“Point at It with Your Smartphone”: Assessing the Applicability of Orientation Sensing of Smartphones to Operate IoT Devices

Heinrich Ruser[1](✉) and Ilan Kirsh[2]

[1] Bundeswehr University Munich, Neubiberg, Germany
heinrich.ruser@unibw.de
[2] The Academic College of Tel Aviv-Yafo, Tel Aviv, Israel
kirsh@mta.ac.il

Abstract. The built-in orientation and motion sensors of smartphones along with their wireless communication abilities are utilized to control connected IoT devices from any place in a room, by pointing at them with the smartphone in the hand. The information of which device is targeted will be derived from the user’s actual location, the spatial orientation of the smartphone and pre-knowledge regarding the positions of devices. Chosen devices are remotely operated with simple mid-air gestures performed with the smartphone. The feasibility of this cost-effective approach is assessed by user experiments. The continuous readings of the smartphone’s inclination, rotation and magnetic field sensors are recorded with a dedicated freeware app. An algorithm combines the sensor readings to deliver the actual spatial orientation. Our preliminary experiments with different smartphone models and several users show that pointing at defined positions and performing gestures with a smartphone in the user’s hand can be accurately sensed without latency and with small deviations of the orientation measurements in the range of up to 5 degrees, indicating the feasibility of this novel approach.

Keywords: Human-centered computing · Human computer interaction · Pointing devices · Universal remote control · Smartphone sensors · User experience

1 Introduction

A way to control technical devices in the living environment in a simple, consistent and intuitive manner would be highly desirable. One of the most intuitive and natural ways to address a visible object in front of the user but beyond the range of touch is to point in its direction with a finger, hand or arm [1]. In order to ‘control’ the targeted device the user would move a finger, hand or arm ‘in the air’ to perform basic gestures [2].

As experience shows, we are surprisingly good at pointing with a finger, hand or arm at visible targets in specific spatial directions. User studies (e.g. [3–6]) revealed that despite the parallax between eye and hand, the spatial accuracy of pointing at a target with a finger or arm of the dominant hand is usually below 10° both in horizontal and vertical dimensions.

C. Stephanidis et al. (Eds.): HCII 2021, LNCS 13095, pp. 115–131, 2021.
https://doi.org/10.1007/978-3-030-90963-5_10

The question is how to use this ability of precise direction pointing to control remote devices, i.e. how to address the devices and communicate with them in a flexible, natural and seamless way. We propose to use commodity smartphones for both, selecting a device from a distance and remotely controlling its basic functions.

The state-of-the-art of using smartphones as remote controls in a smart home environment with different appliances (TV sets, music boxes, lights, blinds, fans, etc.) accessible via Bluetooth or WLAN (so called IoT devices) is to launch the device-specific app and select and operate the device via the touchscreen. However, in many situations these display-based smartphone apps are perceived as not very user-friendly: It can be annoying to find the right app on the smartphone followed by maneuvering through menus and touching designated small and slippery ‘software buttons’ on the touchscreen. With dry, wet, cold or trembling hands, without glasses, being in a dark room or in a hurry or for users with visual impairments or hand movement disorders, this all may be challenging. In addition, the effort is often perceived as disproportionate to the simplicity of the task at hand, like dimming a light or switching off the TV.[1]

Smartphones are truly ubiquitous – most people carry their smartphone at almost all times wherever they are. Besides being a versatile and trusted always-on communication device, a smartphone is also a powerful measuring device, capable of sensing its environment with a variety of built-in-sensors. The ability to sense its spatial orientation can be used for our purpose.

We propose a new type of remote controls: phone-pointing remote apps [7]. Unlike touch-based apps, phone-pointing remote apps can be used with the phone's screen turned off. The interaction scheme is simple and intuitive: With a smartphone in the hand, the user points towards a visible but remote IoT device in a room to select it. Subsequently, holding the smartphone in the hand and performing some specific hand or arm motions ‘in the air’ the user remotely operates the chosen device, triggering several basic functions, like dimming a ceiling light, lifting window blinds or increasing the volume of a TV set or radio. The control information is sent to the chosen device via wireless communication between the smartphone and the wirelessly accessible IoT device.

Whereas phone-pointing to remotely operate IoT devices appears simple and straightforward, its underlying concept requires the determination of the smartphone’s orientation and its localization in a room while using it as a pointing device. From the localization of the smartphone (and the user) and the spatial orientation, the ‘pointing projection’ will be calculated as a straight line in absolute 3D coordinates. Finally, positions along this projection will be matched with a list of known positions of devices to be remotely operated in that room in order to appoint the selected device.

To the best of our knowledge, there are no such implementations yet. In [7], the system architecture needed to turn ordinary smartphones into highly available, cost effective gesture-based remote controls is laid out. In the present paper, the applicability of the new approach of targeting different IoT objects from various user’s

[1] Voice user interfaces, while easier to use, can show annoying performance drops due to disturbances from ambient noise or unclear pronunciation. Moreover, they meet reservations from the hesitation to speak to a technical device altogether and raise concerns regarding data privacy issues.

positions based on the information about the direction of pointing the smartphone in the user's hand is addressed.

The remaining of the text is organized as follows. In Sect. 2, related work regarding the usage of smartphones as pointing devices, for orientation sensing and for indoor localization is presented and discussed. In Sect. 3, the operating principles for phone-pointing remote controls are shortly described. In Sect. 4, we present results from experiments to verify the applicability (in terms of accuracy, repeatability and latency) of standard smartphone sensors for effective direction and motion sensing. The results and consequences are discussed in Sect. 5. A conclusion and an outline of future work are given in Sect. 6.

2 Related Work

As the information about the orientation and the localization of the smartphone in the user's hand are prerequisites for the novel approach, related work in these fields is briefly revisited.

2.1 Smartphones Used as Pointing Devices

A variety of approaches have been applied to allow for user interaction by pointing using smartphones, whether relying on direct pointing with attached flashlights or laser pointers, indirect pointing based on camera images or inertial sensing, using accelerometer and gyroscope sensor measurements.

