Computers in Industry 98 (2018) 153–164

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Vanishing point detection for visual surveillance systems in railway platform environments

Katy Tarrit[a,*], Julio Molleda[b], Gary A. Atkinson[c], Melvyn L. Smith[c], Glynn C. Wright[d], Peter Gaal[a]

[a] Neuroscience Department, University of Rochester, Rochester, New York 14642, USA
[b] Department of Computer Science and Engineering, University of Oviedo, Gijon 33204, Asturias, Spain
[c] Centre for Machine Vision, Bristol Robotics Laboratory, University of the West of England, Bristol BS16 1QY, UK
[d] Aralia Systems Milnwood 13 North Parade, Horsham, West Sussex RH12 2BT, UK

ARTICLE INFO

Article history:
Received 10 September 2017
Received in revised form 6 March 2018
Accepted 13 March 2018
Available online xxx

Keywords:
Visual surveillance
Image registration
Image reconstruction
Adaptive edge detection
Vanishing point detection

ABSTRACT

Visual surveillance is of paramount importance in public spaces and especially in train and metro platforms which are particularly susceptible to many types of crime from petty theft to terrorist activity. Image resolution of visual surveillance systems is limited by a trade-off between several requirements such as sensor and lens cost, transmission bandwidth and storage space. When image quality cannot be improved using high-resolution sensors, high-end lenses or IR illumination, the visual surveillance system may need to increase the resolving power of the images by software to provide accurate outputs such as, in our case, vanishing points (VPs). Despite having numerous applications in camera calibration, 3D reconstruction and threat detection, a general method for VP detection has remained elusive. Rather than attempting the infeasible task of VP detection in general scenes, this paper presents a novel method that is fine-tuned to work for railway station environments and is shown to outperform the state-of-the-art for that particular case. In this paper, we propose a three-stage approach to accurately detect the main lines and vanishing points in low-resolution images acquired by visual surveillance systems in indoor and outdoor railway platform environments. First, several frames are used to increase the resolving power through a multi-frame image enhancer. Second, an adaptive edge detection is performed and a novel line clustering algorithm is then applied to determine the parameters of the lines that converge at VPs; this is based on statistics of the detected lines and heuristics about the type of scene. Finally, vanishing points are computed via a voting system to optimize detection in an attempt to omit spurious lines. The proposed approach is very robust since it is not affected by ever-changing illumination and weather conditions of the scene, and it is immune to vibrations. Accurate and reliable vanishing point detection provides very valuable information, which can be used to aid camera calibration, automatic scene understanding, scene segmentation, semantic classification or augmented reality in platform environments.

1. Introduction

Security often requires reliable and robust video analytics software from visual surveillance systems to monitor unstructured indoor and outdoor environments. Visual surveillance is a broad research field in computer vision that has become very active in recent years [1–3] for security and many other applications. Apart from the necessity of ensuring high levels of security in public areas and facilities, the development of computer vision devices at decreased costs as well as their miniaturization and integration in lots of environments have accelerated the use of visual surveillance systems. The increasing power of standard computing platforms, with general purpose GPU capabilities, allows computer vision tasks—implemented as software layers or modules—to be executed in real-time over the scene action provided by surveillance cameras. Video surveillance systems can be used indoors or outdoors. Applications of these systems range from security concerns, such as crime protection, prevention and forensics, to management, such as traffic and infrastructures.

Image resolution of video surveillance systems is limited by a trade-off between several system requirements, such as sensor and lens cost, transmission bandwidth and storage space, among

* Corresponding author.
E-mail address: Katy_Tarrit@urmc.rochester.edu (K. Tarrit).

https://doi.org/10.1016/j.compind.2018.03.005

others. However, the acquisition rate of modern video surveillance cameras makes it possible to reconstruct an enhanced image from a set of low-resolution images when the computer vision modules of the system are expected to provide accurate outputs. Multi-frame-based reconstruction techniques require a precise alignment of the set of original images to provide an enhanced image, that is, precise subpixel image registration is required.

Most man-made environments are composed of numerous buildings, roads, streets and objects that can be represented by simple volumes such as cubes or basic surfaces such as planes. These volumes and surfaces are themselves formed from elementary geometrical elements such as straight lines. These lines, when projected onto an image, intersect at vanishing points (VPs) and define the perspective of the scene. Moreover, lines are common in the type of environment being considered here (i.e. railway and underground station environments). Such environments make it easy to extract lines and VPs as they appear many times in the captured images. However, due to the large number of straight lines in scenes such as railway or underground stations, VP detection can be problematic. This is partly because of external elements such as variation in lighting, noise, distortion from the camera and occlusions. In addition, these scenes also contain a large number of people, luggage, etc. which make the analysis task even more complex. This implies that any VP detection method needs to be specifically adapted to the type of scene geometry in order to offer a robust and effective system. In this paper, we aim to adapt the detection method to railway scene surveillance, an especially common security scenario, but the principles can be tuned for other environments also.

