The 5th Annual IEEE International Conference on
Cyber Technology in Automation, Control and Intelligent Systems
June 8-12, 2015, Shenyang, China

Virtual Keyboard: A human-computer interaction device based on laser and image processing

Xiaolin SU, Yunzhou ZHANG*, Qingyang Zhao, Liang GAO
College of Information Science and Engineering
Northeastern University
Shenyang, China
zhangyunzhou@ise.neu.edu.cn

Abstract—**Keyboard is a major kind of input device for computer and intelligent device, especially those portable equipment. Aiming at more intelligent human-computer interaction, a virtual keyboard system is presented on the basis of laser and image processing. The virtual keyboard consists of infrared laser module, keyboard pattern projector, embedded system and a single image sensor. The keyboard pattern can be implemented through the projector. A CMOS image sensor is used to collect the sequential frames of fingertips keystroke. Every keystroke can be detected accurately by image processing including morphology principle and ellipse fitting. Experiment results show that the virtual keyboard system can work reliably with high-speed response and high accuracy.**

Key word- Portable equipment, virtual keyboard, laser picture, embedded system, image processing

I. Introduction

As the key input device of computer and intelligent device, keyboard plays an irreplaceable role in many areas such as data input and the transmission of signal control etc. However, traditional keyboard is bulky and is inconvenient to carry, failed to meet the demand of mobile terminal. The appearance of Virtual Laser Keyboard presented a new solution for input on portable device [1][2].

Various new techniques had been adopted on the design schemes of virtual keyboard such as gesture recognition and brain-computer interface. Reference [3] developed a novel drag-type mode based on touch screen techniques to take the place of the keystroke in smart phones. Reference [4] discussed a brain - computer interface virtual input system. It acquired eye-gazed characters through electrooculogram(EOG) signals, gaining the angle of sight, and electromyogram(EMG) signals, recording brow muscle activities for verifying. Reference[5] is also based on EOG signals, facilitating the interaction between eye movements and portable device. . The sampling frequency could reach 176 Hz. And the classification accuracy reached 95%. Reference [6] came up with a multi-touch support virtual input system. Different from brain-computer interface, special gloves and other methods, visual-based virtual keyboards do not require redundant equipment to wear, and keep small in size, thus it has broader future.

Compared with the above-mentioned virtual keyboard, the virtual keyboard system designed in this paper is more like the QWERTY. Not only our system is carried conveniently, but also users' typing custom is considered which meets the need of quick typing. Because of the infrared and image processing technology, there is no extra implements on Virtual Laser key board for users. Reference[7] presented a system (V-Touch) based on Haar-like feature with a single camera and a laser module which is an obvious low cost hardware setup. Reference[8] also realized typing without traditional keyboard, using laser and image processing technology.

Reference[9] established a virtual keyboard system based on three-dimensional computer vision, obtaining hand depth information in order to determine the position of fingertips. Similarly, reference[1] proposed a virtual keyboard, using gesture recognition for tracking fingertips movement and identifying signal input. All over the keyboard layout can be typed. Reference [10] achieved detecting the fingertips in 2D images from a pair of cameras, and then the fingertip position in 3D space is determined using the triangulation method.

Some similar and excellent work has been done. Compared with the system with 3D technology, our system only utilizes one camera. So its cost is lower. In this paper, we research how to get a clear fingertip image, examine the correct rate and analyze the cause of errors.

In this paper, the virtual keyboard system combines laser, infrared ray with image processing technology. And the system uses embedded-system to build a compact and portable input device. Image algorithms, system accuracy, timeliness and stability were verified by actual test.

II. Virtual Keyboard system Architecture

Virtual keyboard system mainly consists of four modules: a keyboard pattern laser projector, infrared ray laser module, embedded-system and a CMOS image sensor. Fig.1 shows virtual keyboard system architecture.

Keyboard pattern is projected on the horizontal plane by the projector, forming input virtual keyboard layout that allows users to "press key". In normal indoor lighting conditions, the projection technology can clearly project full-size keyboard pattern on any plane, allowing users to operate and type as easy as a traditional keyboard.

