Foot Interaction Concepts to Support Radiological Interventions

Benjamin Hatscher¹, Maria Luz¹, Christian Hansen¹

Computer-Assisted Surgery Group, University of Magdeburg, Germany¹

Abstract

During neuroradiological interventions, physicians need to interact with medical image data, which cannot be done while the hands are occupied. We propose foot input concepts with one degree of freedom, which matches a common interaction task in the operating room. We conducted a study to compare our concepts in regards to task completion time, subjective workload and user experience. Relative input performed significantly better than absolute or rate-based input. Our findings may enable more effective computer interactions in the operating room and similar domains where the hands are not available.

1 Introduction

Minimally-invasive radiological intervention is a growing field in which a variety of diseases can be treated without open surgery. During such interventions, the physician in the operating room (OR) relies on an interventional angiography system as imaging modality to navigate a catheter or needle tip to the desired pathological structure inside the patient's body (Odisio and Wallace, 2014). The acquired images are not only used for navigation, but also stored and employed for later reference during the intervention. Furthermore, preoperative datasets from other imaging modalities, e.g., computer tomography (CT) or magnetic resonance imaging (MRI), can be accessed. This requires the physician to interact with interventional imaging software while operating. In clinical routine, navigating medical images is a cumbersome task due to the provided input methods. A common solution lies in using plastic-sheathed controls like joysticks and buttons. Surgeons therefore need to change their position when these devices are out of reach (Hübler et al., 2014). Alternatively, interaction tasks are delegated to a medical assistant located inside the OR or outside in a non-sterile control room. Verbal or gestural communication is error-prone, can lead to time-consuming misunderstandings and may even require physicians to scrub out and in again to interact with non-sterile workstations (Grange et al., 2004; O'Hara et al., 2014). To this end, intraoperative touchless interaction has been subject of research for many years. A literature review of touchless interaction in the OR provides an overview which shows a focus on optical gesture recognition (Mewes et al., 2017). Limitations in the OR such as the close proximity of other team members at the operating ta-

Veröffentlicht durch die Gesellschaft für Informatik e. V. 2017 in M. Burghardt, R. Wimmer, C. Wolff, C. Womser-Hacker (Hrsg.): Mensch und Computer 2017 – Tagungsband, 10.–13. September 2017, Regensburg. Copyright (C) 2017 bei den Autoren. https://doi.org/10.18420/muc2017-mci-0174 ble, restricted movement of the upper limbs due to sterility considerations and partial occlusion of the surgeon's body can interrupt line-of-sight for optical gesture recognition systems. Furthermore, the problem of touchless control when both hands are occupied by holding medical instruments has to be considered (O'Hara et al., 2014).

To tackle these problems, we investigated methods to interact with medical image data with the feet while keeping the hands free for other tasks. Foot pedals are already established in clinical routine, but suffer from the lack of visual control and the need to stand on one foot (Wauben et al., 2006). Therefore, our approach focuses on heel rotation since literature suggests it is the less exhausting method of foot interaction. Three concepts which apply different input techniques are presented and combined with adequate visual feedback. To evaluate our approach, we present a quantitative user study to compare these concepts with regards to task completion time, subjective workload and user experience. We contribute to human computer interaction (HCI) by determine adequate foot input concepts when standing. This work may enable more efficient interaction in the OR as well as other domains such as industrial manufacturing where hands-free interaction is required.

2 Related Work

When using foot interaction in confined spaces and in a standing position, gestures are limited and therefore need an adequate input device and input technique, matching the requirements of the task (Jacob, 1997). Basic ideas of foot input prototypes for the OR have been outlined but not evaluated so far (Fitzke et al., 2015).

For seated positions, participants preferred dragging the foot over lifting it. The most comfortable interaction method was short, in-place movement such as pivoting around the heel (Simeone et al., 2014; Velloso, Alexander, et al., 2015). Aligning with these findings, participants preferred heel rotation over dorsiflexion, plantar flexion and toe rotation. Additionally, external rotation of the foot was preferred over internal rotation (Scott et al., 2010). Foot movement in general quickly becomes exhausting when a posture has to be held. Resting positions therefore are suggested for seated as well as for standing foot interaction (Simeone et al., 2014; Alexander et al., 2012). For activation or deactivation of a function, toe lifting is considered suitable (Simeone et al., 2014) and preferable over sliding gestures (Jalaliniya et al., 2013).

