

Table S1. Taxonomy showing which input modalities and interaction techniques can be enabled by the inside-out (smartphone only) tracking approach for distant display. Indicated with • are tracking abilities enabled by TrackPhone and which currently lack from known related approaches. Indicated with •• are tracking abilities enabled by TrackPhone, which currently lack from known related approaches and which we investigated in our user study.

Inside-out | Smartphone only

Device-tracking input

Smartphone Touch

► Touchscreen finger tracking to:

Interact by **tapping** with the finger for:
→ selecting (clicking) [19,41,40,2,••]
→ absolute 2D cursor placement [19,21,4]

Interact by **two-finger tapping** with the finger for:
→ switching between input states (clutching) [23]

Interact by **pressing and holding** the finger down for:
→ switching between input states (clutching) [19,100,104,2,••]
→ automatic 3D object translation along its normal [41]

Interact by **swiping** with the finger for:
→ left-right item scrolling in a horizontal cover flow [98,97]
→ pushing a 2D object into a direction [99]
→ controlling a target-agnostic 2D cursor [59]
→ transferring content to the distant display [107]

Interact by **one-finger dragging** for:
→ zooming a remote view [103,2]
→ panning remote 2D content [1]
→ absolute 2D cursor control [19,21,40,5,23]
→ relative 2D cursor control [19,99,101,59,5,23,••]
→ absolute depth-control of a 3D cursor [24]
→ relative depth-control of a 3D cursor [24]
→ translating 2D objects [99]
→ translating a 3D object on two-axes [20,42,102,105,45]
→ translating a 3D object using virtual proxies [46]
→ rotating 3D object using an arcball [100,46]
→ rotating 3D object on one-axis [2]
→ sketching 2D strokes [41,107]
→ controlling a game character [103,105]

Interact by **two-finger dragging** for:
→ 2D translation of 2D objects [41]
→ translating a 3D object on one-axis [102,29]
→ translating a 3D object on two-axis [29,70]
→ translating a 2D objects covered by other 2D obj. [99]
→ rotating 3D on one-axis [29]
→ 2D rotation, scaling and translation of 3D objects [100]

Interact by **two-finger twisting** for:
→ rotating 3D on one-axis [29,70]

Interact by **two-finger pinch/spread** gestures for:
→ uniform scaling of 2D objects [41,40]
→ uniform scaling of 3D object [20,41,46]
→ translating a 3D object on one-axis [45,70]

Interact by **three-finger pinch/spread** gestures for:
→ scaling 2D objects [41]

Smartphone Translation (position)

► SLAM tracking to:

Interact by changing the phone's **XY world-position** for:
→ absolute 2D cursor control [2,••]
→ text-selection [2]

Interact by changing the phone's **XYZ world-position** for:
→ absolute 3D cursor control [2,••]
→ translating 3D objects [2,100,117]
→ pointing with a screen-centered crosshair [117]

► Optical flow tracking to:

Interact by changing the phone's **Z world-position** for:
→ uniform scaling of 3D objects [118]

Interact by changing the phone's **XY world-position** for:
→ relative 2D cursor control [114,25,3]
→ rotating 3D objects [118]
→ pointing with a screen-centered crosshair [4]

Interact by changing the phone's **XYZ world-position** for:
→ translating 3D objects [108]
→ drawing and editing 3D strokes [118]

► Marker-based camera tracking to:

Interact by changing the phone's **XY world-position** for:
→ pointing with a screen-centered crosshair [3]

Interact by changing the phone's **XYZ world-position** for:
→ relative 3D cursor control [114]

Phone Keys and Joysticks (legacy)

Smartphone Tilting (orientation)

► IMU tracking to:

Interact by changing the **yaw** of the phone for:
→ steering a in game vehicle [111]

Interact by changing the **pitch** of the phone for:
→ zooming [1]
→ a single-axis rotation of 3D objects [21]

Interact by changing the **roll** of the phone for:
→ horizontal panning in 2D [1]
→ item left/right scrolling in a horizontal cover flow [1]
→ rotating 3D object on one-axis [1]
→ rotating a 2D object [111]

Interact by changing the **yaw and pitch** of the phone for:
→ relative 2D cursor control [109,24,40,105,25,59]
→ absolute 2D cursor control [59]
→ controlling a absolute target-agnostic 2D cursor [59]
→ rotating a 3D object around two-axes [102,45]
→ controlling a 2D marking-menu [19]
→ controlling video and graphical effects [116]

