Simple Gesture Recognition using a PIR Sensor Array

(PREPRINT)

***Piotr Wojtczuk, **Alistair Armitage,**
***T. David Binnie, ***Tim Chamberlain**

*School of Engineering and the Built Environment
**School of Computing
Edinburgh Napier University
Edinburgh,
EH10 4DT, UK
Tel.: +44 (0) 8452 60 60 40
p.wojtczuk@napier.ac.uk,
a.armitage@napier.ac.uk, td.binnie@napier.ac.uk,

*** Scottish Microelectronics Centre
The Kings Buildings
West Mains Road
Edinburgh
EH9 3JF, U.K.
Tel.: +44 (0) 131 650 7009
tim.chamberlain@pyreos.com

Received: /Accepted: /Published:

Abstract: Most current work on gesture recognition focuses on the analysis of complex video sequences. In this paper we present an alternative approach that is intended for simple gesture control using a relatively inexpensive pyroelectric array detector. The detector is manufactured using standard wafer processing techniques. It consists of a 16 element passive infrared sensor array that responds to changing infrared signals, such as are generated by a hand moving at a distance of some tens of centimetres in front of the array. There is quite a large variation in the responsivity of the pixels within the array, but despite that it is relatively easy to use differential signals from the array or to apply a simplified version of an image processing algorithm to track movement in front of the detector. We have developed a prototype system that can recognise hand movements in different directions in front of the detector. This has allowed us to develop a demonstrator system that can be used to control, for instance, a PowerPoint presentation by gesture.

Keywords: touchless interfaces; dynamic hand gesture; gesture recognition; infrared motion sensors; passive infrared sensors

---

1. Introduction

Non-contact operation of electrical appliances is becoming increasingly desirable. This is particularly the case in public areas where hygiene is important. Public washrooms increasingly have taps, cisterns, soap dispensers, hand dryers and doors activated by simple hand proximity/movement detectors. The advantages of removing physical contact from the activation and control of devices range from limiting the chance of infections transmitted through touch and better energy efficiency to simple convenience. In other environments such as medical facilities and food processing areas, strict hygiene rules are essential and direct hand contact with appliances is best avoided if at all possible [1]. These environments present a wide spectrum of opportunities for non-contact control, as reducing the number of touched surfaces positively suppresses risk of cross contamination and infection. Strict rules apply in operating rooms, for example, where doctors' hands must remain sterile throughout a procedure [2]. There are other, rarer cases where a cultural or individual preference may be not to make physical contact with an appliance. In all these circumstances touch-free control is beneficial.

Existing non-contact hand proximity/motion switches provide the binary functionality of an on/off operation. They are commonly based in single or dual cell pyroelectric infrared sensors with a simple lens system [3], although capacitive sensing can also be used [4]. Thus the degree of human control over the electrical appliance is severely limited.

In this paper details are given of a simple pyroelectric array device that detects simple hand movement gestures thereby allowing greater non-contact control capability with little increase in size, cost or power consumption. Numerous touchless alternatives to existing tactile switches are envisaged. This paper gives an overview of existing methods of gesture detection. Further, the problem of hand gesture recognition is discussed and our implementation is explained.

2. Related Work

Gesture recognition systems can be categorized into three broad groups [5]: those that use a handheld pointing device, those that require wearable sensors and those with no user devices but operate using the analysis of 2D images (this can be generalised as external sensing). For general application of gesture recognition sensors, it is important that the system requires no user held or worn devices.

Research into gesture feature extraction from image sequences is the most active. Most common implementation uses a video camera to capture a wide scene, from which hands are extracted and tracked, for example [6] and [7]. Such approach gives great flexibility in terms of detection distance and detection precision and scope. Commercially, systems operating on a video processing basis have been developed by GestureTek [8], EyeSight [9] and Edge 3 Technologies [10]. However, it is not practical to implement such systems ubiquitously, for example as switches, due to size, cost and, in many cases, energy consumption.

Byung-Woo Min et al. [11] classify hand gestures into static (hand posture) and dynamic (hand movement). Dynamic gestures are generally more intuitive for the user, as the direction of motion or its shape can reflect the resulting action, e.g., up/down hand stroke for increasing/decreasing brightness in a room. Moreover, it is easier to capture this type of gesture with a low resolution and low cost imager.

