

Research Article

A Crowdsensing-Based Real-Time System for Finger Interactions in Intelligent Transport System

Chengqun Song,^{1,2} Jun Cheng,^{1,2} and Wei Feng^{1,2}

¹Guangdong Provincial Key Laboratory of Robotics and Intelligent System, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

²The Chinese University of Hong Kong, Shatin, Hong Kong

Correspondence should be addressed to Jun Cheng; jun.cheng@siat.ac.cn

Received 3 July 2017; Revised 8 August 2017; Accepted 7 September 2017; Published 12 October 2017

Academic Editor: Zhaolong Ning

Copyright © 2017 Chengqun Song et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Crowdsensing leverages human intelligence/experience from the general public and social interactions to create participatory sensor networks, where context-aware and semantically complex information is gathered, processed, and shared to collaboratively solve specific problems. This paper proposes a real-time projector-camera finger system based on the crowdsensing, in which user can interact with a computer by bare hand touching on arbitrary surfaces. The interaction process of the system can be completely carried out automatically, and it can be used as an intelligent device in intelligent transport system where the driver can watch and interact with the display information while driving, without causing visual distractions. A single camera is used in the system to recover 3D information of fingertip for hand touch detection. A linear-scanning method is used in the system to determine the touch for increasing the users' collaboration and operability. Experiments are performed to show the feasibility of the proposed system. The system is robust to different lighting conditions. The average percentage of correct hand touch detection of the system is 92.0% and the average time of processing one video frame is 30 milliseconds.

1. Introduction

The World Health Organization in 2017 years reported that approximately 1.25 million people die every year as a result of a road traffic crash [1]. Distracted driving is a major factor of road traffic crash. The distraction caused by mobile phones, vehicle mounted central control display, and navigation display is a growing concern for road safety. Drivers watching the displays while driving slows braking reaction time and makes it difficult to keep in the correct lane and to keep the correct following distances.

In order to reduce the distractions caused by watching display as driving, a projector-camera interactive system can be used to project the display image on transparent glass on the vehicle while users can create and send messages by bare hand touching the projection content directly without any physical clicking devices. People can watch the display information while driving, without causing visual distractions.

The increasing social networking services based on mobile devices are a remarkable trend in mobile

computing [2, 3]. Social networking services can be fused with real-world sensing. Crowdsensing is one kind of fusing [4]. Crowdsensing, as a special form of crowdsourcing [5, 6], leverages human intelligence/experience from the general public and social interactions to create participatory sensor networks to solve specific problems collaboratively [7].

Projector-camera interactive system is also with potential applications to projection-based augmented reality (AR) [8] which directly project augmenting information upon objects in the real world. The vehicle mounted projection-based AR can project traffic information onto the glass screen and overlap the reality. Through the crowdsensing method, drivers can report specific traffic information, such as outside buildings, and store information, road conditions in remote areas, accidents, and so on by the AR interactive system. Then, their information can be automatically aggregated and uploaded to other users in real time.

For the projector-camera interactive systems, many single camera vision-based researches have been carried out. Shah et al. realized finger clicking detection using fingertip path

based on Camshift tracker [9]. However, the method detects clicks using a delay-based scheme, which is not suitable for applications that require fast response. Dai and Chung proposed a touch detection method in which the projection is embedded with imperceptible structured light [10]. He and Cheng realized touch detection method in which a kind of self-adaptive structured light is encoded and embedded into the projection [11]. However, special synchronized and high-speed camera and projector must be used to lock the phase of the embedded structured light. Hu et al. realized a kind of bare finger touch projection interaction by a novel approach based on button distortion [12]. However, the touch precision much depended on the size of button. Cheng et al. proposed a projector-camera interactive system in which touch was detected using a matched pairs of feature points by a projected white circle around the fingertip [13]. However, the additional projected white circle and the shadow of the finger on the projector screen will reduce the accuracy of foreground extraction. Park et al. implemented and evaluated a touch interface for a projection surface in which a depth camera was used for hand recognition and for extracting the hand area from the surrounding scene [14]. But the depth camera is expensive and its precision is affected by surrounding illumination. Zhou et al. proposed a method enabling estimation of the finger depth information using perceptible black-and-white stripe pattern [15]. However, the perceptible black-and-white stripe pattern on the projection screen disturbed the normal display image.

The information cues of the finger's shadow has been used in many kinds of projector-camera interactive systems. Song et al. proposed a handwriting recognition system in which the finger and its shadow were tracked using a camera. Tracks of the finger were used to indicate the direction and location of a document to simulate typical operations [16]. Xu et al. introduced an interactive system based on the shadow derived by the projector [17]. Shadows can provide a simple interface between human and computer systems. Huang et al. proposed a FSM grammar to recognize finger gestures and a shadow-based fingertips detection method [18].

Cai et al. proposed a fingertip touching approach based on the geometric relationship between fingertip and its shadow to estimate the distance from fingertip to projection surface [19]. While these methods require ideal illumination to create shadows that can be well recognized, Dung et al. provided a touch system that used the distance between the finger and its shadow to detect the touch-timing and location [20, 21]. However, an infrared camera and an infrared ray source are necessary in the system which makes the system complicated and costly.

In this paper, a real-time projector-camera interactive system based on the crowdsensing is proposed in which a triangulation method and a linear-scanning method are used to determine the touch for increasing the users' collaboration and operability. These approaches can enhance the robustness and accuracy of the system. The interaction process of the system can be completely carried out automatically, and it can be used as an intelligent device in intelligent transport system. The rest of the paper is organized as follows. In Section 2, the overview of the system is outlined, and

our proposed method for the system including modified multiframe difference method, linear-scanning method, and fingertip detection method is presented. Some experimental results are given in Section 3. Concluding remarks are discussed in Section 4.

