Head Curve Matching and Graffiti Detection

Jing Wang
j.wang2@hud.ac.uk
Zhijie Xu
z.xu@hud.ac.uk
Michael O'Grady
m.ogrady@hud.ac.uk

Computer Graphics, Image and Vision (GCIV) Research Group
University of Huddersfield
Huddersfield, UK

Abstract

Vandalism is a type of the so-called low intensity crimes. However, the nuisance and damage caused by it could cost dearly to the public to clean up. It is also technically difficult in the past to detect and prosecute such a crime, for example, graffiti making. This paper presents an automatic and live graffiti detection system using CCTV cameras and innovative vision algorithms. After a brief introduction on the problem domain, the paper proceeds to proposing a system prototype through highlighting its essential operations such as human shape and head curve recognition for triggering the follow-up operations of graffiti detection. The research methodology, the devised algorithms, and the experiment designs are then explained in details. The experiments in this programme are based on real CCTV footages recorded at a bus station. The result shows that the proposed system can detect graffiti-making actions and to locate graffiti occurring areas effectively under most live illumination conditions and is a practical solution for the design of intelligent real-time CCTV surveillance systems.

Keywords: graffiti detection, human recognition, curve matching.

1 Introduction

Traditional CCTV surveillance systems for detecting vandalism require constant human operator-based monitoring and intervention. Even for the off-line style reviewing task once vandalism act has occurred, it is an even more painstakingly slow and tedious manual job.

During the past decade, remarkable advances have been made on intelligent surveillance systems and their applications based on innovative image processing and pattern recognition techniques. This paper proposes an automatic live graffiti detection method that is capable of recognizing graffiti occurring areas from live video feeds. In this project, the implemented prototype system (mounted at a bus station) can successfully detect graffiti and graffiti-making activities through remove noise signals introduced by passing pedestrians and vehicles as well as adapting to changing outdoor illumination conditions.

One common assumption for graffiti detection is based on its time-endurance feature due to the adhesive paint coating. Therefore, the most convenient and effective real-time solution for graffiti detection is to compare a pre-determined "clean" scene with the live feeds from an image or video device. A generic process pipeline for such a concept is illustrated in Figure 1. In this flowchart, the chosen algorithms will first recognize and separate the static background scene from any foreground moving objects. The process then moves on to track-and-filter any foreground human-shape areas based on any human detection techniques in order to avert noise caused by awaiting passengers and passing pedestrians.

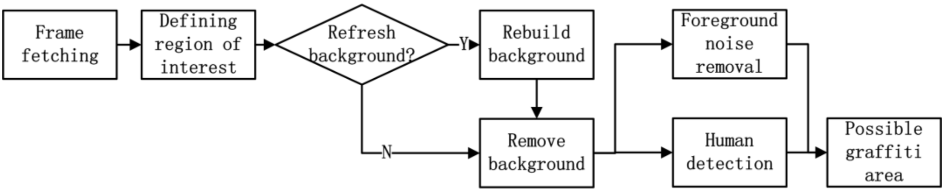

Figure 1: A generic graffiti detection process pipeline

As shown in Figure 2 (a snapshot of a typical UK bus stop), the low resolution CCTV surveillance system produces frames which are subject to serious occlusion problems. The classic solutions under strictly controlled conditions such as those reviewed by Moeslund and Granum [1, 2] cannot be readily deployed with satisfactory result. In this project, innovative algorithms have been developed to tackle the occlusion problems through detecting and removing human figures based on curve mapping. This new approach is rooted on the early work from Zhao and Nevatia [3, 4] which proposed generic human tracking and modelling solution under complex application situations. This project has further improved on a head tracking algorithm by adopting the Mean-Shift (MS) based edge detection technique to reconstruct curves for 2D open curve matching.

Figure 2: A sample frame from a bus stop video clip

To systematically elaborate the research and development works involved in the programme, this paper has been organized in the following order: Section 2 provides the design and implementation details of the graffiti detection system prototype Section 3 .reviews the theoretical foundation for human tracking and head curve matching. The test results are further analyzed with potential improvements discussed in Section 4, and Section 5 concludes the work.

