International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 10, No. 12 (2017), pp.75-94
http//dx.doi.org/10.14257/ijsip.2017.10.12.06

Queueing Detection System Using an Integrated Pedestrian Flow and Optical Regression

Reda Shbib and Mohamad Merhi

School of Engineering, Lebanese International University
Brunel University London
reda.shbib@liu.edu.lb, mohamad.merhi@brunel.ac.uk

Abstract

In this paper, an effective design of an automated queue detection system based on a gathered statistical data, has been proposed. In addition, different metrics have been taken into consideration when implementing this algorithm, such queue size, and the number of people entering the queue, exiting and serving. Therefore, a pedestrian flow analysis algorithm has been well implemented using a proposed virtual mark, that identifies the flow of pedestrian within specific area in an image scene

Keywords*: Counting people, pedestrian flow, queue detection*

1. Introduction

Controlling queues is critical for a range of several profitable uses. Entertainment business always have need of handling queues of people in the turn for cycles, airports need managing travelers waiting to finish their safety check, and supermarket must aid clients waiting in the queue to finish their payment after purchase. Generally, once these queues are mostly big, with several service till, it is tough for staffs to quantitatively asses the queue behavior such as entering frequency of people, the entire length of the queue, growing level and dwell period. Furthermore, if these queues of persons are not effectively monitored by the managers, it generally affects the contentment of clients, which can lead to an unsuccessful resource distribution and hurt of incomes. Accordingly, there is a need, not only to observe and collect the statistical information of the queue in in instantaneously, however to effectively distinguish tendencies that present in queues throughout the time. Automatic queue analysis gives one means to queue approach methods that are capable of accomplishing this

2. Related Work

The key work is to define a queue by its entrances criteria, handling of service and the total number of serving the customers. Several solutions have been presented to describe a queue for example, the notation of M-D-1 denotes a queue structure with a Poisson entrance process, a deterministic service period, and usage of only one service channel. Donald, has distinct three important principles of a queue:

There main three criteria's have been defined by [1]. Time spent which is defined as the period of time that a person stays in a certain location, Moreover, another measurement is the accumulation of customers that has been defined as the number of persons in the specified queue. Finally, the Idle period of the operatives which represent the ratio of time that any operator is idle, or otherwise the entire system is empty of people.

Received (August 25, 2017), Review Result (November 23, 2017), Accepted (December 1, 2017)

ISSN: 2005-4254 IJSIP

A queuing checking system that combines existing growths in people counting methods has been presented by [2]. The proposed approach makes use of an adaptive background modeling techniques to recognize the pixels that are belongs to the foreground which is related to persons within the Field of View (FOV). The proposed system assumes that the total number pixels that are belonging to the foreground are directly constitute the identified people's number in the FOV. Furthermore, the developed system has utilized a counter area, which is produced at for each service counter in the FOV.

This queue detection system has been applied and evaluated at an airport. The system comprises of two main units: Queuing detection unit and a counter unit. The primary module permits the system to count the traveler's total waiting in a specified area at the airport. Then, the other unit has been used to estimate the average waiting time service per passenger at the bureaus and compute the total number of active service counters.

Two kinds of camera including different FOV have been installed in the system application (above and oblique). The system has been made to be strong against light deviation and customer's variation in appearance including travelers having their stuff and baggage.

Furthermore, the proposed system has combined the two units (Queueing detector unit + Counter unit) to estimate the waiting time of the passenger. This has been accomplished by multiplying the passenger's number that are waiting by the average timing of the customer during serving.

A weighted area function has been applied to the numbers of travelers based on their location in an image frame. The queue dimensions' modules are built on the resulting equation:

$$N = \sum_{i=1}^{a} \sum_{j=1}^{b} \varphi_{(i)} A_{(i,j)} \tag{1}$$

Where N represents the total number of travelers in the (FOV), a is representing the zone's height and b is representing its width. N has been represented as a weighted area of A where $A_{(i,j)}$ is the foreground and $\varphi_{(i)}$ characterized as the weighted of area function.

The accuracy level of the counter unit is dependent on the cameras' position. The location of the camera played an important role in determining the accuracies. It is reported that the overhead camera has shown more accuracy level results that the other one. In order to avoid this problematic, the author has recommended the usage of probability measurements in to improve the counting results, which take in quasi-calibration method. For example, the maximum lengths of a passenger have been distinct by its minimum dimension of its head

The results have shown a satisfactory level in approximating on the total number of travelers in the waiting area, and were unsuccessful in some situations to precisely evaluate the service time. Moreover, the proposed approach hasn't worked well when there is a changing in some situations such as crowded queue structures with several camera viewpoints. In some cases, some of the main faults occur when the service desk (whether if it is functioning or not) has been wrongly estimated.

Additionally, a number of related studies have been proposed regarding automated queue monitoring. Among such studies, individuals crowd observing and queue estimation has been presented to monitor and calculate the number of people in a given queue of an ATM machine[3]. Another system has been presented for audience tracking to measure the queue. These study-works are nearly related and the established queueing concepts deal with complications which are the same [4]. Both methods are based still images, which are updated each forty and twenty second correspondingly. The lighting variations have affected the system when employed in a daytime and in night.

To achieve the background subtraction, Gaussian Mixture Model (GMM) has is applied. GMM is usually utilized in segmentation and foreground extraction, which is efficient in locations where lighting variations environments, occurs. Then, after the foreground detection and the segmentation of the is accomplished, both methods have applied a pixel based analysis and SVM to achieve a classification of the length of the queue

In the same classification method, the other implemented system has used the entire pixel's foreground in the waiting part. Yet, the idea of utilizing pixels can be beneficial in minor crowd, in many circumstances in the number of individuals is high, the extracting procedure might be impractical, lead to some errors in the segmentations process. Both systems did not consider the full criteria's and factors of the queue and have measured only of the number of persons in the waiting part.

Counting might similarly be achieved by means of person's detection approaches [5]– [6], based on person centric features, For instance. features which are describe the entire individual, using the (HOG) descriptor of the entire individual [7]. In addition, the deformable part based approach (DPM) [8] similarly constructs an HOG descriptor of an entire individual, by means of a further flexible layout model for the three-dimensional relationship among HOG parts at many scales. Although these results in an approach which is better adapted to moveable positions of a particular individual, it may have difficulties in identifying partially occluded individuals in crowds.

