OPEN ACCESS
AMERICAN JOURNAL OF SOCIAL ISSUES AND HUMANITIES

Distributed Monitoring System Based On Weighted Data Fusing Model

[1]Reda Shbib, [2]Shikun Zhou, [3]David Ndzi and [4]Khalil Al-Kadhimi

School of Engineering,University of Portsmouth, UK.
[1]reda.shbib@port.ac.uk, [2]Shikun.zhou@port.ac.uk, [3]david.ndzi@port.ac.uk, [4]Khalil.al-kadhimi@port.ac.uk

Abstract
This paper presents a novel people counting system based on weighted fusing technique using camera and infrared sensors. Conventional techniques are based image vision data to count people. However, these vision-based people counting methods frequently fail due to irregular illumination change and crowded environment and some geometrical limitation that may affect the accuracy level of these systems. In this paper, a dynamic data fusion technique is provided in order to provide more accuracy results.
Keywords
Counting system, Data Fusion, Image Processing

Introduction

Because of the rapid development of the business and commercial economy, lots of large shopping centres and public services (train stations, building, and underground) are usually crowded. In addition, most of crowd public places need large number of persons in order to provide various types of services and usually costumers look for much better services. Thus, human recourses sectors should have a clear visions and strategies in order to provide an efficient management, either personal, security or assets managements.

Knowing the number of people for a specific location provides vital information for business decision makers. It might be used as an indicator for purchasing popularity of chain stores or Shopping centre, the number of people over a period of time (year, month) can be used to predict the maximum crowded period over the year, furthermore, it is possible to estimate the number of visiting persons at the entrance of a stadium in order to use this information for a security purposes. . Effective counting system of people is extremely vital and it has high value for many applications

Infra- red sensors

Even though the traditional manual techniques for counting people can provide and efficient accuracy level, but it can cost too much amount of efforts and time. Currently, many automated counting system that uses shaft devices, has been proposed in order to solve this problem, which can reduce the high cost of human intervention with a satisfactory level of accuracy. However, one of its disadvantages is that just one person has to pass through at a time, which is not suitable for counting pedestrian in a crowd.

(IR) Infrared Wireless Sensors are the most dominant techniques use by counting people systems. Infrared lines are being interrupted in order to count the number of persons. The motions and the direction of people are recognized based on the interference of two or more sensors, however, if many persons (in case of beam) passed the entrance at the same time it can affect on the accuracy level. The new technologies of high processing computer has led to a new surveillance video systems based on video and image processing techniques which provides automatic monitoring and controlling systems

Many researches and techniques have been developed over the last few years in the area of image processing with the purposes of getting more reliable and accurate counting- people systems. Furthermore, new methods are introduced and interested in this filed. These developments have significant impacts and advantages such as:

- The images provides has a high quality and the level of accuracy is extremely higher than infrared sensors
- Marinating of video surveillance system is easy and economical
- The collected data obtained from a crowd people will help to perform remote managements in a real time

There are many points that must be taken into consideration in such systems. As an example, camera monitoring the passageway in a large place such an airport or underground station, or set of cameras located in a shopping centre which covers different regions within the centre, or camera in a busy street. These entire sensors camera offers a big amount of images which might be used to count the number of persons at different interval of times; however the situations of the surrounding environment where the sensors are embedded can be challenge in determining the accuracy of the system

Counting passing people using image processing is used to estimate how many persons in input images. Lots of information such the number of people and their movement and position based on image processing can be very useful in reduce the observation cost, in addition, the information about pedestrian are very important in many types of applications.

This paper is organized into the following sections. Section 3 provides the literature survey of counting techniques and their limitations and some related works on people counting system. Section 4 provides an insight into the proposed fusion techniques of results coming from two-different sensors. Finally section 5 concludes the paper.