Infrared ray generated by infrared ray laser module covers

Supported by National Natural Science Foundation of China (No. 61471110) and China Postdoctoral Science Special Foundation (No. 2014T70263).

the level surface, about 1mm above the keyboard-pattern, to form the infrared curtain.

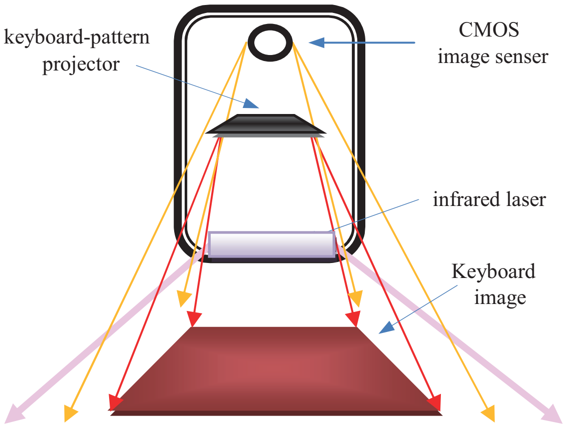

Fig. 1. Virtual keyboard system architecture

CMOS image sensor inclined downwards, pointing and shooting towards keyboard projection pattern. CMOS image sensor is equipped with IR filter to filter out visible light, to allow only infrared to pass. When the user “presses key”, the CMOS image sensor captures the infrared reflection spot, then embedded-system processes the image. At last spot position is mapped back to the character. Fig.2 shows the flowchart of embedded-system processing image.

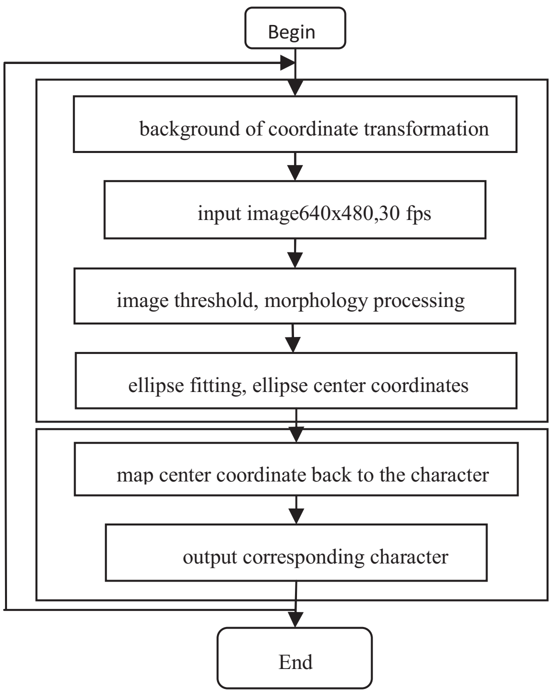

Fig. 2. Flow Chart of Virtual keyboard image process

Embedded-system board core is K60 microprocessors, integrated ARM Cortex-M4 core and DSP digital signal processor with 512K Flash, 512K Flex memory, 128K SRAM, and 16K Cache, suitable for smaller-scale digital image processing and peripheral interface control. Its frequency is up to 150MHz, with optional single- precision floating-point unit and four configurable resolution high-speed 16-bit ADC. Its 32-channel DMA, for peripherals and memory, can reduce the CPU load, speed system through put. UART supports I2C and SPI, can be directly applied to CMOS image sensor interface. GPIO is easy to access and control the infrared and laser modules.

III. Keystroke Extraction and Recognition

A. *Camera calibration and coordinate mapping*

To locate “keystroke”position, the system needs to work with computer image coordinate system. In this paper, the background has been converted from the objective world coordinate system to the computer image coordinate system.