Pointing with wearables has been widely investigated in the context of in-air remote interaction with large displays and screens [8, 9]. Examples are commodity devices like the *AirMouse* or special-purpose controllers like *XWand* [10] or *MagicWand* [11] based on orientation data provided by built-in inertial sensors. Usually, the goal has been to design a precise pointer that is usable simultaneously as an input device ('point-and-select', 'zoom-and-pan', 'drag-and-swipe' selected objects on the screen, etc. [12]) and also an output device (get information from the selected objects and other 'feedback'). A *PointerPhone* [13] was realized by attaching a laser pointer to a mobile phone and using a static camera to track the bright laser dot on the remote screen. Software buttons and fingertip gestures on the phone's touchscreen were proposed to address and manipulate the object pointed at on the remote screen. Re-calibration of the screen's position and orientation is required each time the positions of the camera or the screen change.

Phone-pointing techniques like *SmartCasting* [14] or *TiltCasting* [15] based on the smartphone's built-in inertial sensors have been proposed to be used in Augmented Reality (AR) applications to project the current smartphone display image via ray-casting into the surrounding 3D space. In AR/VR applications wearing a Head-Mounted Display (HDM), smartphones can be useful as virtual joysticks: a good motor control and dexterity are often expected from the users, as well as short selection times [16].

An interesting observation regarding the pointing accuracy of users in front of large screens has been repeatedly made (e.g. [17]): On average, pointing errors were larger when the user was in closer distance to the object displayed on the screen than if the screen was in a larger distance, presumably due to the larger parallax between the eye-to-target line and the arm-to-target line. However, since in closer distance the same object appears larger, the pointing accuracy is less demanding in this case and the larger errors are tolerable.

2.2 Smartphone Sensors for Orientation Sensing

Modern smartphones are equipped with a large number of miniaturized sensors which can be grouped into two categories:

- Position and motion sensors (magnetometer, accelerometer, gyroscope) measuring the strength of the surrounding magnetic field as a 3-component vector and the linear acceleration and rotational velocity of the smartphone along three axes;
- Environmental sensors (front and back cameras, thermometer, barometer, hygrometer, photometer) taking images and measuring the temperature, air pressure, air humidity and illumination.

The measured accelerations along all three axes (azimuth, roll, pitch, see Fig. 1) can be used, for example, to determine the smartphone's absolute 3D orientation (i.e., by what angle the phone is tilted) and to identify time-dependent movements such as rotation, swing or shake.

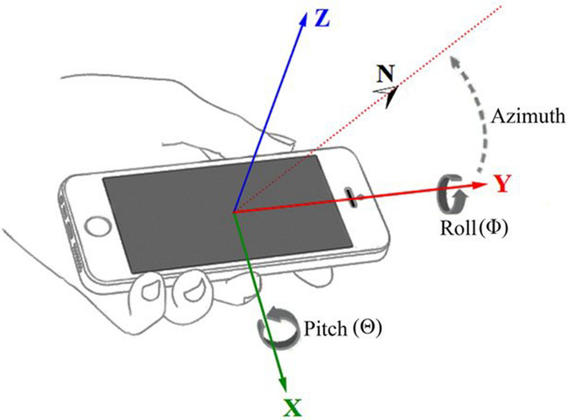

Fig. 1. Three fundamental axes and rotation vectors relative to a smartphone, with azimuth – the rotation angle about axis z (the gravity vector), pitch – the rotation angle about axis x and roll – the rotation angle about axis y. [18]

The angular velocities resulting from the gyroscope measurements describe the change in the rotation angle (pitch, roll and yaw) of the smartphone over time and therefore its relative, time variant orientation in 3D. If the initial orientation is known, the absolute orientation can be determined from it [19]. In most smartphones, the accelerometer and the gyroscope are combined in a miniaturized inertial measurement unit (IMU) [20–22].

In the absence of strong magnetic fields in the phone's proximity, the magnetometer's 3D measurement of the strength of surrounding magnetic fields is dominated by the Earth's magnetic field. From the combination of all 3 components, the absolute global orientation of the smartphone with respect to Magnetic North is derived (compass function) [23]. In Fig. 2, the azimuth angles calculated from measurements of the magnetic field sensor are shown when rotating the smartphone in the horizontal plane (used for calibration) demonstrating the high accuracy and repeatability.

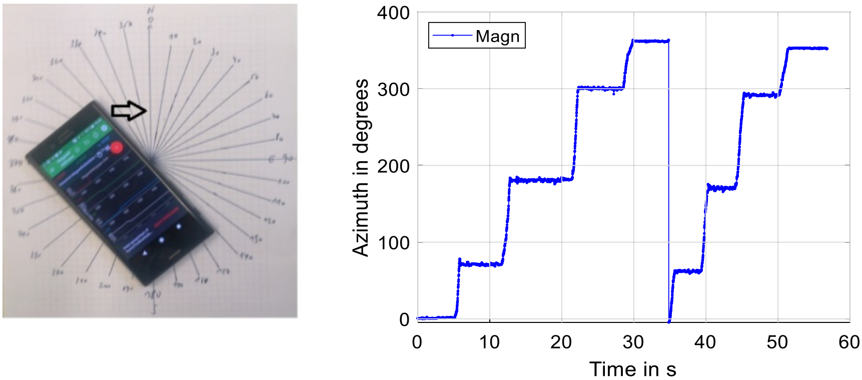

Fig. 2. Exemplary results from different orientation angle measurements using the smartphone's internal 3D magnetic field sensor.

With their constraints in size, weight, and power consumption, the miniaturized smartphone sensors have – compared to laboratory equipment – a restricted sensitivity and measurement range and are affected by noise and external disturbances to a larger extent (e.g. influences from electronic subsystems adjacent to the sensors, the Wi-Fi transceiver, battery, the package material, etc.) as well as can show long-term drifts in their measurements.