Image Processing and Counting Techniques Constraints

There have been a many of techniques has been developed in order to solve the counting people problems, they can be classified for three major categories:

1. People detection and tracking
2. Feature trajectories clustering
3. Low-level features regression

Feature trajectories clustering Methods: These techniques are based on feature trajectories clustering (Ya-Li and Pang, 2011) . First, some of features which can be tracked such corners are extracted, then A feature trajectories is formed by tracking the extracted feature frame by frame and then trajectories are clustered based on a similarity measurements, at last, each cluster represent one person. This type of techniques are specially fit for counting people flow in crowded places, for some degree, they can decrease the effects of occlusion and they are for somehow impervious to the variation angles of cameras.

Reference (Morris and Trivedi, 2008) has stated two main disadvantages for techniques of counting of feature trajectories. The first problem that, it needs a complex trajectory management ("*handling broken feature tracks due to occlusions or assessing similarities between trajectories of different length"*)

Second, in crowded scenes, it is often the case that some of coherently moving features do not fit to the same person. Therefore, matching the number of persons to the number of trajectory clusters can lead for many errors.

Low-level features regression Method: Regression-based crowd counting has been first developed and applied for monitoring underground platforms. These methods typically work by:

1) Background Subtraction.(Chan et al., 2008, Ryan et al., 2010)
2) Extracting several features from the foreground pixels, such as (Total area, perimeter, edge, orientation of edges, textures) (Paragios and Ramesh, 2001), (Siu-Yeung et al., 1999), (Davies et al., 1995), edge count ,or texture (Marana et al., 1998)

Estimating the density of the crowd or counting crowd using regression methods, such as piecewise linear (Regazzoni and Tesei, 1996) ,or with neural networks methods (Marana et al., 1998). Counting goal can be achieved without any detection or tracking process. However, these methods only offer an estimation of crowd density. If textures from the background and person clothing are complicated, these techniques may not be well effective.

Related works

Camera-based algorithm uses the individual shape and texture information. An algorithm has been proposed by (Chih-Chang et al., 2011) In which the silhouette of person is extracted and tracked even though the computational complexity is comparatively low, miss counting error has been reported when occlusion is occurred when tracking a large number of persons

A real-time system has been developed by Z Tao et al.(Tao et al., 2008) A box based feature extraction has been used in combination with a background extraction in order to segment the individual. Errors result has been recorded when the background and the detected persons share same colures in addition; the background subtraction techniques couldn't extract each person successfully from the crowd.

Currently several and significant results has been achieved by directly detecting persons from images has attempted to extract several features from a static camera has been developed. Haar wavelets, SIFT-like features, Histogram of Oriented Gradients, these techniques could attain for some degree more accuracy counting level and human detection results in small crowded location. However, the majority of these techniques are time consuming and only has been showed results for a crowd with a small amount of occlusions. High resolution images are also required for such type of techniques.

Using the overhead camera is adopted to simplify the counting task by avoiding occlusions among people and simplifying the human model. However, the overhead camera can only monitor a very limited region, which is not suited for wide area monitoring.(Jaijing et al., 2009)

Some of the main problems of counting systems in video processing are the changing of light, shadow, and crowded environment with unstable background, (Yahiaoui et al., 2010) Furthermore , it appears that many of counting systems are based on the assumption that any moving object are persons which lead to low accuracy . In addition, Some of these system are time consuming and has showed results for small crowded locations with few occlusion.

It is necessary to have robust real-time counting system that are able to report high performance with little manual reconfiguration that have to be sufficiently extensible and adaptable in order to automatically adjusted and be able to cope up with any changes in the environment under surveillance such as the geometry of the scene, activities taking place in the scene, and lighting, location and period of the day.

Crowd Segmentation and Object Detection: The majority of smart surveillance systems provide for object detection as the first module in the pipeline of video processing. Its main intended use is the detection of relevant objects within the scene.

