Journal of AI and Data Mining
Vol 7, No 1, 2019, 1-16
DOI: 10.22044/JADM.2018.5742.1696

Pedestrian Detection in Infrared Outdoor Images Based on Atmospheric Situation Estimation

S. M. Ghazali and Y. Baleghi*

Electrical and Computer Engineering Department, Babol Noshirvani University of Technology, Babol, Iran.

Received 13 June 2017; Revised 04 November 2017; Accepted 06 January 2018
Corresponding author: y.baleghi@nit.ac.ir (Yasser Baleghi)

Abstract
Observation in absolute darkness and daytime under any atmospheric situation is one of the advantages of thermal imaging systems. In spite of increasing trend of using these systems, there are still lots of difficulties in analyzing thermal images due to the variable features of pedestrians and atmospheric situations. In this paper, an efficient method is proposed for detecting pedestrians in outdoor thermal images that adapts to variable atmospheric situations. In the first step, the type of atmospheric situation is estimated based on the global features of the thermal image. Then, for each situation, a relevant algorithm is performed for pedestrian detection. To do so, thermal images are divided into three classes of atmospheric situations: a) fine such as sunny weather, b) bad such as rainy and hazy weather, and c) hot such as hot summer days, where pedestrians are darker than background. Then 2-dimensional double density dual tree discrete wavelet transform in three levels is acquired from input images, and the energy of low frequency coefficients in third level is calculated as the discriminating feature for atmospheric situation identification. Feed-forward neural network classifier is trained by this feature vector to determine the category of atmospheric situation. Finally, a pre-determined algorithm that is relevant to the category of atmospheric situation is applied for pedestrian detection. The proposed method in pedestrian detection has a high performance so the accuracy of pedestrian detection in two popular databases is more than 99%.

Keywords: *Outdoor thermal images, Atmospheric situations, Artificial neural network, Wavelet transform, Pedestrian detection.*

1. Introduction

The color and visibility of images depend upon the illumination with changing intensity, color balance, direction, etc. Furthermore, nothing can be captured in total darkness. Thermal imaging systems, which are called thermography, use reflected heat from environment to identify, track, and verify the activities of people in civil and military industries [1]. Color thermal imaging provides a more illustrative temperature distribution. Although details are hard to distinguish, color thermal imaging is great for evaluating heat leaks, finding zones of high temperatures, fire operations, home inspections, and power line problem detections. Black and white images are easier and quicker to process by the brain and deliver more natural-looking images. These images are great for detection, recognition, and identification of motionless and moving objects, hunting, search, and rescue. Military applications include surveillance, tracking, target acquisition, and night vision; while remote temperature sensing, thermal efficiency analysis, weather forecasting, and environmental monitoring are some of the non-military applications [2]. In the studies of wild animals, thermal imaging can be useful for diagnosis of disease, control of reproductive processes, and analysis of behavior as well as detection and estimation of population size. Thermal imaging has been applied to detect fungal infections in stored wheat. Thermal imaging can be used as a diagnostic tool for electrical joints in power transmission systems [1]. Observation in absolute darkness and daytime under any atmospheric situation is one of its advantages. When thermal characteristics of

pedestrian are different from the background image, the target is detected. In spite of increasing trend of using these systems, there are still lots of difficulties in analyzing thermal images because of changeable features of pedestrians and atmospheric situations. Infrared (IR) radiation is emitted by all objects with a temperature above absolute zero, which is often referred to as thermal radiation. IR images have low signal-to-noise, and hence, cannot provide proper visualization of the captured objects [3,4]. Fig. 1 shows two examples of the same scene captured with both a visual and a thermal camera. Human being is not perceptible in the visible image on the left column but can be detected from thermal images on the right images with insufficient or over illumination [4].

Figure 1. Visible images (left column) and thermal images (right column) of the same scenes that pedestrians are detected in thermal images with insufficient or over illumination [4].

Detecting pedestrians in thermal images is more difficult than locating them in visible images due to the lack of information about color, texture, and shadow. Generally, the difference in intensity value of pedestrians in thermal images compared with non-pedestrians leads to target detection [5]. Researchers have put a lot of effort into detecting and tracking of pedestrians in thermal images in the recent years. Figure. 2 illustrates the main sections of different approaches in this subject. Background image subtraction is one of the most common methods available for pedestrian detection. After subtracting background image and suitable thresholding, motion regions are detected. In some methods, background image is obtained by adaptive threshold without any additional information [7-9,13]. In some other methods, a subset of database images is required to obtain background image [6,14,19]. In block-based methods, the distance of camera from the target and the sizes of pedestrians in images are very important. The recorded image is divided into equal blocks. Then the ratio of suitable criteria such as entropy, energy, mean, and variance are computed for each block of main image [2].