Sensor based touchless solutions have become more popular after the recent success of touch screen interfaces. Mechaless [12] provide a hardware solution. They use an active near infrared arrangement with multiple IR LEDs and a sensor that is used to calculate the distance of an object from each LED by timing pulse to reception interval. They propose use of their touchless interface in wall-mounted gesture switches (increase / decrease a control by up / down hand movement) or in a remote control with a touchless ‘joystick’.

One of many NIR proximity array systems is being developed by Dongseok et al. [13]. A 3-dimensional map of the scene is created by analyzing reflections of emitted near infrared light captured by an infrared array. Each pixel can read range, as with simple infrared range-finders. Interestingly, the research group develops a near-field touchless, free air touchpad-like device, where position of a finger in front of the sensor is tracked precisely and analysed.

An ultrasonic realization of a touchless pointing interface is proposed by Chiu et al. [14]. In this case an array of ultrasonic receivers creates an image-like output based on reflections received from a pointing object.

The related idea of interactive shop windows is introduced by Krumm [15]. Interactive or immersive advertising reacts to people passing by, trying to engage them by showing dynamic content accordingly.

Active arrangements, comprising an emitter and a receiver, usually have the advantage of providing absolute measurements of distance and position. However the presence of the emitter results in higher energy needs than for a passive measurement. For dynamic hand gestures such precision is not necessary and detection can be realised by passively sensing thermal infrared radiation change only.

We believe that our solution is very well suited to the task of near distance, dynamic hand gesture recognition application and provides an alternative to existing propositions, that is more cost and energy efficient due to reduction in number of components or their complexity.

3. PIR Array Based Hand Motion Recognition

The set of basic hand movements in front of a sensor is limited, e.g., horizontal, vertical and diagonal swipes across the field of view, circular motion and in and out movements. At this stage, only linear motion across the field of view is considered. Four directions of swipe were chosen as easy to describe and reproduce: up / down, left / right.

3.1 Design Principle – PIR Array

As identified in the introduction to this paper, a simple, low cost and low energy consumption device to extend the concept of a touchless interface is desirable in many environments.

A PIR Array is an interesting potential solution: the technology has been widely proven in settings such as burglar alarms and automatic light switches. However, only now has it become financially viable for ubiquitous application in the form of an array (previously mostly used a single or dual element occupancy sensor). Dual element detectors have an advantage in that the pixel outputs can be used to generate a differential signal: useful in reducing the effects of drift and noise. This principle is taken further in the idea of the quadrant detector. More widely used in optical detection, the signal from four quadrants is coupled differentially to produce X and Y signals directly proportional to the position of a spot image on the detector.

Passive has several meanings in this type of application. There is no emitter required for illumination of the scene, as the human body is naturally a good thermal infrared emitter [16]. Scenes a temperature above 0 K may be dark in visible spectrum, while still full of thermal energy [17].

A limitation common to all passive infrared detectors exists. The minimum detectable temperature difference produces a band of hand temperature relative to the background temperature that produces too low a signal to noise ratio. When the temperature of the detected object is higher than that of the background, the response of the infrared elements will be positive. When the temperature of the object is lower, the response is inverted. If the temperature of the environment is close to the temperature of the hand, the signal from the sensor can drop to .

The passive array responds to temperature differences; in fact the signal varies with the speed of temperature change. This means that the speed of the swipe motion has an effect on the signal size. This is potentially an interesting area for further work, as the repertoire of gestures that can be recognised can be expanded if the velocity of motion can be measured. However, this is not an easy measure to make: from the detector's point of view, it is the angular velocity that affects the signal size. Therefore it depends not just on the velocity of the gesture, but also on the distance between the detector and hand. For this reason it may be easier to recognise other gestures, such as a pause in the middle of a swipe, since this is not dependent on the separation of source and detector.