Several proposed techniques about detecting moving object; background subtraction and the frame- difference models have been widely used (Naylor and Attwood, 2003) Optical flow algorithms are used in one group of methods. On the one hand, they estimate the movements via calculating the gradient for the whole image or analyze a certain scene to emphasize the characteristic aspect (Tamgade and Bora, 2009). On the other hand, such methods are expensive in the context of computation. Moreover, they require videos of high frame rate and images of high resolution; thus, such methods are not appropriate for their real time processing. Nevertheless, certain tasks cannot be accomplished without optical flow calculation algorithms (Lenz et al., 2011). So, there are two cases when this type of approaches is beneficial. If it is necessary to analyze the video stream from a non-static camera, optical flow makes it possible to compensate for certain local changes while detecting the global motion (Thakoor and Gao, 2004). The other usage is the estimation of a general motion of the crowd when it is difficult to get the information about the movement of single objects in crowd scenes (Fradi and Dugelay, 2012).

The following set of methods helps to implement the detection of objects, the appearance of which is known in advance. In case it is not practical to get a background image and it is necessary to detect a single image or a group of objects, this approach is of great use .The object or its part can also be detected on the basis of certain features of the models supplied (Ryan et al., 2010). It is applicable in case of detecting some registration plate or face. Popular algorithms of such kind use certain cascade classifiers on a dataset provided and Haar-like features (Khac et al., 2009). Another approach known as a Mean-shift based approach is applied in case one is aware of the color distribution of the detected object (Hsu and Hsu, 2004).

Gaussian Mixture Models based method (GMM): The chain of video processing includes the first stage of detecting the moving objects. Afterwards, other processing modules utilize the results of this process. Temporal and spatial information is typically required for generating object binary masks in most algorithms of video segmentation (Hongliang and King, 2007) . Common time-averaging function of frames is not suitable for the systems of surveillance. The video frames have insufficient adapting capabilities, so a different solution is required. Gaussians Model allows the detection of moving objects with the help of background subtraction and spatial segmentation. This useful method is based on modeling mixtures of Gaussians or pixels and updating the model through online approximation (Qin and Yaonan, 2008).It can be used to cope with the change of lighting and to adapt to some variations on the observed scene. The background model can even allow changes in the color of the pixel, being multi-modal. So, it is not a problem to model the sequences of traffic lights or waving of the branches. The purpose of background modeling is differentiating the foreground pixels of the objects that move and model the real time background of the observed scene (Dalka, 2006) .

In this model the value of particular pixel over the time is seen as a mesurment X_t of a stochastic variable. At any time along the current measurment of X_t , The history M_t= { X_1 , X_2 , …, X_{t-1} } is known (Stauffer, 1999).

Therefore , the current history of a particular pixel can be moddeled by mixture of K Gaussian distrubutions. Different colours are supposed to denotes as different Gaussian .The probability to detect the current background pixel X_t is the weighted sum of the K distrubution

$$P(X_t) = \sum_{i=1}^{K} w_{i,t} * \mathcal{F}(X_t, \mu_{i,t}, \Sigma_k)$$

K= number of Gaussian distrubutions

$w_{i,t}$, is the wieght of i^{th} distrubution at time t ,and the $\sum w_i$ =0

Where μ_k is the mean and Σ_k is the covariance matrix of the kth density

Thus , how longer a color is staying in the picture in represented by the probability density function :

$$\mathcal{F}\left(X \middle| \mu_{i,t}, \Sigma_k\right) = \frac{1}{(2\pi)^{\frac{n}{2}} \left|\Sigma_k\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(X-\mu_k)^T \Sigma_k^{-1}(X-\mu_k)}$$

Fig.1: Individual Detection using GMM (video taken from INIRIA Dataset)

It is possible to detect even a shadow that moves simultaneously with the object while implementing the segmentation of the object itself. The shadow is detected with the help of an algorithm of a background removal as a component of a foreground object. The shadow is not as bright as the object itself, but in terms of chromatic composition the shadow is generally the same (Horprasert et al., 1999) . The background subtraction process implies recognizing all new pixels as components of a foreground object and checking them to determine if they make up the moving shadow. Every darker pixel is considered to be a part of a shadow and background scene.