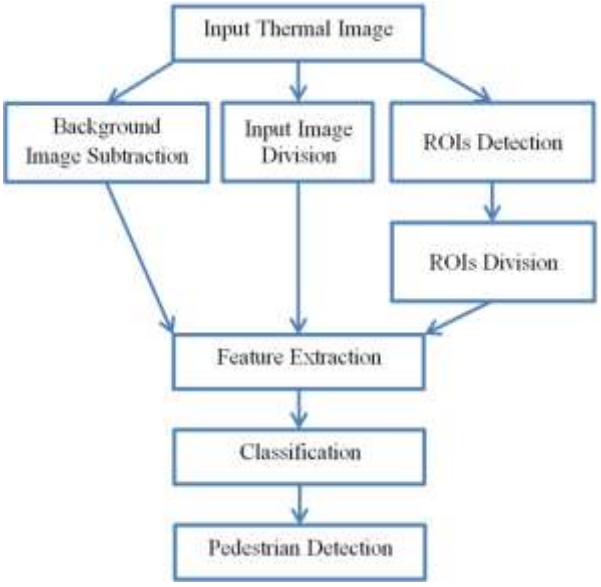

Figure 2. Main sections of recent approaches for pedestrian detection in thermal images.

Pedestrians and non-pedestrians are classified by the classifier that is trained with these features. Generally, in these methods, pedestrians who are located between several blocks are not detected during image division. In order to solve this problem, images should be divided by overlapping blocks [10,11]. In some methods, highly probable pedestrian regions are determined to increase the speed of detection. These methods increase the speed of detection because the numbers of positive and negative samples obtained are fewer, and it also increases the accuracy of detection as a result of omitting unnecessary samples such as overlapping blocks. Then the regions of interest (ROIs) are divided into overlapping blocks. Finally, suitable features are extracted from these sub-blocks, which are used to train the classifier. These methods result in a higher speed and accuracy in comparison with the methods that do not use candidate regions. Feature selection, type of classifier, and ROI detection methods are the most challenging parts of these approaches [14-16]. The features should have the least sensitivity to noise, rotation, translation, and scaling of objects [17,18,21]. Histogram of oriented gradient (HOG) [4,10,16,17,20], gabor coefficients [14], wavelet coefficients [14,15], statistical moments [14, 16], shape distribution histogram (SDH) [14], edgelet descriptors [20], energy [2], entropy [2,7], and

shape descriptors [25] are the main features that are used in thermal image pedestrian detection algorithms. Different classifiers are used for pedestrian detection in thermal images such as support vector machine (SVM) [10,11,15,20,26], adaptive boosting [19,20], modified sparse representation classifier (MSRC) [14], and feed-forward neural networks (FFNNs) [16,23]. In this paper, an efficient method is proposed for detecting pedestrians in outdoor thermal images that adapts to variable atmospheric situations. In the first step, the type of atmospheric situation is estimated based on the global features of the thermal image. Then, for each situation, a relevant algorithm is performed for pedestrian detection. To do so, thermal images are divided into three classes of atmospheric situations. Then 2-dimensional double density dual tree discrete wavelet transform (2D DD DT DWT) in three levels is acquired from input images, and the energy of low frequency coefficients in third level is calculated as the discriminating feature for atmospheric situation identification. FFNN classifier is trained by this feature vector to determine the category of atmospheric situation. Finally, a pre-determined algorithm that is relevant to the category of atmospheric situation is applied for pedestrian detection. This paper consists of 6 sections. Introduction is given in Section 1. Section 2 reviews the related works. The proposed method is introduced in Section 3, which consists of 6 sub-sections: a) Categorizing thermal imaging, b) extracting feature of training images, c) Training FFNN in order to determine the atmospheric situation with extracted features, d) ROI detection, e) Extracting feature of ROIs, and f) Training FFNN in order to detect pedestrians. In Section 4, two databases that are used to evaluate the proposed method are introduced. Parameter of proposed method, experimental results, and discussion are described in Section 5. Finally, conclusion is represented in Section 6.

