Webcam Based Virtual Keyboard Using Shadow Analysis-A Review

Rupali T. Bhole
Department of Electronics and Communication Engineering
SSGBCOE&T, Bhusawal, Maharashatra, India

Prof. Anant S. Bhide
Department of Electronics and Communication Engineering
SSGBCOET, Bhusawal, Maharashtra, India

Abstract-Virtual Keyboard is just another example of today's computer trend of "smaller and faster". Computing is now not limited to desktops and laptops, it has found its way into mobile devices like palm tops and even cell phones. But what has not changed for the last 50 or so odd years is the input device, the good old QWERTY keyboard.
Alternatives came in the form of handwriting recognition, speech recognition, abcd input (for SMS in cell phones) etc. But they all lack the accuracy and convenience of a full-blown keyboard. Speech input has an added issue of privacy. Even folded keyboards for PDAs are yet to catch on. Thus a new generation of virtual input devices is now being paraded, which could drastically change the way we type.
The Virtual Keyboard has been implemented in a number of different forms, based on 3-D optical ranging, gloves, rings, hand gestures based and projection based devices but practical implementation of these devices is not feasible.
Our proposed Virtual Keyboard, being a small, handy, well-designed and easy to use application, turns into a perfect solution for cross platform multilingual text input by using shadow analysis which includes finger shadows to detect finger touches.

Keywords–Human-computer interaction (HCI), Virtual Keyboard (VK), Shadow analysis (SA), edge detection (ED), shadow extraction (SE)

I. INTRODUCTION

Human-computer interaction (HCI) is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them. In these interaction, the input device as interaction way has been replace to a non-physical, unknown device. Keyboard is an input device that has been used since the beginning of computer technology. Its ability to express character with a great speed and accuracy makes this input device has consistent and stable usage. Keyboard has a long development history that consists of the port connector change, size change, and some of it is the change in the input media. To realize a new revolutionary way in interaction, the main idea is by using webcam as a cheap and easily obtained input media. By using webcam as media, it is possible to interact with computer without a real physical-interaction device. Webcam media has many advantages such as the existence of many graphics processing algorithm that can help computation. The webcam is also a very common device, making the development for webcam can progress greatly, as people can easily try and continue the development easily. The main problems in webcam virtual keyboard is how the webcam will be used, and the algorithm used to transform the typing gesture to a keyboard input while still maintaining the comfort that can be found in the physical keyboard. Webcam is an applicable media for virtual keyboard even in its limitation with the help of fast image processing method.

In our implementation, we use a Logitech Webcam and a sheet of paper with the keyboard printed on it. The only unique aspect of the keyboard is that it has four colored endpoints which are used to identify the keyboard. The implementation is based on use of image processing. The objective of this paper is to develop a Virtual Keyboard (VK), using shadow analysis. Traditional QWERTY keyboards provide a minimal but functional interface. However these keyboards are bulky and offer very little in terms of enhancements. In this age of miniaturization, where the size of laptops and desktops is becoming smaller, the traditional keyboard difficult to further miniaturization. Our proposed Virtual Keyboard (VK), with its minimal physical form can provide a solution to this problem.

The rest of the paper is organized as follows. Section II represents our proposed work according to the problem definition. Proposed algorithm has been presented in section III. Concluding remarks are given in section IV.

II. PROPOSED WORK

Input to small devices is becoming an increasingly crucial factor in development for the even more powerful embedded market. Speech input promises to become a physical alternative to tiny keyboard, yet its limited reliability, robustness and flexibility. Render it unsuitable for certain task or environment. Various attempts have been made to provide the common keyboard metaphor without a physical keyboard, to build “Virtual Keyboard”. The Virtual Keyboard has been implemented in a number of different forms, as described by Kölsch, M. and Turk, M [1]. It also gave an overview of the range of input devices and methods for alphanumeric data touch typing as input method and highlighted its benefits. Of these, the ones based on 3-D optical ranging described by H. Du, T. Oggier, F. Lustenburger and E. Charbon [2] .They describe system consists of a pattern projector and a true-3D range camera for detecting the typing events and exploit depth information acquired with the 3D range camera and detect the hand region using a pre-computed reference frame. The fingertips are found by analyzing the hands' contour and fitting the depth curve with different feature models. To detect a keystroke, analyze the feature of the depth curve and map it back to a global coordinate system to find which key was pressed. These steps are fully automated and do not require human intervention. CCD cameras [3] are most significant as they are based primarily on image processing The VK designed makes use of a single CCD camera. Even more significant is the work presented in [4] introduces Play Anywhere, a front-projected computer vision-based interactive table system which uses a new commercially available projection technology to obtain a compact, self-contained form factor, where we propose a system based on a shadow analysis to detect finger touches in our implementation of the Virtual Keyboard. The detailed working along with the execution methodology has been explained deeply in next section.