Besides, Visual tracking might be applied for Queue counting. In [9], Kanade-Lucas-Tomasi (KLT)which is a based Feature Tracker is applied to calculate the tracklets of people for additional crowd behaviour investigation. Yet, the trajectories of obtained tracks turn out to be noisier and separated when the occlusion is huge as in congested situations

Other additional classic method is the employment of the background subtraction followed by a sequences of morphological (erosion followed by a dilation) methods to find an effective shape of the individual for segmentation process (or "processing of blobbing"). As shown in Figure 1. This silhouette would be utilized to approve the blobbing object is definitely a human [10], or to identify the silhouette of the bounding box from where additional individual's features can be obtained for that goal [11]. Because some of morphological procedures do not assure that every individual translate to accurately one segment (or blob), an additional pass should be accomplished where segments which are sufficiently adjacent have been combined together. The outcome is that it is shared to combine segments that do not fit together, additionally as to distinct segments which make up the matching body, as in [12]. Such systems generally execute extra phases after blobbing is completed, in an effort to accurate such abnormalities

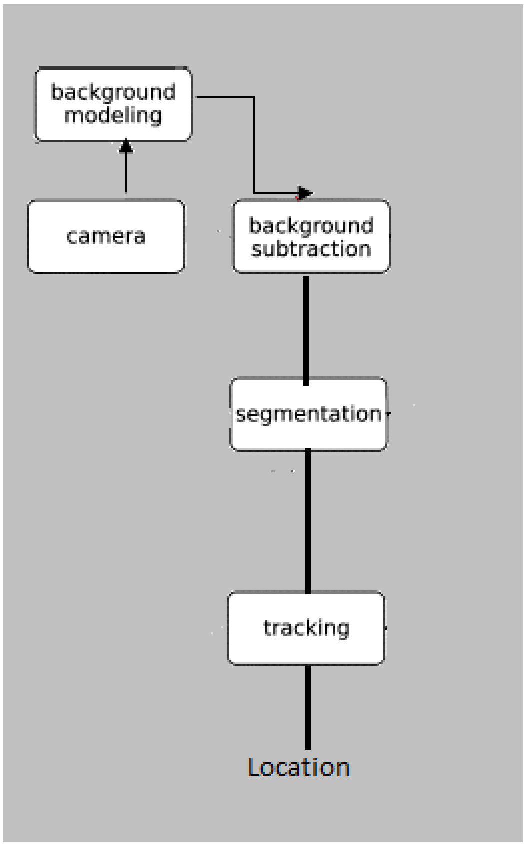

Figure 1. Segmentation and Tracking Process

Alternative method has been suggested in which persons in a queue have been tracked and calculated in a congested condition has been implemented by [13]. The proposed algorithm makes use of a unit for individuals tracking by applying the (MMSE) Minimum Mean Square Error. Double successive frames are matched to obtain the best corresponding zones of two images. The proposed system has been manually tuned and tracks are accomplished by locating the best matching over two succeeding frames with the MSE given by

$$\mathrm{MSE} = \frac{1}{\mathrm{KK'}} \sum_{\mathrm{K}}^{y=1} \sum_{\mathrm{K'}}^{x=1} [\mathrm{F}(x,y) - \mathrm{F'}(x,y)]^2 \qquad (2)$$

Where K and K’ represents the frame’s dimensions, F and F’ are the frames a n and n+1 time correspondingly. The values which have been found by $\mathrm{F}(x,y)$ and $\mathrm{F'}(z,y)$ denotes the Red-Green-Blue color model value that will be applied for evaluation. While trying to catch the best matching, simply a negligible ROI has been searched in the $\mathrm{F'}$ which growths the speed of the tracking with respect to the persons motion.

The established method has shown some unsuccessful results at the time of heavy occlusions are taking place and consequently, a person has to be manually allocated once more. Using Minimum Mean Square Error, the developed method can automatically recognize when fault is increasing, and applying a manual tracks whenever desirable. The achieved tracks attained in this scheme tries to focus on: The person’s number passing, their average number in addition to Estimating the frequency of number of persons for each minute and their movement’s speed.

The developed method has quite a lot of complications concerning the initialization of people, which must be manually done. Besides, numerous mechanisms have to be examined in order to resolve the problematic of occlusions.

The developed system depends simply on individuals tracking and the queue factors have not been fine recognized. The method is grounded on the detection of moving body which lead to small level of accuracy

Chen [14] essentially makes use of infrared sensors and overhead the counter lanes in order to observe the number of the customer and the behavior of the queue. The developed approach is capable to automatically capture a precise data and compute in instantly the average of the length of the queue length, approximating waiting times, cashier spending times and total operation servicing time

QN-MHP [15] is an another approach of queueing management system which is described by the usage of electric devices to be able to manage the movement of clients or peoples waiting to get served. A key feature of this kind of queueing management approach is that, there exist continuously an audio and/or flashing light notification to make the next individual see that he is prepared to get served. Most of the beyond stated structures are known as token managing structures where their objective is to be able to manage the waiting time and decrease rush at the counter desks stands.

3. Proposed Algorithm

The queue management system contains two main elements. First, the suggested system utilizes of an advanced counting unit, and lets the suggested method to figure out the number of persons which are waiting to be served in the in a queue. Following, two “virtual entrances” have been used, situated at the entry and the exit of the queue. The chief idea regarding these virtual entrances will be used for people detection when they walk over them and as well to for approximation the rate of persons arriving and exiting the queue.

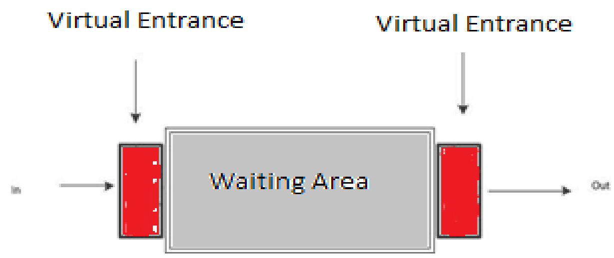

Figure 2. General Structure of the Proposed Algorithm

As illustrated by Figure 2, the people counting unit has been located over the queuing zone, along with virtual entrances positioned at the two ends of the queue. The combination of these modules has been used to analyze queue elements

Current pedestrian-flow monitoring techniques can usually be classified into two groups. First is detection and tracking. These algorithms make use of feature tracking [16], In crowded views where the camera is sufficiently near to the FOV, face detection has been applied for tracking incoming people, [17], in addition to the head and shoulder regions.