To reduce these influences and to suppress outliers and erroneous measurements, state estimation filters such as a Kalman filter and many more advanced algorithms have been applied [24]. Furthermore, the readings from different sensors can be combined and relied on each other and compared to known reference values. Such "software sensors" provide improved, smoothed estimates of the actual acceleration and orientation [25]. The typical update rate is approx. 200 measurements per second for the IMU and 100 measurements per second for the magnetometer, giving a sufficiently dense data stream for post-processing algorithms for many practical applications.

The accuracy of such "software sensors" after refining the raw sensor data was investigated in various studies. [26] confirms the presence of heterogeneities when gathering orientation output data from different smartphone devices pointing in exactly the same direction. For different smartphone models containing different sensor units, the measured orientation showed deviations from references of up to 2.1° and 6.6° for the pitch and roll angles, respectively. The accuracy achievable with smartphone acceleration sensors was shown to be sufficient for a successful recognition of uni-stroke symbols or letters written ‘in the air’ [27, 28].

A common way to improve the estimate of the absolute orientation is to combine (fuse) the readings from the IMU and the magnetometer, under the condition that the perpendicular axes of both sensors are aligned [29]. To ensure this, an occasional *re-calibration* has to be performed to correct for observed misalignments between the axes. Such automatic re-calibration could be performed e.g. at time intervals when the acceleration is measured only in the direction of gravity and the accelerations along the other axes are (almost) zero [30, 31]. Similarly, the (absolute) orientation angle from the magnetometer can be updated by the angular changes observed by the gyroscope. The gyroscope’s output can again be calibrated by comparing the outputs of the accelerometer and the IMU orientation integration algorithm, after arbitrary motions [21].

Manual re-calibration can also be important. Many re-calibration procedures are based on a multi-position approach where the smartphone is moved by hand and held in a few different static positions (recognized by a ‘static detector’), providing correction factors for (systematic) scale and misalignment (bias) for both the accelerometer and gyroscope 3D readings [29]. To calibrate the magnetometer, which is sensible to stationary and transient magnetic interferences from surrounding magnetic fields and metallic surfaces such as elevators, radiators, or concrete reinforcements, it is usually suggested to rotate the smartphone in all possible orientations.

2.3 Indoor Location Techniques with Smartphones

Using smartphones for indoor positioning is attractive for many applications where the ability to independently track people is important, e.g. in large offices or hospitals, factories and warehouses. A large variety of location techniques have been proposed and the obtained localization accuracies along with the effort and cost to achieve them have been extensively studied (see [25, 32, 33] for large in-depth surveys). During the Microsoft Indoor Localization Competition, organized in several rounds over the years 2014–2017, more than 100 teams from academia and industry deployed their indoor location solutions in quasi-realistic environments, allowing to directly compare the achieved accuracies and deployment costs [34].

However, no standard method has been brought up to date that would guarantee a similar accuracy, repeatability and seamless availability in indoor environments that global navigation satellite systems (GNSS) offer outdoors. Often, the average accuracy remains inadequate for many applications. Challenges for accurate indoor localization stem from the lack of a dense grid of absolute references for the built-in smartphones sensors for (occasional) re-calibration, the often complex indoor interior design enhancing multipath propagation or shadowing and the building materials themselves which distort or block radio and satellite signals. A major challenge for many localization-based systems, however, is the requirement for these systems to reliably track pedestrians in a highly dynamic environment, e.g., while they are walking with the smartphone in their pocket. As will be explained in Sect. 3, our application is different in this respect.

Generally, for indoor localization sensing with smartphones infrastructure-based and infrastructure-free approaches can be distinguished. Infrastructure-based approaches rely on the purposeful, optimized deployment of Bluetooth (BLE) beacons, or customized radio-frequency (RF), visible or infrared light sources or ultrasound transmitters. The signals transmitted by those beacons are picked up by appropriate smartphone sensors and translated into positions from the proximity to the closest beacon (BLE) or from the travel time of modulated signals, e.g. Ultra-Wide-Band (UWB), from the source. Often, the deployment costs remain high: Special-purpose hardware needs to be carefully deployed and hardwired or battery-powered in every area where indoor location services are needed. Whereas BLE proximity estimation allows only for a low average accuracy of about 3 to 10 m, most infrastructure-based techniques are reported to achieve localization accuracies of about 2 to 3 m in standard indoor scenarios [33].

Most infrastructure-free approaches focus on exploiting existing Wi-Fi signals from WLAN access points [35, 36], others on ambient FM radio or TV signals, geomagnetic or sound signals. As the source for indoor localization, a received signal strength (RSS) indicator is used, both for (manually) building an RSS distribution map of the ubiquitous signals in the specific indoor environment (a laborious work called 'fingerprinting') and later for finding the actual localization by matching the measured RSS to this map. Generally, the map will not be dense enough or the RSS will be unstable and distorted and hence the achievable accuracy is generally not better than 3 to 5 m [36]. A fine-timing protocol called *Wi-Fi location* includes the time it takes for the Wi-Fi signal to travel, enabling position estimation with improved accuracy of up to about 2 m [37].

Surveys revealed that due to its constant availability and high sensitivity, the best positioning accuracy could be achieved at no extra cost relying on the built-in IMU (e.g. [33, 34]). Based on double-integrating the continuously measures accelerations, the current position of the device is determined by accumulating the path vector from a known starting position (a class of techniques called 'dead reckoning'). However, small acceleration errors can rapidly accumulate to large positional errors of several meters [38]. If frequent re-calibration at reference points could be applied, e.g. at zero-points of the acceleration, the localization accuracy could be greatly improved [30].

Many realizations have been described which combine different technologies (hybrid systems) to improve the accuracy and availability of position estimates [25, 39]. For example, using IMU and WiFi RSS indicator readings and combining dead reckoning and a fingerprinting technique, localization accuracies in the range of 1 m could be achieved [32, 40]. In Table 1, average localization accuracies for selected smartphone sensors and sensor combinations are summarized.