Background modeling provides for obtaining a binary mask that denotes current frame pixels of foreground objects. Object segmentation in its turn is allowed by morphological processing It comprises the discovery of interconnected parts, eliminating objects that are not big enough, morphological closing and holes that are to be filled in certain areas. In addition, morphological reconstruction is applied to implement the algorithm of removing the shadow (Dougherty and Lotufo, 2003).The mentioned procedure implies using a marker and a mask as binary images. The background pixels have the value of zero, while the shadow or moving object pixels have the value of one in the mask. After all the pixels of the shadow are removed, it is possible to obtain a marker.

When the moving objects start or stop, they cause the appearance of ghosts, which are to be detected and removed. It is a challenging part of object segmentation. There are two main conditions of appearance of the ghosts. Firstly, a moving object can get stationary and merge into the background. There is a ghost behind it after it starts moving. Secondly, when a background objects starts moving, like a car that used to be parked, it causes this problem as well. Comparison of similarities between the frame edges and edges of foreground objects can serve as a possible solution (Yin et al., 2008). Such comparison is based on the awareness of the peculiarities of moving objects.

Fig.2: Image backgournd

Fig.3: Image segmentation + backgournd substraction

Fig.4: Peple Detection and tracking

Information Fusion: The proposed system is based on an effective weighted fusion of the information extracted by the IR sensors and the vision image. In the proposed technique, the input Data coming from Infrared-sensor and camera has to be fused in order to improve the accuracy level of the system. Infrared-sensor will provide more information which complements the image obtained in the visible range. The visible images permit a rich content where the detection of persons can however be affected and limited by the changing in lighting and many other factors such occlusions. An intelligent fusion of the information offered by both sensors might decrease the amount of false alarm and the advent of non-detected persons, thus, increasing the performance of an individual detection and monitoring system

Fusion of sensors' information provides many potential advantages as follows (Elmenreich, 2002)

1. Using of different types of sensors in our system would offer redundancy, which, allow the system keeping delivering information needed in a partial failure situation data might be loosed from a sensor, the capability of fault tolerance and efficient functionality.

2. Improving geometrical coverage and the availability of complementary information.
3. Accuracy level of one sensor is confirmed by the measurements of the other sensors, getting cooperative arrangement and improving the confidence of the obtained results
4. Cooperative information would tend to decrease uncertainty interpretations
5. Expanded temporal coverage, data is continuously available.
6. Better objects detection because of less ambiguity offered by the fusion procedure

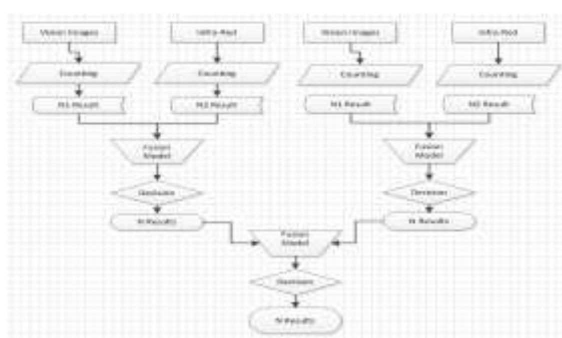

Fig.5: Integrated Image + Infra-Red System

Proposed Data Fusion Model

The proposed model is based on an effective weighted fusion of the information extracted by the IR sensors and the vision image. In the proposed technique, the input Data coming from Infrared-sensor and camera has to be fused in order to improve the accuracy level of the system. Infrared-sensor will provide more information which complements the image obtained in the visible range. The visible images permit a rich content where the detection of persons can however be affected and limited by the changing in lighting and many other factors such occlusions. An intelligent fusion of the information offered by both sensors might decrease the amount of false alarm and the advent of non-detected persons, thus, increasing the performance of an individual detection and monitoring system

Scenario (a)