The second group is optical flow regression. Generally, these techniques apply a statistical model, which is more appropriate for crowded areas.

In most crowded places, people detection, face detection, and tracking are difficult tasks. This is because of the magnitude of the crowd and the distance of the camera from the view. In various situations, these approaches are also very expensive. Consequently, regression-based techniques are the most suitable in crowded locations.

In this section, an integrated approach is proposed: feature points are detected (corners and vertices) and their motion is then assembled over an interval of time. This assembled motion acts as an input for the regression approach. The developed technique is based on the perception that the number of people who are crossing through a virtual door is associated with the feature point's numbers passing the virtual count line. However, the relationship does not have to be linear.

The model will generate a ROI around the virtual count line, and then compute the total sum of the optical flow of these features. Consequently, an integration of feature extraction and regression has been applied to arrive at the final count result.

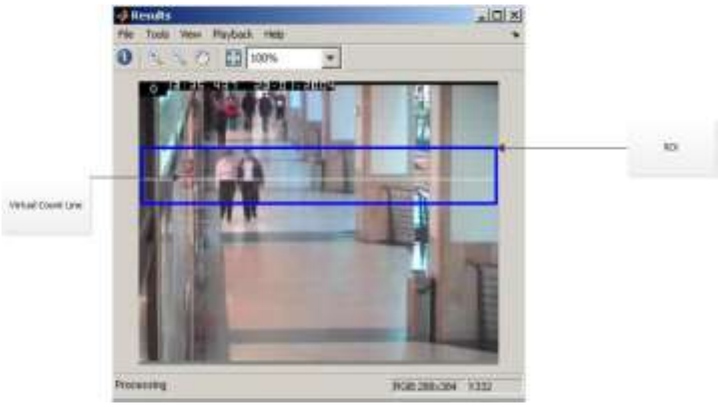

Figure 3. Virtual Counting Line

The proposed virtual door is shown in Figure 3. It involves a virtual line, which is bounded by the ROI, and a trajectory direction (TD), which indicates the direction in which people are walking. The group of pixels within the ROI is represented by G, and the vector direction in the TD is represented by r.

The proposed system assembles optical flow in the trajectory direction (TD) in a discrete group of extracted feature points. Each video chain is split to a set of sub-chains in which the optical flow is assembled. Then the regression to each window is applied separately to count the number of people crossing the virtual door.

The features belonging to each chain are computed as following. The optical flow is represented by o_t at any time T, and the optical flow at a pixel p is represented by $o_{t\,p}$. The component of this flow, which points in the trajectory direction (TD), is denoted as the aligned optical flow, and is calculated by applying the dot product:

$$\widehat{o_{t\,p}} = o_{t\,p} \,.\, TD \tag{3}$$

The group of the extracted features is represented:

$$\boldsymbol{F}\,(f_1, f_2, \ldots, f_{n,t,}) \tag{4}$$

Where $f_{n,t}$ denotes the sort of feature under extracted (edges, corners.) at time t

At every frame of the video sequence, the overall aligned flow is estimated:

$$e_{t\,F} = \sum_{p \in F} \widehat{o_{t\,p}} \tag{5}$$

Then, video is divided to sequences of time chains numbered by k. The group of frames, which are belonging to the kth chain, is represented as C_k, and each chain is considered the same size. Over each time chain C_k, the overall aligned flow is assembled:

$$\delta_{t\,F} = \sum_{t \in C_k} e_{t\,F}$$

\

Even though features are not obviously tracked, this integration will be approximately proportional to the feature's number passing the virtual line since the integration of aligned flow for each feature point is equal to the distance it travels over the virtual door.

A feature point crossing over the virtual door is defined by a sequence of optical flow vectors, the aligned components of which sum to the width w of the ROI.

Thus, each feature moving over the virtual door will contribute a roughly similar quantity of aligned flow, so that the overall aligned flow over a time window, $\delta_{t\,F}$, is relative to the number of features passing the door. Consequently, in the integration of aligned flow during time t, a chain is a measure of the number of crossing feature.

Moreover, the use of flow assemblage, instead of using an explicit tracking of features, will minimise the computational operations.

4. Histogram of Optical Flow

In this part, the functionality of optical flow histograms in improving system execution is illustrated by breaking up the impact of possible noise and true motion. Generally, optical flow does not always have accurate values, as they are subject to error or noise.

Assuming that a feature point is corresponding to a background object that is stable, the exact value of optical flow is 0. However, it can be given a small partial value such as 0.03. Therefore, to separate between noise and the true motion, a histogram that is based on flow size has been used. Aligned optical flow is estimated within various histogram bins:

Pixel p has been allocated to a histogram bin b based on the size of $\widehat{o_{t\,p}}$. A group of pixels belonging to bin b is represented by S_b.

The overall aligned flow for bin b and for feature ***F*** is represented by:

$$e_{t\,F,b} = \sum_{p \in F \cap S_b} \widehat{o_{t\,p}} \tag{6}$$

The overall aligned flow over time chain C_k for bin b and feature ***F*** can be then estimated using the following:

$$\delta_{t\,F,b} = \sum_{t \in C_k} e_{t,F,b} \tag{7}$$

The group of all features and histogram bins $\delta_{t\,F,b}$ are accumulated into a feature vector v_k, which defines the time chain C_k. It is important to mention that, when several features and histogram bins are used, these features are linked together to generate the higher feature vector. Then a regression has been trained to learn the relation among v_k and true values g_k.

5. Features Extraction

In this part a set of features are suggested for use with the proposed approach. A number of image features have been previously applied for pedestrian counting; these were classified into, appearance, magnitude, and discriminant-points features.

Same features selection has been adopted for proposed algorithm. Features have been used in this chapter are as flows:

These selected features are corresponding to the magnitude, and discriminant-points features, which have been used in pedestrian counting approach of Chapter 4. Appearance based features were also used for pedestrian counting; these features are derived from the perimeter pixels of each foreground segment. Foreground segmentation was achieved using adaptive background modelling techniques, which is not employed within the proposed virtual door algorithm. As an alternative, in this part, the proposed algorithm used optical flow, for which there is no corresponding appearance feature. Therefore, two kinds of feature have been used in this section: total image (T) pixels and discriminate points (corners) (D-P).

