HEAD COUNT FOR THERMAL IMAGES IN H1N1 SCREENING AT KLIA

ZALIDA WATI BT AHMAD

**A Thesis Submitted in Fulfillment for the
Requirement Award of the degree of
Bachelor of Electrical Engineering (Electronics)**

**Faculty of Electrical and Electronics Engineering
Universiti Malaysia Pahang**

JUNE 2012

ABSTRACT

In computer vision field, head detection can be considered as one of the most important problems. Nowadays, H1N1 screening at KLIA for flue detection is done manually. Therefore, this project will detect and count the head for H1N1 screening automatically by applying image processing. The head count will be done at KLIA to monitor the total passenger that will undergo H1N1 screening for record. However, the head count of H1N1 screening is complex since crowded people will lead to overlapping head and head detection will be hard to accomplish. A method based on implementation of Hough transform is applied. This project can detect human head effectively in complex background, and provide a new approach of human head detection especially when the colour information of image is limited. Other than that, shape detection for roundness head will be use as an additional method. Besides that, I also use other method which is cropping and mapping method. Thus, by implementation of one of these methods, the head detection can be automatically counted in order to monitor the passenger that undergoes H1N1 screening at KLIA. The expected outcome of this project is a Graphical User Interface (GUI) in MATLAB that takes thermal images of passenger screening at KLIA and output detected the head in the images with total head count.

ABSTRAK

Dalam bidang visi computer, pengesanan kepala boleh dianggap sebagai salah satu masalah yang paling penting. Pada masa kini, saringan H1N1 di KLIA untuk pengesanan virus H1N1 dilakukan secara manual. Oleh itu, projek ini akan mengesan dan mengira jumlah kepala untutk pemeriksaan H1N1 secara automatik dengan menggunkan pemprosesan imej. Kiraan kepala akan dilakukan di KLIA untuk memantau jumlah penumpang yang akan menjalani pemeriksaan H1N1 untuk tujuan rekod. Walau bagaimanapun, kiraan kepala pemeriksaan H1N1 adalah kompleks ekoran jumlah manusia yang terlalu ramai pada satu-satu masa akan menyebabkan kepala bertindih dan pengesanan kepala akan sukar untuk dicapai. Satu kaedah berdasarkan pelaksaan Hough Transform akan digunakan. Projek ini dapat mengesan kepala manusia secara berkesan di latar belakang kompleks dan menyediakan satu pendekatan baru pengesanan kepala manusia terutamanya apabila maklumat warna imej adalah terhad. Selain daripada itu. pengesanan dan pengiraan kepala berdasarkan ‘Shape Detection’ teknik akan digunakan sebagai kaedah alternatif. Kaedah lain yang turut digunakan dalam projek ini ialah kaedah pemetaan. Oleh itu, pengesanan dan pengiraan kepala boleh dilakukan secara automatik untuk memantau penumpang yang menjalani pemeriksaan H1N1 di KLIA. Hasil yang diharapkan daripada projek ini adalah Graphical User Interface (GUI) dalam MATLAB yang mengambil imej haba saringan penumpang di KLIA dan output mengesan kepala di dalam imej termasuk jumlah kiraan kepala.

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Overview

This project is done in order to detect and count the head in thermal images. In human anatomy, the head can be defined as the upper portion of the human body. It supports the face and is maintained by the skull, which itself encloses the brain. Thermal images, or thermo grams, are actually visual displays of the amount of infrared energy emitted, transmitted, and reflected by an object. Because there are multiple sources of the infrared energy, it is difficult to get an accurate temperature of an object using this method. A thermal camera is capable of performing algorithms to interpret that data and build images. Although the image shows the viewer an approximation of the temperature at which the object is operating, the camera is actually using multiple sources of data based on the areas surrounding the object to determine that value rather than detecting the actual temperature. In order to track moving objects in thermal images, it is needed to decide how many moving objects exist and where their positions are. In general, head detection could be used to separate and locate moving humans.

1.2 Objective of Project

The main core objectives for this project are:

i. To detect head in thermal images for H1N1 screening at KLIA.
ii. To count head in thermal images for H1N1 screening at KLIA.