Interact by changing the **yaw, pitch and roll** of the phone:
→ for rotating a 3D objects [20,100,41,110,111,112,103,114,115,108,46]
→ for rotating a 3D scene camera [20,40,103]

Interact by a **throwing gesture** with the phone:
→ to transfer content to the distant display [1, 106]

► Optical flow tracking to:

Interact by changing the **roll** of the phone for:
→ rotating a 2D object [114]
→ steering a in game vehicle [114]

Smartphone Pose (position + orientation)

► SLAM tracking to:

Interact by **moving the phone in world-space** for:
→ absolute 2D cursor control [••]
→ absolute 3D cursor control (virtual-hand) [100,2,••]
→ absolute 3D cursor control (depth ray-casting) [••]

Smartphone User Interface

► Phone's graphical user interface to:

Interact by **UI elements** (e.g. buttons, sliders, list):
→ for controlling a remote user interface [97,119,111,101]
→ for controlling a game character [97,120]
→ for controlling translating 3D objects [121]

Interact by a **peephole user interface** to:
→ project phone-touch to the distant display [3,4]

Interact by a **distributed user interface** to:
→ input data [91,123,124]
→ exchange content [91,123,124]

Body-tracking input

Traveling / Walking

► Simultaneous face- and world-tracking to:

Interact by **walking in world-space** for:
→ body position tracking demo [•]
→ body y-axis orientation tracking demo [•]

Head motion

► Simultaneous face- and world-tracking to:

Interact by **moving your head in world-space** for:
→ absolute 2D cursor control (head-pointing) [••]
→ absolute 3D cursor control (depth head-pointing) [••]

Eye motion / Gaze

► Simultaneous face- and world-tracking to:

Interact by **moving your eyes in world-space** for:
→ absolute 2D cursor control (gaze-pointing) in 2D [•]
→ absolute 3D cursor control (gaze-pointing) in 3D [•]

► Phone hardware keys and joysticks to:

Interact by **pressing** a key for: → selecting ("click") [25,47,31]
Interact by **pressing and holding a key** for: → switching the input state (clutching) [20,42,108]
Interact by **tilting the joystick** for: → relative 2D cursor control [25,47]



Keys [47]



Touch [19]



UI [4,111]



Tilt [20,21]



Translate [114]



Device pose [2]



Body pose [•]

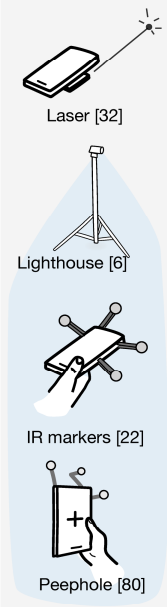



Head pose [•]



Eye-Gaze [•]

Table S2. Taxonomy showing which input modalities and interaction techniques can be enabled by the hybrid (smartphone + additional hardware) tracking approach for distant display. On the bottom, we show related scenarios that could also benefit from our taxonomy and findings.