Figure 4 shows a linear relation between the selected features and pedestrian numbers passing through the virtual door.

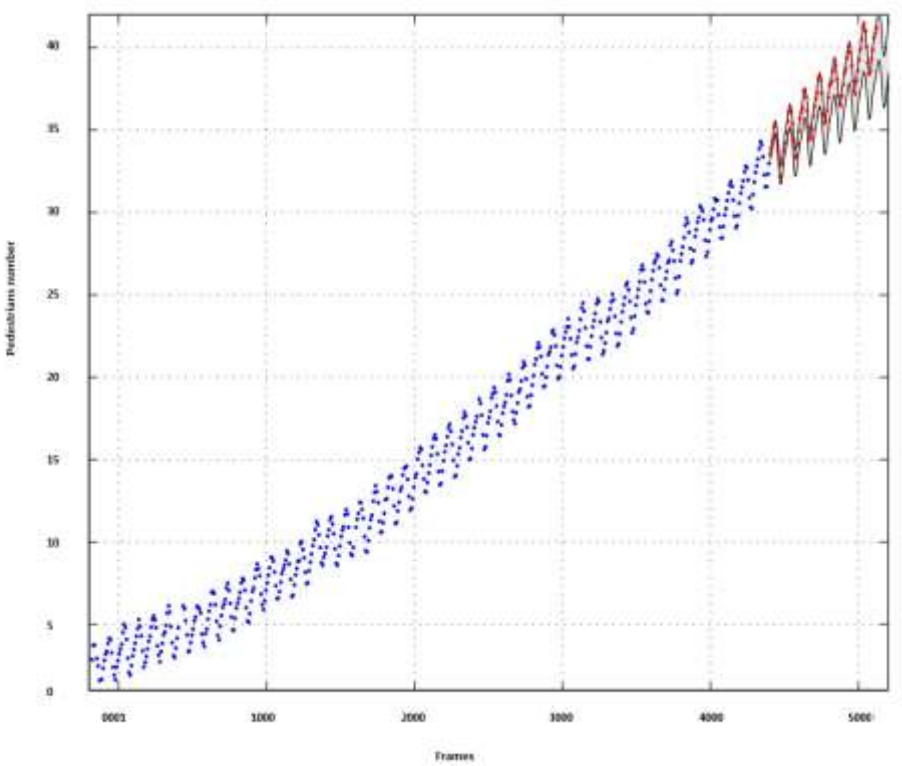

Figure 4. Relationships between these Features and the Number of People Passing

6. Proposed System Training

True value is indexed via GUI to detect the time stamp of individuals moving over the virtual door in addition, these indexing are then allocated automatically to chains so that the individual's number passing through the door during each chain is found automatically. Let the number of annotated people passing through the door during time chain C_k be represented g_k. An equivalent set of features, linked together into the feature vector v_k, is extracted for each chain. Finally These features and targets form the training data set { v_k , g_k }, then, regression model has been applied, in order to learn the relation among the two.

For regression, the Gaussian regression model has been adopted, since this model does not consider any previous hypothesis on the relation between the extracted features and the total size.

The main objectives for calculating the posterior is that it permits us make an estimate for hidden trial cases. This is valuable if we have sufficient prior information about particular dataset to surely identify previous mean and covariance.

- f : represents the function of training cases (x)
- f* : represents the function of the testing set (x')
- $\mu = w(x_i)$: is the means w(x)
- $\mu *$: represents the test means
- ∑ : covariance
- ∑* : is the covariance of the training set
- ∑** : is the covariance of the training-test set

$$f^* \mid f \sim N(\mu^* + \Sigma^{*T}\Sigma^{-1}(f-\mu), \Sigma^{**} - \Sigma^{*T}\Sigma^{-1}\Sigma^*) \tag{8}$$

This is the subsequent distribution for a precise set of testing trials.

$$F \mid D \sim GP(w_D, k_D) \tag{9}$$

$$w_D(x) = w(x) + \Sigma(X,x)^T \Sigma^{-1}(f-w) \tag{10}$$

$$k_D(x,x') = k(x,x') + \Sigma(X,x)^T \Sigma^{-1}\Sigma(X,x') \tag{11}$$

Where ∑(X,x) represent the of covariance between every training trial and x.

The mean and covariance can be initialised in terms of hyper parameters

$$f \sim GP(m,k), \tag{12}$$

$$w(x) = Ax^2 + Bx + C \tag{13}$$

$$k(x,x') = \sigma_y^2 e^{-\frac{(x-x')^2}{2l^2}} + \sigma_n^2 \delta_{ii'} \tag{14}$$

Where the hyper parameters:

$$\theta = \{A, B, C, \sigma_y, \sigma_n, l\}$$

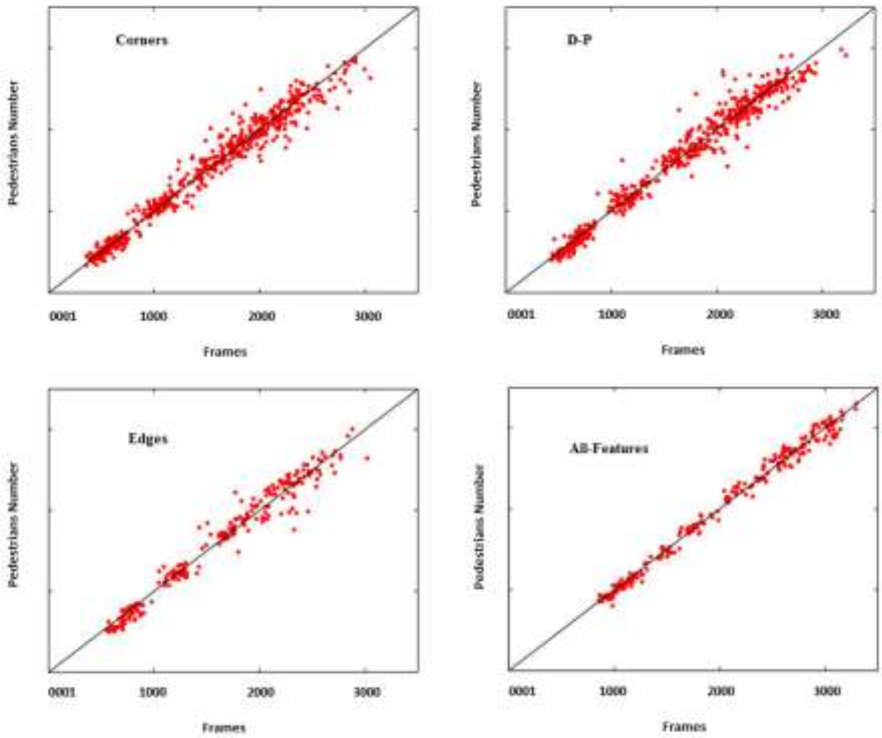

Figure 5. Relationships when using Different Types of Features

Figure 5 shows the relationship between the used features and the equivalent number of people within the virtual 1 dataset. The result of changing the training dataset magnitude has been also studied by applying sub-sets of the initial training dataset. For a particular size of a training dataset, obtained results have been averaged over several sub-sets.

7. Experimental Results and Discussion

In order to evaluate the presented virtual door approach, two datasets have been collected and indexed. The datasets (virtual -1), (virtual -2), (virtual-3) and (virtual-3) are separated into a number of people sequences, having a length of 2 minutes. The proposed approach has been tested using a cross validation technique.

Next, in order to evaluate the introduced queue monitoring model, video has been found from an airport's luggage check-in counter (people waiting in check-in counters at (suvarnabhumi airport, Bangkok, Thailand)

This dataset included of one sequences called queue 1, queue 1 is of size 10 minutes, comprising 30 travellers' arrivals, 10 travellers serviced, and around 20-25persons in the queue.

7.1. Evaluation of the Proposed Virtual Door Algorithm Using Different Datasets

In this section, the effectiveness of the virtual door approach is evaluated. A cross validation mechanism has been applied as follows:

1. The dataset is annotated with ground truth data: the individual's number crossing over the door during each 2 minute sequence is then annotated, in addition to individual's number crossing over the door during each 35 second chain.

2. One sequence is set aside for testing. The introduced virtual door approach has been trained by the ground truth of the other sequences. Then, testing has been achieved on the selected testing sequence.
3. This mechanism is repeated for all sequences within the dataset, by doing a rotation of testing and training sets. Next, MAE, and MRE are identified.

Figure 6. Queue 1 Structure Dataset

The features used in this evaluation are referred to as Total image pixels (T) and Discriminant-points (D-P), three different mixtures of these features have been evaluated. Moreover, the following two parameters have been evaluated: Histogram of optical flow and Chains.

The character `X' in table 1 indicates the use of three histogram bins, ranging between [0- 0:05), [0:05-0:25) and [0:25 - ∞). The exclusion of `X' is an indication that histograms of optical flow are not included.

The character `X' indicates that each 2 minute sequence has been divided into six chains of length 20 seconds, which are evaluated separately (during both training and testing). The exclusion of `X' is an indication that each sequence has been analysed globally. The use of chains is analogous to the use of low-level features in Chapter 4.

This gives rise to 12 different integrations of features and parameters. For each cross validation test, the feature-parameter integrations are ordered from 1 (top) to 12 (poorest) in terms of Mean Absolute Error, and Mean Relative Error.

The collected results of (virtual -1) are arranged in (Table 1, 2, and 3). The top-ranking feature vector is ((D-P), T) with histograms and chains have been used. The exclusion of histograms causes an increasing in MRE from 7.25% to 36.27%and the exclusion of chains causes an increasing in MRE from 7.25% to 84.36%. The MRE is between 7.25% and 10.81% for the best effectiveness of feature sets.

Table 1. Evaluation of Features, Histograms and Chains (virtual -1)

Features	Histogram	Chains	MAE	MRE
(T)	X	X	2.682	64.1%
(T)	X	Omitted	46.965	86.80%
(T)	Omitted	X	9.301	65.04%
(T)	Omitted	Omitted	10.741	49.31%

Table 2. Evaluation of Features, Histograms and Chains (virtual -1)

Features	Histogram	Chains	MAE	MRE
(D-P)	X	X	2.910	10.81%
(D-P)	X	Omitted	35.42	77.88%
(D-P)	Omitted	X	10.312	68.00%
(D-P)	Omitted	Omitted	11.753	59.66%

Table 3. Evaluation of Features, Histograms and Chains (virtual -1)

Features	Histogram	Chains	MAE	MRE
(D-P), T	X	X	2.664	7.25%
(D-P), T	X	Omitted	47.164	84.36%
(D-P), T	Omitted	X	5.29	36.27%
(D-P), T	Omitted	Omitted	52.93	97.16%

Figure 7 shows the estimated values of the introduced approach (using the ((D-P), T) feature set) against the ground truth for two-selected sequence in (virtual -1); the estimate queue size is sufficiently accurate.

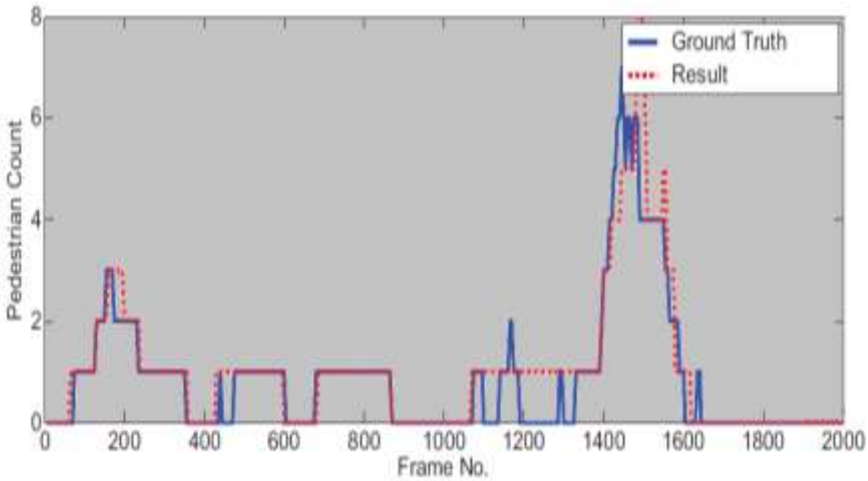

Figure 7. Cross Validation Evaluation (virtual -1)

7.2. Evaluation of Virtual 2

Virtual 2 has been validated on data from other viewpoints in the recording collected from at (suvarnabhumi airport, Bangkok, Thailand). (A+M+D-P+ C) have reached the best performance with an MAE of 1.21 and MRE = 10.11%. The use A+M+D-P has reached the better level in terms of the mean relative error. For these sequences, the MRE is 11% and the including of histogram and chains are vital for good system efficiency (Table4).