2 System Prototype Design

For testing the devised algorithms mentioned in Figure 1, a graffiti detection system is prototyped in this research and is composed of following functional modules.

2.1 Region of interest (ROI) identifier

Graffiti are normally painted on flat surface such as walls for grabbing attention. For example, in a partially covered bus stop, the glass panels or walls can be primary target. Often, these areas can be defined by regular triangular or quad polygons in the image field, as illustrated in Figure 3.A. In this project, the ROI for a bus stop is defined manually by connecting the corners of the glass panels using straight line segments.

2.2 Background refreshment

This project introduces a statistical "median" value for background refreshment through removing static settings from the current frame. In a real world implementation, a single "standard" background will not be ideal for solving all segmentation problems due to the changing illumination conditions and shadows. This problem is resolved in the pilot by deploying a regular background refreshment mechanism (e.g. every other hour). This mechanism ensures a new background will be produced and used to comparing with any live feeds.

2.3 Foreground noise removal

As illustrated in Figure 3.B (Absolute difference frame by using "current" background) and 3.C (Binary image of Figure 3.B), background refreshment improves a newly fetched frame but will not be sufficient for differentiating newly painted graffiti area from moving objects in the foreground. However, a graffiti area will be much more stable than any dynamic foregrounds, which can be removed by measuring time endurance of any suspicious "targets" remaining at same position longer than a pre-determined number of frames. This mechanism is realized by implementing buffer-like structure and a stability threshold and for removing the "short-lived" foreground objects as shown in Figure 3.D.

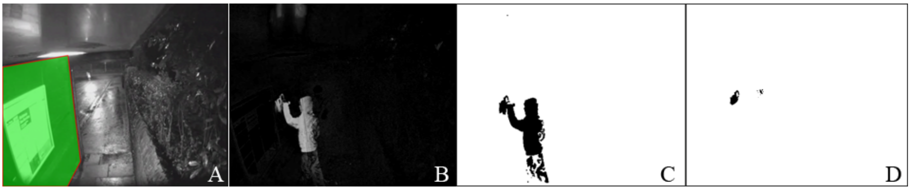

Figure 3: One example of graffiti detection process.

2.4 Human noise removal

The above foreground noise removal process can filter out most of the moving foreground objects such as passing pedestrians and vehicles. But in a bus stop, passengers might remain largely still for relative long time in the so-called "awaiting" mode, although small posture changes are often observed in an "oscillation" style. By increasing the stability threshold (more sensitive), it might (often not) resolve the problem but subjecting a deteriorating which seriously affecting the system's "real-time" performance. For example, in Figure 3.D, noise coming from human body, which is at the right side of real graffiti area, cannot be filtered clearly only by removing the "short-lived" foreground objects.

Due to these noise are usually coming from human body, a more advanced and appropriate approach is to distinguish human shape and filtering them out as noise directly. Details of this human detection technique will be presented in Section 3.

The gradual refinement by removing background/foreground noise and human shape subtraction has highlighted suspected graffiti area in each frame. This algorithm ensured the flexibility and robustness of such a task-specific vision system and its real-time application potentials.

3 The Implementation of Human Noise Removal Based-on Head Curve Matching

The non-rigid human body is difficult to track using a traditional vision system due to the inherent difficulties in defining a generic mathematical model and the settings of filtering thresholds to abstract and represent the entire human body. Different from the Haar-like feature-based method introduced by Viola and Jones [5] which carries out the classification operations by applying multiple thresholding and learning processes, a more direct method for tackling this challenge is to track only the head shape of a human body and then to combine the findings with other filtering results for other possible body-part shapes based on spatial distances. Compared with the complex task of modelling an entire human body, defining and tracking a human head is a much simpler job since it is a relatively rigid part of the human body. More importantly, all people seem to show an identical head and shoulder contour, an "Ω" shape, even when being oriented to different directions excepted a perfect side view as shown in Figure 4.