2. Related works

S. Rajkumar and Chandra Mouli applied median filter on input thermal images, and then divided the image obtained into 15*15 blocks in their proposed method [2]. They calculated the ratio of entropy, energy, and combination of these criteria of blocks to the whole image. Then if the ratio was more than one, that block would be detected as pedestrian. They showed that using the combination of energy and entropy resulted better than using one of them. The proposed method cannot recognize pedestrian well in bad atmospheric situations such as a foggy weather. It also recognizes things such as electric pole and car as pedestrians whose intensity is like pedestrians'. Simplicity in performing this method is one of its advantages. Jianfu Li et al. detected ROIs to increase the speed of pedestrian detection [15]. To do so, they used the statistical adaptive intensity-oriented projection method, which contained horizontal and vertical projections. A flexible threshold was used to eliminate noise and compensate for low brightness pixels. ROIs were decomposed with 2-dimensional double density dual tree continuous wavelet transform (2D DD DT CWT) on three levels, and wavelet entropy of 96 high frequency sub-bands were considered as feature to train SVM classifier. Low accuracy in pedestrian detection on hot days and the foggy images are its disadvantages, and high accuracy in pedestrian detection in good atmospheric situations and its generalizability to other databases are some of the advantages of this method. S. Rajkumar and Chandra Mouli subtracted background images, adaptively [7]. To do so, the first peak of the histogram (P_e) was determined for each image, and then if the intensity of each pixel was less than P_e, it would become zero and the other values of pixels would become their difference from P_e. By doing so, the image intensity would decrease. Therefore, they used high-boost filter to increase the image intensity. The image obtained was made binary using a calculated threshold from variance, entropy, and mean of pixel values. Finally, white regions were considered as pedestrians in binary image. Low accuracy in pedestrian detection in bad atmospheric situations and on hot days compared with other methods mentioned in this paper is its disadvantage, and obtaining background image adaptively is one of its advantages. Wei Li et al. used HOG features to detect pedestrians in thermal images [10]. In their proposed method, the input image was divided into 20*25 sub-blocks, and then again, each of these sub-blocks was divided into smaller 4*5 sub-blocks (cells). SVM classifier was trained with the HOG feature vector of each cell. They used 3 geometric features including the mean contrast, standard deviation, and ratio of bright pixels to total pixels to obtain better results. Low accuracy in pedestrian detection in bad atmospheric situations is one of its disadvantages; a lot of train samples were used to remove this problem. High accuracy in pedestrian detection in good atmospheric situations is the advantage of this method. Chunwei Yang et al. detected ROIs with binarized normed gradients (BING) [28] to

increase the speed of pedestrian detection [16]. Then these ROIs were divided into overlapping blocks, and again each block was divided into non-overlapping cells. Some of these overlapping blocks were considered as positive (pedestrian) and the others were considered as negative (non-pedestrian) samples. Afterward, the appearance feature and HOG feature of each cell were calculated from these samples. As to extraction of appearance, the gray value of each pixel of the resized candidate was obtained, and then was normalized by norm 2. Then normalized pixels were reshaped to a one-row matrix and used as appearance features of the candidate. Next FFNN with one hidden layer was trained with these features, in which extreme learning machine (ELM) [27] was used as the learning method of FFNN. ELM is a new learning method for single-hidden layer neural network, where the hidden layer neurons can be randomly generated independent from the training data and application environment. ELM has proved a fine ability in terms of non-linear regression and classification [27], and has overcome some challenging issues such as slow learning speed and poor computational scalability. Low accuracy in pedestrian detection in bad atmospheric situations and on hot days is the disadvantage and motionless pedestrian detection with high accuracy being its advantage. Xinyue zhao et al. used the shape distribution histogram (SDH) features and contour saliency map (CSM) to detect pedestrians in thermal images with MSRC [14]. In their proposed method, primarily, background image was subtracted using mean intensity of training images. Then CSM was calculated for the foreground image. Afterwards, the most salient contours were determined using the Otsu thresholding method. In this method, a morphological operator [12,21] was used to reduce the noise, and the foreground connected region was extracted and fitted with a rectangle bounding box. They utilized a 16*30 box for bottom half the image and an 18*36 box for upper half the image. Then the distance between random points on the thinned CSM of objects in the candidate regions was used to obtain the SDH feature, which could exactly describe the pedestrian characteristics. Finally, pedestrians were detected by the MSRC classifier that was trained with the obtained features. High accuracy in pedestrian detection in all atmospheric situations and using a few positive samples for training are its advantages. This method does not perform well in motionless pedestrian detection and noisy images.

3. Proposed method

The main challenges of pedestrian detection in thermal images include A) Low contrast between intensity value of pedestrian and other objects such as trees, electricity poles, animals, and cars (Figs. 3-a and 2-f), B) Coverage of pedestrian body with thick clothes, umbrella, etc. (Figs. 3-b and 2-d), C) Variable intensity value of background (like road) in comparison with the pedestrian intensity. For example, pedestrian is usually brighter than background; however, in sunny days, the vice versa is correct. This phenomenon can be seen in Fig. 3-c, and D) bad weather conditions such as rainy and hazy situations that lead to low intensity value of pedestrians in thermal images (Figure. 3-e).