2. Crowdsensing-Based Interactive System

2.1. System Overview. The proposed crowdsensing-based real-time finger interactions system for intelligent transport system consists of a computing device, a projector, a camera, and vehicle sensing devices. For different tasks such as the data collection of outside buildings and store information, road conditions in remote areas, and accidents, specific distributed crowdsourcing platform is established in which incentive mechanism with privacy protection is adopted to attract users to participate in the data collection. A hybrid incentive approach is used in this system for crowdsensing data collection. Monetary reward incentive method is used for the data collection of outside buildings and store information. Entertainment and gamification incentive method is used for the data collection of road conditions in remote areas. Virtual credit incentive method is used for the data collection of accidents.

The collected data is uploaded to the cloud service platform after being preprocessed. The cloud service platform analyzes and processes all the uploaded data. Other users download the information from the cloud service platform and project the information on the front screen of the car by projection-based AR mode. A registration method is used in the projection-based AR system based on projective reconstruction technique using natural features. At first, four points are specified to build the world coordinate which is used to superimpose virtual objects. Next, the live video natural feature is tracked using the Kanade-Lucas-Tomasi feature tracker [22]. The corresponding projective matrix is estimated using the tracked natural features in the image sequence [23]. Then, the registration matrix is computed by transforming the four specified points by the projective reconstruction technique for AR.

The hand casts its shadow on the display screen because of the projector light source, and the movements of the hand and its shadow are tracked by the camera. The system extracts the area of the hand and its shadow from the projection image and estimate the fusion degree of the hand and its shadow when the user interacts with the projector screen. When the hand and its shadow have been detected as being fused completely, the location of the fingertip will be detected and used to determine the location of the touch.

2.2. Foreground and Shadow Extraction. There are three approaches for detecting moving foreground target in image sequences: optical flow [24], background subtraction [25], and frames difference [26]. The optical flow method uses the brightness change of the image to extract the motion information of the object. The motion vector of each pixel is used to generate a complete vector field. However, the projection images changed randomly, which makes the optical flow method not suitable in our projection interactive application.

In addition, the optical flow method is quite complex and has poor noise immunity, so it cannot be applied to real-time processing. The background subtraction method use the subtracting between the current frame and a reference image called background to detect the foreground target. The background is found by means of an image selection process. The frames difference approach has the same principle with the background subtraction method, but the previous frame is treated as the reference image. Because of its small computation and high real-time performance, the frames difference method is more commonly used in foreground detection. However, the detection results of frames difference are prone to image smear and holes. In this paper, combining the background subtraction method, we proposed a modified multiframes difference method to detect the foreground and its shadow in the system.

At first, the camera captured image is geometric calibrated to the corresponding projection image in order to obtain the location relationship between the projection image and the camera view. The process can be carried out automatically using geometric calibration on camera captured image. A 3×3 matrix H homography [27] can be used to describe the mapping relationship between a point in the camera view and the projection image:

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix}. \quad (1)$$

The main steps of geometric calibration are described as follows [28]. (1) Project a chessboard to the display screen. (2) Capture the projected images using the camera. (3) Detect the chessboard corners in projection image and captured image, as shown in Figure 1. (4) Estimate the homography between the projection image and captured image using the corresponding detected corner locations in the two chessboard image. To estimate the homography, only 4 points of one rectangle in chessboard are necessary. In this paper as shown in Figure 1, 40 corner points are used to achieve higher accuracy.

Next, the foreground and its shadow will be segmented from the calibrated camera captured image. If image is projected onto the screen directly, the captured current image should be equal to the reference captured front image. If there is a moving foreground in front of the screen, the surface albedo will change. In order to increase the robustness of extracting the foreground and its shadow, block searching is used to obtain the albedo change of pixel $[x, y]$ which is computed by estimating a set of albedo ratios $a[x \pm i, y \pm j]$:

$$a[x \pm i, y \pm j] = \frac{I_k(x \pm i, y \pm j)}{I_{k-1}(x, y)} \quad (2)$$

$(i = 1, 2, \dots, m. j = 1, 2, \dots, n),$

where I_k is the current frame image and $I_k(x \pm i, y \pm j)$ is the gray value of pixel $[x \pm i, y \pm j]$ in I_k . I_{k-1} is the previous frame image and $I_{k-1}(x, y)$ is the gray value of pixel $[x, y]$ in I_{k-1} .

Pixel $[x, y]$ belongs to the foreground region if any one of $a[x \pm i, y \pm j]$ satisfies

$$a[x \pm i, y \pm j] < 1 - s, \quad (3)$$

or

$$a[x \pm i, y \pm j] > 1 + s, \quad (4)$$

where s is a constant of 0.5–0.8 as a tolerant scale of the albedo change.

If the hand movement is fast enough, as described above, the hand and its shadow can be extracted using two consecutive frames' difference. If the hand movement is slow, using the difference of two consecutive frames cannot extract the hand and its shadow completely. So, multiframes $I_k(x, y), I_{k-5}(x, y), I_{k-10}(x, y), I_{k-15}(x, y)$ are used to make differential operation. The extracted foreground and its shadow image can be obtained by

$$T_{kl(l=5,10,15)}(x, y) = \begin{cases} 255, & \text{if } a_l[x \pm i, y \pm j] < 1 - s \text{ or } a_l[x \pm i, y \pm j] > 1 + s \\ 0, & \text{else,} \end{cases} \quad (5)$$

where a_l is the albedo ratio between the current frame image $I_k(x, y)$ and previous reference image $I_{k-l}(x, y)$, ($l = 5, 10, 15$).