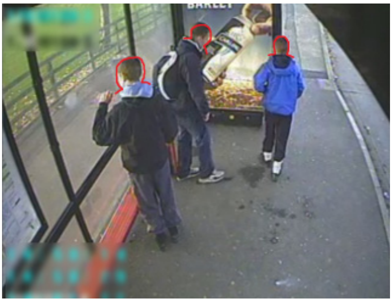

Figure 4: " Ω " contours of human heads

The technique of the "Ω" shape for human detection was first introduced by Zhao and Nevatia in 2002 [3]. In his paper, the locations of human heads were used for human ellipsoid model initialization, which required the pre-requisite knowledge on system calibration and body proportion assumptions. Although the method is fairly robust and effective in a largely static scene, the rotation and scale changes of a head in a more dynamic scene are hard to detect using only a single "Ω" shape template. In this project, experiment work has been carried out to improve this approach based on a 2D open curve matching algorithm. The renovated method can detect human head by matching an "Ω" shape without any pre-request knowledge on head proportions, or scaling registrations. The algorithm is consisted of 4 separate processes and is being illustrated as a flowchart in Figure 5.

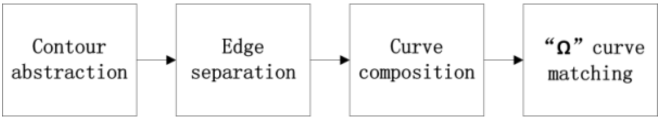

Figure 5: Head curve matching algorithm

3.1 Contour abstraction

Head contours are commonly recognized as a collection of edges extracted from intensity images. Various popular filters can be adopted for this purpose, for example, Tao has chosen the Canny algorithm [6] to highlight the foreground of the human silhouette. In this project, this operation has been improved in terms of flexibility by using edges generated from the Mean-shift (MS) segmentation algorithm [7] (Figure 6.B) and to deploy the Robert edge detection techniques (Figure 6.C). Different from Canny edges, the MS-based edges can separate colour intensities if the MS segmentation result is marked by different values. Edge pixels sharing the same feature value stand for a group and could be a geometric entity such as a curve. Multiple sets of those edge pixels can be defined as "edge groups" which can be readily used for further curve analysis as shown in figure 6.C

3.2 Edge separation

This process separates different edges into different groups by assessing their intensity values (set as labels), through manipulating the threshold values. As shown Figure 6.D, a threshold value has been set to highlight the edge groups possessing the same intensity value. Actually, different groups holding a same label might be identified by pixel connectivity. Figure 6.E shows all identical edge groups from Figure 6.D.

3.3 Curve composition

A labelled edge group is just a container for a set of pixels. The "Ω" contour matching operation requires a unified coordinates system to enable the linking of pixels into curves by exercising curve composition algorithms.

In this research, a composition algorithm is devised through evaluating "neighbourhood" edge pixels based on minimum cost paths method introduced by Hart et al. [9]. While in operation, the technique first applies the so-called "hit-miss transformation" [10] and Lookup table (LUT) to locate the end points in one edge group. Distinguishing the so-called "start" and "end" points is not necessary due to the symmetric "Ω" contour shape. Details of this operation can be found from [11]. The process then projects the curve from the start position towards the end by following the direction determined by the distributions of edge groups. The coordinate series of this curve can then be assembled by recording the traces.

3.4 " Ω " curve matching

For the final step of "Ω" curve matching, this research has extended the correlation curve matching algorithm introduced by Cui et al. in 2009 [12]. The basic theory of the method is to compare two curves by correlating and evaluating curvature similarities through employing a curvature integral to weight-labelling the correlation, which significantly reduces the problem caused by scaling and rotational transformations.

