SCIENCE CHINA
Information Sciences

• MOOP •

June 2018, Vol. 61 064101:1–064101:3
https://doi.org/10.1007/s11432-018-9385-7

MagicMark: a marking menu using 2D direction and 3D depth information

Fei LYU[1,2*], Rui XI[1], Yuxin HAN[1] & Yujie LIU[1]

[1]*School of Digital Media and Design Arts, Beijing University of Posts and Telecommunications, Beijing* 100876*, China;*
[2]*Beijing Key Laboratory of Network Systems and Network Culture, Beijing University of Posts and Telecommunications, Beijing* 100876*, China*

Received 31 January 2018/Accepted 13 March 2018/Published online 8 May 2018

Citation Lyu F, Xi R, Han Y X, et al. MagicMark: a marking menu using 2D direction and 3D depth information. Sci China Inf Sci, 2018, 61(6): 064101, https://doi.org/10.1007/s11432-018-9385-7

Menus are important interactive components of user interfaces, broadly applied to command exploration and selection [1]. For large displays, RapMenu and touchless circular menus explored the design space of tilt and 2D direction [2, 3]. However, as one of the important input modalities, 3D depth information has not been fully utilized for menu control in large displays.

Among a wide variety of menu techniques, marking menu [4] provides two menu selection mode: novice mode to perform command selection with a visible menu, and expert mode to activate a command with a straight mark. Different from traditional stroke shortcuts, marking menu can support the seamless transition from a novice user to an expert user [1].

To extend the selection capability of large screen interactions, we propose MagicMark, a hierarchical menu combining 2D direction and 3D depth information. With 3D depth inputs, MagicMark requires less display space than the other hierarchical menus. MagicMark is also a marking menu that includes both novice mode and expert mode. Figure 1(a) illustrates the novice mode in which a user selects the main menu with 3D depth information, and then selects the submenu item with 2D direction information. Figure 1(b) shows the expert mode in which users make 3D "zig-zag" compound marks to activate the menu item.

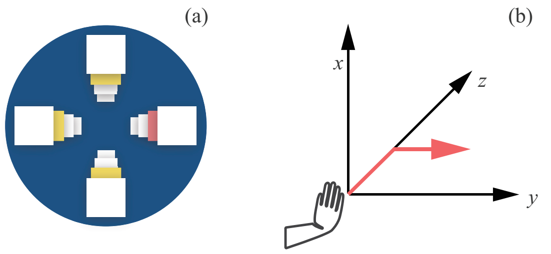

Figure 1 (Color online) MagicMark. (a) Novice mode; (b) expert mode.

MagicMark technique. MagicMark is a pop-up menu and allows users to make menu selection through two interactive techniques: depth-based main menu selection and direction-based submenu selection. Figure 1 shows MagicMark and its gestures in a 3D Cartesian coordinate. The origin of the coordinates is the position of a user's dominant hand when the menu is activated. The z-axis represents the depth and its value corresponds to the arm stretching distance from the original position. This distance controls the main menu selection. The waving gestures on the x-y plane can be used to select the submenu items according to the gesture direction. The gesture recognition is based on Kinect's skeletal tracking technique. We collect the coordinates of the required skeleton points (hand and shoulder center) in the Kinect skeletal space for gesture recognition.

* Corresponding author (email: feilyu@bupt.edu.cn)

• Depth-based gesture recognition. The depth-based interaction is calculated by the following algorithm. First, a user needs to straighten the dominant arm as far as possible to complete the calibration. As the user fully stretches the arm, the arm length and the calibrated menu depth of MagicMark are calculated according to

$$\begin{aligned}\text{menuDepth} &= \text{armLength}\\ &= Z_{h_{\text{Calibration}}} - Z_{s_{\text{Calibration}}},\end{aligned} \tag{1}$$

where $Z_{h_{\text{Calibration}}}$, $Z_{s_{\text{Calibration}}}$ are the z-axis value of user's hand and shoulder, which correspond to skeleton points of hand and shoulder center respectively; armLength is the length of user's arm; menuDepth is the calibrated menu depth of user.

