Infrared Physics & Technology 53 (2010) 439–449

Contents lists available at ScienceDirect

Infrared Physics & Technology

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

Detection of pedestrians in far-infrared automotive night vision using region-growing and clothing distortion compensation

Ronan O'Malley *, Edward Jones, Martin Glavin

Connaught Automotive Research Group, Electrical and Electronic Engineering, National University of Ireland, Galway, University Road, Galway, Ireland

ARTICLE INFO

Article history:
Received 21 May 2010
Available online 24 September 2010

Keywords:
Pedestrian detection
Automotive
Infrared
HOG
Region-growing
Driver assistance

ABSTRACT

We present a night-time pedestrian detection system based on automotive infrared video processing. Far-infrared or thermal night vision is a technology well suited for automatic detection of pedestrians at night as they generally appear warmer than the background. However, the appearance of a pedestrian in IR video can vary dramatically depending on the physical properties of the clothing they wear, the time spent adjusting to the outside environment, and the ambient temperature. We highlight the difficulties of detection in low temperatures (below 8 °C) when pedestrians typically wear highly insulating clothing, which can lead to distortion of the IR signature of the pedestrian. A pre-processing step is presented, which compensates for this clothing-based distortion using vertically-biased morphological closing. Potential pedestrians (Regions of Interest) are then segmented using feature-based region-growing with high intensity seeds. Histogram of Oriented Gradients (HOG) features are extracted from candidates and utilised for Support Vector Machine classification. Positively classified targets are tracked between frames using a Kalman filter, adding robustness and increasing performance. The proposed system adapts not just to variations between images or video frames, but to variations in appearance between different pedestrians in the same image or frame. Results indicate improved performance compared to previous HOG–SVM automotive IR pedestrian detection systems, which utilised stereo IR cameras.

1. Introduction

Automotive safety continues to be an important factor in vehicle design and technology. Recently the focus has extended beyond the occupants of the vehicle and has turned towards other road users, particularly so-called Vulnerable Road Users (VRU). In the EU, new regulations are placing the onus on automotive manufacturers to protect pedestrians by introducing stringent new passive protection standards [1] on a staggered basis from 2011 to 2019. These include provisions such as specifications for collision tests and geometric constraints for vehicle bonnets and windscreens. However, manufacturers may avoid a large section of the requirements if they equip the vehicle with a collision avoidance system. This is based on a European Commission study that found that pedestrian protection can be significantly improved by a combination of passive and active measures, especially adoption of active brake-assist systems.

Almost one-third of road fatalities occur during the hours of darkness [2], despite the greatly reduced traffic volume. It has been estimated that per vehicle mile, the road fatality rate is 3–4 times higher in darkness than in daylight [3]. Hence, night-time is of particular importance to automotive safety systems. Far-infrared technology is greatly suited to the task of night-time automotive pedestrian detection, as thermal radiation from humans peaks in the 8–14 μm far-IR spectral band, and no illumination is required. It has also been found that far-infrared yields better response times for drivers detecting pedestrians than near infrared [4]. A comparison of images of a pedestrian at 35 m in the visible and far-infrared spectrum is shown in Fig. 1.

Some techniques widely used for automotive pedestrian detection with visible cameras can also be applied in the case of IR systems. However, due to fundamental differences between visual and thermal imagery, many of the conventional techniques are not directly applicable [5]. The first step in most systems is to determine Regions of Interest (ROIs) in the image, which are examined more closely and classified as "pedestrian" or "non-pedestrian". It has been established that pedestrians are usually warmer and hence appear brighter than the background [6]. Therefore, image thresholding is a common starting point for extracting pedestrian candidates.

The technique outlined in [7] defines a bright pixel threshold as the difference between maximum image intensity and a set constant. In [8] a static threshold is derived by performing Bayes Classification on a set of templates known to contain pedestrians.

* Corresponding author.
E-mail addresses: ronan.omalley@nuigalway.ie (R. O'Malley), edward.jones@nuigalway.ie (E. Jones), martin.glavin@nuigalway.ie (M. Glavin).

1350-4495/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.infrared.2010.09.006

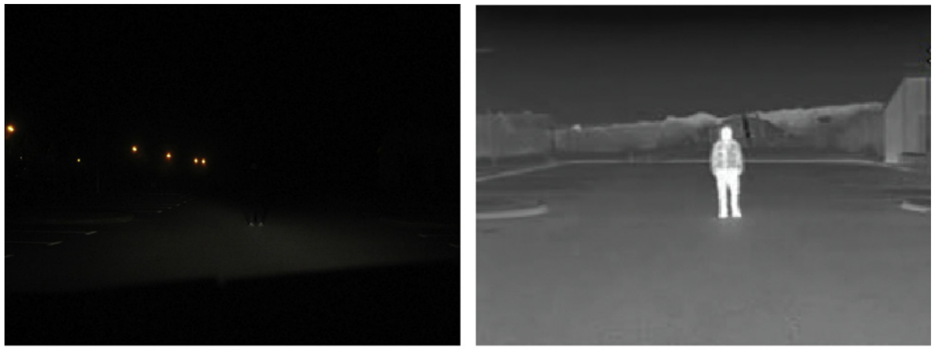

Fig. 1. Examples images of a pedestrian at 35 m with dipped headlights with minimal ambient lighting in (a) the visible spectrum and (b) far-infrared.