Another important factor is the input technique. Techniques to use foot postures for continuous input can be divided into discrete rate-based, absolute and relative approaches. Rate-based input continuously modifies a parameter as long as a certain condition is met, like joysticks or a car's accelerator pedal. Custom foot pedals deliver feedback and can be used for 1D and 2D-tasks (Klamka et al., 2015) but may be triggered involuntarily when the foot is at rest (Velloso, Schmidt, et al., 2015). Pressure distribution performed well for rate-based interaction in a non-mobile setup (Sangsuriyachot and Sugimoto, 2012) but needed to be trained to the user when investigated for daily situations (Fukahori et al., 2015). Kicks were found advantageous over holding a foot posture, since the foot could be rested between subtasks (Alexander et al., 2012). Absolute input maps foot movement to one or more parameters in a direct fashion. Applied to 1D and 2D tasks, the difficulty in reaching small targets was reported as the biggest challenge

when using foot movements under the desk (Velloso, Alexander, et al., 2015). This aligns with the reported median selection error of 11.77° for dorsiflexion, plantar flexion, heel- and toe rotation (Scott et al., 2010). Relative input allows repositioning of the feet, similar to using the mouse wheel or scrolling on a touch screen. Large trackballs were found sufficient for non-accurate spatial tasks (Pakkanen and Raisamo, 2004). Foot mice (Balakrishnan et al., 1999; Fitzke et al., 2015) and relative mapping under the desk was described but not yet evaluated (Simeone et al., 2014).

3 Design Considerations

In this work, we present methods to interact with medical images. Based on the advantages reported by literature, we focused on heel rotation and ball lifting/tapping as input method. Unlike most of the research conducted on foot interaction so far, our approach is developed with a specific scenario in mind. Therefore, our methodology is to describe the task to be fulfilled, identify restrictions and develop suitable foot interaction techniques.

3.1 Intra-operative Interaction Task

One of the most common intra-operative interaction tasks is scrolling back and forth through a stack of images. From such a data set, a specific image might be selected as an overlay to facilitate further interventional steps (Hettig et al., 2017). In a similar manner, preoperative volume data from imaging modalities such as CT or MRI is provided as a stack of slices. Therefore, interacting with an image stack covers a wide range of situations and requires only one degree of freedom (DoF). Depending on the kind of data set, the number of images may differ significantly, which requires a method to allow for an arbitrary number of slices to be navigated.

3.2 Restrictions and Requirements

During radiological interventions, the physician has to operate different instruments and devices and therefore needs both hands most of the time. The feet, especially in a standing position, are used constantly to maintain a stable stance. Lifting a foot when using a foot pedal is possible but uncomfortable, since it disturbs the balance. Furthermore, there is no visual control over the pedal; it may get lost and is hard to find under the table (Wauben et al., 2006). Additionally, during radiological interventions heavy lead aprons have to be worn to shield against radiation. The weight makes it even harder to maintain unstable positions. Requirements for interaction methods in the OR can be summarized as follows:

- The space to carry out the task needs to be minimal.
- Interference with a stable stance must be kept low.
- · Feedback has to be provided without looking directly at the feet.
- The input modality always needs to be via the foot.

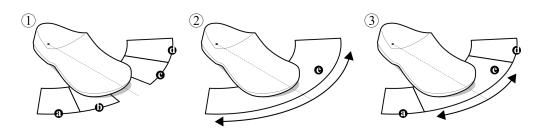


Figure 1: Foot scrolling concepts for manipulating one DoF via heel rotation. Discrete buttons (1) increment (a,b) or decrement (c,d) the current image as long as being stepped on. The outmost buttons (a,d) change the position by one at a rate of 0.2 seconds, the innermost (b,c) at 0.8 seconds. Foot scrolling (2) allows for continuous change of the current image every 10° by rotating the foot (e) while the foot tip is on the floor. By lifting, the foot can be repositioned without changing the current image. Step and scroll (3) combines the discrete buttons for fast rate-based input (a,d) and an area in between for foot scrolling (e).

4 Foot Interaction Concepts

In the following, we present three interaction approaches for one DoF where only rotation around the heel and lifting the tip of the foot are utilized. All concepts are designed in such a way that the foot never has to be held tip-up for long amounts of time. To establish consistent behavior, the mapping between the medical image stack and foot rotation is always realized in a fashion that scrolling back (decrementing) the image stack involves rotation towards the left (i.e., turning the right foot in) and scrolling forward (incrementing) is towards the right (i.e., turning the right foot out). Lifting the foot disables all virtual foot controls, but the visual representation of the foot position follows the foot, similar to hovering with a computer mouse. For simplicity, we focused on the right foot, but all the concepts could be mirrored.