Background subtraction method is combined with multiframes difference method for the proposed interactive projection system, which is basically similar to the frame difference method, but the reference image is background image. The projection image is used as the background image in the modified multiframes difference method. According to the difference between the background image and the geometric calibrated camera captured image, the extracted foreground and its shadow image can be obtained by

$$T'_k(x, y) = \begin{cases} 255, & \text{if } a_b[x \pm i, y \pm j] < 1 - s \text{ or } a_b[x \pm i, y \pm j] > 1 + s \\ 0, & \text{else,} \end{cases} \quad (6)$$

where a_b is the albedo ratio between the current frame image $I_k(x, y)$ and background image.

The multiframes difference results $T_{kl(l=5,10,15)}(x, y)$ are used to perform AND logical operations with $T'_k(x, y)$, respectively,

$$\begin{aligned} D_1 &= T_{k5}(x, y) \cap T'_k(x, y), \\ D_2 &= T_{k10}(x, y) \cap T'_k(x, y), \\ D_3 &= T_{k15}(x, y) \cap T'_k(x, y). \end{aligned} \quad (7)$$

Then, perform OR logical operations on $D_1, D_2,$ and D_3 to obtain D .

The extracted foreground and its shadow result D are processed by morphological erosion and dilation [29]. Because of the noise interference, there is often a small amount of color

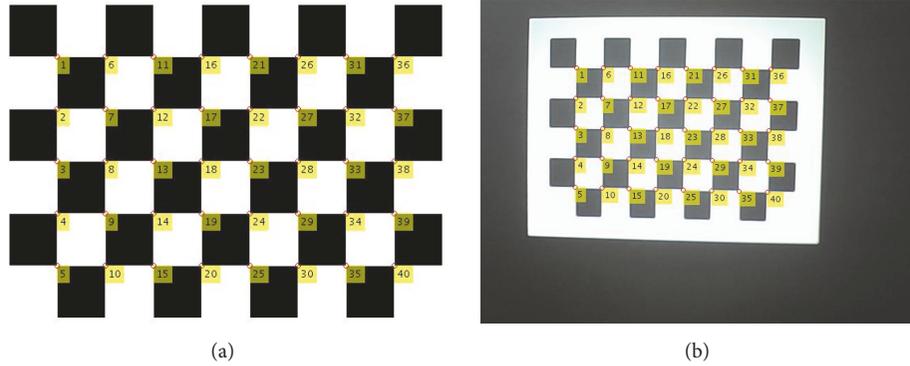


FIGURE 1: (a) The chessboard in projection image and (b) the chessboard in captured image. The detected corners of chessboard are marked with circle for both the chessboard in projection and captured image.

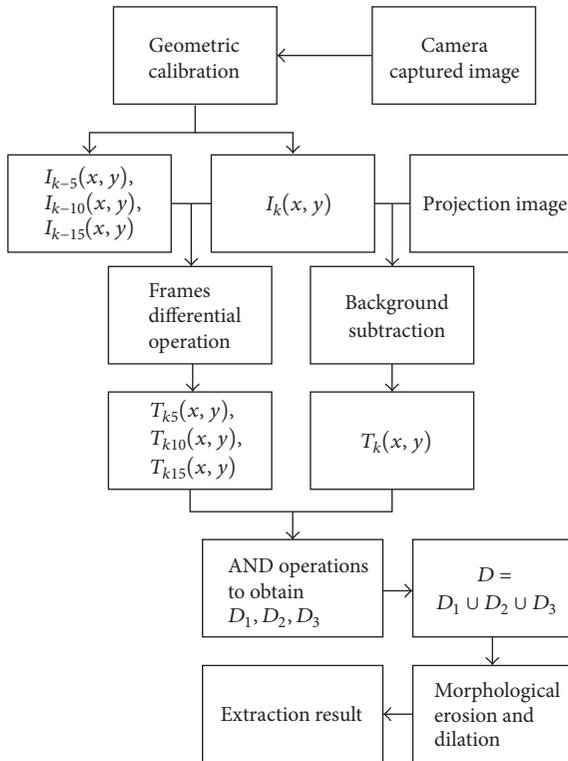


FIGURE 2: Flow chart of the foreground and shadow extraction.

or gray similarity between the target and the reference image. Image D often has many isolated points and holes, which will interfere the detection of foreground and its shadow. Through erosion and dilation treatment, isolated spots are removed and holes are filled.

The flow chart of the modified multiframes difference method as described above for foreground and shadow extraction is shown in Figure 2. Figure 3(a) is the captured current image with foreground and its shadow; Figure 3(b) is the captured previous 10th image with foreground and its shadow. Figure 3(c) is the captured previous 15th image with foreground and its shadow. Figure 3(d) shows the foreground

and its shadow extraction result using the modified multi-frames difference method.