Table.1 Average localization accuracies achieved with selected smartphone sensors and sensor combinations (adapted from [32, 33, 40]).

	Sensor	Sensor information (technique)	Approx. sampling rate per sec.	Average localization accuracy
Infra-structure-free	Magnetometer	Orientation (compass)	100	2…10° (static) (= 0.2…0.7 m in 4 m dist.)
	Accelerometer + gyroscope (IMU)	Orientation (azimuth, pitch) (dead reckoning)	200	1…3° (static) (= 0.1…0.2 m in 4 m dist.) 1…2 m
	IMU + Magnetometer	Orientation (compass + PDR)	100	0.3…1 m (static) 1.5…2.5 m (dynamic)
	Camera	Image series (optical flow)	20	2…3 m
	Barometer	Relative height (air pressure)	10	0.3…1 m
	GNSS	Global position	1	5…50 m
Infra-structure-based	WiFi	RSSI (Fingerprinting)	0.5	2…10 m
	Bluetooth	RSSI (Time difference of arrival)	1	3…10 m
	Photosensor	Position of light sources (illumination)	10	0.3…2 m
	Acoustic, Ultrasound	Distance to walls (Time of flight)	20	0.2…0.5 m

3 Operation Principles

We propose a universal, gesture-based remote control for operating electronic devices in the living environment, which would be very easy, almost intuitively to use. The built-in orientation and motion sensors of smartphones (magnetometer, accelerometer, gyroscope) along with their wireless communication abilities are utilized to control connected IoT devices in a room by pointing at them with the smartphone in the hand. The information of which device is targeted will be derived from the user's actual location, the spatial orientation of the smartphone and pre-knowledge regarding the positions of devices to be remotely operated.

A device would be marked as selected when three states are registered by the app: 1) The smartphone is pointed at a certain point in space for a longer while (i.e. the 3D coordinates of the smartphone's orientation are almost stable during a period of 1…2 s), 2) the projection of this orientation (roughly) matches with a 3D position from a stored list of devices' positions, and 3) the inclination (pitch angle) data show a short tilting or 'ticking' moment, which would stem from a natural and intuitive hand gesture by the user to confirm this selection. This 3-stage confirmation is to avoid any device to be accidentally operated by motions of the smartphone while using it for other purposes than as a remote control.

Selecting the right devices requires the app to find out the smartphone's (and hence the user's) approximate location in an indoor environment. Compared to many dynamic position-tracking applications where even approximate indoor position location finding with smartphones might be difficult (see Table 1), this problem would be facilitated here since during the process of device selection many parameters (e.g. the user's position, the smartphone's orientation) change very slowly and the number of potential positions a certain device might be selected from would be limited (by the size of the room, furniture, etc.) and/or known from previously recorded observations and teachings. Hence, the promising results from other studies for the localization accuracy obtained for IMU or magnetometer readings for static situations give positive indications regarding the feasibility of our approach.

After selecting a device by pointing at it with the smartphone, the user can specify an operation using movement gestures with the arm or hand while still holding the smartphone. A gesture is carried out with the hand-held smartphone describing 'a trajectory in the air'. For example, an upward forearm movement or a quick tilting of the smartphone (as to make a 'tick') can signal turn-on, whereas a downward movement can signal turn-off. Similarly, a clockwise rotation can signal volume-up, whereas a rotation in the opposite direction, counterclockwise, can signal volume-down. Only the relevant, intended part of the trajectory is evaluated for recognition; the delivery and final movements are discarded. By assigning the trajectory features to one of the predefined gesture classes, the gesture is automatically recognized. Knowing the basic gestures to remotely operate the main functions of the selected device, the recognized gestures can be automatically decoded by the app into device-specific commands (see Table 2).

Table.2 Example of a simple gesture lexicon and possible associated functions.

Gesture:				Could be associated with:	
No.	Acronym	Action	Trajectory	Device	Function
1	ON	Point at it, tilt	—	Lamp	Switch on
2	OF	Point at it, tilt	—		Switch off
3	UP	Move hand or arm straight bottom-up	Line bottom-up	Blind	Lift/open
4	DN	Move hand or arm straight top-down	Line top-down		Lower/close
5	LR	Move hand or arm straight left-to-right	Line left-to-right	Radio	Volume higher
6	RL	Move hand or arm straight right-to-left	Line right-to-left		Volume lower
7	CR	Turn/rotate hand or arm to the right	Circle or ellipse clockwise	Heater	Warmer
8	CL	Turn/rotate hand or arm to the left	Circle or ellipse counterclockwise		Colder

The exact gestures can vary from one implementation to another and could be customized by the user [41, 42]. The device-specific 'meaning' of the gestures performed with the smartphone in the user's hand will be sent wirelessly to the selected IoT device. The use of device-specific and also personalized gestures can be 'learned' by the app in a similar way: Users would demonstrate selections of devices by pointing at them from relevant locations and demonstrate gestures to remotely operate different functions of these IoT devices by arm and hand movements.

4 Experimental Results

The accuracy and consistency of orientation measurement were evaluated via offline calculation of experimental data, collected using a dedicated freeware app[2]. An algorithm combined and filtered the sensor readings to deliver the actual 3D orientation angles.

To make users consecutively point at specific positions in a room, a test room was prepared for the experiments with different numbered post-its ('markers') attached to different positions on a wall, in various spacings from 15 cm to 60 cm, horizontally and vertically, in heights from 0.5 m up to 2 m above the floor.

The user experiments were conducted as follows: Every user takes a position marked on the floor in a certain distance to the wall and starts recording the measurements of the built-in sensors via the dedicated app by touching the start/stop button on the smartphone's touchscreen. The user then points at a visible marker holding the smartphone with the upper edge (i.e. along the y axis, see Fig. 1) towards the marker for a short time, and possibly shortly ticks the smartphone to indicate the selection of an (imaginary) device at the marker's position. No "dry run" was performed for familiarizing the user with either the smartphone interface or the task. The user may then choose to point at additional markers consecutively in time. The position and orientation sensor readings are continuously recorded with a rate of approx. 100 measurements per second.