Let's start with a basic scenario with the following design parameters

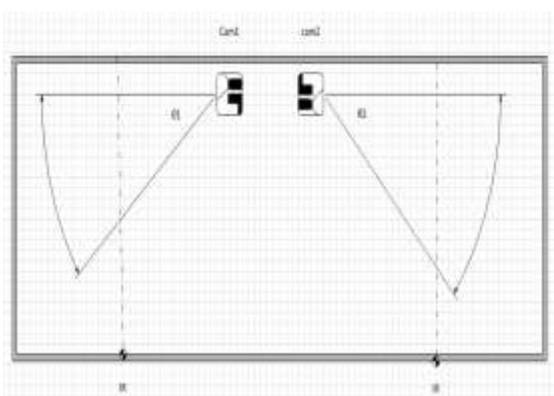

Fig. 6: Basic scenario with 4 sensors

- *4 sensors: 2 cameras + 2 Infrared Sensors*
- *Let i = number of sensors (NoS)*
- *N i is the number of people detected by first cam*

We are making the following assumptions:

1. *Θ1 = θ2 = θ*
2. *All sensors has 100% accuracy*

Given the above scenario and parameters: the total number of pedestrians (N) can be given by the following equations:

$$\boldsymbol{N = (\alpha_1 N_1 + \beta_1 N_1) + (\alpha_2 N_2 + \beta_2 N_2)}$$

Where α_i *and* β_i *are the weights for the video and Infra-Red sensors respectively which follow the table below*:

i	α_i	β_i
1	1	0
2	0.5	0.5
3	$^1/_3$	$^1/_3$
4	0.25	0.25

The value of α_i *and* β_i *is given by* $\alpha_i = \beta_i = \frac{1}{NoS}$ *as all the sensors has the same accuracy and covering the same area.*

For *i number of sensors, the total number of people can be given by the following equation:*

$$N = \sum_{i=1}^{n} \alpha_i Ni + \beta_i N_i$$

Identifying α and β with different accuracies and geometric challenges

Identifying α and β With Different Accuracies and Geometric Challenges

The main idea of the proposed model is to calculate the weights associated with outputs of each sensor start from identifying and grouping various factor affecting the accuracy of each sensor as well as their impact level of counting .Identifying α and β depends on many factors and on the situation when collecting the data. Fig 2 shows some of the factors that can affect the accuracy of sensors

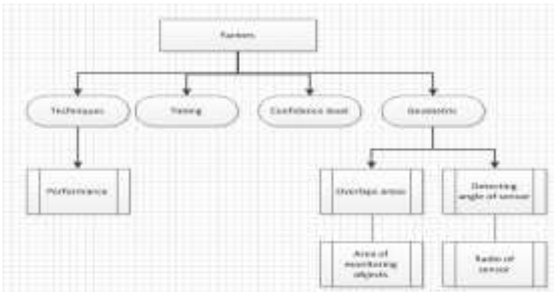

Fig.7: Factor affecting accuracy

Fuzzy Logic

The point of fuzzy logic is to map an input space to an output space, and the primary mechanism for doing this is a list of if-then statements called rules. All rules are evaluated in parallel, and the order of the rules is unimportant. The rules themselves are useful because they refer to variables and the adjectives that describe those variables. Before we can build a system that interprets rules, we have to define all the terms we plan on using and the adjectives that describe them.

To say that the accuracy level is good, we need to define the range that the accuracy's level can be expected to vary as well as what we mean by the word accuracy. The following diagram provides a roadmap for the fuzzy inference process. It shows the general description of a fuzzy system on the left and a specific fuzzy system

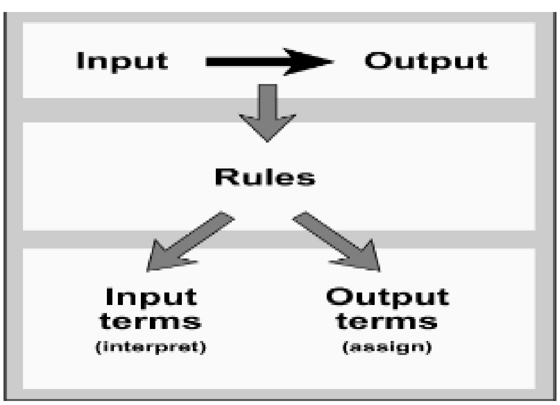

Fig.8: Fuzzy inference

Using Fuzzy inference method we can interprets the values in the input vector (factors) and, based on some set of rules, assigns values to the output vector (weights)

Identifying weights using Fuzzy Inference system

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned.