2.3. Touch Detection. The principle of touch detection is based on the triangulation [30] in our interactive projection system. A linear-scanning method is proposed to detect the triangulation of the finger and its shadow without separating the hand from its shadow, which can increase the robustness of the system and reduce the computation cost. Before detecting the triangulation fusion degree of the finger and its shadow, the hand in the extracted image is demarcated to improve the accuracy and processing speed of the fusion detection for the disturbance from the rest of the captured image can be avoided. The pixels of the foreground and its shadow can be detected on the right edge of the extracted images if the user is on the right side of the projector display screen, while they cannot be detected on the left edge of the extracted images, as shown in the left parts of Figures 4(a) and 4(b). Then, a section including the hand from the far-left pixels line of the foreground and its shadow in the extracted image to the right, as shown in the right parts of Figures 4(a) and 4(b), is intercepted as the demarcated area for fusion detection of the finger and its shadow. The pixels of the foreground and its shadow can be detected on the left edge of the extracted images if the user is on the left side of the projector display screen, while they cannot be detected on the right edge of the extracted images, as shown in the left parts of Figures 4(c) and 4(d). Then, a section including the hand from the far-right pixels line of the foreground and its shadow in the extracted image to the left, as shown in the right parts of Figures 4(c) and 4(d), is intercepted as the demarcated area for fusion detection of the finger and its shadow.

A vertical scanning line is used to scan the demarcated area to estimate the fusion degree. If a vertical line is detected on which a 0-pixel segment exists and the pixel value at both ends of the segment is 255, the finger and its shadow have not completely fused, as shown in the right parts of Figures 4(a) and 4(c). Otherwise, if no such vertical line is detected, the finger and its shadow have fused, as shown in the right parts of Figures 4(b) and 4(d).

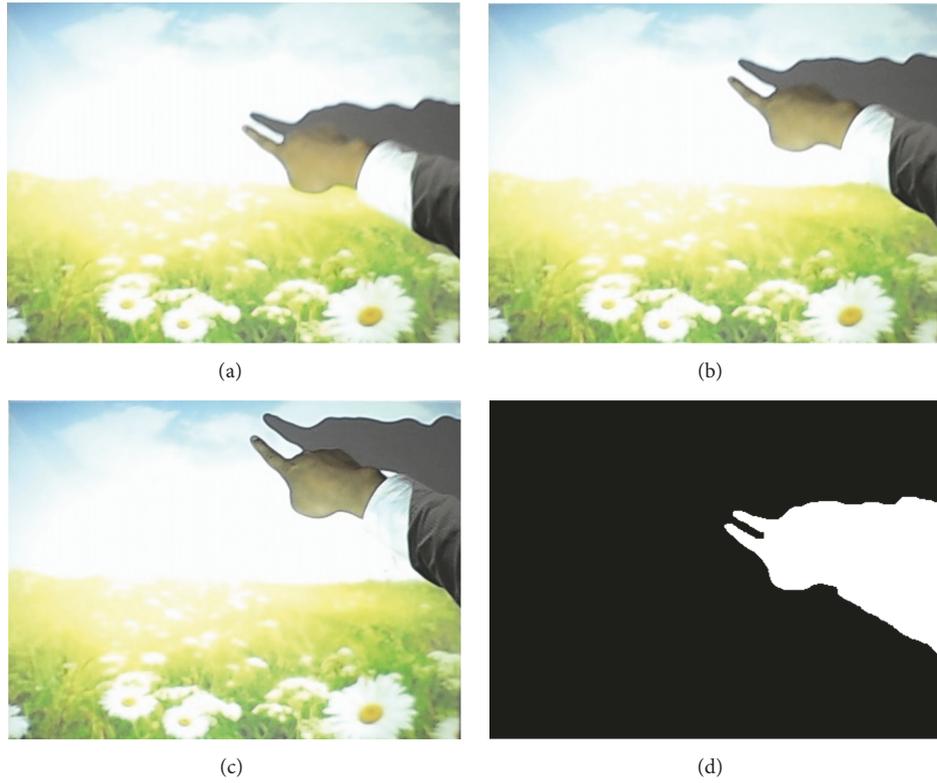


FIGURE 3: Foreground and shadow extraction. (a) The captured current image. (b) The captured previous 10th image. (c) The captured previous 15th image. (d) Foreground and its shadow extraction result.

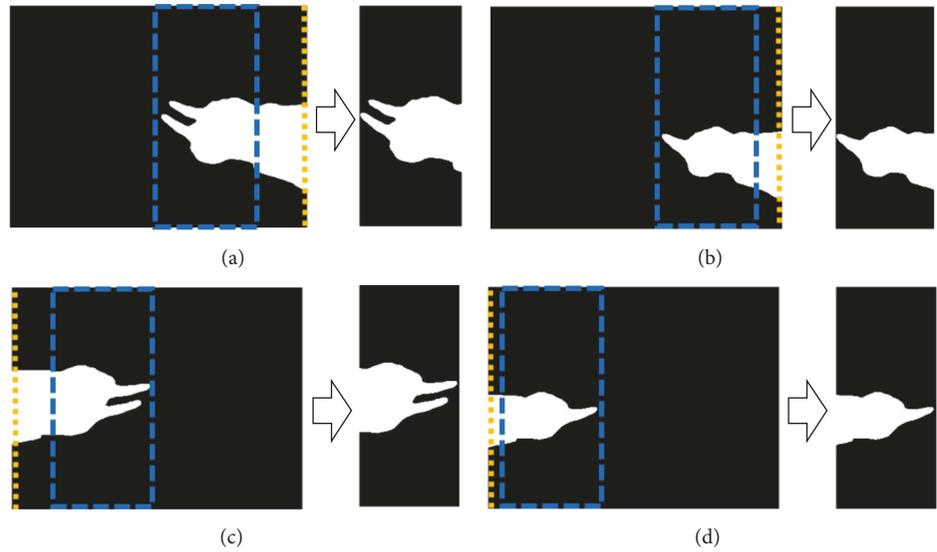


FIGURE 4: Schematic diagram of the linear-scanning method. (a) The finger has not fused with its shadow when the user is on the right. (b) The finger has fused with its shadow when the user is on the right. (c) The finger has not fused with its shadow when the user is on the left. (d) The finger has fused with its shadow when the user is on the left.