Table 4. Evaluation of features, histograms, and Chains (virtual –2)

Features	Histogram	Chains	MAE	MRE
(T)	X	X	2.421	18.62%
(T)	Omitted	X	8.652	54.81%
(T)	X	Omitted	2.121	11.353%
(T)	Omitted	Omitted	12.842	73.71%
(D-P)	X	X	3.334	21.22%
(D-P)	Omitted	X	8.443	44.76%
(D-P)	X	Omitted	1.886	13.01%
(D-P)	Omitted	Omitted	12.423	45.66%
(D-P), T	X	X	2.111	36.44%
(D-P), T	Omitted	X	3.567	19.22%
(D-P), T	X	Omitted	1.808	12.28%
(D-P), T	Omitted	Omitted	7.001	33.43%
A+M+D-P	X	X	1.21	11.02%
A+M+D-P+ C	X	X	1.14	10.11%

Figure 8 shows the estimated values of the introduced approach (using the ((D-P), T) feature set) against the ground truth for two-selected sequence in (virtual -2)

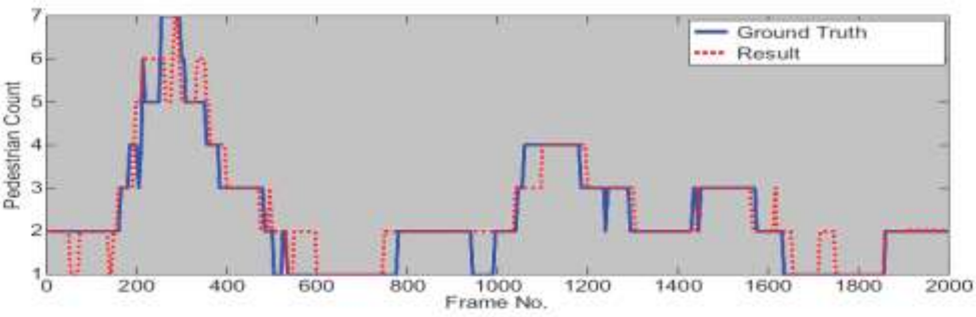

Figure 8. Cross Validation Evaluation (virtual -2)

7.3. Evaluation on Queue 1 Dataset

In this section, the introduced queue-monitoring model has been tested evaluated using the queue structure described before. This structure is shown in Figure 9. This structure include of one sequences called queue 1, queue 1 is of size 10 minutes, comprising 30 travellers' arrivals, 10 travellers serviced, and around 20-25persons in the queue.

Figure 9. Queue 1 Structure

In order to validate the operation of the pedestrian counting approach developed in chapter4 in a queue situation, we have applied cross validation mechanism, by training the system on the first 600 frames and tested on the remaining frames.

Table 5 shows the error metrics of the proposed framework on queue 1. The best performance reached an MRE of 9.56%.

Table 5. Queue Size Estimate using the Proposed Pedestrian Counting Approach (queue1)

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 1 Dataset	Queue size	0001-2500	2.884	9.054	11.4%
Queue 1 Dataset	Queue size	2501-4000	19.932	13.488	16.2%
Queue 1 Dataset	Queue size	4001-6500	11.517	21.795	21.3%
Queue 1 Dataset	Queue size	6501-9000	1.42	21.350	9.56%

Queue size parameter is tested in this section. Obtained results are compared in Table 5 on different frame sequences. The individual's features have been normalised to compensate for perspective

From the Table 5, it can be seen that the count performance has been enhanced after the using of both histogram and chains. In fact, the relationship between the number of people and the number of features used becomes linear as shown in Figure 5. Different types of passengers have been attempted in this experiment. The obtained results show that best performance of the system was between frame 4001 and 6500 when the queue is almost stable and people are not moving very fast. However, when the queue is too large, many zones without people has been identified as passengers (luggage case), which has caused a false count estimates

7.4. Evaluation on Queue2 Dataset

Another Dataset called Queue 2 size 15 minutes, comprising around 58 travellers' arrivals, 10 travellers serviced, and around 11 persons in each queue channel (Figure 10).

Figure 10. Queue 2 Structure

Table 6 shows the error metrics of the proposed framework on queue 2. The best performance reached an MRE of 11.33%.

Table 6. Queue Size Estimate using the Proposed Pedestrian Counting Approach (queue2)

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue2 Dataset	Queue size	0001-2500	1.380	3.423	**11.33%**
Queue2 Dataset	Queue size	2501-4000	1.552	4.119	14.25%
Queue2 Dataset	Queue size	4001-6500	2.750	14.350	16.26%
Queue2 Dataset	Queue size	6501-9000	1.986	8.573	12.55%

The MRE column reports the error metrics for the proposed pedestrian counting approach on queue1 dataset. Generally, it can be noted that the true value and the estimations agree fine. For the queue 1 dataset that includes up to 25 people, the mean relative error was less than 12%, and the average mean relative error was 14.615%. For queue 1 dataset, which holds up to 58 travellers, the best performance reached the MRE of 11.33% and the average absolute error was 13.59%.

Although queue size estimate in both queue1 and queue 2 datasets practically agree rationally fine, upon closer analysis, it noted that errors are exist in some cases. These errors are referred to two key causes. First, some errors happened because of inadequate calibration. Radial distortion is the base to some errors and can be irregular or follow various forms; the most frequently faced distortion is radially symmetric when passengers were close to the borderline of the view (the high density-crowd situation). Then, unexpected lighting variations due to light coming from the ceiling were occasionally has not been totally compensated, which give rise to remaining errors (particularly queue 2 dataset). Yet, the error's amount has been insignificant enough to have slight influence on the global waiting time.

