Counting People Using Blobs and Contours

Shafraz Subdurally, University of Mauritius, Réduit, Mauritius
Devin Dya, University of Mauritius, Réduit, Mauritius
Sameerchand Pudaruth, University of Mauritius, Réduit, Mauritius

ABSTRACT

Counting the number of people in public locations has become imperative in surveillance applications for the good management of public space. The automatic counting of people can indeed help carry out the above tasks better and faster. In this paper, the authors propose two systems for counting people from images. Their proposed methods are based on the observation that heads are significantly more visible than any other features and are thus more easily distinguishable. The proposed systems use blobs and contour detection respectively to count the number of people. The results obtained from each system are very reliable. The average head detection rate of the systems is 82 and 84 percent respectively.

Keywords: Blob, Contour Detection, Crowd, Observation, Public Space

1. INTRODUCTION

People have always attended their preferred political party’s meeting, supported their favourite team at a sports event, demonstrated a particular idea during a rally or simply done some shopping during the new year’s eve. Manual counting of people in such places can often be difficult since they are not arranged in a specific order and they may be in constant motion. The aim of counting people is numerous and a few conventional ways of counting people include using laser beams, infra-red sensors, thermal sensors (Lefloch, 2007), etc. Counting people in a particular location can help in decision making such as to see if there is a need to enlarge the location area for greater pedestrian’s safety (Anthony C. Davies, 1995). Events held on streets like concerts and public meetings make counting people more difficult. People counting systems have evolved in recent years from manual head counts to computer vision neural learning but still most recent systems do not perform efficiently where there is a large number of people with occlusions (Lim Eng Aik, 2009). Huge accumulation of people in such events poses real life threats to the people (Damian Roqueiro, 2006). A slight mismanagement or panic may put at risks many innocent lives. As a matter of fact, it is imperative that such locations where people are likely to be accumulated must be well designed or at

DOI: 10.4018/ijcvip.2013040101
least well monitored for their safety. One way of avoiding such a mishap is to count the number of people present there. Assuming that the number of people should not exceed a particular value, the system may raise an alarm if there is an over accumulation of people. However, counting people in crowded environments still remain an elusive task.

People count can be used to monitor the limit of occupancy – ensuring that the people count do not exceed a safety level for if the number of people exceeds the capacity, people may get crushed or the structure they are standing on may even collapse. Manual counting of people has failed for different reasons. Firstly, people tend to lose concentration when counting and thus the counts are prone to errors. Secondly, it is a very time consuming task as well as difficult when people’s positions are not fixed. Thus, there is a pressing need for an automated system that can detect and count people in crowded environments with reasonable accuracy.

This paper proceeds as follows. Section 2 provides an in-depth literature review where state-of-the-art approaches to estimate crowds size are described. In Section 3, our proposed systems are explained in detail. The solutions proposed are then tested and evaluated in Section 4. Section 5 concludes the paper with a view on the weaknesses of the proposed systems and ideas for future work.

2. LITERATURE REVIEW

In Dan Kong (2006), a learning-based method for counting pedestrians in crowds has been proposed. The system uses edge orientation and blob size histograms. This is done by applying background subtraction and edge detection to each frame. The features are then extracted (Figure 1).

The points in the image and the point on a 3D plane are related by a plane perspective transformation called homography. Having obtained the features, a homography is calculated between the ground plane and the image plane coordinates for the region of interest (ROI). For feature normalization, a density map measuring the relative size of persons and a global scale measuring camera orientation have been used (Figure 2).

Two training methods are used: Linear Fitting and Neural Networks. There two types of training because feature size and pedestrian counts do not have a simple linear relationship. As such, Neural Networks is used for the case of occlusions in the crowd and counting of people using Linear Fitting is used based upon edge and blob histograms.

In Danny B. Yang (2003), the system makes use of a group of image sensors connected through a network to estimate crowd density (Figure 3).

The first step is to compute this projection from the silhouettes measured by the sensors through background subtraction. The projection is a set of polygons. Figure 4 shows some top views of the resultant projections obtained:

![Figure 1. Features of Pedestrians: (a) original image, (b) foreground mask image (c) edge detection map, (d) the edge map after the ‘AND’ operation between (b) and (c)](image-url)
Encoding Human Motion for Automated Activity Recognition in Surveillance Applications

Calculating Absolute Scale and Scale Uncertainty for SfM Using Distance Sensor Measurements: A Lightweight and Flexible Approach