When the fingertip and its shadow are detected as being completely fused using the linear-scanning method, the location of the fingertip should be found in the camera captured image to obtain the touch location on the projection

image. According to the characteristic of the user, when the user is on the right side of the projector display screen, the far-left pixel of the foreground in the extracted image is as the fingertip, as shown in Figure 4(b). When the user is on

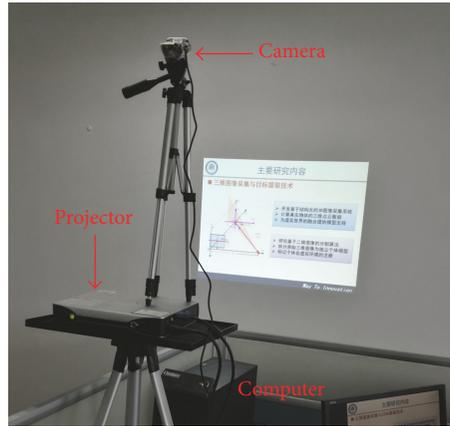


FIGURE 5: The proposed projector-camera interactive system.

the left side of the projector display screen, the far-right pixel of the foreground and its shadow in the extracted image is as the fingertip, as shown in Figure 4(d).

3. Result and Discussion

An simulation experiment is conducted to evaluate the performance of our proposed projector-camera interactive system. The experiment platform, as shown in Figure 5, includes a Lenovo computer with 3.6 Ghz CPU and 4.0 G RAM, a 640 × 480 resolution camera, a SONY projector with resolution of 1440 × 1050, and a projector screen. Lps is 1.5 m, and Lcp is 0.5 m.

Figure 6 shows the experimental results of the foreground and its shadow extraction under different lighting. Figures 6(a) and 6(b) are the foreground and its shadow extraction results in natural lighting with bright and dark projection, respectively. Figures 6(c) and 6(d) show experimental results in LED (Light Emitting Diode) lighting with bright and dark projection, respectively. It can be seen that our method is robust in the foreground and its shadow extraction.

The touch depth accuracy is studied in the experiment. When a touch has been detected, the maximum vertical distance from the fingertip to the screen plane is defined as maximum effective touch depth (METD). 9 points on the screen are selected as our METD detection points which are represented with red asterisks as shown in Figure 7. Figure 8 illustrated the METD detection results at the selected 9 points on the projection screen in natural lighting condition with a bright projection. It can be seen from Figure 8 that most of the METDs are below 10 mm except one of 12 mm. The average METD is 8.3 mm. The results demonstrate that our proposed method is effective and it will lead to an accurate fingertip depth detection during interaction.

The randomness and accuracy of touch position on the projection screen are verified in the system. As shown in Figure 9, 220 random points are measured after the finger traverse through the screen. Most of the touch position deviation on the X - Y screen plane is below 8 mm only except one of 21 mm. The average touch position deviation is 3.3 mm.

The results imply the touch position detection is accurate and unaffected whether the points locate in the center region or at the edge of the screen during the interaction.

Experiment is conducted for touch detection under different lighting conditions, as shown in Table 1. In the experiment, when the METD is below 8 mm and the touch position deviation is below 8 mm, the touch detection is supposed to be correct in normal projector interactions. In Table 1, the touch detection accuracy rates of our method are shown in comparison with Song et al.'s [16] and He and Cheng's [11] method. Among 200 sampled frames, we show the touch detection accuracy rates experimental results in bright and dark projection on natural lighting condition and LED lighting condition, respectively. The touch detection accuracy rate in bright projection on natural lighting condition is the same between our method and Song et al.'s method; both are 98.0%, which is higher than 96.0% of He and Cheng's method. The touch detection accuracy rate in dark projection on natural lighting condition is also the same between our method and Song et al.'s method; both are 88.0%, which is higher than 86.0% of He and Cheng's method, while the touch detection accuracy rate of our method is much higher than Song et al.'s and He and Cheng's method in bright or dark projection on LED lighting condition. The average percentage of correct detection of our method (92.0%) is higher than that of their methods (88.0% and 88.5%).

The execution time of the blocks of the projector-camera interactive system is shown in Table 2. The average time of processing one video frame is 30 milliseconds; the total time variance is 0.6; the total time range is 4. We can detect the touch-timing and location at 33 frames per second. That is to say, our method can achieve real-time computing in an ordinary computing system.