Due to the production processing, the cameras’ lens will introduce distortion. Because the camera is close to shooting plane, the radial distortion might easily lead to serious distortion on image edge. Fig. 3 shows obvious barrel distortion on the edge of keyboard pattern. So radial distortion can not be ignored and must be corrected.

Radial position adjustment formula of a 2D point in camera coordinate system is as follows:

$$\begin{cases} u_d = u\left(1+k_1 r^2 + k_2 r^4\right) \\ v_d = v\left(1+k_1 r^2 + k_2 r^4\right) \end{cases} \tag{1}$$

$u(u,v)$ is the original position of distortion point, (u_d, v_d) is the position after correction, k_1 *and* k_2 were distortion coefficients, r is the distance between distortion point and camera optical center.

Camera captures 10 chessboard calibration plate images from different angles..We use Matlab calibration toolbox to get camera parameter for calibration,distortion factor, focal length, skew and so on, and we establish the relation between objective world and computer image. After camera radial distortion correction, preliminary result is shown in Fig.4.

During shooting, the positional relationship between keyboard pattern projector module and the camera will also bring distortion. The display of the keyboard pattern projector is difficult to remain strictly horizontal, while camera plate also has certain dip to horizon. These also introduce distortion. Above conditions can affect accuracy of the "keystroke" edge positioning.

Fig. 3. Keyboard image

Fig. 4. Preliminary correction result

$$\begin{bmatrix} t_i \cdot u_d \\ t_i \cdot v_d \\ t_i \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{2}$$

$\begin{bmatrix} u_d \\ v_d \end{bmatrix}$ is coordinates after correction. $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$ is perspective transformation matrix. $\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$ is original image coordinates.

Perspective transformation does not require overall scaling. Define i = 1. Perspective transformation homography matrix is calculated by acquiring 4 points and constructing 8 equations. Preliminary result is corrected through perspective transformation homography matrix. Fig.5 shows the result after perspective correction. The edge of keyboard is nearly horizontal, which could improve recognition accuracy.

In this paper, four vertex coordinates of one "key" are saved in the map. Thus, 88 vertex coordinates are needed for 22"key". Matlab can get these computer image coordinates (u, t). The mark"*" in Figure 5 means the vertex needed.

Fig. 5. Afterperspective correction

B. "keystroke"recognition processing

The spot produced by typing the "key" is recorded as processing data for recognition processing and locating of the spot. In addition, every frame should be corrected before processing.

The images collected by the camera must be preprocessed at the beginning. All images should be clear and free of noise. The key of identifying character is spot shape and position. In this paper, we remove the pixels under the threshold, which is because the pixel above the threshold is very likely to be the spot, yet the pixel under the threshold may be background.

Threshold process could effectively remove the noise in background or generated by the camera. There are some bright spots near the main spot, mainly due to the uneven brightness reflected by fingers.The bright spots might have a severe impact on fingertip locating. It is necessary to introduce mathematical morphology operations for processing.

In mathematical morphology operations, open operation can effectively remove bright details in small size compared with the structural elements. The process is corrodingand dilating, successively. Open operation not only draws the outline of the image smoothly, but also plays a role in filter to remove small protuberances.

In this paper, open operation is adopted to the threshold figure. After that, a fingertip figure is acquired as shown in Fig.6. In this figure, noisy points and hot spots disappear completely. At the same time，the edge of the fingertip is smooth and clear. Fingertip recognition is usually realized by algorithms such as gray centroid method, minimum to process square ellipse fitting and Hough transform,etc. As the spot formed by "typing" is elliptical and data is large. Thus an efficient way, the minimum square ellipse fitting, is adopted in this paper.