As shown in Figure 7.A and 7.B, the operation from calculating the curvature defined by points outlining the distribution of an "Ω" curve $c(s)$. After noise removal using the B-spline approximation [13], the absolute curvature $|\kappa|$ and integrating of curvature $K(s)$ (shown in the Figure 7.C) at any arc length s is given by

$$|\kappa(s)| = \|\ddot{c}(s)\| = \|[\ddot{x}(s), \ddot{y}(s)]\|, \tag{1}$$

$$K(s) = \int_0^s |\kappa(t)| dt, \tag{2}$$

where "$\cdot\cdot$" denotes the second derivative.

The final step of the process is to correlate the template and the pattern curves by using the integrals of the curvature mapping distributions: curvature w.r.t. integral of curvature, as shows in Figure 7.D. Based on the normalized cross correlation method by Lewis [14], the similar curvature (matching distance) between the template "Ω" curves and the extracted pattern can be calculated by

$$v(u) = \frac{\sum_{i\in\Omega}[p(i)-\bar{p}][t(i-u)-\bar{t}]}{\sqrt{\sum_{i\in\Omega}[p(i)-\bar{p}]^2 \sum_{i\in\Omega}[t(i-u)-\bar{t}]^2}}, \tag{3}$$

where p and t denote pattern and template curves, respectively. "$\bar{\ }$" denotes mean value and i denotes the point on the curve in Eq.3, the template curve slides along the pattern curves and u is the offset of one curve to the other. The range of v is [-1,1]. Larger value of $v(u)$ means more identical of the two curves.

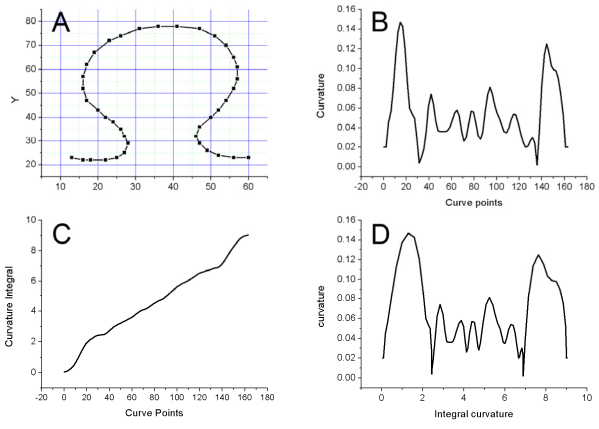

Figure 7: Curvature features of template "Ω" curve

Due to the over-segmentation problem often occurred in MS applications, a number of curves appeared as short and separated pieces in this project especially when the background is complex. For example, as shown in Figure 5.C, the curve outlining the head and shoulder is actually formed by two pieces. In this project, this problem has been resolved by applying the Cui's curvature feature-based "whole-to-part" matching method, which is valid even when one curve is partially occluded [12].

4 Test and Evaluations

The system prototype introduced in Section 2 and 3 have been developed in the LabVIEW program running on a PC with an AMD Athlon 2.62GHz CPU and 2G RAM. A series of experiments have been carried out to test the "Ω" shape recognition performances and MS-based curve reconstruction, which are the key functions of the system.

The first experiment is to assess the effectiveness and efficiency of the "Ω"-based head detection approach using the devised open curve matching techniques. As discussed in Section 3, the human head and shoulder curve shows a unique "Ω" shape after the MS-based curve reconstruction. To determining an appropriate generic head template can facilitate the rapid head detection from live video feeds, the "Ω" template used in this programme is based on the mean value calculated from 30 samples from different video clips.

For evaluating the open curve matching algorithm, a number of artificial curves (as shown in the Figure 8) have been tested. As shown in Table 1, the "Peak correlated index" highlights the locations of the matched curve segments. A close study of the result also shows that the "similarity" index calculated by the Eq.3 has accurately reflected the likeness between the two curves. In addition, using this technique, curves subject to similarity transformations such as rotation and scaling can also be recognized correctly based on the distributions of the curvature and the integral values that are invariant to similarity transformations.