1.3 Scope of Project

In order to achieve this project, there are several scopes that had been outlined. However, this project would focus on image processing of human head detection. Besides detect the head, this project also will count the total amount of passenger that undergoes H1N1 screening at KLIA for surveillance and monitor the passenger for record by implementation of Matlab software.

1.4 Problem Statement

The surveillance in H1N1 screening at KLIA for flue detection is done manually. Therefore, this project will detect and count the head in thermal images automatically. However, the head count of surveillance H1N1 screening is complex since crowded people will lead to overlapping head and head detection will be hard to accomplish. We can apply image processing by using Hough Transform for circle detection since head detection is difficult to detect in thermal images. The head detection can be automatically counted in order to monitor the passenger that undergoes H1N1 screening at KLIA.

1.5 Significances of Study

The significances study for this project is mainly for surveillance and monitoring in crowded area. Human detection is important in visual surveillance system. A number of surveillance applications require the detection and tracking of people to ensure security, safety and site management. Examples include the estimations of queue length in retail outlets, the monitoring of entry points, bus terminals or railway stations. Although person detection and counting systems are commercially available today, there is a need to further research to address the challenges of real world scenarios.

1.6 Thesis Outline

Chapter 1 contains explanation on the introduction of the project which consists of the overview, problem statement, objectives, significances of study and scope of head detection project.

Chapter 2 is more focused on literature review which exposed the research and studies of this project done by the previous researcher.

Chapter 3 views the method which is implementation of Matlab programming for image processing that is used in completing this project.

Chapter 4 presents the result, analysis, and discussion of the project.

Chapter 5 concludes the overall of the project which includes the problem and future development.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter includes the history of tracking head method, counting people, surveillance system and image segmentation.

2.2 Head Shape

Stan Birchfiled [1] modelled the head contour with an oval, and then let the model search in the entire state-space. He proposed that the location of the head can be determined by calculating the gradient and mean of the elliptical contour. However, his method worked well only for frontal extracting

Hosub Yoon [2] proposed that all human heads have the same shape "Ω" in spite of any viewing direction. The detection of shape "Ω" utilized gray-scale, colour and edge information, then the ellipse model was used to detect the head area. This method is relatively stable and it suitable for a variety of head shapes, but it is less accurate when the colour information is limited. Besides that, this method is also less accurate since crowded people will lead to overlapping head in thermal images.

The methods above are mainly based on frontal detection. When some information is missed such as image colour, the result is not desire. Head detection will be complex in crowded area due to the overlapping of head in thermal images.

.

2.3 TRACKING HEAD

Shao Ming-Ming [3] proposed an algorithm based on the existence of two concave points between head and neck contour. The researcher detected the two concave points in order to achieve the purpose of head detection. This technique was applied according to the theory of circle detection using Hough Transform. This method was often applied to extracting a single frontal head only.

In advance, Mokhatian F [4] proposed the CSS (curvature scale space) theory. In his theory, he assumes that the contour is composed by a number of boundary points (x,y). The difficult problem is, in order to track moving objects in thermal images, it is needed to decide how many moving objects exist and where their position are $[5] - [7]$. In thermal image that consist of human scene ,human silhouettes can be easily extracted from the background regardless of lighting conditions and colors of the human surfaces and backgrounds [8]-[11].

Vertical projecting histogram was often used to locate head of moving human in visual surveillance system. Vertical projecting histograms are computed by projecting the binary foreground region along the major axis. But when people are occlude,or do not stand vertically, the high peaks of vertical projection histogram do not correspond to people's head [6].

Liyuan Li [12] fused the motion cues from background subtraction and the temporal differencing, and used a fuzzy reasoning technique to detect and locate motion objects from vertical projection histogram. However, when the peaks of the histogram are merged, it is difficult to separate these peaks.

The curvature scale space (CSS) theory method is effective compare than other methods because the experimental results show that the algorithm can accurately locate the head regions. This method calculates the curvature of each pixel on the curve and then it match the pixel value with the ratio of head size to detect the head.