Results of the proposed module on queue1 and queue2 datasets are shown in Figure 4-12. The counting module operates by segmenting a crowd into clusters and estimating the number of pedestrians in each cluster, as shown in Figure 11

Figure 11. Queue Size Estimate using the Proposed Pedestrian Counting Approach (Queue1)

Figure 12. Queue Size Estimate using the Proposed Pedestrian Counting Approach (Queue2)

7.5. Evaluation of Virtual Door Algorithm

The virtual gate module has been assessed by counting number of entrances (Figure 12(b)) and number of services (Figure 12). The feature set `(D-P), T ' has been selected; time chains of length 20 seconds have been used; and three histogram bins have been used

The dataset is divided into four sequences of length 3 minutes, and the cross validation mechanism has been used. Table 6, 7 & 8 show the error metrics for each of the virtual door (number of entrances and services) over these four sequences.

Figure 13 illustrates all three components running at the same time. The number of entrances E_t, and number of persons being served P_t , have been calculated directly using the virtual door approach while the queue size was calculated using the pedestrian counting approach.

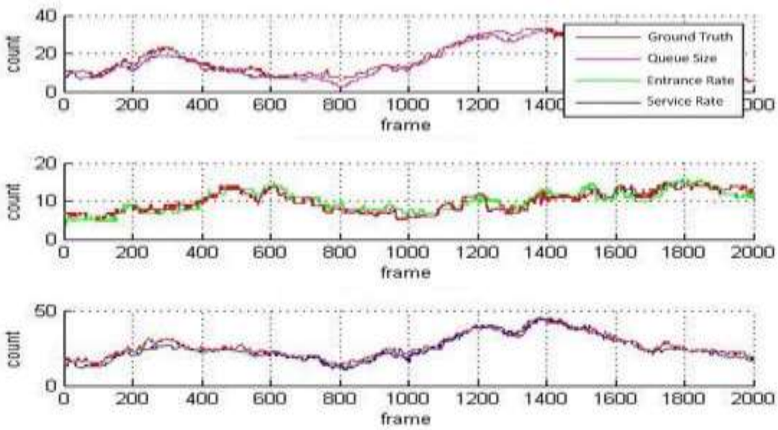

Figure 13. Results of the Proposed Framework Involving Three Parallel Modules on Queue 1

The best estimation results are reposted for the Entrance parameter reported in table 6, in which number of the passenger entering the queue is considered. The MRE is around 6.9%. The percentage of the other sequence frames has reported an error of 22.35%.

One reason of such errors in the entrance parameter is the sudden increase of passengers ‘number. For instance, the sequence 1400-1800 (Table 6) in the test set of queue 1. A large number of pedestrians have arrived and thus, some of their luggages were considered as background. However, considering the leaving parameter, some passengers have unexpectedly moved away. So between 1000-1400 frames sequences (Table 7), the estimations had not been effectively completely yet and a MRE of 31.72% has been reported. Consequently, it has caused some wrong frequency estimations are some cases and an MRE of 14.16% has been reached (Table 8). The motion of non-human items such luggage also resulted in wrong estimation. However, the ground truth of the number of passenger is relatively analogues to the estimates in all three parameters (Figure 9). Hence, the motions of some luggage have caused few error percentages.

Table 6. Cross Validation of Entrance Module on Queue 1

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 1 Dataset	Entrance	200-600	1.655	4.985	6.89%
Queue 1 Dataset	Entrance	600-1000	2.086	6.429	7.805%
Queue 1 Dataset	Entrance	1000-1400	3.258	16.224	9.808%
Queue 1 Dataset	Entrance	1400-1800	3.605	20.709	22.351

Table 7. Cross Validation of Leavening Module on Queue 1

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 1 Dataset	Leavening	200-600	3.642	29.622	9.37%
Queue 1 Dataset	Leavening	600-1000	4.596	38.561	10.11%
Queue 1 Dataset	Leavening	1000-1400	6.412	78.320	31.72%
Queue 1 Dataset	Leavening	1400-1800	5.821	55.012	24.23%

Table 8. Cross Validation of Services Module on Queue 1

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 1 Dataset	Frequency Service	200-600	1.312	3.205	10.33%
Queue 1 Dataset	Frequency Service	600-1000	1.101	8.258	9.22%
Queue 1 Dataset	Frequency Service	1000-1400	2.451	11.841	14.43%
Queue 1 Dataset	Frequency Service	1400-1800	2.323	12.605	14.16%

7.6. Validation of Parameters on Queue2

Table 9 shows the error metrics for each of the virtual door (number of entrances and services and leaving) over these four sequences on queue 2 Dataset

Table 9. Cross Validation of Entrance Module on Queue 2

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 2 Dataset	Entrance	Sequence 1	2.588	3.096	11.25%
Queue 2 Dataset	Entrance	Sequence 2	5.217	20.711	36.16%
Queue 2 Dataset	Entrance	Sequence 3	1.231	3.115	16.43%
Queue 2 Dataset	Entrance	Sequence 4	1.612	3.165	18.11%

Validation of the queue parameters have tested the strength of entrance, leaving and frequency based counting on queue 2 dataset. For all parameters, the queueing model and feature vector has been trained using 1500 frames of the annotated frames. Estimates have been then calculated on the remaining frames 15 min of each set. Qualitatively, the virtual door tracks the variations in passenger flow fairly well. Some errors tend to appear when few passengers exist in the FOV. Such errors are realistic, since that; there are no training samples with such few passengers in Queue2. This problem can be simply solved by applying more training samples.

In some cases, in more challenging situation some errors happens when passengers hols their bags and moving very quickly to the counter desks. Again, such errors (frequency service up to 17% of MRE) are realistic as there are very few samples of passengers fast walking having their bags in the training dataset. These challenges can be fixed by either: 1) applying some mixture components when segmenting the image to label detect high speed walking passengers or 2) identifying outlier (such bags) that have different shape or movement from the passengers. In both scenarios, the process of segmentation is not as simple due to the different view perspective; passengers walking in the foreground regions pass at the similar speed as luggage presented in the background regions.