3.2 Research Questions

The primary aim of this study is to investigate whether it is feasible to use a passive infrared array for hand gesture recognition, and if so, under what conditions, and what range of gestures can be recognized. The secondary aim is to compare the cost effectiveness and power efficiency of the passive infrared system with existing solutions using active infrared or visible cameras.

4. Implementation

We propose a system to realize the task of simple dynamic hand gesture recognition at short distances.

4.1 Prototype

The implementation described in this paper is a proof-of-concept prototype. Existing signal conditioning integrated circuit and PC interface board and software provided by the sensor manufacturer are used for signal capture. Signals are processed using the computer for ease of development. In future developments a migration to an embedded system is anticipated in order to provide a low footprint, low power solution.

4.2 Capture Device

A thin-film, passive infrared sensor, 4 pixel x 4 pixel array is enclosed in a package with a broadband infrared filter. This thin-film sensor is manufactured using standard, high throughput wafer processing techniques. Therefore a sensor die is a customisable MEMS structure with pixels patterned on a membrane. Such an approach provides an integrated, robust array with consistent physical properties.

Thin-film infrared elements have low thermal mass compared with bulk PZT elements. The result is higher optimal frequency of operation, well suited for tasks such as gesture recognition, where a detected object moves quickly across the field of view.

An infrared-blocking film is mounted on the underside of the filter, with a pinhole centred above the array, as seen in 0

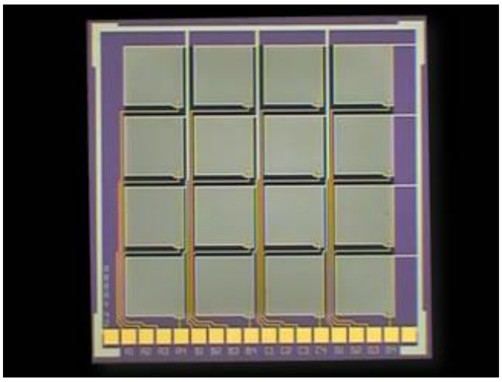

Fig. 1.4x4 PIR Sensor Array (Source: [18]).

This optical arrangement produces a low resolution infrared image of the scene in front of the sensor within a field of view of 60^{o} up to 20 cm away from the sensor.

Electrical signals from the infrared array are amplified, low pass filtered and digitised. The images are captured using a dedicated interface board. The digitized frames are sent to a computer and pre-processed to provide a stable and uniform output. This involves per-pixel low pass filtering and normalisation of signals across the array.

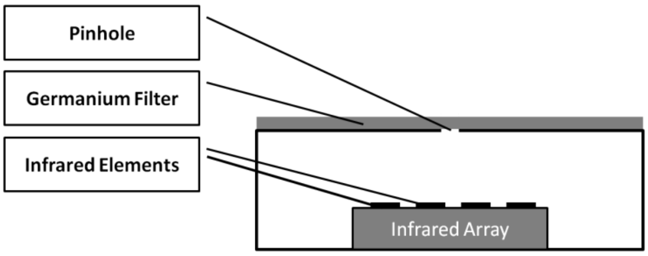

Fig. 2. Optical arrangement.

4.3 PIR Sensor Array

The PZT infrared array sensing elements transform thermal radiation into electrical charge based on the pyroelectric principle.

The transducer operates in the same manner as standard security PIR motion detectors, where the change in incident flux arises from the motion of a thermal infrared source across the field of view. This has an advantage in that the background subtraction is inherent in the transducer itself: stationary infrared emitters do not cause a change in incident flux, and therefore are not detected. However, it also becomes a limitation in that an object that moves into the field of view and then becomes stationary will soon become invisible. This is not a problem in a system designed to recognize simple ‘swipe’ type gestures where motion, and the direction of the motion, are the dynamic features that are to be recognized.

4.4 Pixel performance

Each pixel is a pyroelectric detector sensitive to changes in incident IR radiation. In testing, the change in temperature is provided by using a chopper to mechanically interrupt the IR beam from a calibrated source. In the application area, the change in radiation comes from the motion of the target: a stationary target cannot be detected. However, in the intended application of simple gesture recognition, this is not a problem as it is dynamic gestures that are being detected.