2.4 SURVEILLANCE

Apart from implement head detection at H1N1 screening at KLIA, various techniques have been applied by past researchers to construct fast and reliable person detectors for surveillance applications. For example, Nakjima et al [13] use Support Vector Machines to this problem. Besides that, Gravrila [14] use different method which is a tree based classifier to represent possible shapes of pedestrian. Giebel et al. [15] use dynamic point distributions models. C. Schmid [16] presented the idea of learning part detectors using Ada-Boost and set of weak classifiers. This method is an alternative by to modelling the appearance of an entire person by designing detectors for specific objects part and combine the result those.

A learning approach is then being used to combine the set of weak classifiers to body parts detectors which are then combine using a probabilistic person model. All these approaches require a fair amount of training data to learn the parameters of the underlying model. Although these classifiers are robust to limited occlusions they are not suitable to segment a group of people into individuals.

In order to overcome this, there is one possibility of segmenting a group of people which is by use the information of various different camera views. For example, A. Mittal et al. [17] presented M2-tracker explicitly assigns the pixels in each camera views to a particular person using colour histograms. On the other hand, Zhao and Nevatia [18] make clever use of the fact that they know the camera calibration and can find possible head locations using a head detector. The locations of all individuals in the scene are estimated by maximizing an observation likehood using Markov Chain Monte Carlo. Their result clearly shows that it is extremely helpful to know the location of the ground plane and the camera parameters. The head detector is based on edge information. Under certain imaging conditions it can be challenging to extract clean edge maps. However, in order to overcome this, J. Rittscher et al. [19] developed a model based segmentation algorithm that simultaneously estimates the position and size.

Group segmentation alone is not sufficient to count the number of people entering or leaving a site. It is necessary to extract the direction of travel. Therefore, the idea of a virtual gate was introduced. Yang et al. [20] propose a people counting method that makes use of different views.

In surveillance application, most of the method use in order to surveillance human is by implementation of video form instead of static image. This is because the surveillance is much easier to be done in video form compare than an image form.

Therefore, the above method is less appropriate to be implemented in my project since my project use thermal images.

2.5 DETECT HUMAN IN THERMAL IMAGES

There exist a few approaches for detecting humans in thermal images in the literature. Bertozzi [21] introduced a pedestrian detection method as a part of a driver assistance system. The algorithm is divided into three parts. (1) Candidate generation. The input thermal image is processed to locate warm symmetrical objects with a specific size and aspect ratio. (2) Candidates filtering. The candidates may contain poles, road signs and buildings, which also have symmetry characteristic. These false positive objects can be filtered by analyzing the shape of the vertical histogram in each search window. (3) Validation of candidates. Morphological characteristics of a human are extracted to form a model. Each filtered search window is compared with the model for validation.

The weakness of this method is that human should be hotter than its background. Since human appearance can vary considerably in thermal images due to temperature variations, Davis and Keck presented a two-stage template-based method [22], which takes advantage of the invariance of edge information. In the first stage, human contours are obtained by creating Contour Saliency Map (CSM) [1] of thermal images. CSM represents the belief of each pixel belonging to an edge contour of a person [22]. Then a screening template is produced by averaging the human samples cropped from the CSM images. Last, a multi-resolution screening procedure is applied to obtain candidates. In the second stage, four Sobel filters with different angles are applied to the human samples to get four projected edge images. An Adaboost classifier is trained with the projected images and is applied to new input images. This method proves that edge is a robust feature for object detection in thermal images.

Recently Kai and Arens [23] proposed a local-feature based pedestrian detector on thermal data. In the training phase, they used a combination of multiple cues to find interest points in the images and use SURF (Speed Up Robust Features) [24] as features to describe these points. Then a codebook is created by clustering these features and building Implicit Shape Model (ISM) [23] to describe the spatial configuration of features relative to the object centre. In the detection stage, SURF features are first located in each image. Then the matching between the features and the codebook is conducted to locate object centre. The challenge of this detector is whether a high performance can be achieved when local features are not obvious, for example, in thermal images of poor quality.