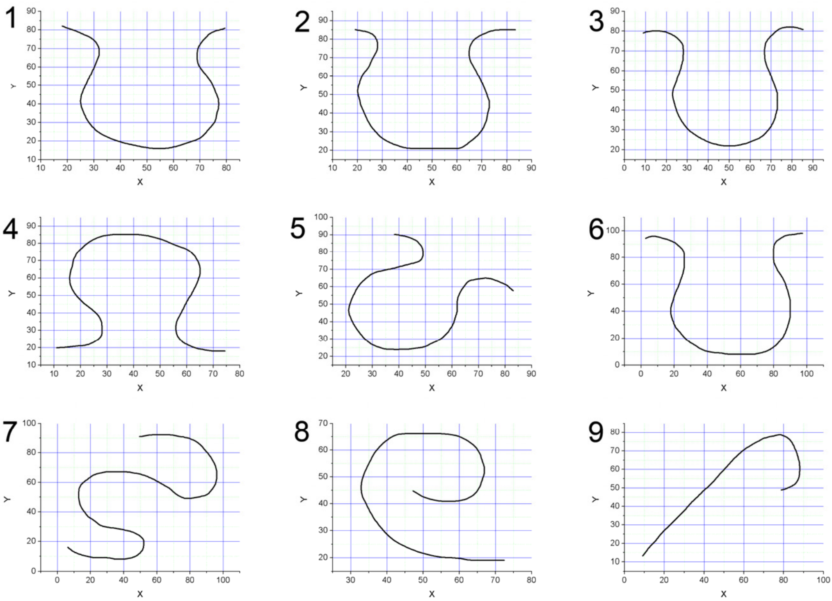

Figure 8: Artificial curve samples

After deploying this algorithm on real world surveillance video record, figure 9 illustrates the Recall-Precision-Curves of the newly developed "Ω" method and Zhao's approach. Compared with Zhao's head curve matching approach based on the Canny edge detection and the head proportion assumptions, this improved curve matching approach has presented better performance.

Curve index	Peak correlated index	similarity
1	13	0.79
2	24	0.82
3	7	0.91
4	9	0.81
5	12	0.82
6	19	0.87
7	65	0.62
8	104	0.43
9	84	0.14

Table 1: 2D artificial open curve matching using the pattern from Figure 6

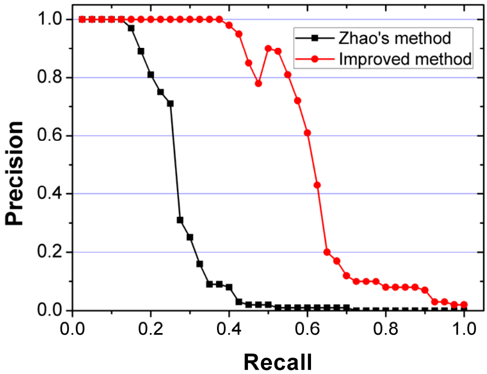

Figure 9: Detection performance by using various methods on 30 minutes CCTV record

The second experiment is evaluation of the entire graffiti detection system. After removing the frame background by applying the median background reconstruction, a number of real-world CCTV clips recorded from a bus stop have been tested in this experiment. Figure 10 shows the example snapshots and the detected graffiti area, while the Table 2 summarizes the detection results with a 72-hour stability test. Under an uncontrolled recording condition, many background noise sources could be introduced into the video such as waving leaves and branches, passing vehicles, and the changing lighting

Figure 10: Selected detection results

conditions.

Feature	Description
Detection time consumption	710 ms
Background Reconstruction time consumption (per frame)	3 ms
False positive rate	9.8%
False negative rate	0.17%
Human detection rate	91.1%
CPU consumption	43%
Video format	AVI
Video compressed filter name	Indeo (R) V4.5
Image format	U32 RGB
Image size	200 by 164
Image channel	Luminance
Main Interface RAM consumption	98Mb
Background Reconstruction RAM consumption (per frame)	69Kb
Configuration file format	DAT
ROI file format	DAT
Configuration file size	<200 bytes
ROI file size	<800 bytes
Program size	568Kb

Table 2: Stability test and system performance results

5 Conclusions

This paper has reported a pilot project on investigating and developing a live graffiti detection system in a real application scenario. During the research, an integration of existing and devised frame-based video/image processing techniques has been implemented, which has improved the stability and accuracy from the current systems. The system prototype can successfully and efficiently identify newly appeared graffiti areas and likely offenders within complex image scenes.