III. PROPOSED ALGORITHM

In this paper a shadow based analysis is used to acquire depth information from a 2-D image. The algorithm shows the steps for design of this project. [1]The VK implementation flow is shown in the flow graph. Initially, the keyboard is detected using reference points. Once the keyboard is detected, the user's hand is detected using color segmentation. This process is also useful in extracting the hand's shadow. If the hand is detected, the process flow moves onto edge detection which is used to identify the finger tips.

A. *Detection of Keyboard-*

The endpoints of the VK are identified using color differentiation. The endpoints of the keyboard are colored blue, thus on thresholding, these points can be easily identified. The area of interest i.e. the location of the VK in the image is then defined.

B. *Detection of Hand-*

Initially, a large collection of hand images was created. These images were observed for their RGB (Red Green Blue) values in the area of interest, i.e. the hand. It was noted that, in these hand regions the red component was higher than the other two components. These observations were consistent across all the images that were tested. There may be certain regions on the hand such as the fingernails and/or veins where this observation might not be true. However, this does not affect the overall result as a significant portion of the hand follows the expected pattern and the hand is appropriately detected. In order to remove the abnormalities in the hand regions, we use image enhancement techniques. Finally, the detected hand regions are threshold to white, while the rest of the image is made black

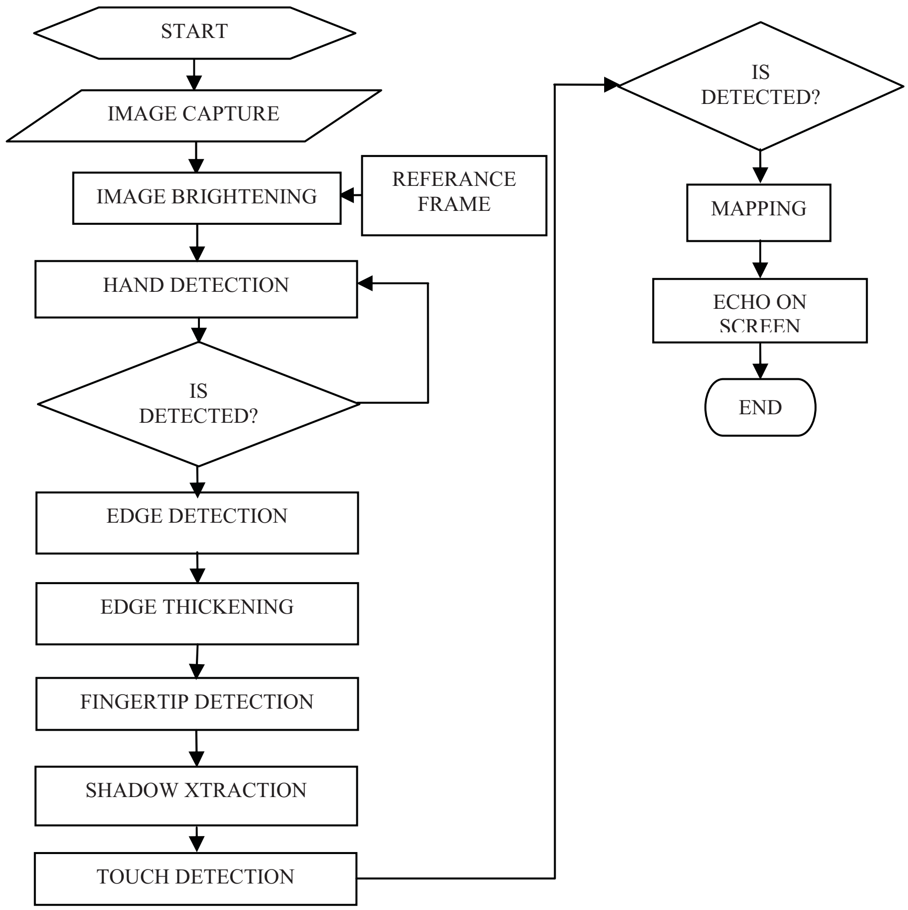

Figure 1. Flowchart for implementation of Virtual Keyboard

C. Detection of Edge-

For edge detection we use the Sobel technique, which is found to give better results when compared to the Canny, Prewitt and Zero-Cross. Edge detection of the hand is required in order to detect the fingertips. The edge obtained is then thickened in order to remove discontinuities [9]. This enables a complete traversal of the hand edge which is described in the next step. Edges characterize boundaries and are therefore a problem of fundamental importance in image processing. Edges in images are areas with strong intensity contrasts a jump in intensity from one pixel to the next. Edge detecting an image significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image. There are many ways to perform edge detection. However, the majority of different methods may be grouped into two categories, gradient and Laplacian. The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image. The Laplacian method searches for zero crossings in the second derivative of the image to find edges. An edge has the one-dimensional shape of a ramp and calculating the derivative of the image can highlight its location.
In digital image, the so-called edge is a collection of the pixels whose gray value has a step or roof change, and it also refers to the part where the brightness of the image local area changes significantly.

The advantage of Sobel edge operand is its smoothing effect to the random noises in the image. and because it is the differential separated by two rows or two columns, so the edge elements on both sides have been enhanced and make the edge seems thick and bright. Sobel operator is a gradient operator. The first derivative of a digital image is based on a variety of two-dimensional gradient approximation, and generates a peak on the first derivative of the image, or generates a zero-crossing point on the second derivative. Calculate the magnitude and the argument value of the image horizontal and vertical first-order or second-order gradients, at last calculate modulus maxima along the angular direction and obtain the edge of the mage. But when the mage has lots of white Gaussian noises, it is very difficult to get the peak value of the first derivative; the reason is because that the noise points and the useful signals mix up.

