Initial Plan

Project Title: Detecting Crowd Characteristics that are Indicative of Disorder and Violence in City Centre Locations from Video Data

Author: Kaelon Lloyd
Supervisor: Prof. David Marshall
Moderator: Prof. Paul Rosin

Module Number: CM3203
Module Title: Large One Term Individual Project (61)
Credits Due: 40

Project Description

Around all major city environments a multitude of close circuit television (CCTV) cameras are placed strategically in order to provide video surveillance. The purpose of these cameras is to provide real-time playback of footage in order to identify disorderly or illegal activity in public environments.

The video steam of each camera is sent directly to large array of monitors inside a police or security centre for observation. When disorderly behaviour is identified nearby officers will be alerted and directed towards the scene; if no nearby police personal are present then a team will be dispatched.

The response time of the police force can make a huge difference in their effectiveness at providing public safety and so it's imperative that any illegal or disorderly conduct be identified as soon as possible.

Observation of the CCTV footage is done by human personnel. The job of the observer can be very demanding due to the high number of real-time video streams being monitored at once. Failing to identify a situation can lead to negative outcomes and so it's a requirement that observers stay attentive at all times.

This project will be used to develop assistive software that can aid the observer in order to make the job of an observer less intensive and to hopefully reduce the number of missed incidents. This project does not propose to replace the human observer, but to only act as a notification program to draw human attention to a specific video steam suspected of displaying disorderly or violent conduct.

The aim of the project is to devise and implement a video analysis system that can learn crowd characteristics in order to automatically detect disorderly or violent behaviour from

CCTV footage. Given the context of the situation it is important that the system developed can operate in real-time with reasonable accuracy.

A small selection of CCTV footage from inside Cardiff City centre has been provided by South Wales Police. The provided footage depicts various scenes of violence and scenes of non-violence; this data will be used to benchmark the performance of the system.

Project Aims and Objectives

Aim: Locate more instances of CCTV camera footage depicting both disorderly and non-disorderly behaviour for training and testing.
Objectives:

- Investigate data set repositories for videos of inner city CCTV footage
- Investigate data set repositories for videos that resemble inner city CCTV footage
- Obtain CCTV footage from public media websites and annotate the data
- Modify all videos obtained outside of the initial South Wales Police data set such that they have identical frame rates.

Aim: Remove noise data from video files
Objectives:

- Implement a method of video stabilization to remove any jitter caused by weather conditions
- Image correction for lens distortion; each camera may require different correction parameters
- Resolve colour constancy issues where applicable
- Remove unnecessary data contained on video frames such as timestamps

Aim: Create a user-interface and program structure for effective system testing and evaluation
Objectives:

- Generate multiple graphs and figures that describe the effectiveness of scene classification
- The program structure should be modular to allow for rapid prototyping and easy extension
- Provide a method of displaying a target video with classification results overlain for an intuitive way of visualising system performance

Aim: Extract meaningful features for scene description
Objectives:

- Extract crowd dynamics using motion estimation techniques
- Identify regions of interest for image region extraction
- Normalize temporal features to account for variation of sampling rates across multiple video files
- Research methods of spatial feature normalization to provide accurate classification against scenes at different scales
- Derive suitable feature descriptors from extracted information for system training and testing
- Information extracted temporally must be sampled in windowed segments to reduce the scope of information
- Explore temporal correlations between actions in order to create a more descriptive feature set
- Reduce spatial correlation between features for successful matching between similar scenes with spatial diversity.

Aim: Develop a system fork capable of performing classification in real-time
Objectives:

- Perform code optimization by taking full advantage of MATLAB functions
- Implement feature extraction and processing methods that favour speed over robustness
- Research the prospects of implementing hardware acceleration

Aim: Investigate and implement methods of machine learning and classification
Objectives:

- Research machine learning techniques
- Generate sample data to test the performance of machine learning techniques that may be used in the system.

Work Plan

Supervisor Meetings

Meetings have been scheduled to take place on Friday every week at 12 noon to discuss project status.

Final Report Deliverables

The following is a list of deliverables that I aim to include in the final report:

- Software demonstrating program capabilities
- Weekly status reports describing advancements made during each week
- Description of matters discussed in supervisor meetings and how they effected the course of development
- Documentation explaining
    - Chosen approach to solving the problem with analysis of each core component of the system
    - Code nomenclature
    - Design styles and interface choices
    - Methods of testing used during development
    - How results progressed after each stage of testing
- A conclusion stating the final projects ability to solve the provided problem
    - Statistical, objective point of view
    - Personal, subjective point of view
- List of unsuccessful aims and objects with reasons of failure
- Further Improvements

Week 1 (27/01/2014): *Perform project research and develop an initial Plan*

Week 2: *Obtain more video data and normalize sample rates*

The data supplied by South Wales Police contains few instances of disorderly behaviour and so to achieve adequate testing more CCTV footage is required; footage can be found on research data set repositories or public media sites such as YouTube.

Week 3: *Research and implement noise reduction methods*

Data provided by South Wales Police displays useless information on frame and is also highly susceptible to camera shake. These are two issues I will need to address from the provided data; self-obtained videos may need further correction.

Week 4: *Research and implement motion extraction*

This time will be spent researching the best way to extract crowd dynamics and how to best represent the data in a feature vector.

Week 5 - 6: *Research and implement a method of feature classification*

At this stage the aforementioned crowd dynamics will be available for classification; using the derived feature vectors a suitable method of classification can be found and implemented.

Week 6: *Create an intuitive interface for outputting classification results*

Suitable diagrams and graphs need to be generated to reflect results from the classifier so that performance evaluation can be done accurately and efficiently.

Week 7 - 8: *Explore spatial and temporal correlations in features extracted from crowd dynamics*

Look into improving the features extracted from crowd motion by identifying temporal and spatial correlations between features.

Week 9 - 10: *Research and implement methods of region extraction and description*

Find a method of representing small segments of a video frame to identify objects; attaching visual descriptors to motion information should provide more discrimination between scenes found in CCTV footage.

Week 11: *Implement video playback with classification overlay*

Being able to display how the system classifies video as its playing will be an intuitive method of visualizing results.

Week 11*: Perform bug fixes*

Week 11 - 12: *Exhaustive testing with evaluation of results*

Week 13 + 14: *Produce the Final Report* **(06/05/2014: Hand in Date)**

Work Plan

Task	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Produce final report													███	███	
Exhaustive testing with evaluation of results											███	███			
Perform bug fixes											███	███			
Implement video playback with classification overlay											███				
Research and Implement methods of region extraction and description									███	███					
Explore spatial and temporal correlations in features extracted from crowd dynamics							███	███							
Create an intuitive interface for outputting classification results						███									
Research and implement a method of feature classification					███	███									
Research and implement motion extraction				███											
Research and implement noise reduction methods			███												
Obtain more video data and normalize sample rates		███													
Write Initial Report	███														