Variation of α

- Input 1 = Accuracy
- Input 2 = Angle of coverage

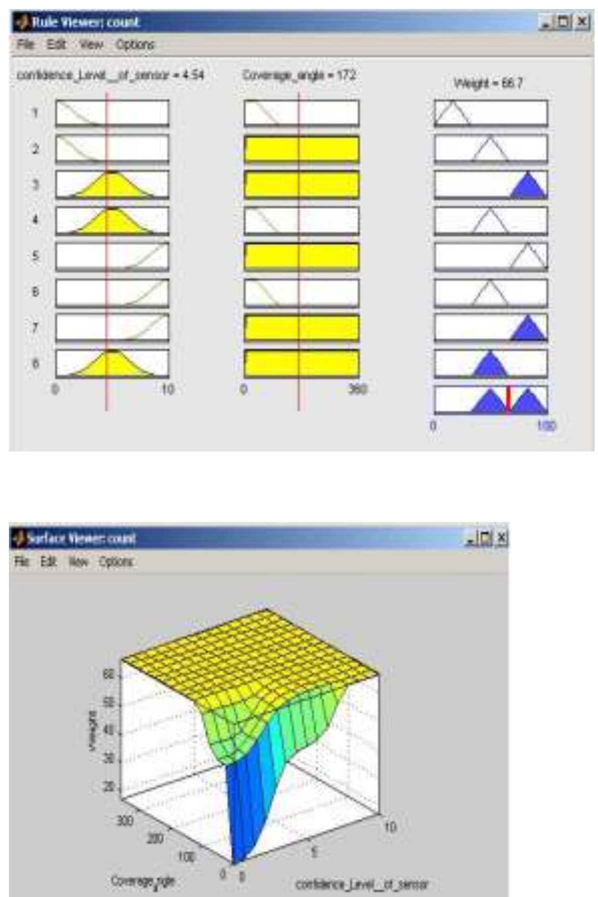

Fig .10: Variation of weights

Conclusion

We proposed the fusion sensor based people counting technique, which is robust to the illumination change and crowded condition. The proposed data fusion model consists of determining main weighted parameter that is able to cope with the variation of accuracy level and the geometric limitations in order to provide one weighted value. Future work is how to determine the value of the sensor weights based on the situation assessment and the parameter that may affect the accuracy of the sensors.