The large majority of past research into VP detection is only applied to relatively simple volumes such as cubes, parallelepipeds, or other environments such as roads or architectural images. These such environments are often termed "Manhattan" as the lines tend to be mostly parallel or perpendicular to each other. In addition, the images are generally captured with good illumination conditions and where the perspective is sufficiently extreme that the extraction of lines and VPs is relatively straightforward. The main difference between previous work and this contribution is the adaptation of well-established techniques to a different and more complex environment, where geometrical elements such as lines and VPs are so numerous that they create noise and occlusions. Furthermore, complications arise due to non-Manhattan lines such as escalator rails. For the majority of this paper, we assume Manhattan geometry, i.e. the imaged scene contains three mutually orthogonal directions. However, we also conduct preliminary experiments to show that the method has scope to locate VPs considering non-Manhattan lines near the end of the paper.

This paper presents a novel method for VP detection, improving upon our earlier works described in [4,5], that is fine-tuned for railway and underground station environments, a common application of CCTV. Unlike much other work from the computer vision community, the proposed approach takes advantage of using real CCTV data acquired by the video surveillance system in a multi-frame image processing method, as well as the geometry of the environment in train and metro platforms. Fig. 1 shows four different examples of images acquired by video surveillance systems in rail platform environments. As can be seen, the main planes and objects are delimited by straight lines denoting strong perspective effects, which is a common feature in these platforms.

The contribution of this paper is two-fold, first, we present an optimized super-resolution technique to enhance image quality in a railway platform setting. Second, we propose a novel VP detection method that is fine-tuned to railway station environments and uses the enhanced image as input. Moreover, the novelty of this work lies in the use of specific a priori knowledge

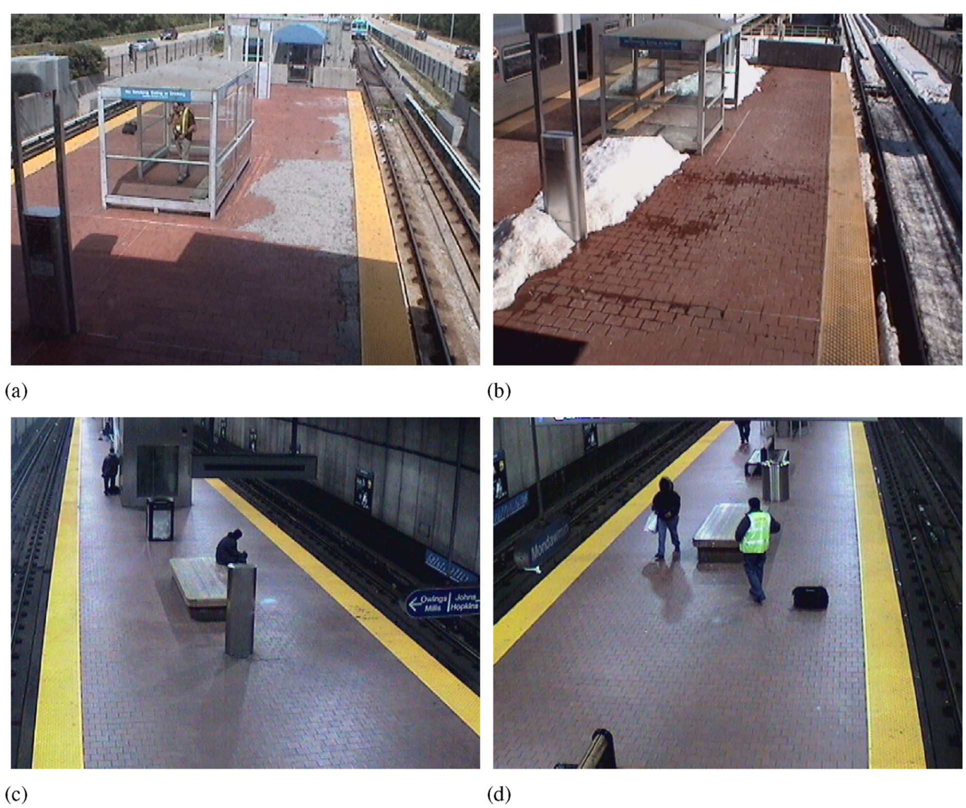

Fig. 1. Examples of platform environments denoting strong perspective effects: (a and b) outdoor rail platforms in different weather conditions; (c and d) indoor rail platforms.