BackPat: One-Handed Off-Screen Patting Gestures

Karsten Seipp
Goldsmiths College, University of London
London SE14 6NW
k.seipp@gold.ac.uk

Kate Devlin
Goldsmiths College, University of London
London SE14 6NW
k.devlin@gold.ac.uk

ABSTRACT
We present *BackPat* – a technique for supporting one-handed smartphone operation by using pats of the index finger, middle finger or thumb on the back or side of the device. We devise a novel method using the device's microphone and gyroscope that enables finger-specific gesture detection and explore efficiency and user acceptance of gesture execution for each finger in three user studies with novice *BackPat* users.

Author Keywords
Back-of-device; pat; thumb; movement; sound; off-screen

ACM Classification Keywords
H.5.2 User Interfaces–Interaction styles: .

INTRODUCTION AND PREVIOUS RESEARCH
Using a phone with only one hand is the preferred mode of operation for many users [5]. However, input via the thumb is often limited. Research to overcome this and enrich the one-handed input vocabulary can be divided into three groups:

Motion as input: JerkTilts [8] uses the accelerometer to define a set of eight quick jerk gestures performed with the wrist. However, this technique seems unsuitable for continuous input, as moving the whole device could be tiring and requires the user to refocus on the interface after each jerk. ForceTap [3] uses the accelerometer to determine a tap's force on the screen to enrich input, but does not solve problems such as interface occlusion and reaching distant targets. TapPrints [7] infers tap location on the front of the device by analysing accelerometer and gyroscope data, but does not show if and how this can be achieved on the device's back.
Sound as input: [6] and [2] use the different sounds of the finger nail, tip, pad or knuckle on the device's screen, extracted via a stethoscope. However, these gestures are largely unsuitable for one-handed input.
Back-of-device interaction: Wobbrock et al. [13] recommend complementing thumb input on the front of the device with index finger input on the back of the device. In [1] researchers use an additional touch pad on the back to improve bimanual input, whereas [12] shows how a button on the back of a phone can be used to show contextual information, but does not explore its capabilities for continuous input. The Unifone [4] uses additional, touch-sensitive hardware on the side of a phone to use squeezes of the device for input. Although a prototype, the system seems prone to inadvertent operation, especially if sensors were attached to both sides for ambidextrous users or if the users corrected their grip. TimeTilt[10] shows how the phone's accelerometer can be used to detect a tap on the back for switching a mode, whereas [9] uses a sound created by tapping the back of the phone to control voice services. Finally, [14] shows how taps on a tablet's corners can be detected using internal sensors. However, the researchers only use sound volume and device motion for classification, but not frequency analysis (FA) and do not explore the technique's applicability to one-handed interaction or finger differentiation.
Altogether, previous work uses either additional hardware, is not suitable for continuous input, does not solve the problem of interface occlusion or does not exhaust the sensors' potential. Thus, it remains unclear to what extent sound volume, sound profile and motion can be combined to enrich one-handed input without external hardware. Also, which finger is most suitable for a technique using these properties on the device's back or side, which applications can benefit, and can the technique be used for continuous input while addressing problems of interface occlusion and thumb mobility?

DESCRIPTION
To approach these questions, we present *BackPat* (BP): A technique for supporting one-handed input by using pats of either the index finger, middle finger or thumb on the back or the side of the device – not to *replace* but to *supplement* existing on-screen input via the thumb. This way, users can choose their preferred input method and use the BP gestures for more functionality or facilitation of hard-to-perform tasks. The gestures (Fig. 1) we refer to in this paper are:

- BP-index: Using the index finger to pat the upper part of the device's back.
- BP-middle: Using the middle finger to pat the middle outer part of the device's back.
- BP-thumb: Using the thumb to pat the device's side.

For activation, the user long-taps the screen with their thumb. Subsequent "patting" of the device's back with index or middle finger, or of the device's side with the thumb, will be interpreted as input by the system. To perform a patting gesture, users can either lift their thumb off the screen and then pat the back of the device, or leave it pressed down while patting. This is subject to user preference and app configuration. A

MobileHCI 2014, September 23–26, 2014, Toronto, ON, Canada.