References

1. Chan, A. B., Liang, Z. S. J. & Vasconcelos, N. Year. Privacy preserving crowd monitoring: Counting people without people models or tracking. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, 23-28 June 2008 2008. 1-7.
2. Chih-Chang, C., Hsing-Hao, L. & Chen, O. T. C. Year. Tracking and counting people in visual surveillance systems. In: Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, 22-27 May 2011 2011. 1425-1428.
3. Dalka, P. 2006. Detection and segmentation of moving vehicles and trains using Gaussian mixtures, shadow detection and morphological processing. MG\&V, 15, 339-348.
4. Davies, A. C., Jia Hong, Y. & Velastin, S. A. 1995. Crowd monitoring using image processing. Electronics & Communication Engineering Journal, 7, 37-47.
5. Dougherty, E. & Lotufo, R. 2003. Hands-on morphological image processing.
6. Elmenreich, W. 2002. An Introduction to Sensor Fusion. Austria: Vienna University of Technology,.
7. Fradi, H. & Dugelay, J.-L. Year. People counting system in crowded scenes based on feature regression. In: Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European, 27-31 Aug. 2012 2012. 136-140.
8. Hongliang, L. & King, N. N. 2007. Automatic video segmentation and tracking for content-based applications. Communications Magazine, IEEE, 45, 27-33.
9. Horprasert, T., Harwood, D. & Davis, L. S. 1999. A statistical approach for real-time robust background subtraction and shadow detection
10. Hsu, C.-C. & Hsu, C.-C. 2004. An intelligent interface for customer behaviour analysis from interaction activities in electronic commerce. Proceedings of the 17th international conference on Innovations in applied artificial intelligence
11. Ottawa, Canada: Springer Springer Verlag Inc
12. Jaijing, K., Kaewtrakulpong, P. & Siddhichai, S. Year. Object detection and modeling algorithm for automatic visual people counting system. In: Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2009. ECTI-CON 2009. 6th International Conference on, 6-9 May 2009 2009. 1062-1065.
13. Khac, C. N., Park, J. H. & Jung, H.-Y. 2009. Face Detection using Variance based Haar-Like feature and SVM. World Academy of Science.
14. Lenz, P., Ziegler, J., Geiger, A. & Roser, M. Year. Sparse scene flow segmentation for moving object detection in urban environments. In: Intelligent Vehicles Symposium (IV), 2011 IEEE, 5-9 June 2011 2011. 926-932.

15. Marana, A. N., Costa, L. F., Lotufo, R. A. & Velastin, S. A. Year. On the efficacy of texture analysis for crowd monitoring. In: Computer Graphics, Image Processing, and Vision, 1998. Proceedings. SIBGRAPI '98. International Symposium on, 20-23 Oct 1998 1998. 354-361.
16. Morris, B. T. & Trivedi, M. M. 2008. A Survey of Vision-Based Trajectory Learning and Analysis for Surveillance. Circuits and Systems for Video Technology, IEEE Transactions on, 18, 1114-1127.
17. Naylor, M. & Attwood, C. I. 2003. Annotated digital video for intelligent surveillance and optimised retrieval (ADVISOR). iniria laboratories . France.
18. Paragios, N. & Ramesh, V. Year. A MRF-based approach for real-time subway monitoring. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, 2001 2001. I-1034-I-1040 vol.1.
19. Qin, W. & Yaonan, W. Year. Background subtraction based on adaptive non-parametric model. In: Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on, 25-27 June 2008 2008. 5960-5965.
20. Regazzoni, C. S. & Tesei, A. 1996. Distributed data fusion for real-time crowding estimation. Signal Processing, 53, 47-63.
21. Ryan, D., Denman, S., Fookes, C. & Sridharan, S. Year. Crowd Counting Using Group Tracking and Local Features. In: Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on, Aug. 29 2010-Sept. 1 2010 2010. 218-224.
22. Siu-Yeung, C., Chow, T. W. S. & Chi-Tat, L. 1999. A neural-based crowd estimation by hybrid global learning algorithm. Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on, 29, 535-541.
23. Tamgade, S. N. & Bora, V. R. 2009. Motion Vector Estimation of Video Image by Pyramidal Implementation of Lucas Kanade Optical Flow. International Conference on Emerging Trends in Engineering and Technology.
24. Tao, Z., Nevatia, R. & Bo, W. 2008. Segmentation and Tracking of Multiple Humans in Crowded Environments. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 30, 1198-1211.
25. Thakoor, N. & Gao, J. 2004. Automatic Object Detection In Video Sequences With Camera In Motion
26. Ya-Li, H. & Pang, G. K. H. 2011. People Counting and Human Detection in a Challenging Situation. Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, 41, 24-33.
27. Yahiaoui, T., Khoudour, L. & Meurie, C. 2010. Real-time passenger counting in buses using dense stereovision. Journal of Electronic Imaging, 19, 031202-031202.
28. Yin, F., Makris, D. & Velastin, S. A. 2008. Time efficient ghost removal for motion detection in visual surveillance systems. Electronics Letters, 44, 1351-1353.