In “Head Location Based on Fuzzy Weighted Projection Histogram in Infrared Thermal Sequences” journal, a fuzzy weighted projection histogram of thermal images is proposed by Yang Xuan using position and thermal information of pixels. The high peaks are sharper and more separated in the fuzzy weighted projection histogram. Experiments show that this method effective and feasible to count and locate the objects in thermal images [32].

However, most above method is less efficient because all the above method cannot overcome the overlapping head in thermal images and head detection will be hard to accomplish accurately except Head Location Based On Fuzzy Weighted Projection Histogram In Infrared Thermal Sequences” journal. This journal is the most appropriate and efficient method compare than other method because it can detect human head effectively and able to overcome overlapping head in thermal images especially in crowded scene. This method will match the position of the head with the thermal value of head in thermal images.

2.6 COUNTING PEOPLE IN CROWD

There has been a lot of work in the field of crowd estimation and people counting. Some of the earlier work in this area has relied on heavily confined environments. For instance, Terada *et al.* [25] count people going through a gateway that only allows for a small number of people to go through at the same time. The stereo cameras used in this system are mounted overhead in order to avoid occlusions. Work by Velipasalar *et al.* [26] use a similar setup of ceiling–mounted cameras in order to avoid the problem of occlusions. The camera view is also very narrow and does not cover a large scene, unlike what we are trying to achieve. Work realized by Zhao and Nevatia [27], [28] segments and tracks humans in crowded scenes using a human model composed of ellipses corresponding to the different parts of the body. This model helps keep track of individuals and is thus capable of giving a count of people in the scene.

In [29], Rabaud and Belongie describe a method of counting crowded moving objects. Their counting technique is based on clustering a set of features in a video sequence and estimating the trajectories of these different detected clusters over time. The object counting is then performed based on these trajectories. Kong *et al.* [30] present a viewpoint– invariant way of counting people in a crowd. The key idea in this work is to use feature histograms in conjunction with feature normalization to make the algorithm viewpoint invariant. In [31], Kilambi *et al.* present a technique to count groups of pedestrians utilizing camera calibration information. They project the foreground blobs onto different planes and use an area–based heuristic to estimate the number of people in group. All of the above mentioned counting methods rely on effective background subtraction.

Counting people in crowd also use implementation of video in order to surveillance people. It use camera to record the scene and then use various method to detect and counting the human. All of the above method is also not appropriate to be use in my project since my project use thermal images as the data collection.

2.7 IMAGE SEGMENTATION

There are many algorithms used for image segmentation, and some of them segmented an image based on the object while some can segment automatically. Nowadays, no one can point out which the optimal solution is due to different constraints. In [33], a similarity close measure was used to classify the belonging of the pixels, and then used region growing to get the object. Unfortunately, it required a set of markers, and if there is an unknown image, it is hard to differentiate which part should be segmented. Linking the area information and the color histogram were considered for building video databases based on objects [34]. However, the color information has to be given first, and it is not useful for the life application. A genetic algorithm adapted the segmentation process to changes in image characteristics caused by variable environmental conditions [35], but it took time learning. In [36], a two-step approach to image segmentation is reported. It was a fully automated model-based image segmentation, and improved active shape models, line-lanes and live-wires, intelligent scissors, core-atoms, active appearance models. However, there were still two problems left. It is strong dependency on a close-to-target initialization, and necessary for manual redesign of segmentation criteria whenever new segmentation problem is encountered. The authors in [37] proposed a graph-based method, the cut ratio is defined following the idea of NP-hard as the ratio of the corresponding sums of two different weights of edges along the cut boundary and models the mean affinity between the segments separated by the boundary per unit boundary length. It allows efficient iterated region-based segmentation as well as pixel-based segmentation.