Figure 3. Samples of thermal images in different atmospheric situations [29].

The above-mentioned challenges have resulted in the atmospheric situation-dependent performances of the recent works in different sceneries. The main idea of this paper concentrates on an adaptive approach that selects the appropriate algorithm for each atmospheric situation. For this reason, an efficient algorithm is proposed based on identification of atmospheric situations. A block diagram of the proposed method is shown in Fig. 4. In the proposed algorithm, firstly, the training set input images are classified into three classes: fine, bad, and hot. Then suitable features are extracted from images of each class. An FFNN with 6 neurons in the hidden layer is trained with these features to identify the type of atmospheric situation of test image. In the next step, a relevant algorithm for each class with adaptive criteria is used to increase the accuracy of detection. In all algorithms, in order to increase the speed of detection, ROIs are detected via edge detection and morphological operators. ROIs are then divided into equal overlapping blocks. Afterwards, the blocks that are overlapped more than 70% with the pedestrian block in the groundtruth are classified as pedestrian. In all algorithms, suitable features are extracted from blocks, and are used to train FFNN with 20 neurons in the hidden layer. In the test phase,

initially, the atmospheric situation of each image is determined, and then the algorithm of the corresponding class is used to detect the pedestrians.

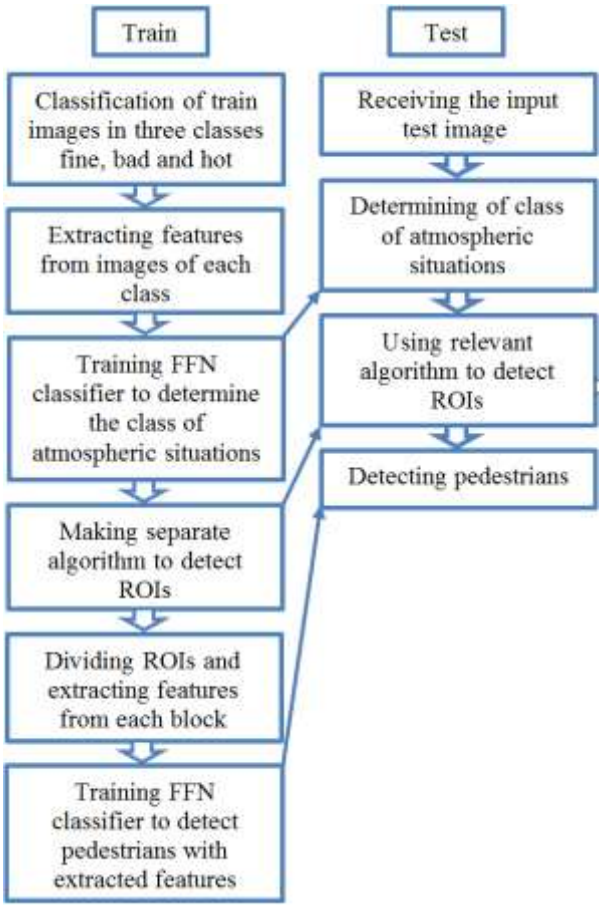

Figure 4. Block diagram of proposed method.

3.1. Categorizing training input thermal images

According to the research works carried out on thermal images, they can be classified into three classes. Class 1 includes images that are recorded in fine weather such as rather cloudy time, sunny or nights with good atmospheric conditions. Class 2 includes images that are recorded in bad weather such as foggy, rainy, and hazy weather. Class 3 includes images that are recorded in hot weather such as hot summer days, on which pedestrians are darker than background images. Table 1 shows the atmospheric index and weather conditions for 10 sets of OSU dataset [29].The images of sets 2, 4, 5, and 6 are set to class 1 (fine weather), the images of sets 1, 7, 8, 9, and 10 are assigned to class 2 (bad weather), and the images of set 3 are classified in class 3 (hot weather). In class 3, UV index is high, and the difference between dew point and temperature is high as well. The visibility in class 2 compared with class 1 is low. Pedestrians in class 1 are lighter, and have a better contrast compared with pedestrians in class 2. In class 3, mostly, pedestrians are darker than the background image.

Table 1. Atmospheric index and weather conditions for 10 sets of OSU thermal images [29].