Hybrid Smartphone + X	
Device-tracking inptus	
Smartphone Touch <p>Interact by tapping with the finger for: → selecting (clicking) [22,125,48,23,29,121,66,6,32,126]</p> <p>Interact by double-tapping with the finger for: → switching between input states (clutching) [126] → selecting (clicking) [22,125] → deselecting (reverse clicking) [22,125]</p> <p>Interact by pressing and holding with the finger for: → switching between input states (clutching) [22,125,23,29,66,6,104,45] → grabbing 2D objects [22,125] → grabbing 3D objects [33]</p> <p>Interact by one-finger dragging for: → zooming [22,125,66] → relative 2D cursor control [6,126] → rotating 3D object using an arcball [33] → performing strokes gestures [66,32,33]</p> <p>Interact by two-finger dragging for: → rotating 2D objects [6,38,32] → translating 2D objects (panning) [38,33]</p> <p>Interact by two-finger pinch/spread gestures for: → uniform scaling of 2D objects [6,32,33] → uniform scaling of 3D objects [33]</p> <p>Interact by swiping with the finger for: → triggering pre-defined application commands [22,125]</p>	Smartphone Pose (postion + orientation) <p>► Lighthouse camera and phone attached laser tracking to:</p> <p>Interact by moving the phone in world-space for: → absolute 2D cursor control (ray-casting) [33] → absolute 2D cursor control (ray-casting) [32,31,129]</p> <p>► Lighthouse camera and IR marker tracking to:</p> <p>Interact by moving the phone in world-space for: → absolute 2D cursor control (ray-casting) [22,125,48,128,30,66,32] → absolute 3D cursor control (virtual-hand) [29,121] → translating and rotating 3D objects [29]</p> <p>► Lighthouse camera tracking to (e.g. Kinect):</p> <p>Interact by moving the phone in world-space for: → absolute 2D cursor control (ray-casting) [104]</p> <p>Smartphone Tilting (orientation)</p> <p>► IMU tracking to:</p> <p>Interact by changing the roll of the phone for: → rotating a 2D object [6,32]</p> <p>► Lighthous camera and IR marker tracking to:</p> <p>Interact by changing the yaw, pict and roll of the phone: → for rotating a 3D objects [29,127]</p> 
Smartphone User Interface <p>► Phone's graphical user interface to:</p> <p>Interact by UI elements (e.g. buttons, sliders, list): → for controlling a remote user interface [131,128,127]</p> <p>Interact by a peephole user interface: → for controlling a remote user interface [127,128,80] → to project phone-touch to the disntat display [48]</p> <p>Interact by a distributed user interface to: → inptut data [128,127]</p>	Body-tracking inptus <p>Traveling / Walking</p> <p>► Lighthouse camera and IR marker tracking to:</p> <p>Interact by walking in world-space for: → body position tracking monitoring [35,39]</p> <p>Interact by leaning the body for: → manipulating a 3D camera viewport (traveling) [45]</p> <p>Head motion</p> <p>► Lighthouse camera and IR marker tracking to:</p> <p>Interact by moving your head for: → absolute 2D cursor control (head-pointing) [6,23]</p> <p>Eye motion / Gaze</p> <p>► Eye-tracking glasses to:</p> <p>Interact by moving your eyes in world-space for: → absolute 2D cursor control (gaze-pointing) [6,38]</p> <p>► Lighthouse eye-tracking to (e.g. Tobii):</p> <p>Interact by moving your eyes in world-space for: → absolute 2D cursor control (gaze-pointing) [126,130]</p> 
Smartphone Translation (position) <p>► Phone attached laser tracking to:</p> <p>Interact by changing the phone's Z world-position for: → zooming [33] → scrolling [33]</p> <p>► Lighthouse camera and IR marker tracking to:</p> <p>Interact by changing the phone's XYZ world-position for: → for translating 3D objects [127]</p> <p>► Lighthouse camera tracking to (e.g. Kinect):</p> <p>Interact by changing the phone's XYZ world-position for: → for translating 3D objects [110]</p>	
Extras	
A list of domains that are out of the scope of this work could however directly benefit from TrackPhone to initially enable or enhance their existing interactions.	
Distant display + X <p>Vision-based lighthouse tracking for: bare-handed [195, 245, 7, 196, 197, 68, 198, 200, 199., 201], finger-gestures [202], body pose [34, 200, 203, 204, 245, 195, 205], body outline [8, 206, 207, 208], body skeleton [36], eye-gaze [68, 209, 7, 37, 210, 211, 212], head-based [199, 213, 214], tangible object [201, 100, 15, 215, 216, 71] interactions; Drone-based tracking [217]; Wearable devices: wearable eye-tracking glasses for gaze interaction [218, 219, 220, 221], smartwatches [195, 222, 223, 224, 225], shoulder-worn system [226], bands [228, 227], textiles [230, 231], shoes [232], rings [233]; Handheld devices: wands [76, 34, 234, 235, 236, 237], handheld remote controllers [43, 238], laser pointers [239, 240, 31, 241], pens [242, 243], joysticks [238]; On-Display inputs: Large touchscreens [244, 245, 99, 246], phones in physical contact with the distant display [247, 248], speech [249]; 3D user interfaces: 3D user interfaces [250, 251], fish-tank user interfaces [76, 252, 77, 253, 75], caves [254].</p>	
Smartphone only Smartphone+ X <p>Interactions: touch [132, 133, 134, 73, 136], gaze [137, 52, 138, 139, 140, 141, 142], head [143, 54, 55, 144], phone's casing [145, 146, 147], foot [148], hardware buttons [149], barehanded gestures [150, 151, 152, 153], finger in mid-air [154], head-mounted displays [155, 156, 136, 52, 157], cross-device interactions [44, 158], smartwatch [160, 161], pen [50, 162, 51, 163, 164]; User interfaces: handheld augmented-reality [50, 162, 165, 13, 167, 168, 169, 55], spatially-aware [170, 171, 172, 173, 174, 153, 175], peephole [176, 178, 179, 180, 181, 182, 183, 184, 185], magic lens [186, 187, 188, 189, 190, 191], phone-projector [192, 193, 194].</p>	