A region-growing style threshold is implemented in [9] using two static threshold values. The lower of the thresholds is restricted to areas spatially connected to seed regions resulting from the higher threshold. In [10], region-growing is conducted until edge pixels are reached to extract vehicular targets from thermal night vision video. A threshold value is defined from the mean and maximum image intensity values in [11], while in [12] a threshold value is chosen as the last local minimum of the image histogram before the saturation point. However, these types of global image threshold, that adapt to the properties of the current frame, fail to address the potential differences in appearance between multiple pedestrians in the same frame. A thresholding technique in this area should extract each pedestrian candidate individually from the background.

While it has been observed that pedestrians appear brighter than the background, segmentation based on thresholding is not a trivial task. A threshold value that is too low may extract background regions as well as the candidate pedestrian, distorting the candidates shape and intensity characteristics. On the other hand, a threshold that is too high may fragment the candidate into numerous sub-regions, creating a new task of attempting to group these regions to form potential candidates, using such techniques as active contours [13,14]. Xu et al. label such fragmented candidates "body-ground candidates". In fact, this phenomenon is particularly common in low temperatures, when heavy clothing tends to insulate body-heat, particularly the torso. The result of this is that it can be difficult to determine an optimum threshold to extract the pedestrian, and in some cases it may be impossible to extract a pedestrian by thresholding as the surface temperature of well-insulated clothing nears the temperature of background objects.

The number of non-pedestrian ROIs generated from the thresholding step can be reduced by filtering the regions identified, based on the expected physical features of a pedestrian. Very small ROIs may indeed be pedestrians, but it classifying them can be challenging, especially given the relatively low-resolution of automotive IR sensors. Objects can be filtered according to aspect-ratio [11] since pedestrians are generally expected to be taller than they are wide. This technique is used widely but with different bounding parameters, some more stringent than others. In [7], the concept of "inertial" is introduced to identify pedestrians.

Horizontal and vertical grey-level intensity profiles have been analysed [7] to form a shape-independent approach composed of two steps . Their method first segments the image into vertical strips by finding the local minima of a thresholded horizontal image intensity profile. These strips are then further segmented in the vertical. As humans display bilateral symmetry, grey-level symmetry and edge symmetry of pedestrians can also be used to segment Regions of Interest as demonstrated in [15]. Edge density can also be analysed as pedestrians in FIR images are usually much brighter than the background, and there can be a sharp change in image intensity at their edges [15]. The gradient operator can also be used to aid detection [12,16].

Stereo has also been proposed for infrared pedestrian detection, and there are many advantages to utilising stereo sensors for this task [9,16–18]. Stereo systems have demonstrated robust detection with simple techniques, as disparity can be used to effectively find ROIs. However, multiple thermal imaging cameras are not a viable option for many automotive designers as cost, power consumption and physical space are significant factors.

Models and templates of pedestrians are frequently matched against images [19,20] and a probabilistic template is created in [8]. This template matching is commonly performed using cross-correlation.

Night vision systems have become an important safety feature in some road vehicles. Commercial automotive IR pedestrian detection systems are starting to appear on the market, e.g. the system described in [6]. This system uses several features to distinguish relevant (pedestrian) objects from non-relevant hot or warm objects. A motion based cue is primarily used. A hot or warm object that changes shape at a rate above a certain threshold is considered relevant. This is combined with size and shape assumptions, template matching and image profiles. However, this approach relies on the assumption that pedestrians appear as entirely warmer than the background. While it is acknowledged that clothing can distort the thermal signature of a pedestrian, no special consideration is given to dark body segments caused by low temperatures and heavily insulating clothing.

Statistical classification is the process by which ROIs are placed into groups, in this case pedestrian or non-pedestrian. This decision is based on characteristics (features) determined from a training set of manually labelled images. This is conventionally achieved by means of a statistical classification method such as Support Vector Machines (SVM) [11,16], Artificial Neural Networks (ANN) [21] or Boosting [22]. It was found in [11] that grey scale classification was more successful than binary classification, as the binary candidates were too sensitive to shape. Experimental results showed that a single classifier for all types of pedestrians (along road, across road) performed better than the application of multiple classifiers. Other types of specialised statistical classification are explored in [23].

The choice of feature for classification of pedestrians is crucial to system performance. Suard et al. have successfully used Histogram of Oriented Gradients (HOG) features for pedestrian classification in automotive stereo thermal imagery [16,18]. HOG are used most effectively with an SVM classifier. Alternatively, the grey scale and binary images themselves are used as feature