ACM 978-1-4503-3004-6/14/09 ...$15.00.
http://dx.doi.org/10.1145/2628363.2628396

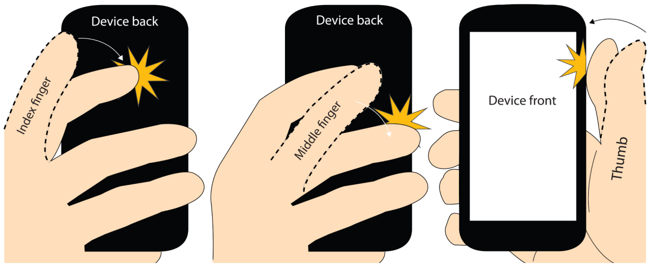

Figure 1. The 3 *BackPat* gestures: BP-index, BP-middle, BP-thumb.

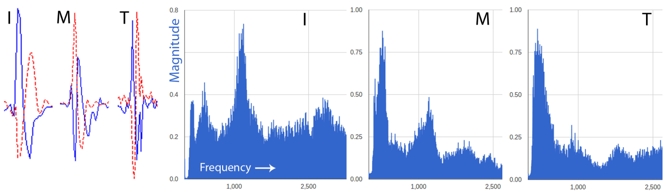

Figure 2. LEFT: The typical gyroscope patterns of BP-index (I), BP-middle (M) and BP-thumb (T) when holding the device with one hand. The full line represents the angular velocity around the x-axis, the dashed line the angular velocity around the y-axis. For BP-index and BP-middle the angular velocity around the z-axis is not noteworthy, but is very high for pats of the thumb on the side of the device, supporting correct pat detection of the otherwise rather similar patterns T and I. RIGHT: Averaged frequency (x-axis) magnitudes (y-axis) of each pat between 0–2500Hz. A pat of the thumb shows a characteristic profile between 0–1200 Hz, the middle finger between 0 and 2300 Hz – partially resembling the thumb – and the index finger between 400 and 2700 Hz.

demo of the technique [11] showed great user interest and learnability, but in this paper, we detail the gesture detection and evaluate each gesture's performance and user preference in three applications.

Finger-specific gesture detection

To configure our gesture detector, we recorded gyroscope and sound properties of pats made with the index and middle finger of each hand from six users (3 F, mean age 32, SD 3.74). The gyroscope data shows characteristic patterns for each pat (Fig. 2), allowing easy detection. However, gestures can be falsely detected by inadvertent movement. Thus we chose the sharp, brief changes in the volume of the microphone to act as gesture delimiters, which are detectable even while talking in close proximity to the phone. If the volume input rises above a certain threshold, we compare the volume values in a short window before and after the peak. If a pat (a high rise and fall in a very short time) is detected we analyse the gyroscope data to determine the finger used. This approach has two advantages. First, by analysing a short window around the volume peak, the pat can be easily distinguished from background noise, providing a relatively reliable delimiter. Second, the windowing delay allows enough time for the characteristic velocity changes to be interpreted. In a following study with the same users we adjusted the detection parameters for index and middle finger pats and added the gyroscope characteristics for pats of the thumb on the side of the device.

To further improve gesture detection, we examined the sound created by each pat. For this we recorded the patting sounds for all gestures from three users (1 F, mean age 30.3, SD 4.1). A Fast Fourier Transform of the data and visualisation of the magnitude of the frequencies gives a distinctive image

Pat sound	Speech	Music	Thumb	Index	Middle
Thumb	0.00 – 0.14	-0.10 – 0.20	0.58	0.37	0.44
Index	0.10 – 0.30	0.00 – 0.31	0.37	0.48	0.35
Middle	0.10 – 0.30	0.10 – 0.30	0.44	0.35	0.56

Table 1. PCC range for the pat sounds compared to a recorded parliamentary speech (Speech) and to a Jungle tune (Music). Measurements were taken twice per pat sound and sound source. PCC range is based on the rounded average of the lowest and highest PCC measured during 60s of playback. Table also shows mean PCC of pats by six users (2 F, mean age: 31.8, SD:3.7) compared to the pat data.

	Pk (L)	Pk (T)	G (L)	G (W)	FA (L)	FA (T)	All (L)
T	87%	87%	85%	75%	77%	35%	91%
I	98%	98%	83%	70%	78%	63%	73%
M	97%	85%	87%	68%	78%	73%	80%

Table 2. Percentage of correctly interpreted pats by six users (1 F, mean age: 33.2, SD:4.5, 10 pats/module) for each module (Gyro (G), Peak (Pk) and Frequency Analysis (FA), separately in % under lab conditions (L = sitting, low noise level), with recorded talking (T) at -0.6 to -0.3 db in the background, while walking (W), and all modules active (All) with equal weighting. Column All shows the importance of a tiered approach over an equally weighted one, as overall accuracy can be lowered.

for each pat (Fig. 2). We used the averaged data as a comparative basis for a detection algorithm which extracts three frequency ranges we defined to be representative of the pats (Fig. 2) from the microphone input and calculates the Pearson Correlation Coefficient (PCC) in relation to the comparative data. The highest value designates the "winner" and registers a pat of the respective finger when found to be above or equal to 0.5. Tab. 1 shows that a differentiation of index, middle finger and thumb patting gestures is possible due to the characteristic audio profiles created by the different angles and locations in which each finger connects with the device when patting. However, talking partly overwrites the frequencies of the pats and can lower the PCC down to 0.32, which is too low for reliable detection, as this is close to the PCC observed when comparing the pats to speech or music (Tab. 1).