The data recording is continued until the user presses the smart/stop button on the touchscreen again. All recorded data are stored in a data sheet which after stopping the recording is ready to be sent from the smartphone to the computer to be further processed and visualized.

At this early stage of our investigations, the post-processing is limited to baseline reduction, phase-wrapping and motion detection, as to highlight the quasi-static orientation measurements (pointing) and differ them from the dynamic states (= motion between pointing instants; gestures). So far, no further automatic re-calibration has been carried out.

Figure 3 shows results from continued azimuth and pitch angle measurement with a smartphone held in the user's hand. Here, the user turns from West (270°) to South (180°) and back 4 times. In Fig. 3 the effect of tilting the smartphone while pointing (as to switch an imaginary device on or off) is also shown.

[2] Physics Toolbox Sensor Suite, https://www.vieyrasoftware.net/.

Figure 4 shows the potentially high pointing precision and accuracy. The test person consecutively swapped between pointing at two markers in a small distance to each other (15 cm horizontally). The absolute displacements of the smartphone's pointing positions were calculated from the known distance from the person's position to the markers (approx. 2.5 m here).

Figures 5 shows results from continued azimuth and pitch angle measurements when pointing the smartphone consecutively at different markers (their positions indicated as crosses in the 2D area plot). Below, the histograms of the azimuth and pitch measurements are given, showing a high precision (mean and maximum width of the histogram) and accuracy (deviation from the marker's azimuth positions).

Accuracies in the range of $\pm5°$ can be achieved in detecting the direction of pointing with a smartphone. This accuracy range indicates the consistency of the user

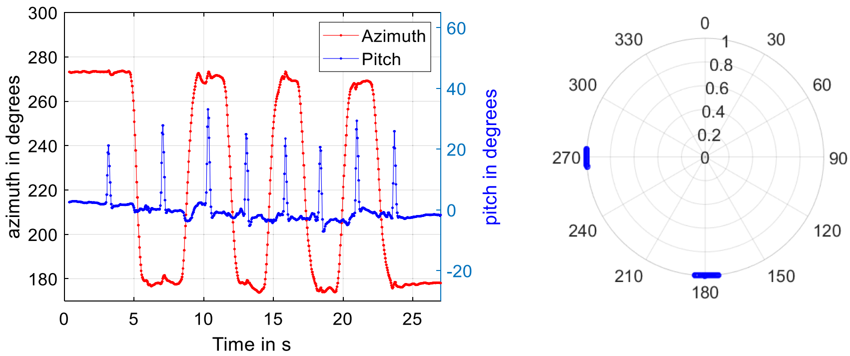

Fig. 3. (Left) Results from continued azimuth and pitch angle measurement with a smartphone held in the user's hand. Here, the user turns from West (270°) to South (180°) and back 4 times. Every time when pointing at these directions, the user slightly tilts the smartphone to indicate the wish to operate a device. (Right) Polar view of the azimuth angles when the smartphone is not moved and rests in the hand.

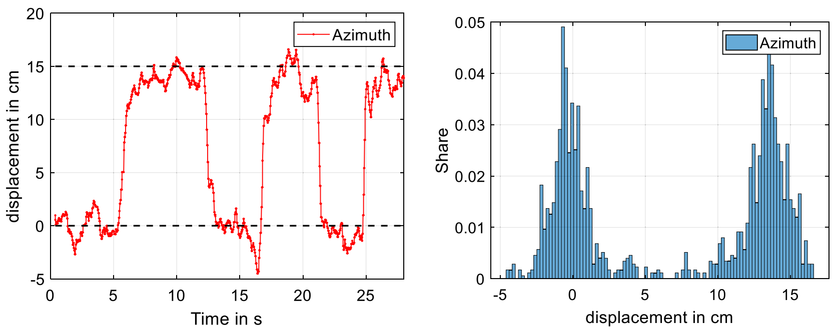

Fig. 4. Exemplary results from continued azimuth angle measurements with a smartphone in the user's hand, as a function of time (left) and histogram (right), when repeatedly pointing at two reference markers separated horizontally by 15 cm, from a distance of 2.5 m.