However, the majority of different methods may be grouped into two categories: Gradient based Edge Detection: The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image. Laplacian based Edge Detection: The Laplacian method searches for zero crossings in the second derivative of the image to find edges. An edge has the one-dimensional shape of a ramp and calculating the derivative of the image can highlight its location [12]. An edge in a digital image is a boundary or contour at which a significant change occurs in some physical aspect of an image, such as the surface reflectance, illumination or the distances of the visible surfaces from the viewer. Changes in physical aspects manifest themselves in a variety of ways, including changes in color, intensity and Texture Edge always indwells in two neighboring areas having different grey level. It is the result of grey level being discontinuous. Edge detection is a kind of method of image segmentation based on range non-continuity. Image edge detection is one of the basal contents in the image processing and analysis, and also is a kind of issues which are unable to be resolved completely so far [13]. When image is acquired, the factors such as the projection, mix, aberrance and noise are produced. These factors bring on image feature is blur and distortion, consequently it is very difficult to extract image feature. Moreover, due to such factors it is also difficult to detect edge. The method of image edge and outline characteristic's detection and extraction has been research hot in the domain of image processing and analysis technique. Detecting edges is very useful in a number of contexts. For example in a typical image understanding task such as object identification, an essential step is to an image into different regions corresponded to different objects in the scene. Edge detection is the first step in the image segmentation. Edge feature extraction has been applied in many areas widely. In order to gain more legible image outline, firstly the acquired image is filtered and denoised. In the process of denoising, wavelet transformation is used. And then different operators are applied to detect edge including Differential operator, Log operator, canny operator and Binary morphology operator. Finally the edge pixels of image are connected using the method of bordering closed. Then a clear and complete image outline will be obtained.

As we all know, the actual gathered images contain noises in the process of formation, transmission, reception and processing. Noises deteriorate the quality of the image. They make image blur. And many important features are covered up. This brings lots of difficulties to the analysis. Therefore, the main purpose is to remove noises of the image in the stage of pre-treatment. The traditional denoising method is the use of a low-pass or band-pass filter to denoise. Its shortcoming is that the signal is blurred when noises are removed. There is irreconcilable contradiction between removing noise and edge maintenance. Yet wavelet analysis has been proved to be a powerful tool for image processing [15]. Because Wavelet denoising uses a different frequency band-pass filters n the signal filtering. It removes the coefficients of some scales which mainly reflect the noise frequency.

Then the coefficient of every remaining scale is integrated for inverse transform, so that noise can be suppressed well. Wavelet analysis widely used in many aspects such as image compression, image denoising. Edge is a part of an image which contains significant variation of intensity. Edges describe the boundaries between an object and the background in an image, which helps in segmentation and object recognition [16]. Edge detection is useful in image segmentation which identifies whether a line or an edge is present or not and depicts them in a suitable way. It is defined as the process of identifying and locating sharp discontinuities, boundaries of objects or textures depicted in an image (i.e. edges). These discontinuities are rapid variations in pixel intensity which characterize objects boundaries in an image [17]. The goal of edge detection is to extract the important features like lines, corners, curves etc. from the edges of an image. The main motive of edge detection is to discard the unnecessary information and preserves the essential information and thus it reduces the amount of data which needs to be processed. The important characteristics of edges are position of subarea, amplitude and direction.