User also needs to pull back the forearm as much as possible, with the palm facing the front and the upper arm against the body, to initiate MagicMark. The hand position is the starting position of MagicMark. Then, the user pushes the hand out horizontally to control the menu. We track the motion of user's dominant hand and calculate the depth menu selection by

$$\Delta Z = Z_{h_i} - Z_{h_0}, \tag{2}$$

$$\Delta\text{Item}_i = \left\lfloor \frac{\Delta Z}{|\text{menuDepth}|} \times m \right\rfloor + 1, \tag{3}$$

where ΔZ is the difference between the current position and the starting position of the hand in z-axis; m is the number of menu items, Item_i represents the current menu item where the hand is on.

• Direction-based gesture recognition. Waving gesture is used for converting the main menu to the submenu. Meanwhile it selects the submenu item based on the direction of the gesture. To identify whether a waving gesture is intentional or unintentional, we use three adjacent position $(x_{e_1}, y_{e_1}, z_{e_1})$, $(x_{e_2}, y_{e_2}, z_{e_2})$, $(x_{e_3}, y_{e_3}, z_{e_3})$. e stands for a time stamp of a position. An intentional waving gesture is determined when the following conditions are met.

$$\begin{aligned}&\frac{\sum_{i=e_1}^{e_3} |\sqrt{(x_{i+1}-x_i)^2 + (y_{i+1}-y_i)^2}|}{e_3 - e_1}\\ &> \text{thresSpeed}_{xOy},\end{aligned} \tag{4}$$

$$\frac{\sum_{i=e_1}^{e_3} |z_{i+1} - z_i|}{e_3 - e_1} < \text{thresSpeed}_z, \tag{5}$$

where thresSpeed_{xOy} is the threshold that detects the waving gesture on x-y plane; thresSpeed_z threshold ensures that the depth of hand does not change much during the waving gesture.

• Interaction state transition. The design of MagicMark can be summarized as the transition of three states. The main feature of this interactive design is that the transitions do not require additional confirmation gestures. Take a 4×4 menu as an example (Figure A1). First, the user holds his/her dominant hand with an activation gesture to bring up the menu (state 1). The dominant hand moves forward or backward to select the depth level in main menu (state 2). When the user makes the waving gesture that determines the depth of the main menu, an item in the submenu is selected based on the gesture direction: up, down, left or right (state 3). By detecting and thresholding the movement of the user's hand, the menu item of the submenu is selected and the corresponding command is activated.

• Novice mode and expert mode. MagicMark offers two modes. The user will transfer from the novice mode to the expert mode smoothly.

In the novice mode, the user first needs to bring up the MagicMark visual menu by the activation gesture and then select the main menu by controlling the movement of the hand in the depth direction, when a participant stretches his/her hand to a certain depth, the color of the corresponding item in this depth will change. The user then selects the submenu by waving the hand up, down, left, or right directions in the x-y plane. When this selection is confirmed, the selected item will change to red, and the selection completes. We call this interaction process "push-and-wave". When the selection is completed, the menu command is executed, and the MagicMark menu disappears.

In the expert mode, the user no longer needs to bring up the menu display. According to the memory of the depth and direction of a menu, the user can directly select the target menu item by using push-and-wave gesture. This interactive technique provides experienced users a similar use of shortcuts and allows seamless transition from novice mode to expert mode without extra training.

Experiment. To validate our hypothesis that MagicMark is an efficient and effective menu technique, we conducted an experiment to compare MagicMark and the traditional linear menu. We chose traditional menu as a contrastive technique because it was proved to be faster or less error-prone than other free hand techniques [5].

• Participants and apparatus. Ten participants, 6 females and 4 males, aged between 19 and 24, were recruited. They were all right-handed, and 3 of them had experience using depth cameras. The experiment used a tablet PC, a Microsoft Kinect, and a projection screen with a resolution of 1024×768 pixels. The camera was placed 1.2 m above the floor. According to the official manual, the distance between the camera and the user was set to 1.2 m.