As a hand is swept in front of the detector, the infrared radiation first increases then decreases (assuming the hand is at a higher temperature than the background). The pixels are sensitive to the differential of the radiation with respect to time; this leads to a double peak, first in one direction then the other. In the system we used, the raw signals after initial amplification appear inverted. This is simply an artefact of the electronics and has no effect on the measurements. The signal is digitised with 10 bit resolution, leading to an output of a number in the range 0-4095. In theory, the output with no changing input should be in the middle of this range, but offsets in the electronics mean that the quiescent level may differ from this slightly. Figure [3] shows the output of a single pixel as a hand is moved in front of the detector. The inversion means that the first peak is negative as the hot target first appears; this is followed by a positive peak as the hot target moves away, leaving the pixel exposed to the cooler background.

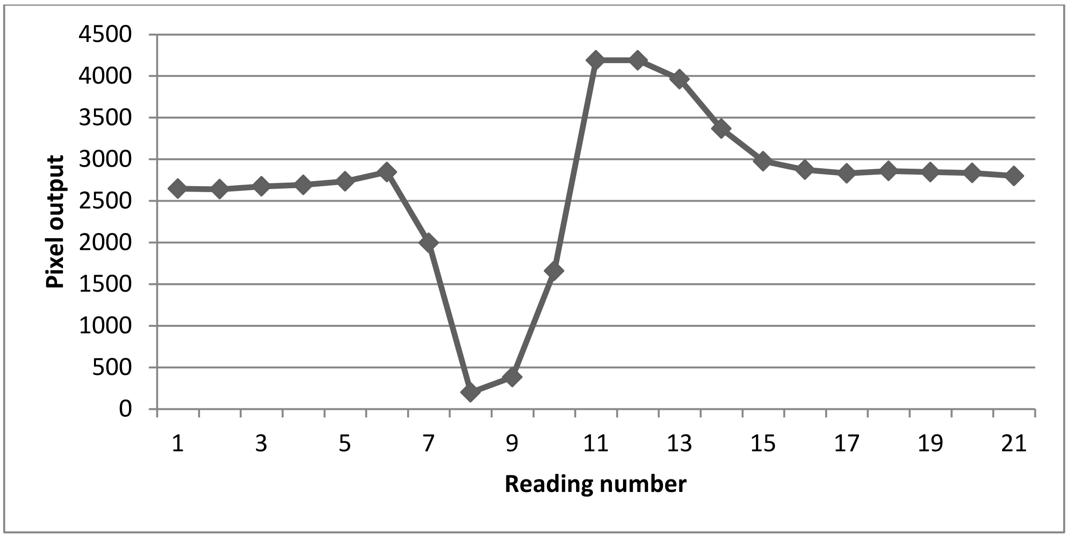

Fig. 3. Output from a single pixel in response to a hand swipe

4.5 Array characterisation

The thin-film sensor array has the advantage that it is manufactured using standard wafer processing techniques. This will result in a device that can be manufactured in large quantities at low cost. However, it also means that there are some limitations in the performance of the sensor, both in terms of the sensitivity, but also in the amount of noise and other error sources.

The first limitation of the device is in the variation between pixels. This is largely caused by the thermal conduction from the pixels at the edge to the surrounding substrate. This has the effect of reducing the sensitivity of the edge pixels. Furthermore, the fact that one edge of the detector is the bond-pad side of the die which has a wider frame and hence acts as a 'stronger' heatsink, means that the thermal characteristics are not symmetrical round the four edges of the square array. This can be seen in figure [4], which shows the relative sensitivity of the 16 pixels in an array. These measurements were made using a calibrated thermal source at a temperature of 500K, chopped at 10 Hz.