The segmentation simplifies and changes the illustration of a scene into more meaningful way and hence it is easier to detect the image. Discontinuity and similarity are generally the two basic properties on which image segmentation algorithm is based upon [18]. There are an extremely large number of edge detection operators available like Sobel, Robert, Prewitt, Zero cross, Canny etc. Structure of edge, direction of edge and noise conditions are the factors which are involved in selection of different edge detection operators. Poor concentration and refraction effects can moderately change the object boundaries which can lead to problems like noise susceptibility, false edge detection and high computational time. With the help of edge detection, more perceptible information can be taken out by knowing the physical and geometrical changes in an input image. Edge detection is most common approach for finding discontinuities in amplitude values. By applying first and second order derivatives, these discontinuities can be identified.

The edge detection can be categorized in two different parts as follows: Gradient Edge Detection: Gradient is the first order derivative used in digital image processing. Edges are detected by looking for the maximum and minimum value in the first derivative of the image. E.g. Sobel, Robert. Laplacian Edge Detection: The second order derivatives in image processing are generally computed using Laplacian. To find the edges in an image, the Laplacian method looks for zero crossings (i.e. places where the sign of Laplacian changes) in the second order derivative [11]. An edge is the one-dimensional ramp shaped and the position of edge can be identified by computing the second order derivative of the image. E.g. Zero cross.

These steps in edge detection are described as follows:

Filtering: Images are generally distorted by some haphazard variations in intensity values, known as noise. Various types of noise like Gaussian, impulse, salt and pepper noise frequently occur in an image. Filtering or smoothing suppresses as much noise as possible, without destroying the edges.

Enhancement: Enhancement emphasizes upon pixels where there is a significant variations in local intensity values and is generally done by calculated the magnitude of gradient vector.

Detection: It detects the edges by using the thresholding criteria that which edge pixels should be discarded as noise and which should be retained.

Localization: It identifies the proper position of an edge in an image. Edge linking and edge thinning are usually essential in localization.

D. Detection of Tip-

In this stage we intend to find all the finger tips maximum of 10 finger tips present in the image taken by the webcam. The hand edge obtained above is a thin single lined edge which has lots of discontinuities. These discontinuities make it difficult to traverse along the edge of the hand. To overcome this problem, the thin single lined edge is dilated using a structuring element to give a thick edge. On this thickened edge our algorithm is run to find the finger tips. We use the following priority structure as given in reference paper while traversing the edge to determine the next pixel to be tested for touch. Edge thickening is essential as the edge obtained using Sobel may not always be perfect and continuous.

8	2	5
4		1
7	3	6

(a)

7	3	6
4		2
8	1	5

(b)

Figure 2. (a) Priority structure 1 for moving down hand edge. (b) Priority structure 2 for moving up hand edge

E. Extraction of Shadow-

Ideally, background subtraction should detect real moving objects with high accuracy and limiting false negatives (not detected) as much as possible. At the same time, it should extract pixels of moving objects with maximum possible pixels, avoiding shadows, static objects and noise. In the detection of shadows the foreground objects are very common, producing undesirable consequences. For example, shadows connect different people walking in a group, generating a single object (typically called blob) as output of background subtraction. Background subtraction is a popular technique to segment out the interested objects in a frame. This technique involves subtracting an image that contains the object, with the previous background image that has no foreground objects of interest. The area of the image plane where there is a significant difference within these images indicates the pixel location of the moving objects. These objects, which are represented by groups of pixel, are then separated from the background image by using threshold technique.

The original image captured from the camera is dynamically brightened depending on the lightning conditions. This eliminates all background 'noise' such as shadows due to multiple light sources and keyboard's characters; leaving behind only the hand and the umbra of the shadow. The hand obtained is then subtracted from this image. This gives us only the shadow of the hand.

F. Detection of Touch-

A small region around the fingertips is scanned for shadow; the tips are white while the shadow regions are black. The ratio of the white to black pixels is determined. If the ratio of the area of non-shadow region to the area of the shadow region exceeds a particular threshold we can say that touch has occurred. The threshold is determined using an extensive set of test cases.