4.1 Discrete Buttons

Corresponding to the finding of (Zhong et al., 2011), the *discrete buttons* concept (DB) focuses on buttons (Fig. 1.1) which are easily selectable given the reported foot angle selection error and range by (Scott et al., 2010). Each button occupies 20°, which allows five buttons to fit in an area of -40° to 60° in front of the foot. Interaction is realized with a rate-based approach. To allow scrolling through image stacks of arbitrary length, two buttons for different rates (0.8 and 0.2 seconds) are provided for each direction. According to the principles listed above, the buttons left of the center decrement the currently shown image, and the ones on the right increment it. The center is left without function, to provide a position to put the foot at rest.

4.2 Foot Scrolling

To account for the arbitrary size of medical image data sets, *foot scrolling* (FS) connects to the relative input technique. By rotating the foot inwards or outwards, the current position inside the image stack is incremented or decremented every 10°, depending on the direction of movement (Fig. 1.2). At the end of the interactive area, the tip of the foot has to be lifted, rotated in the

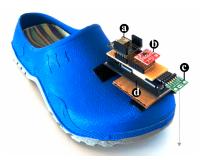


Figure 2: Hardware prototype mounted on an OR shoe with Velcro fastener consisting of a microcontroller with Bluetooth stack (a), gyroscope (b), distance sensor (c) and power supply (d).

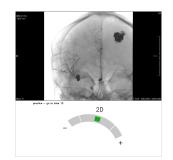


Figure 3: Graphical user interface showing medical image data on top and visual feedback for foot interaction at the bottom. A green cursor indicates the foot's position and shrinks to 60% of its size when lifting the foot.

opposite direction and placed back on the floor. In contrast to the first concept, there is no dedicated area to rest the foot, since it can be rested anywhere without triggering a function.

4.3 Step and Scroll

A combination of both former concepts called *step and scroll* (S) integrates rate-based and relative input. It consist of two buttons at the outermost positions of the interactive area for fast, rate-based scrolling described in the first concept (Fig. 1.3). The area in between consists of a scrollable area similar to the second concept. It is possible to switch seamlessly between both kinds of control elements, where the function activated is determined by the exact position of the cursor's center.

5 Prototype

According to the interaction concepts, the prototypical interface measures two parameters: the foot orientation, and whether the tip of the foot is lifted or lowered. The setup consists of offthe-shelf sensors which are connected to a microcontroller via I²C-Bus (see Fig. 2). The device is mounted on an OR-Shoe. Foot orientation is gathered by reading the Z-axis of a three-axis gyroscope (MPU-9250, InvenSense, San Jose, CA, USA). A downward-facing time-of-flight distance sensor (VL6180X, STMicroelectronics, Geneva, Switzerland) is utilized to determine the height of the foot tip in relation to the ground. Orientation and tip-to-floor distance are read by a microcontroller with an integrated Bluetooth low energy stack (RFD22102, RFduino Inc., Hermosa Beach, CA, USA) and sent to a computer. To account for different floor colors which influence the distance sensor, calibration of the system is done by setting the value read at a resting position as zero. For robust tap detection, we use a bi-level threshold with an upper threshold of 10 mm and a lower threshold of 5 mm. Orientation is computed by adding up the gyroscope readings. For simple sensor offset correction, the average value at a resting position over one minute was gathered and is subtracted from each gyroscope reading.

Dependent variables	s df	F	t	р	sig	η^2_{part}	d	Effect
Subjective workload	1 2, 18	5.93		0.01	*	0.4		large
FS vs. S	9		3.67	0.02	*		1.16	large
DB vs. S	9		1.59	0.44	n.s.		0.5	medium
FS vs. DB	9		1.88	0.28	n.s.		0.6	medium
User experience								
Overall	1,49, 13,41	6.37		0.02	*	0.42		large
FS vs. S	9		4.99	< 0.01	*		1.58	large
DB v. S	9		2.48	0.11	n.s.		1.11	large
FS vs. DB	9		0.76	1.00	n.s.		0.24	small
Task completion tim	ne 2, 18	12.25		< 0.01	*	0.58		large
FS vs. S	9		5.34	< 0.01	*		1.69	large
DB vs. S	9		1.70	0.37	n.s.		0.54	medium
FS vs. DB	9		3.06	0.04	*		0.97	large

Table 1: Summary of the test statistics for subjective workload, user experience and task completion time for concepts discrete buttons (DB), foot scrolling (FS) and step and scroll (S).

