Automation in Construction 38 (2014) 39–45

Contents lists available at ScienceDirect

Automation in Construction

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

Optical flow techniques for estimation of camera motion parameters in sewer closed circuit television inspection videos

Mahmoud R. Halfawy *, Jantira Hengmeechai

Infrastructure Data Solutions, Inc. (IDS), 2 Research Drive, Suite 150-E, Regina, SK S4S 7H9, Canada

ARTICLE INFO

Article history:
Accepted 18 October 2013
Available online 30 November 2013

Keywords:
Sewer inspection
Closed circuit television inspection
Computer vision
Optical flow
Motion tracking

ABSTRACT

This paper discusses a novel approach for automated analysis and tracking of camera motion in sewer inspection closed circuit television (CCTV) videos. This approach represents an important building block for any system that supports automated analysis and defect detection of CCTV videos. The proposed approach employs optical flow techniques to automatically identify, locate, and extract a limited set of video segments, called regions of interest (ROI), which likely include defects, thus reducing the time and computational requirements needed for video processing. Tracking the camera motion parameters is used to recover the operator actions during the inspection session, which would provide important clues about the location and severity of the ROI. Techniques for estimating the camera travelling distance, position inside the sewer, and direction of motion from optical flow vectors are discussed. The proposed techniques were validated using a representative set of sewer CCTV videos obtained from the cities of Regina and Calgary, Canada.

1. Introduction

During the past four decades, closed circuit television (CCTV) sewer inspection systems have been used as the primary tool for inspecting the internal surface of non-man-entry sewers. Despite the emergence of several new sewer inspection technologies (e.g., digital side scanners), CCTV technology is expected to remain an important inspection tool in any sewer condition assessment program [18]. However, CCTV technology exhibits a number of limitations. Some of these limitations are due to the high level of subjectivity inherent in the assessment process, which depends largely on the operator's interpretation, experience, and skill level. Also, long inspection sessions may lead to operator fatigue, and potentially erroneous interpretation and assessment of the sewer condition. Although these limitations can be partially mitigated through adopting best practices in operator training and certification, they cannot be entirely eliminated [13]. Subsequently, CCTV inspections are often required to be reviewed and audited off-site to ensure the quality, accuracy, and consistency of the inspection process [5]. To overcome these limitations of CCTV inspections and to enable more efficient off-site reviews, there is a wide recognition of the need to develop systems that can support automated analysis of CCTV videos.

Automated CCTV video analysis systems would also provide a valuable tool for re-evaluating archived videos and extracting valuable historical sewer condition information. Ironically, while municipalities possess thousands of hours of archived CCTV videos, they often lack adequate information required for developing reliable deterioration models to predict sewers' future condition. An automated CCTV video analysis system would support efficient and cost-effective analysis of archived videos to provide valuable historical condition information needed for developing predictive models and renewal plans.

There have been several research projects dealing with automated analysis and detection of defects in sewers. Most of these projects have proposed image processing and pattern recognition techniques to locate defects in still images, mostly collected using digital side scanning systems (e.g., the sewer scanning evaluation technology, SSET). However, very few research projects have addressed the analysis of videos produced from conventional CCTV systems. For example, Kolesnik and Baratoff [11] proposed an algorithm for circle extraction and conic fitting to support 3D interpretation of sewer circular structures from a sewer video record. Their algorithm calculates the distance of joints and pipe ends to the robot, given prior knowledge of the sewer diameter. Pridmore et al. [15] and Cooper et al. [4] proposed an algorithm to recover the camera position and orientation from CCTV images by finding the vanishing point from mortar lines of brick sewer images. Kannala et al. [10] proposed an approach to automatically reconstruct the 3D shape of a surveyed pipe from video captured using a calibrated fish-eye camera. Sarshar et al. [16] proposed service-oriented architecture and image processing algorithms to support automated evaluation of archived sewer CCTV videos. Several researchers, such as Moselhi and Shehab [14] and Guo et at. [6], have also proposed automated defect detection algorithms. However, most of the proposed algorithms were developed

* Corresponding author. Tel.: +1 306 790 1415.
E-mail addresses: Mahmoud.Halfawy@infrastructure-data-solutions.com (M.R. Halfawy), Jantira.Hengmeechai@infrastructure-data-solutions.com (J. Hengmeechai).

0926-5805/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.autcon.2013.10.016

to process still images collected manually from CCTV inspections, and have not addressed issues of camera motion and video analysis to automatically identify and extract still images used for defect detection.

Unlike side scanning systems, where a 2D image representing an unfolded view of the sewer internal surface is provided, automated analysis of conventional CCTV videos poses several unique challenges. Some of these challenges are due to the high computational requirements to process CCTV videos, complex motion sequences of the camera, different structure and constituent segments of images depending on the camera orientation, and non-uniformity of illumination conditions inside the sewer. This paper proposes a novel approach to address one of these challenges, namely the analysis and tracking of camera motion in CCTV videos. The proposed approach represents an important building block for any system that supports automated analysis and defect detection in conventional CCTV videos.