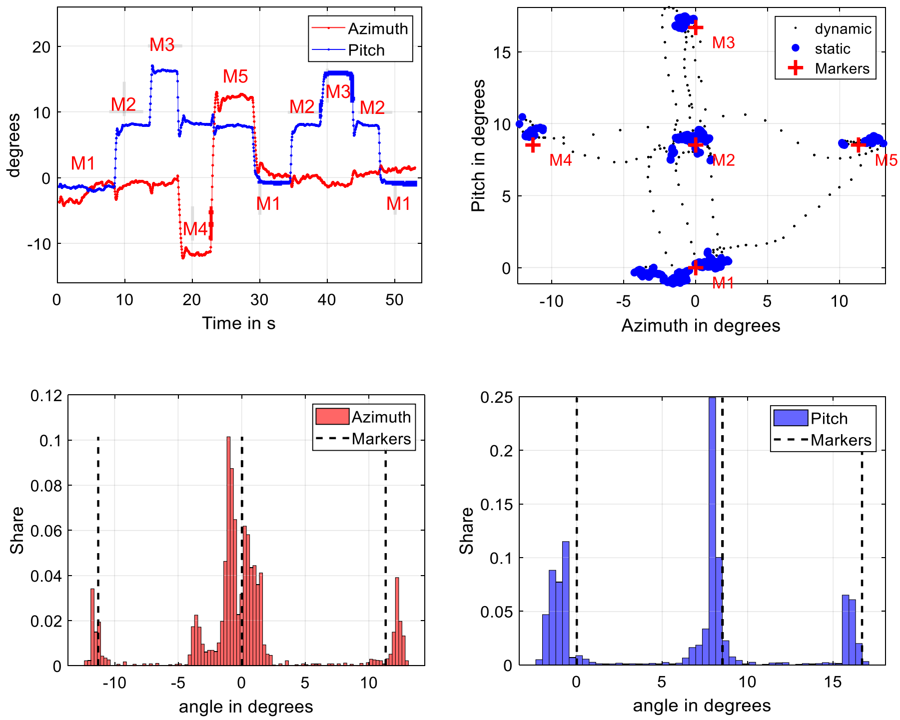

Fig. 5. Exemplary results from continued azimuth and pitch angle measurements with a smartphone in the user's hand, repeatedly pointing at 5 reference markers separated horizontally and vertically by 30 cm, from a distance of 2.5 m: Azimuth and pitch as a function of time and 2D plot of the orientation measurements (upper row), histograms of static azimuth and pitch measurements compared to the positions of markers (lower row).

operations rather than merely of the sensor measurements. For most applications, this accuracy is expected to be sufficient to select different devices located in a room by freely pointing at them with a smartphone in the hand.

In Figs. 6 and 7, results for recorded dynamic azimuth and pitch values along the smartphone's motion trajectory are shown when simple linear and circular gestures

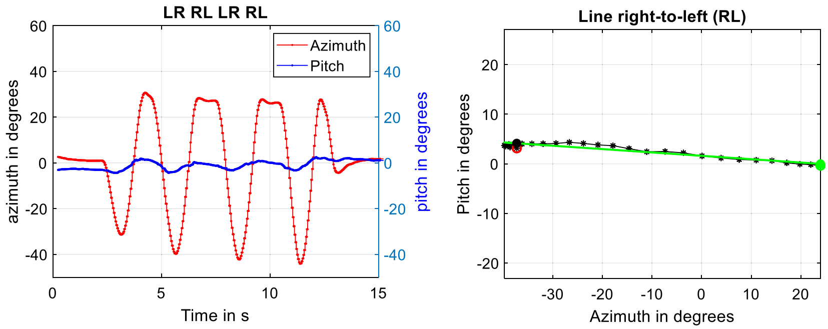

Fig. 6. (Left) Results from continued azimuth and pitch angle measurements with a smartphone in the user's hand, performing several left-right movements (Right) Trajectory in a chosen time interval and recognized gesture (in green). (Color figure online)

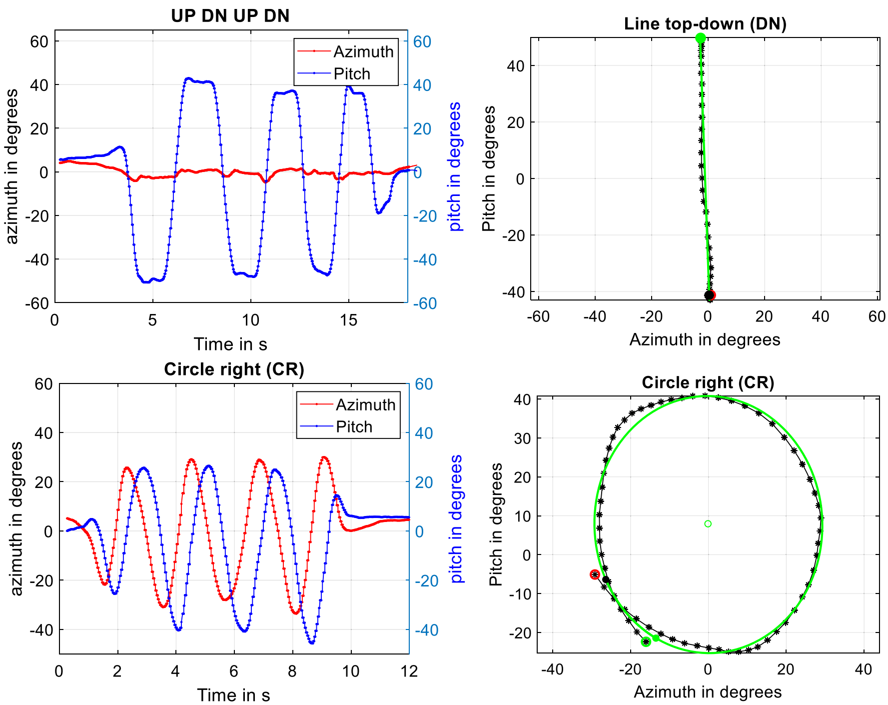

Fig. 7. Results from continued azimuth and pitch angle measurements with a smartphone in the user's hand, performing several straight up-down movements (top row) and clockwise circles (bottom row). In the right column, the performed trajectory of a chosen time interval along with the recognized gesture (in green) are shown. (Color figure online)

from the gestures lexicon are performed. To contribute those trajectories to several gesture classes, approximation fits of straight lines or ellipses are calculated.

5 Discussion

The proposed solution for a simplified seamless remote control is based on built-in smartphone functionalities and does not require additional hardware. Using a phone-pointing remote app, users would be able to select devices by physically pointing their smartphone at them, and then use hand or arm movement gestures, while holding the smartphone, to operate these devices, without needing to turn on the phone screen. Phone-pointing remote apps will use standard smartphone sensors, including the magnetometer and the g-force meter (or the accelerometer) to identify 3D pointing directions during device selection, and the accelerometer and the gyroscope to recognize movement gestures while holding the smartphone. As the published results revealed and our experiments confirm (see Sect. 4), the smartphone's spatial orientation is sensed by the built-in position sensors with high accuracy.