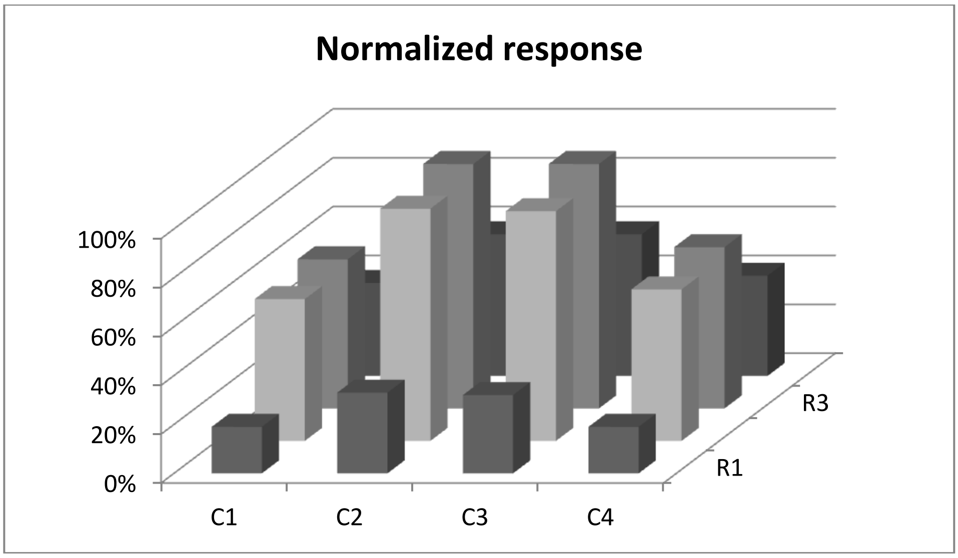

Fig. 4. Relative sensitivity of pixels in one array

There is a much smaller variation in the noise than in the sensitivity. For the same pixels as in figure [4], the variation of the normalised noise was at most 18% (with a standard deviation of 5%), considerably less than the variation in the sensitivity. The variation in the sensitivity can, to a certain extent, be calibrated out as it is a fixed systematic error. However, there are limits to this as the sensitivity of the edge pixels can be adjusted for with an increased gain, but this also amplifies the noise. The signal to noise ratio of the edge pixels is significantly poorer than for the centre pixels.

As the intended application is the detection of hand motions, it is of interest to look at the response of the array to a hand moved across the detector's field of view. Figure [5] shows the response of three pixels; it demonstrates the relatively poor performance of the pixels at the edge, and particularly the corner.