As Tab. 1 suggests a minimum PCC of 0.38 for reliable detection, we decided to define the following thresholds for our accuracy test to improve detection under non-lab conditions: Thumb: PCC >= 0.45, index and middle finger: PCC >= 0.38. We changed the minimum PCC for the thumb to 0.45 as opposed to 0.38 to reduce false identification as a middle finger pat, which is similar (Fig. 2). Tab. 2 shows the average detection rates per pat and module in various conditions. Peak and gyroscope pattern detection are the most reliable modules, but the reliability of the latter is impacted by walking. Good frequency and peak detection is possible under lab conditions, but is reduced with background noise. While pat detection based on FA might be improved with frequency filtering, per-user calibration, a larger sample size or contact microphones as recommended by [6], we consider it not sufficiently reliable in a real life environment when used as the sole classifier. Therefore, we allocated a minor role to the FA: If a pat has occurred (peak) and a finger could not be reliably determined (gyroscope) the FA is used as a fallback method if the PCC is sufficiently high. This way, the FA can support detection especially when the gyroscope is impacted by walking (Tab. 2). In return, gyroscope analysis allows more reliable detection in a noisy environment than the FA (Tab. 2). If a peak has been incorrectly detected, either the gyroscope or FA pattern has to be characteristic of one of the three pats to reduce false positives. This tiered approach provides a relatively robust gesture detector which can be used with either right or left hand, without extra hardware and per-user calibration. However, the FA is likely to require calibration per device, as different casings may create different sounds. To gain an impression of the technique's performance, we conducted three user studies in a quiet office, using a PCC threshold of 0.5 for the FA module:

User study: Text selection

Text selection can be challenging when operating the phone with only one hand. We wanted to find out if BP could improve selection speed and which BP gesture is the most efficient and preferred by users. In particular, we compared the performance of the BP gestures with "normal mode" (i.e. moving the thumb left or right to drag the selection bracket).

A BP gesture extended or reduced the selection one word per "pat". The study was conducted using a text field 24 lines high, a font size of 27 pixels (px) and line height of 32 px, using the Roboto Regular font in black on a HTC Sensation XE with a resolution of 540 x 960 px at 256 PPI. The text field filled the whole screen, with six words per line (Fig. 3), each in the form of "xxxxx", surrounded by spaces. Tasks started in the horizontal and vertical centre of the field, where users had to select 0.5 lines, 1 line and 1.5 lines of text in either direction using each technique three times, recorded in milliseconds (ms). Selection of the correct words had to be maintained for 500ms to be considered successful. Errors, such as choosing the wrong start point or overshooting by one word, resulted in users restarting the test round. 20 users took part (5 F, mean age: 25.4, SD: 4.57, 18 right-handed, 2 left-handed). The study was counterbalanced by mode (normal/BP), task, and finger. We scanned for outliers using scatter plots and a rule of thumb looking for values significantly smaller or greater than three times the SD. Two participants were removed due to missing data, one was removed due to being unable to hold the phone, leaving 17 cases.

Results and Discussion

A Greenhouse-Geisser corrected ANOVA showed a main effect of amount of text to select $F(1.81, 28.95) = 16.00$, $p < .001$; a main effect of mode, $F(1.83, 29.24) = 6.19$, $p = .007$; and an interaction of mode and amount, $F(4.23, 67.66) = 10.78$, $p < .001$. A Wilcoxon test shows that when selecting 0.5 lines of text beginning in the centre and ending at the edge (Fig. 3), any BP method is faster than normal mode. As shown in Tab. 3, BP-index is the fastest technique – significantly faster than moving the thumb left ($Z = 3.62$, $p < .001$) or to the right ($Z = 2.68$, $p = .007$) in normal mode. BP-index is also significantly faster than BP-thumb, and BP-middle is faster than normal mode. However, selecting 1 line of text that starts in the middle of a line and ends in the middle of the following line (Fig. 3) is fastest when moving the thumb in normal mode. While faster than the normal mode of operation in some cases, for longer selection tasks BP-thumb is less suitable as it requires change in the grip of the phone.