(a) the 48thframe (b) the 104thframe

Fig. 6. Eliminating noise and bright spots fingertip

The elliptic curve equation is:

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + 1 = 0 \tag{3}$$

The variance sum is

$$F^2 = \sum_{i=1}^{N} \left(Ax_i^2 + Bx_i y_i + Cy_i^2 + Dx_i + Ey_i + 1 \right)^2 \tag{4}$$

Calculate partial derivatives of each coefficient in the (3) and make sure that each formula equals zero. A system of equations with 5 equations and 5 unknown numbers is formed. Thus Elliptic equations can be determined after the equations being solved. Ellipse center coordinates(x0, y0) can be determined by:

$$\begin{cases} x_0 = \dfrac{BE - 2CD}{4AC - B^2} \\ y_0 = \dfrac{BD - 2AE}{4AC - B^2} \end{cases} \tag{5}$$

During ellipse minimum quadratic fitting, the coordinates of known outline are required to determine the parameters in elliptic curve equation, and the center point of spot could be determined afterwards.

After the ellipse minimum quadratic fitting, we use the sequence of boundary points, the outline of spot which represents a fingertip typing on the keyboard is determined. After the above processing steps, it is calculated that the coordinate of ellipse center in the 48th frame is (294,453) and its corresponding key is ”U”; in the 104th frame, the coordinate is (354,399), and corresponding to “H”.Add the ellipse acquired above to the figure of keyboard pattern, processing result of “keystroke” is obtained.

In Fig.7, the fitted ellipse center falls into the region of the key “T” , exactly the same as the target key in the test. Although the fitted ellipse of spot does not fall in the region of the target key entirely, the ellipse center does.

Fig. 7. Sketch map of the keystroke in the 104th frame

Although the fitted ellipse of spot isn’t in the region of the target key completely, the ellipse center is in the region of the target key. A PC3030K camera is adopted in this test with resolution ratio of 640x480 and with a frame rate of 30fps. The flowchart of virtual keyboard system’s information is shown in Fig.8.

Results are as follows: The character displayed on the LCD screen is the right “keystroke”. We do the experiments in 5 groups. Users “press” certain “keys” 20 times every 5 seconds in each experiment.Table I shows experiment results of 5 groups. We also typed 20 “keys” all over the laser keyboard in 10 seconds. Tab.II shows keyboard experimental results overall.

Statistical results in the Tab.II indicates that the recognition rate exceeds 96%. The recognition rate is high enough to meet general use. However, a low error rate still exists. A declining recognition rate may happen in two ways: one is no response, while the other is incorrect recognition.

When we strike the keyboard in high frequency, the “keystroke” is easy to deviate, which means users cannot get feedback just as striking a real keyboard. We strike the “T key” repeatedly and then collect the coordinate deviation statistics. Expected result is shown in Tab. III

TABLE I. EXPERIMENT RESULTS

character	“5”	“6”	“7”	“8”
times	20*5	20*5	20*5	20*5
accuracy	100%	100%	100%	100%
character	“R”	“T”	“Y”	“U”
times	20*5	20*5	20*5	20*5
accuracy	100%	100%	98%	98%

TABLE II. ALLOVER KEYBOARD EXPERIMENT RESULTS

character	*Number*	*letter*	*space*
times	500	500	500
accuracy	98%	96%	100%

TABLE III.EXPECTED RESULTS

Keystroke position	*Fingertip in the key*	*Part of fingertip in the key*	*Little of fingertip in the key*
expect	“T”	“T”	not“T”
times	20	20	20
accuracy	100%	60%	90%

According to experimental results, the main cause of the error is the “keystroke” deviation, resulting in fitting center falling unexpected region.

If the fingertip almost falls in the expected “keystroke” position, the result turns out to be more accurate. However, when fingertip falls in the position between two “keys”, it is difficult to tell which one is expected. And only if little part of fingertip falls in excepted “keystroke”position, the detected “keystroke” becomes actually an unexpected “key”.

Only if the fitting ellipse center falls in expected “keystroke” position can we obtain the right result. Thus, calculation of the ellipse fitting on the center point can effectively reduce the recognition errors. Besides, a higher accuracy of ellipse minimum quadratic fitting also guarantees virtual keyboard system to reach the demand of practical usage.