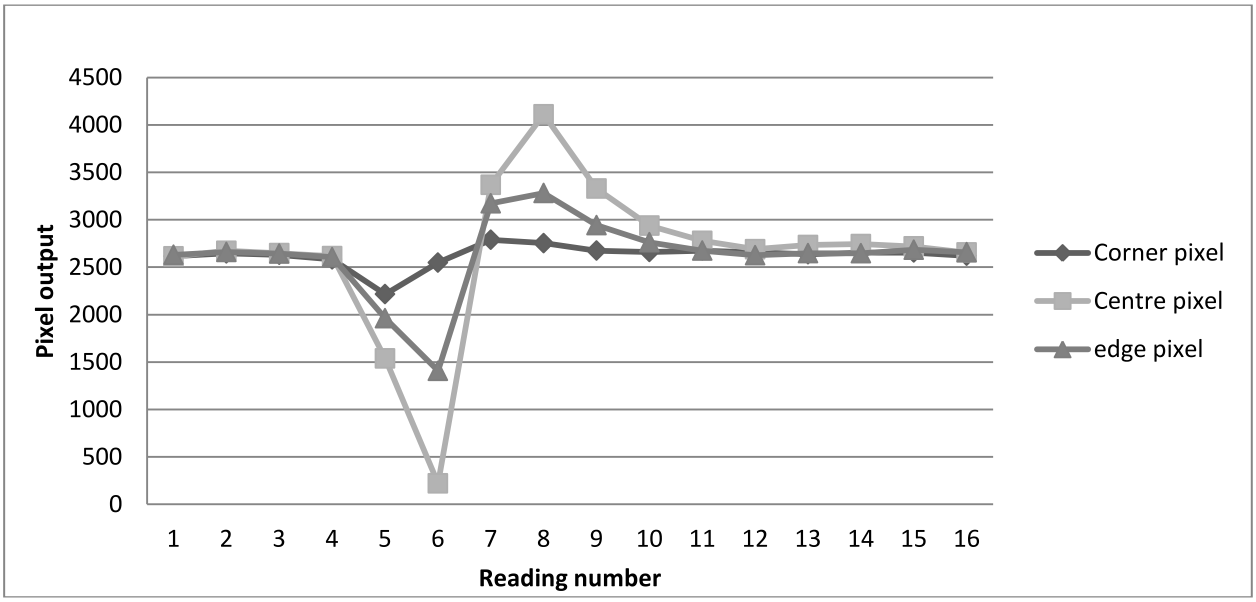

Fig. 5. Variations in pixel response

4.6 Determining direction

The array can be used like a quadrant detector: summing pixels in one section of the array and subtracting pixels in another section of the array to obtain a differential signal. When detecting an image moving across the pixel array, the peaks arrive in different sections of the array at different times. Using the detector as a quadrant detector means that there is a sequence of peaks: + - + for one direction, - + - in the other direction. This can be seen in Fig. 6, where the sum of the right hand pixels is subtracted from the sum of the left hand pixels.

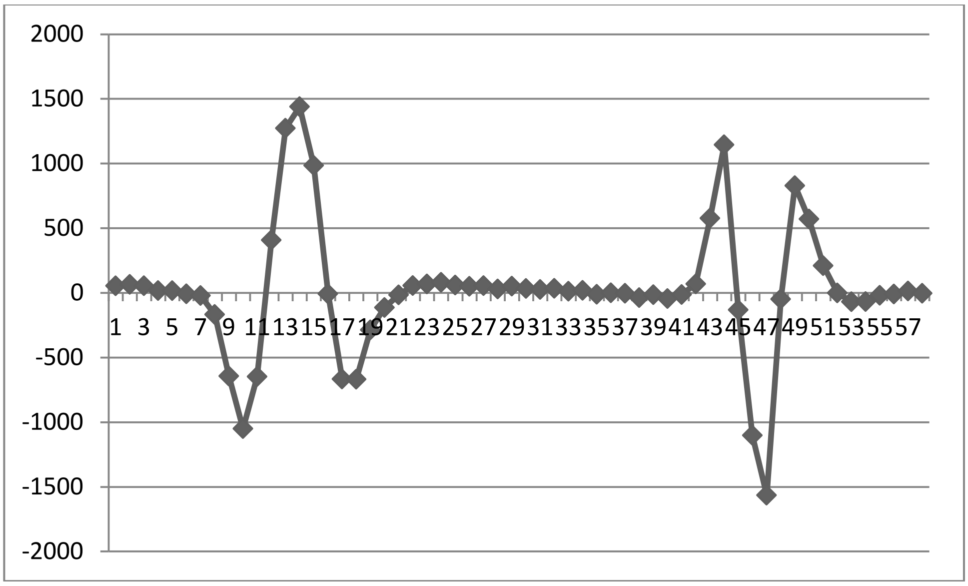

Fig. 6. Using the detector in differential mode to detect a hand swipe moving backwards and forwards across the detector.

Figure 6also shows that there is variability in the signal generated by the movement of a hand in front of the detector (the second peak is not just an inverted form of the first peak). The amount of the change in infrared radiation incident on the detector depends on a number of factors: the relative temperature difference of the hand and the background, the speed of movement, the distance between the array and the hand and the hand's emissivity all affect the signals generated by the array.

4.7 Signal Processing

A tracking algorithm is applied to the frame sequence. A blob detection technique is derived from the Automatic Centroid Extractor [20], also used for a people counting application [19]. Adaptation of the algorithm involved increased processing rate and adjustment for larger object size.

Hotspots are localized and tracked in the image. Analysis of the trail of a detected and tracked object allows the determination of direction of motion. The length of the track is used to determine if it was a valid detection, while the x and y difference between starting and finishing point is evaluated to provide direction information.

5. Initial Evaluation

Functionality of the system was tested in four main directions of hand swipe across the sensor using the blob detection software to deduce the direction of motion. Five participants were asked to perform a set of 100 hand movements each (25 per direction). The detection rate and accuracy of the system response (correct direction detected) were recorded.

Figure [7]. demonstrates the test arrangement and correct detection: a sample hand motion from left to right results in a correct 'right' detection.