Selecting 1.5 lines of text (Fig. 3), with the start in the screen centre and the end at the screen edge, is fastest using the thumb in normal selection mode (Tab. 3). The fastest BP technique is BP-index, which is faster than BP-thumb ($Z = 2.81$, $p = .005$. Bonferroni-Holm correction applied to each test, starting with a divider of 10). For larger amounts of text, normal mode outperforms BP as it allows quick jumping between lines of text which otherwise would have to be "patted" down word by word. Thus, we recommend complementing normal text selection via the thumb with BP: Users can cover large areas of text by moving their thumb over the display and fine-adjust their selection with a few pats. The better performance of BP-index corresponds with the results of the user feedback who judged BP-index to be easier to perform than BP-middle and BP-middle to be easier than BP-thumb.

M	0.5	*SD*	1	*SD*	1.5	*SD*
BP-T	3535	*818*	5395	*1643*	5702	*2448*
BP-I	2669	*366*	3521	*583*	4642	*642*
BP-M	2971	*649*	3775	*710*	4495	*1400*
N-R	3663	*1384*	3406	*1575*	3273	*1048*
N-L	4059	*2548*	3622	*1479*	3474	*1894*

Table 3. Rounded median task times (left column) and *SD (right column)* of the text selection user study for each mode (M: BP and normal left and right (N-L, N-R)) in ms for 0.5, 1 and 1.5 lines.

Figure 3. LEFT: The text selection tasks marked in red. To allow beginning the selection tasks from the middle of the line, three words were broken down into "xxx" and an "i" was appended to one. MIDDLE: List selection task where the user has to select the grey list elements starting at item 5. RIGHT: The target positions of the Pat-into-place study.

User study: Multiple selection

Another application of BP is multiple selection of list items. Users create an initial selection using their thumb, and subsequent BP-index or BP-thumb gestures extend the selection upwards. Using a BP-middle gesture will shrink the selection or extend it downwards. This configuration was deemed logical by users. We asked 24 users (6 F, 21 right-handed, 2 left-handed, 1 ambidextrous) to select either three, six or eleven consecutive list items (Fig. 3) three times. Selection started mid-list and mid-screen and had to be performed up and down, with task completion time recorded in ms. This way, we compared the performance of BP-index, BP-middle, BP-thumb and moving the thumb up and down using direct tap (normal mode). The study was counterbalanced by mode (BP/normal) and task. The data had no outliers.

Results and discussion

An ANOVA showed a main effect of amount, $F(1.28, 29.52) = 353.21$, $p < .001$; a main effect of mode, $F(1.96, 45.18) = 8.10$, $p = .001$; and an interaction of amount and mode, $F(4.54, 104.51) = 10.40$, $p < .001$ (G-G correction applied).

Here, BP shows potential for reducing selection time for six items or more. A Wilcoxon test (Bonferroni-Holm correction

M	3	*SD*	6	*SD*	11	*SD*
BP-T	1236	*646*	2436	*1466*	3680	*1975*
BP-I	1073	*548*	1837	*534*	3071	*991*
BP-M	1728	*2071*	2764	*1732*	4352	*3147*
N-U	939	*482*	2544	*1276*	5704	*1139*
N-D	920	*393*	2439	*815*	5660	*1520*

Table 4. Rounded median task times (left column) and *SD (right column)* for each mode (M: BP and normal up and down (N-U, N-D)) in ms for selecting 3, 6 and 11 items in a list.

applied starting with a divider of 10) showed that when selecting 11 items in a list, BP-index is the fastest approach, being faster than BP-thumb (Z = 3.11, p = .002), BP-middle (Z = 3.51, $p < .001$) and normal mode (Z = 3.97, $p < .001$). BP-thumb is also faster than normal mode. When selecting six items, BP-index is again the fastest technique (Tab. 4), being significantly faster than using the thumb downwards in normal mode (Z = 3.34, p = .001); faster than using the thumb upwards in normal mode (Z = 3.26, p = .001); faster than using BP-middle (Z = 4.09, $p < .001$); and faster than using BP-thumb (Z = 3.26, p = .001). BP-thumb is also faster than moving the thumb up in normal mode (Z = 3.97, $p < .001$). When selecting three items (Tab. 4), normal mode outperforms BP mode. The fastest BP method is BP-index, which is significantly faster than BP-middle, Z = 3.26, p = .001). BP-middle is the slowest, being significantly slower than BP-thumb (Z = 3.20, p = .001). For selecting only three items, direct selection seems fastest due to the cost of grip adjustment.