4. Conclusion

In this paper, we present a real-time projector-camera interactive system based on the crowdsensing which enables user to transform any flat surface into a virtual touch panel and interact with computer by finger touching. The

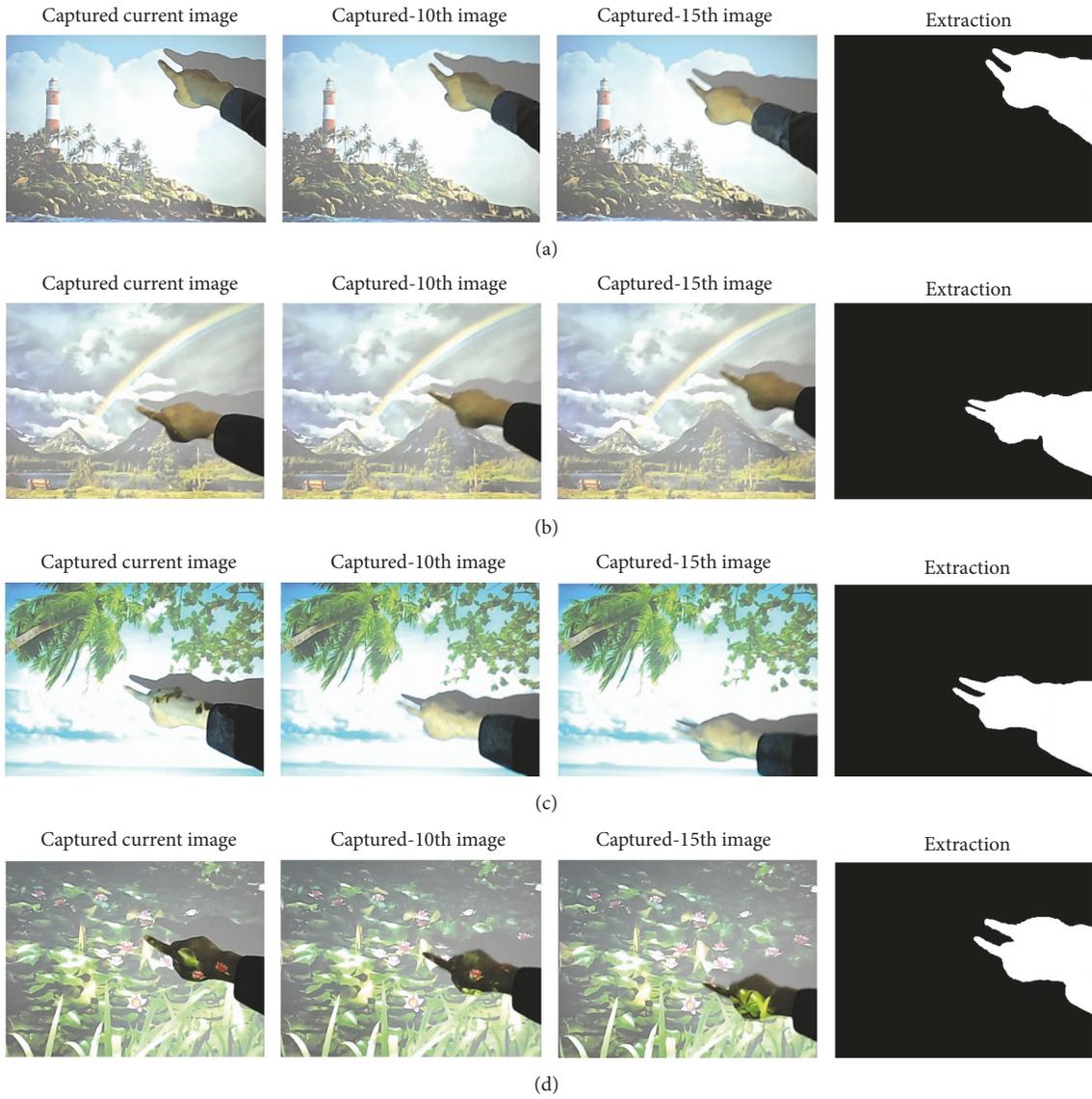


FIGURE 6: The foreground and shadow extraction results under different lighting. (a) is in bright projection on natural lighting, (b) is in dark projection on natural lighting, (c) is in bright projection on LED lighting, and (d) is in dark projection on LED lighting.

interaction process of the system can be completely carried out automatically; it can be used as an intelligent device in intelligent transport system where the driver can watch and interact with the display information while driving, without causing visual distractions. A single camera is used in the system to recover 3D information of fingertip for hand touch detection by the finger and its shadow triangulation. Foreground and shadow are extracted based on a modified multiframes difference method which has both the advantage of the adaptability for projection dynamic environment in the frame difference method and advantage of obtaining complete target information in background subtraction method. Therefore, the interactive system can hardly be affected by the lighting environment, which leads to an accurate interactive gestures extraction and makes the system more adaptable. The fusion degree of the hand and its shadow is detected using a linear-scanning method which avoided the errors caused

by inadequate separation of the finger and its shadow and increased the robustness of the system. Finally, a simple and effective method for detecting the position of the fingertip is proposed according to the user's habit of using the projection interactive system. The experiment results indicate that the projector-camera interactive system can achieve a robust and effective performance using a single camera.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 61403365 and

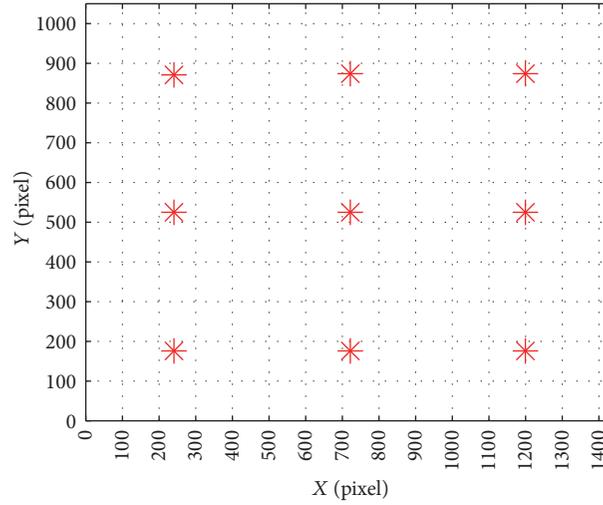


FIGURE 7: Selected 9 points on the projection screen for touch depth detection.

