA) - right hand (RH)
```

Gesture Annotation

Here is our key for how we label video gestures beginning with the general syntax. Categories are enclosed in <>. Some of these could be omitted depending on the particular case. For example, orientation is omitted unless explicitly noted as required in the documentation of corresponding motion/relation/pose.

```
Plural part: <motion>, <orientation>, into <relation>,
<orientation>, into <pose>, <orientation>;

Single part: <motion>, <orientation>, into <pose>, <orientation>;
```

Since we are labeling video segments, i.e. sections of video corresponding with gestures, there will be cases where multiple annotations, separated by ';' may appear on a single video segment's annotation.

The options for body parts are limited, including only:

Body Parts	
Part	**Description**
`body`	Used to describe large motions of the person's upper body

head	Used to describe head motions, e.g. a nod or shake
arm	Refer to 'arms', 'left arm' or 'right arm'. Syntactically represented as 'arms', 'RA' (right arm) or 'LA' (left arm). These should be set to describe large intentional arm motions. For example, a person bringing their hands together is described as moving their arms together.
hand	May be modified as 'hands', 'left hand' ('LH') or 'right hand' ('RH'). Used to describe the orientation, pose and motion of one or both hands. Note that we are not using 'palm' as a key word. Instead, if wishing to describe a pose with palms facing each other we simply say 'hands: facing'

A note, we are not presently addressing facial expressions. At some point down the road expression should be dealt with using automated face tracking and expression recognition.

Body Part Orientations	
Orientation	**Description**
up	Generally used to describe orientation relative to gravity. Examples include 'arms: up;' which would mean arms held up high, or 'left hand: up, claw;' describes a person holding their hands in a 'claw' pose (see below for poses) with fingers pointed upwards.
down	The opposite of 'up'.
front	Used to denote arm motion or hand orientation clearly in the direction that the person's body is facing (their front).
back	The opposite of 'front'; motion or orientation in the direction that the person's back is facing.
left	For objects with a front and back, for example a hand, left describes a hand with its front (palm) pointing toward the subject's left (see longer description of left and right below).
right	Opposite of 'left'.

Note: relations often need an axis rather than a direction. For example, people might place a gap between their hands, and we need to note the orientation of the gap. Use the first term above for each axis, i.e. 'up' (not 'down'), 'left' (not 'right') and 'front' (not 'back').

There are more special terms to capture various poses observed in gestures.

Body Part Poses	
Pose	**Description**

`claw (hand only)`	Describes a relaxed hand pose where fingers are moderately curved, as is the thumb, as though in position to grasp an object. Must be paired with an orientation.
`L (hand only)`	Describes a configuration of the hand, typically in conjunction with the second hand, where fingers and thumb form an L shape. A stereotypical use of this pose would be a person holding up two hands to form a rectangular window through which to look. This is the only annotation where letter case matters: use a capital L.
`x (hands only)`	Describes two hands crossed, making an 'X' shape. Must be paired with an orientation. This can be facing up, front, or left (horizontal)
`point (hand only)`	Describes the common hand pose with one finger extended. This pose is tightly bound to intention expressing the goal of having another person attend to the object indicated. Must be paired with an orientation.
`thumbs (hand only)`	The most common case is the common thumbs up gestures will be described as: 'RH: into thumbs, up;' Therefore, the opposite is 'RH: into thumbs, down;'. In general, it must be paired with an orientation.
`one … five (hand only)`	Counting with fingers; a person holding two fingers up would be described as: 'RH: into two, front;' Clearly the max is five which includes the thumb. Must be paired with an orientation.
`fist (hand only)`	Hand pose with fingers balled up.
`open (hand only)`	A hand held with fingers and thumb all in a common plane (flat) and with fingers and thumb spread wide apart. Must be paired with an orientation.
`closed (hand only)`	A hand held with fingers and thumb all in a common plane (flat) and with fingers either touching each other or close to touching. If in doubt, the label 'close' is to be used. Must be paired with an orientation.
`fence (arm only)`	Place an arm across the table creating a barrier or to suggest a point of alignment. Must be paired with an orientation.
`inch (hand only)`	Hold the hand with the thumb and forefinger about an inch apart (to signal a little bit)
`to face (hand only)`	Pose associated with touching the face with the hand
`to hip (hand only)`	Pose associated with putting hands on hips or in pockets

To describe motions, there are even more terms than for either orientation or pose.

Body Part Relations	
Relation	**Description**
`facing (hand only)`	Generally used to describe hands where palms are pointed toward each other. Must be paired with an orientation. Note this is an orientation described in relative terms; one hand relative to the other.
`opposed (hand only)`	The opposite of 'facing'. Must be paired with an orientation.
`interleave (hand only)`	Two hands with fingers interleaved.
`hold (hand only)`	Two hands joined in some manner of holding. Generalization of interleave; interleave is a special case.
`gap (arms only)`	Two hands close together, but with a gap/space in between them, denoting a gap between two objects. Hands must not be touching. Must be paired with an orientation.
`contact (arm only)`	Two hands close together, but touching, denoting two objects together as a join. Must be paired with an orientation.
`crossed (arms only)`	Two arms, crossed.