We recommend using BP as a complementary method: Small selections should be performed using direct tap, whereas larger selections can benefit greatly from using the BP technique. With BP-middle being the slowest BP technique, we recommend BP-index for extending a selection upwards and BP-thumb for extending a selection downwards. This correlates with the evaluation of the user feedback, who preferred BP-index over BP-thumb and BP-thumb over BP-middle.

User study: Reaching distant targets

In a third study we examined BP's applicability to facilitating interaction with distant targets, which we termed *Pat-into-place* (PIP). Users can perform a BP gesture to move targets into the thumb's reach by touching the screen and subsequently patting the device. This will move targets at the top or bottom of the screen to the level of the thumb, using either BP-index or BP-middle (Fig. 3). While selecting targets directly was faster than using PIP, users felt that PIP made reaching the elements faster and easier than direct access.

CONCLUSION AND FUTURE WORK

This paper has contributed the following over previous work:

- Synthesis of audio profiles, volume changes and gyroscope data can be used to produce three novel off-screen patting gestures to support one-handed input, allowing reaching distant targets as well as continuous input without interface occlusion and extra hardware or per-user calibration for both hands. Using these properties in a tiered approach, gestures can be detected even if one component fails.
- The quantitative performance of thumb, index and middle finger for performing these off-screen patting gestures and user preference for each. This varies by application, but users generally prefer index finger gestures.

This paper shows that *BackPat* can help to improve one-handed interaction, either as a means of direct input or as a "facilitator". However, future work will involve extensive comparison to existing techniques to better judge its impact.

REFERENCES

1. Baudisch, P., and Chu, G. Back-of-device interaction allows creating very small touch devices. In *Proc. CHI'09*, ACM (2009), 1923–1932.
2. Harrison, C., Schwarz, J., and Hudson, S. E. TapSense: enhancing finger interaction on touch surfaces. In *Proc. UIST'11*, ACM (2011), 627–636.
3. Heo, S., and Lee, G. Forcetap: Extending the input vocabulary of mobile touch screens by adding tap gestures. In *Proc. MobileHCI'11*, ACM (2011), 113–122.
4. Holman, D., Hollatz, A., Banerjee, A., and Vertegaal, R. Unifone: designing for auxiliary finger input in one-handed mobile interactions. In *Proc. TEI'13*, ACM (2013), 177–184.
5. Karlson, A. K., and Bederson, B. B. Studies in one-handed mobile design: Habit, desire and agility. Tech. rep., Comp. Sci. Dept., Univ. of Maryland, 2006.
6. Lopes, P., Jota, R., and Jorge, J. A. Augmenting touch interaction through acoustic sensing. In *Proc. ITS'11*, ACM (2011), 53–56.
7. Miluzzo, E., Varshavsky, A., Balakrishnan, S., and Choudhury, R. R. Tapprints: Your finger taps have fingerprints. In *Proc. MobiSys'12*, ACM (2012), 323–336.
8. Oakley, I., and Park, J. A motion-based marking menu system. In *CHI'07 E.A.*, ACM (2007), 2597–2602.
9. Robinson, S., Rajput, N., Jones, M., Jain, A., Sahay, S., and Nanavati, A. TapBack: towards richer mobile interfaces in impoverished contexts. In *Proc. CHI'11*, ACM (2011), 2733–2736.
10. Roudaut, A., Baglioni, M., and Lecolinet, E. TimeTilt: using sensor-based gestures to travel through multiple applications on a mobile device. In *INTERACT 2009*, vol. 5726 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.
11. Seipp, K., and Devlin, K. Backpat: Improving one-handed touchscreen operation by patting the back of the device. In *CHI'14 E.A.*, ACM (2014), 555–558.
12. Stienstra, J., Overbeeke, K., and Wensveen, S. There is more in a single touch. In *Proc. CHItaly*, ACM (2011), 27–32.
13. Wobbrock, J. O., Myers, B. A., and Aung, H. H. The performance of hand postures in front- and back-of-device interaction for mobile computing. *Int. J. Hum.-Comput. Stud. 66*, 12 (Dec. 2008), 857–875.
14. Zhang, C., Parnami, A., Southern, C., Thomaz, E., Reyes, G., Arriaga, R., and Abowd, G. D. BackTap: robust four-point tapping on the back of an off-the-shelf smartphone. In *Proc. UIST'13*, ACM (2013), 111–112.