TABLE 1: Touch detection accuracy rates.

Methods	Types	Frames	Each type's accuracy	Average accuracy	
Song et al.'s method	Bright projection on natural lighting	Correct	49	98.0%	88.0%
		No	1		
	Dark projection on natural lighting	Correct	44	88.0%	
		No	6		
	Bright projection on LED lighting	Correct	44	88.0%	
		No	6		
Dark projection on LED lighting	Correct	39	78.0%		
	No	11			
He and Cheng's method	Bright projection on natural lighting	Correct	48	96.0%	88.5%
		No	2		
	Dark projection on natural lighting	Correct	43	86.0%	
		No	7		
	Bright projection on LED lighting	Correct	45	90.0%	
		No	5		
Dark projection on LED lighting	Correct	41	82.0%		
	No	9			
The proposed method	Bright projection on natural lighting	Correct	49	98.0%	92.0%
		No	1		
	Dark projection on natural lighting	Correct	44	88.0%	
		No	6		
	Bright projection on LED lighting	Correct	48	96.0%	
		No	2		
Dark projection on LED lighting	Correct	43	86.0%		
	No	7			

TABLE 2: Execution time measured.

Block	Average time (milliseconds)	Variance	Range
Foreground and shadow extraction	25	0.8	4
Touch detection	5	0.1	1
Total	30	0.6	4

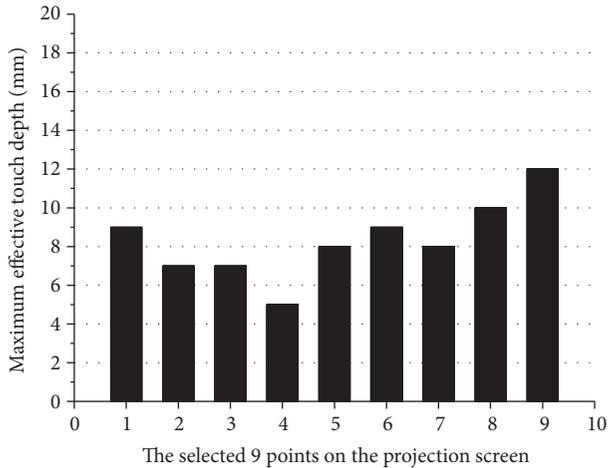


FIGURE 8: Touch depth detection results at the selected 9 points on the projection screen.

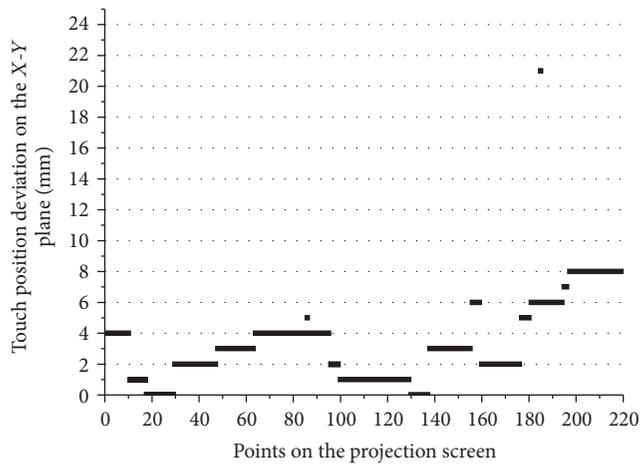


FIGURE 9: Touch position deviation at 220 random points on the X-Y screen plane.

61402458), the China Postdoctoral Science Foundation (Grant no. 2016M602543), the Natural Science Foundation of Guangdong Province, China (Grant nos. 2015A030313744 and 2016A030313177), Shenzhen Technology Project (Grant nos. JSGG20160331185256983 and JSGG20160229115709109), Guangdong Technology Project (Grant nos. 2016B010108010 and 2016B010125003), Shenzhen Engineering Laboratory for 3D Content Generating Technologies (Grant no. [2017]476), Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (Grant no. 2014DPI73025), State Joint Engineering Laboratory for Robotics and Intelligent Manufacturing funded by National Development and Reform Commission (Grant no. 2015581), and CAS Key Technology Talent Program.

References

- [1] World Health Organization, "Road traffic injuries," 2017, <http://www.who.int/mediacentre/factsheets/fs358/en/>.