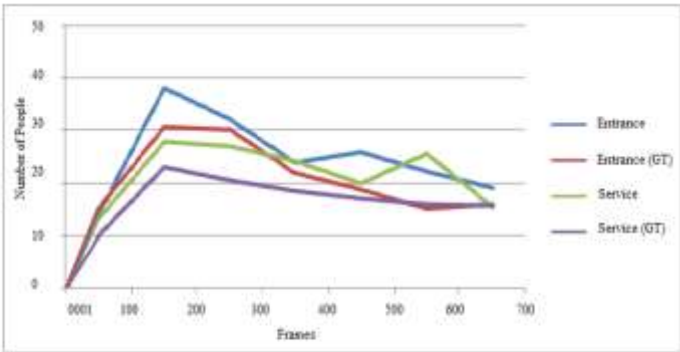

Figure 14. Results of the Proposed Framework Involving Three Parallel Modules on Queue 2

Table 10 presents the estimates error rates for leaving parameters tested on queue 2. In addition, we have reported the total MSE and MRE as an indication of the general performance of the frequency rate. Some points can be concluded. First, the proposed optical flow algorithm has good performance when Gaussian regression model is used. For instance, an MRE of 10.40 is achieved for Gaussian regression. The error rate has been more decreased to 9.65 % by adopting combined virtual door and the applied chains.

Table 10. Cross Validation of Leavening Module on Queue 2

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 2 Dataset	Leavening	Sequence 1	3.121	25.181	10.40%
Queue 2 Dataset	Leavening	Sequence 2	5.645	27.141	11.08%
Queue 2 Dataset	Leavening	Sequence 3	5.265	39.812	33.95%
Queue 2 Dataset	Leavening	Sequence 4	6.080	34.579	27.33%

Table 11. Cross Validation of Services Module on Queue 2

Dataset	Parameter	Frames	MAE	MSE	MRE
Queue 2 Dataset	Frequency Service	Sequence 1	1.113	3.602	9.65%
Queue 2 Dataset	Frequency Service	Sequence 2	1.201	9.101	10.32%
Queue 2 Dataset	Frequency Service	Sequence 3	2.608	14.752	17.21%
Queue 2 Dataset	Frequency Service	Sequence 4	2.505	13.402	16.52%

8. Conclusion

The results above have proved, that the proposed pedestrian counting system developed and the virtual door approach integrated together can be used effectively to monitor queue in complex environment. Some errors have been shown in the pedestrian counting system, this is because the complexity of view, which involves large items and pedestrian carrying their luggage. In addition, some errors have been noted in the virtual door algorithm, this is because the variations in the number of luggage load it by each traveller, blockings between passengers inside the queue, and the distance of the video camera viewpoint. These challenges can be diminished by deploying another camera, which is place nearer to the view, precisely installing only for queue observing purpose instead of all-purpose surveillance. Using an overhead video camera can be practical for this experiment as we can avoid passenger's occlusions.

References

[1] D. Gross, "Fundamentals of queueing theory", John Wiley & Sons, **(2013)**.
[2] V. Parameswaran, "Design and Validation of a System for People Queue Statistics Estimation", Video Analytics for Business Intelligence, Springer**, (2012)**, pp. 355-373.
[3] S. Dhar and T. Rahman, "Case study for bank ATM queuing model", IOSR Journal of Mathematics (IOSR-JM), e-ISSN:, **(2013)**, pp. 2278-5728.
[4] D. Yue, "Analysis of a queueing system with impatient customers and working vacations", Proceedings of the 6th international conference on queueing theory and network applications, ACM, **(2011)**.
[5] X. Zhang, J. Zhang and Y. Zhang, "Pedestrian Detection Using 19-Layer Deep Convolution Neural Network", 2017 4th International Conference on Information Science and Control Engineering (ICISCE), Changsha, China, **(2017)**, pp. 527-531
[6] Z. Ma and A. B. Chan, "Counting People Crossing a Line Using Integer Programming and Local Features", in IEEE Transactions on Circuits and Systems for Video Technology, vol. 26, no. 10, **(2016)**, pp. 1955-1969.
[7] H. Ameur, A. Msolli, A. Helali, H. Maaref and A. Youssef, "Hardware implementation of an improved HOG descriptor for pedestrian detection", 2017 International Conference on Control, Automation and Diagnosis (ICCAD), Hammamet, **(2017)**, pp. 406-410.

[8] A. J. Das and N. Saikia, “Pedestrian detection using dense LDB descriptor combined with HOG”, 2016 International Conference on Information Technology (InCITe) - The Next Generation IT Summit on the Theme - Internet of Things: Connect your Worlds, Noida, **(2016)**, pp. 299-304.
[9] F. Jabar, S. Farokhi and U. U. Sheikh, “Object tracking using SIFT and KLT tracker for UAV-based applications”, 2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS), Langkawi, **(2015)**, pp. 65-68.
[10] M. Peker, “Unattended object detection based on blob tracking”, 2017 25th Signal Processing and Communications Applications Conference (SIU), Antalya, **(2017)**, pp. 1-4.
[11] E. Akagündüz, M. Aslan, A. Şengür, H. Wang and M. C. İnce, “Silhouette Orientation Volumes for Efficient Fall Detection in Depth Videos”, in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, **(2017)**, pp. 756-763.
[12] E. Dibra, H. Jain, C. Öztireli, R. Ziegler and M. Gross, “Human Shape from Silhouettes Using Generative HKS Descriptors and Cross-Modal Neural Networks”, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, **(2017)**, pp. 5504-5514.
[13] G. Mahalingam, “Crowd Motion Analysis from Video”, Proceedings of 2009 NSF Engineering Research and Innovation Conference, Honolulu, Hawaii, **(2009)**.
[14] W.-H. Chen and H.-P. Ma, “A fall detection system based on infrared array sensors with tracking capability for the elderly at home”, 2015 17th International Conference on E-health Networking, Application & Services (HealthCom), Boston, MA, **(2015)**, pp. 428-434
[15] C. Wu, O. Tsimhoni and Y. Liu, “Development of an Adaptive Workload Management System Using the Queueing Network-Model Human Processor (QN-MHP)”, in IEEE Transactions on Intelligent Transportation Systems, vol. 9, no. 3, **(2008)**, pp. 463-475.
[16] S.Ali, and M. Shah, “Floor fields for tracking in high density crowd scenes”, Computer Vision–ECCV 2008, Springer**, (2008)**, pp. 1-14.
[17] T.-Y. Chen, “A people counting system based on face-detection”, Genetic and Evolutionary Computing (ICGEC), 2010 Fourth International Conference on, IEEE, **(2010)**.