6 Evaluation

We conducted a study to compare the performance of the concepts described in section 4 with respect to task completion times, subjective workload and user experience.

Ten right-footed participants (all male) between 25 and 30 years (M = 26.2, SD = 1.8) recruited from our university took part in the study. Prior experience with foot interaction was stated as high by one participant and medium by two participants (on a 5-point Likert scale from 1 = no experience to 5 = very experienced). The remaining participants stated no experience with foot interaction (M = 1.8, SD = 1.4). Shoe size varied from 42 to 49 in EU size (M = 44.1, SD = 2.3).

The study took place in a computer laboratory. The sensor described in section 5 was mounted on a pair of OR-Shoes with Velcro fastener. A patch of linoleum flooring 170 cm x 105 cm in size was mounted on the floor to create a similar friction between shoes and floor as in the OR. In front of the flooring, a 40" monitor with 3840 x 2160 pixel resolution was placed at a height of 140 cm (monitor center), which resembles the Large Display of an Siemens Artis angiography suite (56", same resolution). The distance between participant and screen was approximately 100 cm but was not restricted to a specific distance.

To assess task completion times, an audio signal indicated when to start a task. The participant conveyed verbally when the target slice was reached. Task completion time was measured between the start sound and the last sensor data that indicated foot movement before the participant signaled task completion. To assess the subjective workload we used the NASA-Task Load Index questionnaire without the weighting process, commonly referred to as Raw TLX (RTX) (Hart, 2006). For user experience, the items usefulness, usability, positive and negative

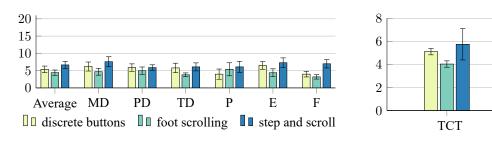


Figure 4: Mean results for the RTX dimensions Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (P), Effort (E) and Frustration (F) with standard error bars. (0 = low/good, 20 = high/poor).

emotions, intention of use and overall rating from the meCUE questionnaire were employed (Minge et al., 2017).

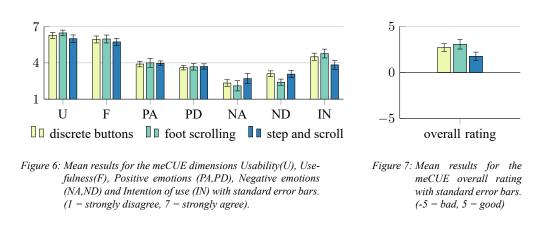
During the study, a demographic questionnaire extended with questions about shoe size and experience with foot interaction was filled out first. After that, the participant selected the bestmatching pair of shoes out of three pairs (size 41/42, 43/44, 45/46), which was then equipped with the sensor. Each participant completed a fixed series of tasks with each concept (withinsubject design). The task sequence was identical for all concepts. The order of concepts was balanced over all participants. For data collection, the first interaction concept was introduced and explained by the instructor. A practice phase consisting of five tasks was completed and followed by ten measured tasks. Participants were asked to navigate to a specific image in a stack of medical images shown on a monitor in front of them. Before each task, they were instructed to set their foot on a central position, the system was reset to a center position and the image at position 9 in the stack was set as current image. The current image number and target image number were shown on the screen all the time. During all tasks, the same data set of radiological images ranging from 0 to 18 was used. After finishing practice and measurement phase, questionnaires for subjective data were filled out by the participant. The data collection procedure was repeated for the remaining two concepts.

The data was analyzed by one-way ANOVA for repeated measures with three levels representing the three interaction concepts described in section 4. If the ANOVA revealed significant results, Bonferroni corrected post-hoc *t*-tests were performed to determine which concepts exactly differ. In addition to the average workload, RTX subscales were analyzed and reported individually, which is a common evaluation variation to pinpoint performance problems (Hart, 2006).