Body Part Motions	
Motion	**Description**
`together`	Used to describe any gesture involving pairs of body parts, i.e. either arms or hands, where they are being brought together. It is not necessary for them to join (see 'join' below) for the motion to be labeled 'together'. Must be paired with an orientation.
`apart`	The opposite of together. Must be paired with an orientation.
`move`	Used to describe any sort of translational movement. Must be paired with an orientation, showing the direction of movement. **Hands cannot "move"**. Instead movement is delegated to the arms.
`rotate`	All parts can rotate. A body rotation must be pronounced and clearly visible. Rule of thumb: more that 30 degrees and clearly observable. For arms, the meaning of rotate is reserved for the

	two arms rotating as a joint motion. Generally, both arms are in motion and the result is an observable rotation about some third point in space. It is possible to have a single arm rotate; think of a person moving their hand through space in a windshield wiper motion. For hands, one or both can rotate, but rotation is local to the hands themselves and the rotation for each hand is centered on each hand independently. Otherwise, the rotation should be described at the level of the arms or body. Note that we are intentionally avoiding clockwise versus counter-clockwise distinctions.
`pivot`	A special case of arms rotating in which one arm and hand remain fixed and the other is rotated around the first. In other words, the arm in motion pivots about the other.
`rub (hands only)`	Most commonly associated with hands, i.e. to rub hands together.
`scratch (hand only)`	Interaction of the hand with some other body part. While seldom intentional in terms of communicating to another person, scratching is a common motion.
`still (body only)`	The absence of motion. This is an important label when labeling a segment of video in which a person is holding a pose without significant motion. Some judgement should be used in decided if a person is holding still since some degree of small accidental motion is normal. The label 'still' should be used if in your judgement the subject is not attempting to move to a new pose.
`nod (head only)`	Used to describe head motion. Should be a pronounced up-down nodding of the head often connected with the intention to convey affirmation.
`shake`	Used to describe a pronounced to-and-fro motion that is repeated more than once rapidly. A common use is to capture head shakes that convey a negation, but it can also be used with hands or arms in relatively rapid repetitive motion usually intended as a servomechanism (this may have a separate label in future). Must be paired with an orientation.
`surround`	A relatively complex hand motion also involving the arm. To illustrate, consider 'RA: surround'. This would be used to describe a subject moving their right arm as though to partially enclose or surround a space. This is a real gesture that has been observed and it is used apparently to draw attention to a region of space.
`tap (hand(s) only)`	A quick, repeated tapping of a body part against something else (as opposed to a shake, which is in the air). Canonical example is a hand tapping against the table.
`beckon`	Curling of arms or hands toward the body. Note the lack of

	orientation. If orientation is needed, consider using `shake`.
enough	The opposite of beckon, pushes arms/hands away and down.
stack (arm only)	A hand gesture involving a rotation of the arm specifically used to designate moving one object on top of another.

Intention Annotation

The syntax for intention annotation is simpler than for gestures.

Intention Special Words	
Word	**Meaning**
servo	Performed in parallel with performer, with the idea that when the signaler stops, the performer will too. Intents without this word are assumed to be command/response.
relative	Direction and/or location is relative to an established referent. Intents without this word are in respect to the table (global).

Intention Categories (Signaler)	
Category	**Possible Values**
#	one | two | ...
object	block(s) | row(s) | column(s) | stack(s) | yourself | pyramid | gap | layer
dir	left | right | forward | backward | middle | between | top | down | diagonal

So far the annotations are with respect to the following keywords. [] indicate optional terms, while <> indicates a category. For example, [<dir>] would represent an optional element of the dir category.

Intention Annotations (Signaler)

label	description
start	Initiate an action or task
wait	Wait for performer to complete action
#	A number. What it refers to must be determined.
[<#>] new <object>	Pickup one or a specified number of new objects.
here [relative]	Designate a target location where some action should take place, often the placement of a block. May be through some form of servoing or mime like gesture.
there [relative]	Designate a target location where a block of interest is located or some block is to be placed. May be through some form of pointing gesture.
this	Establish reference
these	Establish reference
edge <dir>	Marks the edge of the table. Direction specifies which edge
[servo] together	Put two blocks together so they are joined, often with faces touching.
[servo] apart	Take two blocks that are together and move them apart.
stack [relative]	Put two blocks together one on top of the other.
[servo] slide <dir>	Move block(s) sideways, typically across the table top.
[servo] rotate [relative]	Rotate a single block in place.
OK	Signal to performer that some task or action is successful.
yes	Affirmative, but in response to a physical question, not an accomplished task
no	Signal to performer that some task or action is wrong or has failed.
stop	Stop the current on-going action
done	Signal to performer that an action or task is complete and they should stop.

think	A pause by the performer while they think about what to signal next, and/or check the existing block pattern against the target pattern.
fidget	Motions not meant to convey intended actions. Examples include scratching or interleaving fingers.
reference	Establish a reference point
hello	Greet, at the beginning of a conversation
curve	Place blocks as a curve
grab	Grab onto (a block)
talk	For sessions involving speech, when the participant is talking and the body is not moving (no gestures)
out of context	Talking to experimenter
layer	Rotating arm on horizontal plane