BIOSYSTEMS ENGINEERING 169 (2018) 149–164

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/issn/15375110

Research Paper

Development of a rescue system for agricultural machinery operators using machine vision

Yan Zhang [a], Pengbo Gao [a], Tofael Ahamed [b,*]

[a] Graduate School of Life and Environmental Sciences, University of Tsukuba, Japan
[b] Faculty of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, Japan

ARTICLE INFO

Article history:
Received 26 May 2017
Received in revised form
28 January 2018
Accepted 14 February 2018

Keywords:
Agricultural machinery
Rescue system
Driver inattention
Action recognition
Block Hankel-matrix

In this study, an automatic rescue system was proposed to monitor agricultural machinery operators using machine vision. The rescue system was developed to recognise the driver inattention status, that is, the distraction and fatigue by recognising the driver's actions. A Kinect sensor was used to collect image sequences of the operators, and the recognition system relied on the “player extraction” function of the Kinect sensor. A Hankel-based Kernel Mutual Subspace Method (KMSM) was developed to monitor tractor drivers and recognise driver inattention behaviours. To reduce the computational complexity for fulfilling the requirements of recognition, low-dimensional image vectors were used to generate low-dimensional block Hankel matrixes as representations for input action sequences. To evaluate the performance of the proposed KMSM, a driver action dataset was established that included 10 tractor drivers and 5 types of action that denote inattention. The drivers' inattention actions were classified into three danger levels, and the corresponding countermeasures for the actions at each danger level were similarly classified. Both offline and online experiments using similar subjects and different subjects were conducted to evaluate the designed inattention action recognition algorithm. In the offline experiment, the proposed Hankel-based KMSM achieved recognition rates of 91.18% and 86.18% when using similar and different subjects, respectively; and in the online experiment, the proposed method achieved 87.02 and 79.97% when using similar and different subjects, respectively. The average computation time of the Hankel-based KMSM was 0.07 s in the online experiment. Thus, the proposed Hankel-based KMSM method satisfies both the accuracy and the real-time requirements for a driver rescue system.

1. Introduction

With the development of autonomous driving technology, agricultural vehicles for bio-production systems have obtained significant intelligence improvements. In these systems, autonomous applications have recently gained high adoption potential through the improvements to sensors, positioning and navigation performance. Many types of machine vision and positioning sensors have been applied to agricultural navigation, and these may possibly lead to the increased safety of agricultural vehicles for autonomous

* Corresponding author.
E-mail address: tofael.ahamed.gp@u.tsukuba.ac.jp (T. Ahamed).
https://doi.org/10.1016/j.biosystemseng.2018.02.009
1537-5110/

Nomenclature

KMSM	kernel mutual subspace method
MSM	mutual subspace method
PCA	principal component analysis
KPCA	kernel principal component analysis
d	dimensionality of the subspace used for recognition
P	reference subspace registered as the dictionary
Q	input subspace obtained from the input data
u_i	eigenvectors in the reference subspace P
v_i	eigenvectors in the input subspace Q
R^n	n-dimensional feature space
$\overrightarrow{x}_a$	feature vector
C	covariance matrix from k feature vectors $\overrightarrow{x}_a$
λ	eigenvalues of the eigenvectors
X	n-dimensional input patterns
F	high- or infinite-dimensional feature space
$\emptyset$	nonlinear map
α_i	enumeration of coupling coefficients
θ	similarities between the reference subspace P and input subspace Q
K_{ij}	Gaussian Kernel function
σ	A free parameter of Gaussian Kernel
I	sequential images of an action
f_i	image vector
H_I	block Hankel Matrix for sequential images I
h_j	column of the block Hankel Matrix H_I
r	Hankel block parameter for Hankel matrix
S	set of sub-actions generated with sliding windows
s_i	sub-action that includes r action sequence images in a sliding window
G	dimension of vector, where $G = d \times r$
M	number of action classes
N	number of action sequences
J	number of testing sequences
N_{SW}	width of sliding window
N_{AR}	recognised number of inattention actions
TRB	training datasets
TEB	testing datasets

applications (Ahamed, Noguchi, Tomohiro, and Tain, 2016). However, under current conditions, from a safety standpoint, many barriers to the implementation of autonomous driving systems and agricultural vehicle operation exist because—in contrast to passenger car control—in agriculture, autonomous driving must not only adapt to uneven farmland roads and terrain but also conform to the requirements of a variety of tasks.