7 Results

The statistical results of the ANOVA and post-hoc tests are shown in Table 1. The ANOVA revealed significant result for task completion time. As seen from the Fig. 5 the participants

Figure 5: Mean results for the task completion times (TCT) in seconds with standard error bars.



could select the required images fastest with concept FS (section 4.2) (M = 4.03s, SD = 0.88). They needed much longer with concept DB (section 4.1) (M = 5.11s, SD = 0.88), followed by concept S (section 4.3) (M = 5.75s, SD = 1.13). According to post-hoc tests, FS was significantly superior to both, DB and S.

The results for the overall subjective workload as well as for the six RTX dimensions are presented in Fig. 4. The subjects perceived the lowest overall workload working with concept FS (M = 4.42, SD = 2.47), followed by concept DB (M = 5.40, SD = 2.96). The workload was highest by selecting 2D images with concept S (M = 6.67, SD = 3.28). According to one-way ANOVA, this difference was statistically significant. The subsequent post-hoc tests revealed that this is due to the significant difference between FS and S.

The results for the selected meCue dimensions are presented in Fig. 6. The overall rating leads to significant ANOVA results which are presented in Fig. 7. All three concepts were rated positively. However, FS (M = 3.05, SD = 1.62) was rated slightly better than DB (M = 2.70, SD = 1.34) and much better than S (M = 1.75, SD = 1.44). The post-hoc tests indicated significant difference only for FS and S.

8 Discussion

Considering basic restrictions in the OR, we worked out different input concepts, applied them to heel rotation as the most promising input method, realized a technical prototype and evaluated the concepts in a user study. Our evaluation revealed significant excelling results for FS regarding task completion times, subjective workload and user experience. This is a surprising result, since relative input requires repeated movements, which were believed to be exhausting (Simeone et al., 2014) and observed as slower in relation to rate-based approaches, especially when standing (Alexander et al., 2012). The results might be explained by experience due to the similarity to slide gestures on touch screens, in contrast to the unfamiliar setup of multiple foot pedals for fixed rates, possibly in combination with the clinically common but relatively low number of images. Concept S provoked involuntary button activation when the end of the foot scrolling area was reached, which led to worse results compared to other concepts. However, the evaluation focused on tasks with a single DoF which were not integrated in a clinical workflow. Next steps require the comparison with other input methods such as hand gestures as well as clinically established systems. Furthermore, foot interaction during delicate surgical tasks has to be investigated. On the long term, the time span of a real intervention, walking of the user, handing over control and integration of physical foot pedals have to be considered.

9 Conclusion

In this work, we investigated interaction concepts for heel rotation to navigate medical images. A study was conducted to gather first insights about the concepts basic suitability for the task of medical image navigation with one DoF. Foot scrolling seems superior in terms of task completion time as well as subjective ratings, and therefore demands further investigation and improvement to match realistic, more complex interaction tasks.

Acknowledgement

This work is partially funded by the Federal Ministry of Education and Research (BMBF) within the STIMULATE research campus (grant number 13GW0095A).

References

- Alexander, J., Han, T., Judd, W., Irani, P., & Subramanian, S. (2012). Putting your best foot forward: investigating real-world mappings for foot-based gestures. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems* (pp. 1229–1238).
- Balakrishnan, R., Fitzmaurice, G., Kurtenbach, G., & Singh, K. (1999). Exploring interactive curve and surface manipulation using a bend and twist sensitive input strip. In *Proc. of the 1999 symposium on Interactive 3D graphics* (pp. 111–118).
- Fitzke, T., Krail, N., Kroll, F., Ohlrogge, L., Schröder, F., Spillner, L., ... Malaka, R. (2015). Fußbasierte Interaktion mit Computersystemen im Operationssaal. In H. K. Hahn, R. Kikinis, J. Klein, A. Nabavi, & S. Weber (Eds.), CURAC 2015 - Tagungsband (pp. 49– 54). Bremen: Hahn, Horst K. et al.
- Fukahori, K., Sakamoto, D., & Igarashi, T. (2015). Exploring Subtle Foot Plantar-based Gestures with Sock-placed Pressure Sensors. In Proc. of the 33rd An. ACM Conf. on Human Factors in Computing Systems (pp. 3019–3028). CHI '15. New York, NY, USA: ACM. doi:10.1145/2702123.2702308
- Grange, S., Fong, T., & Baur, C. (2004). M/ORIS: a medical/operating room interaction system. In Proc. of the 6th Int. Conf. on Multimodal interfaces (pp. 159–166).
- Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In Proc. of the Human Factors and Ergonomics Society An. Meeting (Vol. 50, pp. 904–908).