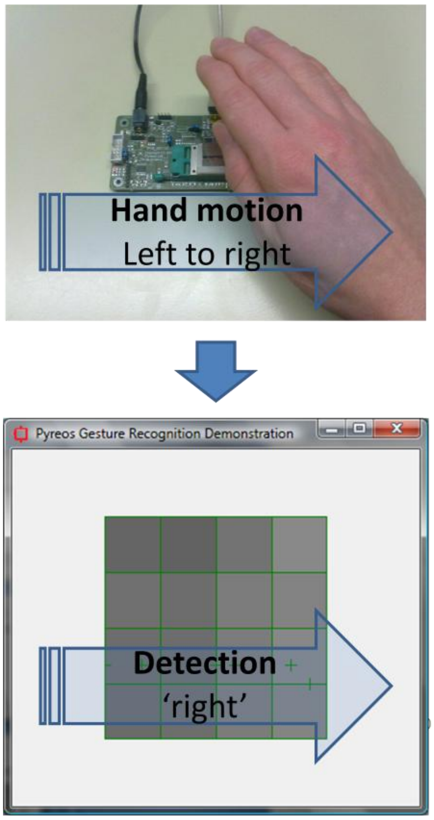

Fig. 7. Hand motion and tracker response [21].

Detection in the distance range of up to 20 cm from the sensor was accurate 92.6% of time overall. 81.4% of gestures were detected at the first attempt. When considering only horizontal swipes (left to right and right to left), it can be seen that the detection rate (87.2%) and accuracy (93.6%) were higher than overall. This is attributed to the test arrangement, where it was more difficult to accurately present clean vertical swipes to the system then it was to produce horizontal ones. This is mostly explained by the mechanics of the arm motion. It is easier to produce a uniform motion side-to-side in the X axis where the bulk of the motion is generated by a simple rotation of the arm at the elbow than motion in the Y direction where the whole arm tends to move towards and away from the body.

Direction	Detection Rate (%)	Accuracy (%)
Up	72.8	93.6
Down	86.4	93.6
Left	88.0	93.6
Right	78.4	89.6
Average	81.4	92.6

Table 1. Initial Evaluation Results

5.1. Application

As an example application of the system, the output of the prototype was used to emulate arrow key strokes on the host computer. This simple test demonstrates the suitability of the system for performing certain tasks on the computer. As an example, the gesture recognizer was used to change slides during a multimedia presentation.

6. Conclusion

With our prototype system we have demonstrated that it is feasible to recognize hand motion direction with a pyroelectric passive thermal infrared array. Our approach shows that the functionality of a simple pyroelectric movement sensor can be expanded to detect differing hand gestures at short range. Several algorithms have been investigated for processing the information from the pixels, either averaging some pixels using a modified version of a quadrant detector, or by using a modified blob detection algorithm adapted from image processing techniques. We have shown that blob detection from a hand waving over a 16 element passive infrared sensor array provides sufficient information to discriminate four directions of hand stroke.

There are a number of limitations to the current system. The response of the array is not ideal, with varying responsivities from different pixels. However, this variation is fixed, and can be allowed for in the processing of the signal from different pixels. Operation under different ambient conditions still needs to be determined: detection relies on there being a perceptible temperature difference between the hand and the background scene.

Optimisation of the optics has not yet been fully explored. The prototype system uses pin hole optics: very simple and inexpensive, but not ideal in terms of radiation gathering properties. It should be possible to improve the signal considerably by the addition of an appropriate lens. However, for a production system the trade off between cost and performance may influence this choice. Higher quality (e.g. zinc selenide) lenses that operate in the mid infrared range are expensive. A plastic Fresnel lens may represent a reasonable compromise between cost and light-gathering for the intended low-cost application area.

Several software algorithms have been investigated, but work remains to choose which is the most effective for the small arrays we currently have available. There is currently a large gap between simple differential systems (whether two or four element) and true imaging arrays with thousands of pixels. The 4 x 4 array in this study is unusual in that it lies at the very boundary of what might be considered an image to which image processing techniques could be applied.