Moreover, in order to understand an image and recognize the represented objects, it is necessary to locate in the image where the objects are [38]. The homogeneity between two pixels and the distance function are included to measure the segmented results $DI = | I(x, y) - I(v,w) |$ [39]. In [38], a confidence level $L_i = \min_j (p_i \; p_j), i, j = 1,...,N$. is used as a new performance measure to evaluate the accuracy segmentation algorithm. In [40], the minimizing function including the approximation mean square error RMSE and the number of distinct region tried to achieve a good segmentation result. In [41], edge detection is a fundamental concept for segmentation because it is easy not only to understand its principle but also to implement. Comparing the basic edge detection method, which is get from Matlab commend, with our algorithm, there are some results as follows:

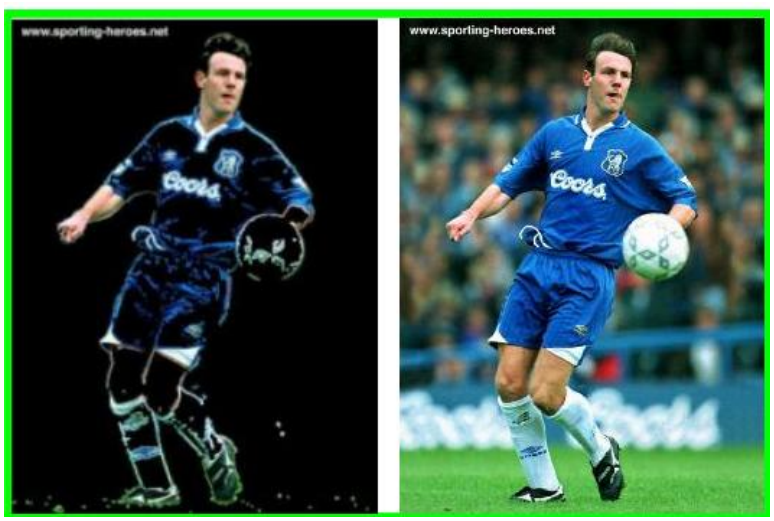

Figure 2.1: (a) Matlab “edge” command (left) (b) Proposed algorithm (right)

Figure 2.2 : (a) Matlab “edge” command (left) (b) Proposed algorithm (right)

Figure 2.1 (a) and Figure 2.2 (a) show the segmented result by Matlab “edge” command, and Figure 2.1 (b) and Figure 2.2 (b) show our proposed algorithm. It is pointed out that although the proposed algorithm extracts some undesired parts into the segmentation result, it represents almost the entire target image while the Matlab “edge” command seems to show the boundary of the target image only. It is because the proposed algorithm considers the color information at the same time in order to recover some disconnected parts while the Matlab “edge” command actually has problem identifying them as part of the target image.

Figure 2.3: (a) Matlab “edge” command (left) (b) Proposed algorithm (right)

In Figure 2.3, there are the segmentation results by performing the Matlab “edge” command (a), and the proposed algorithm (b). It can be shown that the Matlab “edge” command works better than the proposed algorithm because it can correctly extract the whole boundary. However, the algorithm of Matlab can only extract the image with clear vision. In other words, it has problem extracting the boundaries with blurred image and cannot pass the pixel information of those area. On the other hand, the algorithm we proposed can connect some of the disconnected boundary line segments by performing dilation, so our algorithm contains more pixel information than the algorithm of Matlab.

Basically, the proposed algorithm has better performance than the basic edge extraction command due to the consideration of connect the discontinuous points automatically.

CHAPTER 3

METHODOLOGY

3.1 Introduction

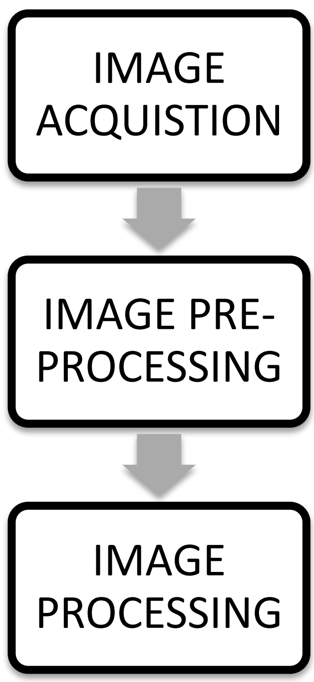

Figure 3.1 : Fundamental steps of Digital Image Processing

The figure 3.1 shows the fundamental steps of digital image processing including image-acquisition, image pre-processing as well as image processing.