- [2] X. Hu, T. H. S. Chu, V. C. M. Leung, E. C.-H. Ngai, P. Kruchten, and H. C. B. Chan, "A Survey on Mobile Social Networks: Applications, Platforms, System Architectures, and Future Research Directions," *IEEE Communications Surveys & Tutorials*, vol. 17, no. 3, pp. 1557–1581, 2015.
- [3] Z. Ning, L. Liu, F. Xia, B. Jedari, I. Lee, and W. Zhang, "CAIS: a copy adjustable incentive scheme in community-based socially aware networking," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 3406–3419, 2017.
- [4] C. T. Calafate, C. Wu, E. Natalizio, and F. J. Martínez, "Crowdsensing and Vehicle-Based Sensing," *Mobile Information Systems*, vol. 2016, Article ID 2596783, p. 2, 2016.
- [5] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, pp. 1–4, 2006.
- [6] Q. Li, P. Yang, S. Tang, M. Zhang, and X. Fan, "Equilibrium is priceless: selfish task allocation for mobile crowdsourcing network," *EURASIP Journal on Wireless Communications and Networking*, vol. 2016, no. 1, pp. 1–8, 2016.
- [7] X. Hu, V. C. Leung, Y. Kwok et al., "SafeDJ: a Crowd-Cloud Co-design Approach to Situation-aware Music Delivery for Drivers," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 12, no. 1s, p. 21, 2015.
- [8] P. Beardsley, J. V. Baar, R. Raskar, and C. Forlines, "Interaction using a handheld projector," *IEEE Computer Graphics and Applications*, vol. 25, no. 1, pp. 39–43, 2005.
- [9] S. A. Shah, A. Ahmed, I. Mahmood, and K. Khurshid, "Hand gesture based user interface for computer using a camera and projector," in *Proceedings of the 2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pp. 168–173, Kuala Lumpur, Malaysia, November 2011.
- [10] J. Dai and R. Chung, "Making any planar surface into a touch-sensitive display by a mere projector and camera," in *Proceedings of the 2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR Workshops)*, pp. 35–42, Providence, RI, USA, June 2012.
- [11] M. He and J. Cheng, "Self-adaptive coding-based touch detection for interactive projector system," in *Proceedings of the 2014 4th IEEE International Conference on Information Science and Technology (ICIST)*, pp. 656–659, Shenzhen, China, April 2014.
- [12] J. Hu, G. Li, X. Xie, Z. Lv, and Z. Wang, "Bare-fingers Touch Detection by the Button's Distortion in a Projector-Camera System," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 24, no. 4, pp. 566–575, 2014.
- [13] J. Cheng, Q. Wang, R. Song, and X. Wu, "Fingertip-based interactive projector-camera system," *Signal Processing*, vol. 110, pp. 54–66, 2015.
- [14] Y.-S. Park, S.-H. Park, T.-G. Kim, and J.-M. Chung, "Implementation of Gesture Interface for Projected Surfaces," *KSII Transactions on Internet and Information Systems*, vol. 9, no. 1, pp. 378–390, 2015.
- [15] Q. Zhou, X. Qiao, K. Ni, X. Li, and X. Wang, "Depth detection in interactive projection system based on one-shot black-and-white stripe pattern," *Optics Express*, vol. 25, no. 5, pp. 5341–5351, 2017.
- [16] P. Song, S. Winkler, S. O. Gilani, and Z. Y. Zhou, in *Proceedings of the IEEE International Workshop on Human-Computer Interaction*, vol. 4796, pp. 49–58, Springer, Rio de Janeiro, Brazil.
- [17] H. Xu, D. Iwai, S. Hiura, and K. Sato, "User Interface by Virtual Shadow Projection," in *Proceedings of the 2006 SICE-ICASE International Joint Conference*, pp. 4814–4817, IEEE, Busan, Korea, October 2006.

- [18] J.-Y. Huang, Y.-Z. Yeo, L. Huei, and C.-H. Tsai, "Image-Based Wearable Tangible Interface," in *Advances in Intelligent Systems and Applications - Volume 2*, vol. 21 of *Smart Innovation, Systems and Technologies*, pp. 455–464, Springer, Berlin, Germany, 2013.
- [19] X. Cai, X. Xie, G. Li, W. Song, Y. Zheng, and Z. Wang, "A new method of detecting fingertip touch for the projector-camera HCI system," in *Proceedings of the 2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 526–529, IEEE, Melbourne, Australia, June 2014.
- [20] L.-R. Dung, G.-Y. Lai, and Y.-Y. Wu, "Shadow touching for interactive projectors," in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1798–1802, IEEE, Vancouver, Canada, May 2013.
- [21] L.-R. Dung and R.-Y. Huang, "A Shadow Touching Technique for Interactive Projector Devices," in *HCI International 2013 - Posters' Extended Abstracts*, vol. 373 of *Communications in Computer and Information Science*, pp. 308–312, Springer, Berlin, Germany, 2013.
- [22] B. Georgescu and P. Meer, "Point matching under large image deformations and illumination changes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 674–688, 2004.
- [23] M. L. Yuan, S. K. Ong, and A. Y. C. Nee, "Registration using natural features for augmented reality systems," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 4, pp. 569–580, 2006.
- [24] A. Agarwal, S. Gupta, and D. K. Singh, "Review of optical flow technique for moving object detection," in *Proceedings of the 2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 409–413, Greater Noida, India, October 2016.
- [25] G. Gemignani and A. Rozza, "A Robust Approach for the Background Subtraction Based on Multi-Layered Self-Organizing Maps," *IEEE Transactions on Image Processing*, vol. 25, no. 11, pp. 5239–5251, 2016.
- [26] K. Wang, L. Xu, Y. Fang, and J. Li, "One-against-all frame differences based hand detection for human and mobile interaction," *Neurocomputing*, vol. 120, pp. 185–191, 2013.
- [27] Z. Zhang, "Computer Vision Technologies for Remote Collaboration Using Physical Whiteboards, Projectors and Cameras," in *Proceedings of the Computer Vision for Interactive and Intelligent Environment (CVIIIE'05)*, IEEE, Lexington, KY, USA, November 2005.
- [28] M. Liao, R. Yang, and Z. Zhang, "Robust and Accurate Visual Echo Cancellation in a Full-duplex Projector–Camera System," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1831–1840, 2008.
- [29] P. Soille, *Morphological Image Analysis—Principles and Applications*, Springer, Berlin, Heidelberg, Germany, 2nd edition, 2004.
- [30] Y. Wang and S. Zhang, "Optimal fringe angle selection for digital fringe projection technique," *Applied Optics*, vol. 52, 7094, no. 29, p. 7098, 2013.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