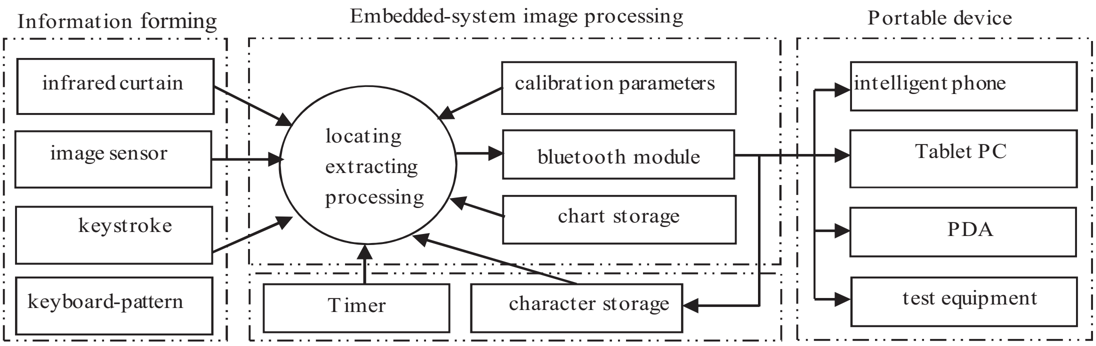

Fig. 8. System information flow chart and core function of main controller

IV. CONCLUSIONS

This paper proposes a virtual keyboard system based on embedded system and image processing. It realizes reliable keystroke gesture recognition and information extraction. The system achieves a high level of images capturing and processing. In terms of real time property, the system reaches 30 fps for keystroke recognition; meanwhile the accuracy rate of keystroke recognition reaches a level of 95%; besides, the start-uptime is no more than 10 s. Experiment results show that the virtual laser keyboard can meet the practical needs with decent results and replace the traditional mechanical keyboard.

REFERENCES

[1] Kölsch M, Turk M,"Keyboards without keyboards: A survey of virtual keyboards, " Workshop on Sensing and Input for Media-centric Systems, Santa Barbara, CA. 2002.

[2] Sarcar S., Ghosh S., Saha P. K., Samanta D. "Virtual keyboard design: State of the arts and research issues,".2010 IEEE Students' Technology Symposium (TechSym), Apr. 2010,pp. 289-299.

[3] Taek young Kwon, Sarang Na, Sang-ho Park., "Drag-and -Type: A new method for typing with virtual keyboards on small touchscreens," IEEE Transactions on Consumer Electronics, , 2014, 60(1): 99-106.

[4] Dhillon H.S., Singla, R., Rekhi N.S., Jha R,"EOG and EMG based virtual keyboard: A brain-computer interface,"2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2009), Aug. 2009, pp. 259-262.

[5] Usakli A.B., Gurkan S. "Design of a Novel Efficient Human-Computer Interface: An Electrooculagram Based Virtual Keyboard,"IEEE Transactions on Instrumentation and Measurement, 2010, pp.2099-2108.

[6] Paul D. Varcholik, Joseph J. LaViola Jr., Charles E. Hughes. "Establishing a baseline for text entry for a multi-touch virtual keyboard,"International Journal of Human-Computer Studies, 2012, pp. 657-672.

[7] Dikovita, C.R.; Abeysooriya, D.P., "V-Touch: Markerless laser-based interactive surface," 2013 International Conference on Advances in ICT for Emerging Regions (ICTer), Dec. 2013: 248-252,

[8] AlKassim, Z., "Virtual laser keyboards: A giant leap towards human-computer interaction," 2012 International Conference on Computer Systems and Industrial Informatics (ICCSII), Dec. 2012:1-5.

[9] Du H, Oggier T, Lustenberger F, et al. "A virtual keyboard based on true-3D optical ranging,"British Machine Vision Conference. 2005, pp.220-229.

[10] Hagara M, Pucik J. "Accuracy of 3D camera based virtual keyboard". 2014 24th International Conference on Radioelektronika (RADIOELEKTRONIKA), IEEE, 2014: 1-4.