To ensure safety during driving, much research has been performed, which can be divided into three categories: first, driving strategy making based on external environment sensing [Matthew, Joseph, Stephen, and Christian, 2016; Zhang, Ahamed, Zhang, Gao, & Takigawa, 2016], second, vehicle control based on its posture detection [Lundahl, Olofsson, Berntorp, Åslund, & Nielsen, 2014; Sun, Cai, Chen, Liu, & Wang, 2016] and third, monitoring a driver's condition [Jung, Shin, & Chung, 2014; Craye & Karray, 2015]. Systems that sense the external environment and detect driver posture inarguably improve the safety of autonomous driving and advance driver assistance; monitoring a driver's condition is both effective and important in ensuring driving safety. Statistics of accidents caused by driver fatigue and distraction have been reported by many countries. One important factor in these accidents, especially on rural roads, is inattention caused by driver fatigue or monotony [Sigari, Pourshahabi, Soryani, and Fathy, 2014]. Moreover, with the aging population and the rapid reduction in the agricultural labour force, the average age of farmers worldwide has risen to approximately 60. One study reported that by 2020, more than 60% of the people engaged in farming in Japan will be older than 60, and 45% will be older than 70 [Ahamed, Tian, Takigawa, & Zhang, 2009]. The poorer physical condition of older farmers causes them to be easily distracted and fatigued, which poses higher health risks while driving [Nilsson, Pinzke, & Lundqvist, 2010]. Thus, a rescue system that can monitor a driver's condition while driving agricultural vehicles in farmland is essential to ensure driving safety.

Monitoring drivers' conditions by measuring physiological characteristics such as their brain waves, heart rate, skin conductivity and pulse rate yields the maximum detection accuracy. However, there are disadvantages when using physiological sensors that require physical contact with drivers that include causing annoyance, noise from the electrical signals obtained from such sensors, and lack of a realistic environmental perception. Less intrusive techniques, including computer vision techniques such as eye- or gaze-tracking have provided good results in judging whether the driver's view deviated from the driving direction [Cyganek & Gruszczyński, 2014; Masala & Grosso, 2014]. At the same time, such techniques provide more information concerning the driver's face, head and body movements that can be even more meaningful than eye motion. Driver video images were used to monitor driver behaviours such as eyelid movements, eye closure percentage, nodding, yawning, gazing, sluggish facial expressions, and sagging posture [Craye & Karray, 2015; Sigari, Pourshahabi, Soryani, and Fathy, 2014; Tawari & Trivedi, 2014]. However, the existing solutions lack sufficient accuracy and speed to function reliably under real vehicle conditions.

Compared with a general RGB camera, the Kinect is a low-cost RGB-D sensor that provides a greater wealth of information. A Kinect includes a colour camera, an infrared (IR) emitter and an IR depth sensor; therefore, it can capture an RGB image and a depth image from which a person's skeletal joints, body (player image), and the area around the driver can be automatically separated from the background of the captured image and cropped. For these reasons, the Kinect has been widely applied to human detection and action recognition [Li, Zhang, & Liu, 2010; Oreifej & Liu, 2013; Zhang, Suryanto, & Fukui, 2014]. Sequential images of human actions contain rich information on body postures and motion. However, the images are easily affected by different viewpoints, illumination conditions and individual characteristics. Although many approaches have been proposed to recognise