2. Camera motion in CCTV videos

CCTV sewer inspection systems involve the use of a remote-controlled forward-looking pan, tilt, and zoom camera mounted on a robot crawler that travels between two manholes. The CCTV equipment is operated by certified operators who are trained to effectively control the camera and interpret videos in real-time to deduce an accurate description of the sewer features and defects. In a typical CCTV inspection, the operator moves the camera forward down the sewer axis at a relatively consistent speed, (often in the range of 2.0–5.0 m per minute) depending on the sewer characteristics (e.g., diameter, cleanliness), the number and extent of defects, and the operator skill level. When a defect or feature of interest is observed, the operator may: (i) slow the forward motion of the camera; (ii) pause the camera and slowly move the camera back and forth; (iii) pan, tilt, or zoom the camera towards the sewer wall or bottom of the sewer to examine or capture the features or defects more closely; or (iv) fully pause the camera to review the images and record observations. Afterwards, the operator resumes the camera's forward motion towards the end-of-sewer.

Identifying different modes of camera motion during CCTV inspection sessions can be used to recover the operator behavior in response to observed defects during the inspection session, which in turn indirectly indicate the presence and severity of defects in the sewer. This hypothesis was observed and validated by examining a large data set of CCTV videos, where a specific pattern of camera motion was found to almost always coincide with the presence of regions of interest (ROI) in the sewer. These ROI would represent the video segments where the sewer looked more suspicious and likely to contain defects or interesting features that attracted the operator's attention. The presence of ROI was found to always correspond to camera pause or slow down events. Also, the duration and the amount of backward or pan/tilt motion of the camera in each ROI was found to correspond to the extent and severity of the observed defects. Therefore, tracking and analyzing the camera motion in the video can be used to recover the operator behavior, and hence, provide important clues about the location and relative severity of the ROI.

The approach proposed in this paper employs optical flow-based techniques [7,12] to analyze the camera motion parameters, and estimate its travelling distance, position inside the sewer, and direction of motion. This analysis is used to automatically identify, locate, and extract a limited set of video segments, called regions of interest (ROI). Further processing and automated defect detection operations can be limited to these identified ROI, leading to significant saving in the time and computational requirements needed for video processing.

3. Analysis of camera motion in CCTV videos using optical flow techniques

Motion analysis and tracking techniques are applied to reveal the dynamics of the camera motion in the image sequence of the CCTV video to identify the ROI. Motion tracking and analysis have been an active area of research in computer vision since early 1980s. Algorithms for detecting motion in a video sequence, commonly known as optical flow, primarily attempt to find matching objects in different frames, and hence enable tracking the motion of these objects or the camera in the video sequence. Studies suggest that the best performance with respect to accuracy can be obtained using differential motion tracking techniques [9]. This study employs a commonly used differential motion tracking technique that was proposed by Lucas and Kanade [12]. Other available techniques include the Horn–Schunck method [7] and the block matching method [1,8].

Optical flow vectors, or motion fields, refer to the motion of corresponding pixels in images sequences. The optical flow algorithm works by first identifying a set of visually significant features that can be tracked from one frame to the next in the video image sequence. The optical flow algorithm then calculates the displacement vectors (or velocities) at the pixels that correspond to these "interesting features" between two consecutive frames. These optical flow vectors indicate both the speed and direction of motion. Since there are no moving objects in CCTV videos, the optical flow vectors would correspond to the camera motion, and therefore, can be used to approximate the camera speed and direction. In that sense, the optical flow algorithm would enable the reconstruction of the trajectory of the camera as it moves inside the sewer during the inspection session.

A typical sewer CCTV video sequence contains tens of thousands of frames that need to be processed and analyzed to detect and locate possible defects. Performing such analysis on each single frame in the video sequence could be very computationally expensive, which renders the process infeasible. However, the vast majority of frames in typical CCTV videos often includes no defects, and therefore, may be eliminated from the analysis process without impacting the accuracy of the detection process. Therefore, the ability to identify and extract a limited set of video segments, or ROI, which would likely contain defects, would significantly reduce the required time and computational resources.

The proposed optical flow techniques were validated using a representative set of CCTV videos obtained from the cities of Regina and Calgary, Canada. Sewers inspected in these videos were of concrete and vitrified clay material, with circular cross-sections and diameters ranging from 250 mm to 900 mm. The CCTV video format was MPEG-2, at a frame rate of approximately 30 frames per second. The resolution of the CCTV cameras was 320×240 or 352×240 pixels. The software was implemented in C++ using the Intel's Open Computer Vision library (OpenCV) [3].

The implementation of the optical flow algorithm employed functions in the OpenCV library for identifying the best features to track and for calculating optical flow vectors for these features. Frames are first captured from the CCTV video in sequence, and the optical flow algorithm is applied on each two consecutive frames. The algorithm used for selecting trackable features was proposed by Shi and Tomasi [17] and implemented in the library using the *cvGoodFeaturesToTrack* () function. The returned set of features (or pixel locations) is then tracked using an efficient pyramidal implementation of the Lucas–Kanade (LK) sparse optical flow algorithm [2], using the library function *cvCalcOpticalFlowPyrLK* ().

Sewer CCTV videos, in general, satisfy the basic assumptions of the LK algorithm, which can be summarized as follows: (1) Brightness of pixels of the same object (e.g., sewer wall) does not change significantly between consecutive frames; (2) Camera motion is relatively slow (i.e., objects or pixels do not move much between frames); and (3) spatial coherence of neighboring points where they generally belong to the same object. Using information of the optical flow vectors, the algorithm automatically detects the video segments where the camera is paused, slowed down, or changed direction (e.g., moved backward or sideways). These segments constitute the ROI that will be considered for further processing and analysis. The following subsections describe the application of the optical flow algorithm to identify and locate the ROI.