The main contribution of this project is the development of a rapid technique for removing foreground noise caused by passengers waiting at a bus stop through head curve matching. To tackle the head curve reconstruction problem in a complex image scene, a MS-base segmentation operation is developed and proven effective when applied on the test video feeds with variant background complexities. As detailed in Section 4, the human head contour represented as edge groups have been extracted by the system accurately. Through applying various morphological analysis methods, edge groups are successfully reconstructed into various forms of curves that have provided the basis for the "Ω" shape-based head detection. Test results show the algorithm and process developed in the system prototype is flexible and robust, which can be adopted for handling other human contour-based detection tasks. Compared with other existing efforts in this application domain, the head curve matching based human detection process has provided a reliable and rapid solution for real-time CCTV surveillance systems. This approach has also avoided the computational expensive and complex process of acquiring the proportion of human bodies and the corresponding similarity transformation often essential to other matching algorithms.

As part of the future work for this research, human gestures will be studied based on human stick-models for real-time gesture analysis to identify other forms of vandalism acts.

References

[1] T.B. Moeslund and E. Granum, A Comprehensive Survey of Computer Vision-based Human Motion Capture. *Computer Vision and Image Understanding*, 81(3):231--268,2001.

[2] T.B. Moeslund, A. Hilton and V. Kruger, A Survey of Advances in Vision-based Human Motion Capture and Analysis. *Computer Vision and Image Understanding*, 104(2-3):90--126,2006.

[3] T. Zhao and R. Nevatia, Stochastic Human Segmentation from a Static Camera. *Motion and Video Computing, IEEE Workshop on*, 43(3):9-14,2002.

[4] T. Zhao and R. Nevatia, Tracking Multiple Humans in Complex Situations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1208-1221,2004.

[5] P. Viola and M.J. Jones, Robust Real-time Object Detection, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 2001, IEEE Computer Society. p. 511-518.

[6] T. Lindeberg, Edge Detection and Ridge Detection with Automatic Scale Selection. *International Journal of Computer Vision*, 30(2):117-154,1998.

[7] D. Comaniciu and P. Meer, Mean Shift: A Robust Approach Toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 24(5):603-619,2002.

[8] L. Gorelick, M. Blank, E. Shechtman, M. Irani and R. Basri, Actions as Space-Time Shapes, in *Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2247-2253, 2007.

[9] P.E. Hart, N.J. Nilsson and B. Raphael, A Formal Basis for the Heuristic Determination of Minimum Cost Paths. *IEEE Transactions on Systems Science and Cybernetics*, 4(2):100-107,1968.

[10] E. Dougherty, An Introduction to Morphological Image Processing. 1992, *SPIE-International Society for Optical Engineering*. p. 122-123.

[11] R.C. Gonzales, R.E. Woods and S.L. Eddins, The Hit-or-Miss Transformation, in *Digtal Image Processing Using MATLAB (2rd Ed.)*. 2009, Gatesmark Publishing: USA. p. 500-503.

[12] M. Cui, J. Femiani, J. Hu, P. Wonka and A. Razdan, Curve matching for open 2D curves. *Pattern Recognition Letters*, 30(1):1-10,2009.

[13] R. Brinks, On the Convergence of Derivatives of B-splines to Derivatives of the Gaussian Function. *Journal of Computational and Applied Mathematics*, 27(1),2008.

[14] J.P. Lewis, Fast Normalized Cross-correlation. *Vision Interface. Canadian Image Processing and Pattern Recognition Society*:120-123,1995.