• Design. We used a within-subject design. Participants were requested to complete selection tasks using two menu techniques. In MagicMark condition, the interface showed a 4×4 2D menu, while in traditional menu, the interface showed a vertical 4×4 linear menu (Figure A2). According to our previous study [6], a 0.8 s dwell-time was used as the confirmation gesture in traditional menu. The order of menu techniques was counterbalanced using the Latin square method. For each menu technique, participants completed 3 blocks, consisting of 16 trials in each block. The 16 trails set all items of a 4×4 hierarchical menu as target items in a random order. Participants were asked to perform all trials as quickly and accurately as possible. There was a five-minute rest between the two menu techniques.

Two dependent variables were calculated: completion time and error rate. The completion time was defined as the time starting from the participant's hand first left the starting region until the submenu item was selected. The error rate was defined as the percentage of incorrect selections in all trials of a menu technique.

Each participant was provided with 5 min practice time to familiar the task. Before each block of trials, participants need to stretch his/her arm as far as possible to calibrate the mapping between his/her arm length and screen pixel.

At the end of the experiment, participants completed a questionnaire to provide their subjective feedback for the two menu techniques. They were asked to rate the speed, error prone, comfortable to use, and easy to learn, on a Likert scale from 1 (worst) to 7 (best). For each participant, the experiment including the selection tasks and the questionnaire lasted approximately 30 min.

• Result. In total, we collected data of 960 trials (16 trials $\times$ 3 repetitions $\times$ 2 menu techniques $\times$ 10 participants). Results from the Wilcoxon signed rank test indicate that MagicMark ($M = 2.43$ s) is significantly faster than traditional menu ($M = 2.74$ s), ($p < 0.001$) (Figure A3). Statistical analysis (via the McNemar test) shows that there is no significant difference in error rate between the two menu techniques ($M = 1.88\%$), ($p = 1.000$). MagicMark is perceived more error prone ($Z = 2.57$, $p = 0.01$) (Figure A4). Differences in other subjective feedback aspects are not significant.

Discussion and future work. We find that MagicMark can significantly reduce the selection time comparing to the traditional linear menu without sacrificing accuracy. We attribute this performance gain to the selection mechanism of the MagicMark. To select a menu item, the MagicMark does not require any additional confirmation but a directional gesture. This is similar to target reverse crossing [7]. It is interesting that although the error rate of the two techniques has no significant difference, the subjective feedback shows difference. The reason might be similar with our previous study of TTR [8], that the experienced participants are familiar with the traditional technique, and underestimate the errors they made.

Our research can be extended in two ways. First, we like to compare the user performance with different menu item sizes using MagicMark. Second, we plan to explore the learning effect of MagicMark, especially the transition from novice mode to expert mode.

Acknowledgements This work was supported by Beijing Municipal Social Science Foundation (Grant No. 16YTC033), National Natural Science Foundation of China (Grant No. 61303162), and Fundamental Research Funds for the Central Universities. We thank Yu GAN for the visual design.

Supporting information Figures A1–A4, videos and other supplemental documents. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Bailly G, Lecolinet E, Nigay L. Visual menu techniques. ACM Comput Surv, 2016, 49: 1–41
2 Ni T, Bowman D A, North C, et al. Design and evaluation of freehand menu selection interfaces using tilt and pinch gestures. Int J Human-Comput Stud, 2011, 69: 551–562
3 Chattopadhyay D, Bolchini D. Touchless circular menus: toward an intuitive UI for touchless interactions with large displays. In: Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces, Como, 2014. 33–40
4 Kurtenbach G P. The Design and Evaluation of Marking Menus. Toronto: University of Toronto, 1993
5 Bailly G, Walter R, Müller J, et al. Comparing free hand menu techniques for distant displays using linear, marking and finger-count menus. In: Proceedings of the 13th IFIP TC 13 International Conference on Human-Computer Interaction, Lisbon, 2011. 248–262
6 Tian F, Lyu F, Zhang X L, et al. An empirical study on the interaction capability of arm stretching. Int J Human-Comput Interact, 2017, 33: 565–575
7 Feng W X, Chen M, Betke M. Target reverse crossing: a selection method for camera-based mouse-replacement systems. In: Proceedings of the 7th International Conference on Pervasive Technologies Related to Assistive Environments, Rhodes, 2014
8 Lyu F, Tian F, Dai G Z, et al. Tilting-Twisting-Rolling: a pen-based technique for compass geometric construction. Sci China Inf Sci, 2017, 60: 053101