Future work on the software algorithms will focus on improving detection reliability, working towards an increase in angular resolution for more gestures to be recognised. Optimization towards embedded system application is required to achieve a low power system suitable for real application. With the development of a dedicated array for this type of recognition, further optimisation of the system is possible.

8. Acknowledgment

This work is supported by the UK Engineering and Physical Sciences Research Council and Pyreos Ltd. under EPSRC CASE Award EP/I50137/1. The authors thank C. Giebeler at Pyreos Ltd. Scotland, UK for his direction and support.

9. References

[1] “WHO | WHO guidelines on hand hygiene in health care” Available: http://www.who.int/rpc/guidelines/9789241597906/en/.

[2] J.P. Wachs, H.I. Stern, Y. Edan, M. Gillam, J. Handler, C. Feied, and M. Smith, “A Gesture-based Tool for Sterile Browsing of Radiology Images,” *Journal of the American Medical Informatics Association*, vol. 15, May, pp. 321-323.

[3] T.R. Winings and R. Samson, *Automated dispenser for disinfectant with proximity sensor*, Google

Patents, 1997.
[4] D.K. Lambert, *Capacitive proximity sensor*, Google Patents, 2004.
[5] M. Ishikawa and H. Matsumura, “Recognition of a hand-gesture based on self-organization using a DataGlove,” *Neural Information Processing, 1999. Proceedings. ICONIP ’99. 6th International Conference on*, 1999, pp. 739-745 vol.2.
[6] Chen-Chiung Hsieh, Dung-Hua Liou, and D. Lee, “A real time hand gesture recognition system using motion history image,” *Signal Processing Systems (ICSPS), 2010 2nd International Conference on*, 2010, pp. V2-394-V2-398.
[7] Liu Yun and Zhang Peng, “An Automatic Hand Gesture Recognition System Based on Viola-Jones Method and SVMs,” *Computer Science and Engineering, 2009. WCSE ’09. Second International Workshop on*, 2009, pp. 72-76.
[8]“GestureTek Technologies Website” Available: http://www.gesturetek.com/.
[9]“EyeSight Mobile Technologies Website” Available: http://www.eyesight-tech.com/.
[10] “Edge 3 Technologies Website” Available: http://www.edge3technologies.com/.
[11] Byung-Woo Min, Ho-Sub Yoon, Jung Soh, Yun-Mo Yang, and Toshiaki Ejima, “Hand gesture recognition using hidden Markov models,” *Systems, Man, and Cybernetics, 1997. “Computational Cybernetics and Simulation”., 1997 IEEE International Conference on*, 1997, pp. 4232-4235 vol.5.
[12] “Mechaless Systems GmbH Website” Available: http://www.mechaless.com/.
[13] Dongseok Ryu, Dugan Um, P. Tanofsky, Do Hyong Koh, Young Sam Ryu, and Sungchul Kang, “T-less□ : A novel touchless human-machine interface based on infrared proximity sensing,” *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, 2010, pp. 5220-5225.
[14] Te-I Chiu, Hsu-Cheng Deng, Shu-Yuang Chang, and Shi-Bing Luo, “Implementation of ultrasonic touchless interactive panel using the polymer-based CMUT array,” *Sensors, 2009 IEEE*, 2009, pp. 625-630.
[15] J. Krumm, “Ubiquitous Advertising: The Killer Application for the 21st Century,” *Pervasive Computing, IEEE*, vol. 10, 2011, pp. 66-73.
[16] C.M. Payne, *Principles of naval weapon systems*, Naval Inst Pr, 2006.
[17] R. Siegel and J.R. Howell, *Thermal radiation heat transfer*, Taylor & Francis Group, 2002.
[18] “Pyreos Website” Available: http://www.pyreos.com/.
[19] T. Chamberlain, “People Counting Demonstration Using a Pyreos IR Detector Array,” 2008.
[20] T.S. Axelrod and T.F. Tassinari, *A VLSI centroid extractor for real-time target tracking applications*, Lawrence Livermore National Lab., CA (USA), 1989.
[21] *Pyreos: Gesture Control - low power, smal size*, 2011.
http://www.youtube.com/watch?v=0YpI3J2lThA [Nov 2011]