In order to use the smartphone as a direct pointing device, the current location of the smartphone in the users' hand has to be recognized. As discussed, deriving the location of the user in a room or home from measurements available to a smartphone can be a challenging problem [33]. Several constraints, specific for the envisioned application, help to keep the orientation errors low, and therefore, allow to be optimistic about the feasibility to determine the user's location with sufficiently high accuracy:

- While using the smartphone as a direct pointer, the user and the smartphone are usually at almost *static* positions, with slow changes of the position and the orientation angle for a period of time. The smartphone is intuitively held in almost horizontal position, simplifying the re-calibration of the acceleration sensor as the gravity component is dominating.
- Typically, the system would be deployed in small to medium sized rooms. As has been shown by several authors (e.g. [6, 17], the average positioning error achieved with algorithms based on inertial and magnetic sensor readings is disproportionally smaller in small rooms than in large rooms. The front size of many targets (i.e. the IoT device to be operated) is typically large compared to typical distances to them and the distance between two adjacent targets is typically large, reducing the requirements regarding the pointing accuracy.
- The position estimation can be improved by integrating a digital building model, containing for example spatial constraints up to non-accessible areas. In many cases it should be sufficient to estimate the most likely positions of users to operate different devices from. The list of these positions can be learned from initial observations of the users in the room and refined by using Machine Learning (ML).

6 Conclusion and Future Work

Commodity smartphones can be turned into ubiquitous, easy-to-use, intuitive gesture-based remote controls by making use of their built-in orientation sensors and wireless communication. Accordingly, they may be used as general-purpose, user-friendly alternatives to existing touch, voice or camera-based smart home interfaces, without the requirement of any extra hardware. We envisage that the proposed approach could be adapted to other contexts of mid-air interactions with multiple devices, beyond the smart home environment, like for example larger-range gestural interactions in workshops or factory halls, technology-enhanced public spaces, etc. with the aim to improve its usability, intuitiveness and experience.

We presented a novel approach to remote controlling, which combines smartphones with hand movement gestures. Unlike conventional remote control apps that use the smartphone touchscreen for input, with this approach, the phone's screen can remain turned off. Users select devices by pointing at them with their smartphone, and then use hand movements, while still holding the smartphone, to operate those devices. Although mobile phone-based interactions with remote screens have been investigated in the past, they have not been considered for applications in everyday tasks.

The promising results of preliminary experiments carried out with several users and with different smartphone models show that the novel remote control allows to accurately point at specific spatial markers without guidance or feedback (like visible dots from laser pointers). This is achievable with sufficient accuracy and repeatability, unaffected by the distances to the targeted device, obstructions along the imaginary line to that target or the ambient light conditions. Based on a high-quality orientation and motion reconstruction, with an improved separation between gesture classes the list of preferred quasi-intuitive gestures can be largely extended (including 'numbers' and other multi-stroke gestures), hence more devices and functions could be added.

As next steps, extensive validation tests to be deployed in the living environments of a larger number of test persons will be conducted and issues towards a real-time implementation of a phone-pointing app will be addressed.

References

1. Clark, H.H.: Pointing and placing. In: Kita, S. (ed.) Pointing: Where Language, Culture, and Cognition Meet, pp. 243–268. Taylor & Francis (2003)
2. Cartmill, E.A., Beilock, S., Goldin-Meadow, S.: A word in the hand: action, gesture and mental representation in humans and non-human primates. Philos. Trans. R. Soc Part B **367** (1585), 129–143 (2012)
3. Wong, N., Gutwin, C.: Where are you pointing? The accuracy of deictic pointing in CVEs. In: Proceedings of the CHI, pp. 1029–1038 (2010)
4. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., Looser, J.: Air pointing: design and evaluation of spatial target acquisition with and without visual feedback. Int. J. Hum. Comput. Stud., 401–414 (2011)
5. Akkil, D., Isokoski, P.: Accuracy of interpreting pointing gestures in egocentric view. In: Ubicomp 2016, Heidelberg (2016)
6. Flanagin, V.L., Fisher, P., Olcay, B., Kohlbecher, S., Brandt, T.: A bedside application-based assessment of spatial orientation and memory: approaches and lessons learned. J. Neurol. **266**(1), 126–138 (2019). https://doi.org/10.1007/s00415-019-09409-7
7. Kirsh, I., Ruser, H.: Phone-pointing remote app: using smartphones as pointers in gesture-based IoT remote controls. In: Stephanidis, C., Antona, M., Ntoa, S. (eds.) HCII 2021. CCIS, vol. 1420, pp. 14–21. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-78642-7_3