G. Mapping-

The final step involves mapping of the fingertips to actual keys. This is a simple 2D mapping based on the information available in the current frame (x, y coordinates) and relative position of the fingertip from the endpoints of the keyboard.

IV. CONCLUSION

The Virtual Keyboard that we propose uses only a standard web camera, with no additional hardware. This paper addresses problems with current Virtual Keyboard implementations and describes a novel technique, namely shadow Analysis to solve these problems. The objective of this paper is to develop a Virtual Keyboard (VK) using only a standard 2D camera without the need for additional specialized hardware.

A virtual keyboard for mobile devices will remove the inherent space constraints and would therefore provide for a full sized keyboard without additional hardware. In implementation, a Logitech Webcam is used and a sheet of paper with the keyboard printed on it. The only unique aspect of the keyboard is that it has four colored endpoints which are used to identify the keyboard. The implementation is based on use of image processing. We propose Shadow Analysis for detection of webcam based Virtual Keyboard (VK). We introduce detection of edge by simple way like Sobel technique.

The VK presented here is only a small application of a larger idea which detects finger touches using a standard 2D camera. Finally we expect that using Image processing implementation of webcam based Virtual Keyboard would enable us to use a full sized QWERTY keyboard without the need for additional physical space or hardware. Moreover, the VK can find applications in gaming, 3d modeling etc

REFERENCES

[1] Y.A , J. Gosalia, A. Kanade, H. Mehta, Prof. N. Shekokar," Virtual Keyboard Using Shadow Analysis",' Third International Conference on Emerging Trends in Engineering and Technology',2010,PP 163-164

[2] B SorioHernanto and IpingSuprianaSuwardi, "Webcam virtual keyboard", 'International conference Indonesia', July 2011, PP 17-19.

[3] Kölsch, M. and Turk, M. Keyboards without Keyboards: A Survey of Virtual Keyboards, Workshop on Sensing and Input for Media-centric Systems, Santa Barbara, CA, June 20-21, 2002

[4] H. Du, T. Oggier, F. Lustenburger and E. Charbon, "A virtual keyboard based on true-3D optical ranging," Proc. British Machine vision Conference (BMVC), Oxford, pp. 220-229, Sept. 2005.

[5] Matsui, N. & Yamamoto, Y. A New Input Method of Computers with One CCD Camera: Virtual Keyboard. Keio University, 2000.

[6] Andrew D. Wilson, PlayAnywhere: a compact interactive tabletop projection-vision system, Proceedings of the 18th annual ACM symposium on User interface software and technology, October 2326, 2005, Seattle, WA, USA

[7] Segen, J., and S. Kumar, Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera in Proceedings of Computer Vision and Pattern Recognition, (1999), 479-485.

[8] Mantyjarvi J., Koivumaki J., Vuori P., "Keystroke recognition for virtual keyboard", Multimedia and Expo, 2002. IEEE International Conference on Volume 2, Issue, 2002 Page(s): 429-432 vol.2.

[9] Samata Gupta and Sumit Ghosh,"Sobel edge detection algorithm", 'International journal of computer science and management Research',2 Feb 2013,vol 2.

[10] Raman Mainiand Dr. Himanshu Aggarwal," Study and Comparison of Various Image Edge Detection Techniques",' International Journal of Image Processing (IJIP)', Volume (3) : Issue (1) ,PP 1

[11] E. Argyle, "Techniques for edge detection",' Proc. IEEE', vol. 59, 1971, PP 285-286.

[12] M.Kalpana, G.Kishorebabu and K.Sujatha," Extraction of Edge Detection Using Digital Image Processing Techniques",'International Journal Of Computational Engineering Research',vol 2, September 2012,PP 1562

[13] Lei Lizhen," Discussion of digital image edge detection method", 'Mappingaviso', 2006, PP 40-42.

[14] Ma an and Zhang Zhihui,"Several edge detection operators comparation, Industry and mining automation", 2004,PP 54-56

[15] Lai Zhiguo, "Image processing and analysis based on MATLAB",'Beijing: Defense Industry Publication', 2007, PP 4

[16] EhsanNadernejad,"EdgeDetection techniques: Evaluations and Comparisons", Applied Mathematical Sciences, 2008

[17] Raman Maini, Dr.Himanshu Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques", International Journal of Image Processing (IJIP)

[18] Gautam Appasaheb Kudale ,Mahesh D. Pawar,"Study and Analysis of Various Edge Detection Methods for X-Ray Images", International Journal of Computer Science and Application, 2010.