- Hettig, J., Saalfeld, P., Luz, M., Becker, M., Skalej, M., & Hansen, C. (2017). Comparison of gesture and conventional interaction techniques for interventional neuroradiology. *Int. Journal of Computer Assisted Radiology and Surgery*. doi:10.1007/s11548-017-1523-7
- Hübler, A., Hansen, C., Beuing, O., Skalej, M., & Preim, B. (2014). Workflow Analysis for Interventional Neuroradiology using Frequent Pattern Mining. In *Proc. of the An. Meeting* of the German Society of Computer- and Robot-Assisted Surgery (pp. 165–168). Munich. Jacob, R. J. K. (1997). Input Devices and Techniques.
- Jalaliniya, S., Smith, J., Sousa, M., Büthe, L., & Pederson, T. (2013). Touch-less interaction with medical images using hand & foot gestures. In Proc. of the 2013 ACM Conf. on Pervasive and ubiquitous computing adjunct publication (pp. 1265–1274).
- Klamka, K., Siegel, A., Vogt, S., Göbel, F., Stellmach, S., & Dachselt, R. (2015). Look & Pedal: Hands-free Navigation in Zoomable Information Spaces through Gaze-supported Foot Input. In Proc. of the 2015 ACM on Int. Conf. on Multimodal Interaction (pp. 123–130).
- Mewes, A., Hensen, B., Wacker, F., & Hansen, C. (2017). Touchless interaction with software in interventional radiology and surgery: a systematic literature review. *Int. Journal of Computer Assisted Radiology and Surgery*, 12(2), 291–305. doi:10.1007/s11548-016-1480-6
- Minge, M., Thüring, M., Wagner, I., & Kuhr, C. V. (2017). The mcCUE Questionnaire: A Modular Tool for Measuring User Experience. In Advances in Ergonomics Modeling, Usability & Special Populations (pp. 115–128). Springer.
- Odisio, B. C. & Wallace, M. J. (2014). Image-guided interventions in oncology. Surgical oncology clinics of North America, 23(4), 937–955.
- O'Hara, K., Gonzalez, G., Sellen, A., Penney, G., Varnavas, A., Mentis, H., ..., Dastur, N. et al. (2014). Touchless interaction in surgery. *Communications of the ACM*, 57(1), 70–77.
- Pakkanen, T. & Raisamo, R. (2004). Appropriateness of foot interaction for non-accurate spatial tasks. In CHI'04 ext. abstracts on Human factors in computing systems (pp. 1123–1126).
- Sangsuriyachot, N. & Sugimoto, M. (2012). Novel interaction techniques based on a combination of hand and foot gestures in tabletop environments. In *Proc. of the 10th asia pacific Conf. on Computer human interaction* (pp. 21–28).
- Scott, J., Dearman, D., Yatani, K., & Truong, K. N. (2010). Sensing foot gestures from the pocket. In Proc. of the 23nd annual ACM symposium on User interface software and technology (pp. 199–208).
- Simeone, A. L., Velloso, E., Alexander, J., & Gellersen, H. (2014). Feet movement in desktop 3D interaction. In 2014 IEEE Symposium on 3D User Interfaces (3DUI) (pp. 71–74).
- Velloso, E., Alexander, J., Bulling, A., & Gellersen, H. (2015). Interactions Under the Desk: A Characterisation of Foot Movements for Input in a Seated Position. In *Human-Computer Interaction – INTERACT 2015* (Vol. 9296, pp. 384–401). Lecture Notes in Computer Science. Cham: Springer Int. Publishing. doi:10.1007/978-3-319-22701-6
- Velloso, E., Schmidt, D., Alexander, J., Gellersen, H., & Bulling, A. (2015). The Feet in Human– Computer Interaction: A Survey of Foot-Based Interaction. ACM Computing Surveys (CSUR), 48(2), 21.
- Wauben, L., van Veelen, M. A., Gossot, D., & Goossens, R. H. (2006). Application of ergonomic guidelines during minimally invasive surgery: a questionnaire survey of 284 surgeons. *Surgical Endoscopy And Other Interventional Techniques*, 20(8), 1268–1274.

Zhong, K., Tian, F., & Wang, H. (2011). Foot menu: Using heel rotation information for menu selection. In 2011 15th An. Int. Symposium on Wearable Computers (ISWC) (pp. 115–116).