8. Jota, R., Nacenta, M. Jorge, J., Carpendale, Sh., Greenberg, S: A comparison of ray pointing techniques for very large displays. In: ACM 36th Graphics Interface Conference, Ottawa (2010)
9. Nancel, M., Chapuis, O., Pietriga, E., Yang, X., Irani, P., Beaudouin-Lafon, M.: High-precision pointing on large wall displays using small handheld devices. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI 2013), pp. 831–840 (2013)
10. Wilson, A., Shafer St.: XWand: UI for intelligent spaces. In: Proceedings of the Conference on Human Factors in Computing Systems (CHI 2003), pp. 545–552 (2003)
11. Ouchi, K., Esaka, N., Tamura, Y., Hirahara, M., Doi, M.: MagicWand: an intuitive gesture remote control for home appliances. In: IEEE International Conference on Active Media Technology (2005)
12. Spindler, M., Schuessler, M., Martsch, M., Dachselt R: Move your phone: spatial input-based document zoom & pan on mobile displays revisited. In: CHI 2014 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2014), pp. 515–518 (2014)
13. Seifert, J., Bayer, A., Rukzio, E.: PointerPhone: using mobile phones for direct pointing interactions with remote displays. In: Kotzé, P., Marsden, G., Lindgaard, G., Wesson, J., Winckler, M. (eds.) INTERACT 2013. LNCS, vol. 8119, pp. 18–35. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40477-1_2
14. Pietroszek, K., Kuzminykh, A., Wallace, J.R., Lank E.: Smartcasting: a discount 3D interaction technique for public displays. In: 26th Australian Computer-Human Interaction Conference on Designing Futures (OzCHI 2014), pp. 119–128 (2014)
15. Pietroszek, K., Wallace, J., Lank, E.: Tiltcasting: 3D interaction on large displays using a mobile device. In: Proceedings of the 28th ACM Symposium on User Interface Software & Technology (UIST 2015), pp. 57–62 (2015)
16. Hartmann, J., Vogel, D.: An evaluation of mobile phone pointing in spatial augmented reality. In CHI 2018 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2018) (2018)
17. Plaumann, K., Weing, M., Winkler, C., Müller, M., Rukzio, E.: Towards accurate cursorless pointing: the effects of ocular dominance and handedness. Pers. Ubiquit. Comput. **22**(4), 633–646 (2017). https://doi.org/10.1007/s00779-017-1100-7
18. Hosseinianfar, H., Chizari, A., Salehi, J.A.: GOPA: geometrical optics positioning algorithm using spatial color coded LEDs. arXiv:1807.06931v1 (2018)
19. Sato, K., Matsushita, M.: Object manipulation by absolute pointing with a smartphone gyro sensor. In: ACM Symposium on Spatial User Interaction (SUI 2019), New Orleans (2019)
20. Alce, G., Espinoza, A., Hartzell, T., Olsson, St., Samuelsson, D., Wallergård, M.: UbiCompass: an IoT interaction concept. Adv. Hum. Comput. Interact. **2018**, Article ID 5781363 (2018)
21. Solin, A., Cortes, S., Rahtu, E., Kannala, J. Inertial odometry on handheld smartphones. In: 21st IEEE International Conference on Information Fusion (FUSION), pp. 1–5 (2018)
22. Gromov, B., Abbate, G., Gambardella, L., Giusti A.: Proximity human-robot interaction using pointing gestures and a wrist-mounted IMU. In: International Conference on Robotics and Automation (ICRA) (2019)
23. Odenwald, S.: Smartphone sensors for citizen science applications: radioactivity and magnetism. Citiz. Sci. Theory Pract. **4**(1), 18, pp. 1–15 (2019)
24. Umek, A., Kos, A.: Validation of smartphone gyroscopes for mobile biofeedback applications. Pers. Ubiquitous Comput. **20**, 657–666 (2016). https://doi.org/10.1007/s00779-016-0946-4
25. Correa, A., Barcelo, M., Morell, A., Vicario, J.L.: A review of pedestrian indoor positioning systems for mass market applications. Sensors **17**(8), ID 1927 (2017)

26. Kuhlmann, T., Garaizar, P., Reips, U.-D.: Smartphone sensor accuracy varies from device to device in mobile research: the case of spatial orientation. Behav. Res. Methods **53**(1), 22–33 (2020). https://doi.org/10.3758/s13428-020-01404-5
27. Kela, J., Korpipää, P., Mäntyjärvi, J., Kallio, S., Savino, G.: Accelerometer-based gesture control for a design environment. ACM Pers. Ubiquitous Comput. **10**(5), 285–299 (2006)
28. Agrawal, S., Constandache, I., Gaonkar, Sh., Choudhury, R.R.: Using mobile phones to write in air. In: International Conference on Mobile Systems, Applications, and Services (MobiSys 2011), pp. 15–28. ACM (2011)
29. Kok, M., Schön, T.: Magnetometer calibration using inertial sensors. IEEE Sens. J. **16**(14), 5679–5689 (2016)
30. Li, X., Li, Z.: A new calibration method for tri-axial field sensors in strap-down navigation systems. Meas. Sci. Technol. **23**(10), 105105 (2012)
31. Salehi, S., Mostofi, N., Bleser, G.: A practical in-field magnetometer calibration method for IMUs. In: Proceedings of the IROS Workshop on Cognitive Assistive Systems: Closing the Action-Perception Loop, pp. 39–44 (2012)
32. Langlois, Ch., Tiku, S, Pasricha, S.: Indoor localization with smartphones. IEEE Consum. Electron. Mag. **10**(17), 70–80 (2017)
33. Nguyen, K.A., Luo, Zh., Li, G., Watkins, Ch.: A review of smartphones based indoor positioning: challenges and applications. IET Cybersyst. Robotics **3**(1), 1–30 (2021)
34. Lymberopoulos, D., Liu, J.: The microsoft indoor localization competition: experiences and lessons learned. IEEE Signal Process. Mag. **09**(17), 125–140 (2017)
35. Ashraf, I., Hur, S., Park, Y.: Indoor positioning on disparate commercial smartphones using wi-fi access points coverage area. Sensors **19**(19), 4351 (2019)
36. Retscher, G.: Fundamental concepts and evolution of wi-fi user localization: an overview based on different case studies. Sensors **20**(18), 5121 (2020)
37. Khalajmehrabadi, A., Gatsis, N., Akopian, D.: Modern WLAN fingerprinting indoor positioning methods and deployment challenges. IEEE Commun. Surv. Tutor. **19**(3), 1974–2002 (2017)
38. Real Ehrlich, C., Blankenbach J.: Indoor localization for pedestrians with real-time capability using multi-sensor smartphones. Geo-spat. Inf. Sci. (2019). https://doi.org/10.1080/10095020.2019.1613778
39. Pascacio, P., Casteleyn, S., Torres-Sospedra, J.: Collaborative indoor positioning systems: a systematic review. Sensors **21**(3), 1002 (2021)
40. Landau, Y., Ben-Moshe, B.: STEPS: an indoor navigation framework for mobile devices. Sensors **20**(14), 3929 (2020)
41. Mezari, A., Maglogiannis, I.; An easily customized gesture recognizer for assisted living using commodity mobile devices. J. Healthc. Eng. **2018**, Article ID 3180652 (2018)
42. Madapana, N., Gonzalez, G., Zhang, L., Rodgers, R., Wachs, J.: Agreement study using gesture description analysis. IEEE Trans. Hum. Mach. Syst. **50**(5), 434–443 (2020)