Set	No. of Images	Temp/Dew Point (F)	Conditions	UV Index/ Visibility (miles)
1	31	45/40	Light rain	1/3
2	28	37/30	Partly cloudy	1/8
3	23	53/28	Partly cloudy	7/10
4	18	41/32	Fair	4/10
5	23	57/37	Partly cloudy	1/10
6	18	53/46	Mostly cloudy	1/8
7	22	68/58	Light rain	1/9
8	24	62/56	Light rain	2/4
9	73	50/47	Haze	0/2.5
10	24	55/79	Haze	2/3

3.2. Extracting suitable features from images of each class

In this step, the extracted features should be able to distinguish the most among the images of three classes. Experiments have shown that low frequency coefficients of 2-D DD DT DWT are able to perfectly distinguish between the three classes. This transform corresponds to a new family of dyadic wavelet tight frames based on two scaling functions and four distinct wavelets ($\psi_{h,i}(t)$ and $\psi_{g,i}(t)$ for i = 1,2). One pair of the four wavelets are designed to be offset from the other pair of wavelets so that the integer translates of one wavelet pair fall midway between the integer translates of the other pair $(\psi_{h,1}(t)\approx\psi_{h,2}(t\text{-}0.5).\ \psi_{g,1}(t)\approx\psi_{g,2}(t\text{-}0.5)$.
Simultaneously, one pair of wavelets are designed to be approximate hilbert transforms (H{.}) of the other pair of wavelets so that two complex (approximately analytic) wavelets can be formed $(\psi_{g,1}(t)\approx H\{\psi_{h,1}(t)\},\ \psi_{g,2}(t)\approx H\{\psi_{h,2}(t)\})$ [24]. 2-D DD-DT DWT is implemented using four over-sampling 2-D DD DWT for the input image in parallel. The rows and columns are filtered using different filter banks. Fig. 5 illustrates the analysis process of two levels of 2-D DD-DT DWT.

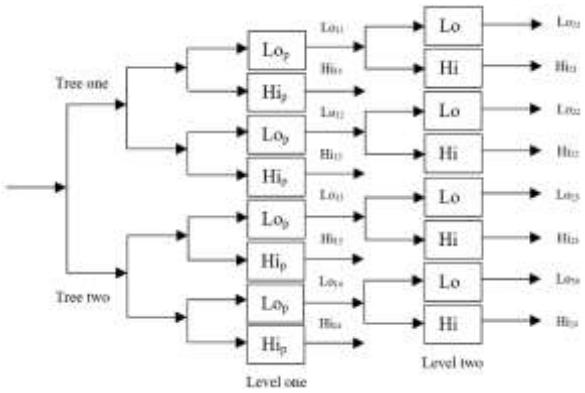

Figure 5. Two levels of 2D DD-DT CWT [15].

Lo_p and Hi_p are the first level of the decomposition filter banks, and represent one scaling filter and eight wavelet filters of 2-D DD DWT. Lo and Hi are the second and remaining levels of decomposition filter banks. Lo_{mn} and Hi_{mn} (m = 1,2; n = 1,2,3,4) represent one low pass sub-band and eight high pass sub-bands separately [15]. To distinguish among the three classes, 2-D DD-DT DWT is calculated for train input images in three levels. Finally, 32 high pass sub-bands for each level and 4 low pass sub-bands are produced for the input image after one level of the transform. In order to reduce the computational cost, 2-D DD-DT DWT is calculated in the central part of the image. For example, in the OSU thermal images of 240*360 pixels, 2-D DD-DT DWT is calculated for pixels (x = 61:180) and (y = 81:240), and four low pass sub-bands with 15 rows and 20 columns are made with these calculations. Figure. 6 shows one of the low frequency sub-bands for the first image of each set of OSU thermal dataset.

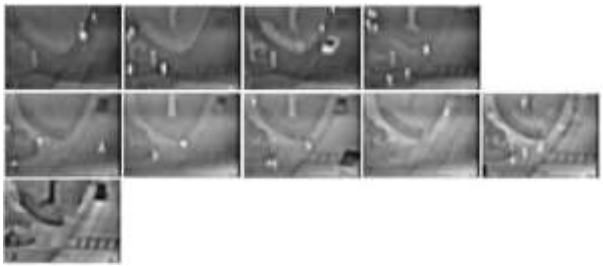

Figure 6. First low pass sub-bands for first image of each set of OSU thermal dataset.

Then the mean energy for every column of each low pass sub-band is calculated. Consequently, a feature vector with dimension of 1*80 is created for each image, and is used to train the FFNN classifier.

3.3. Training FFNN to determine atmospheric situation

In this step, FFNN with specifications of Table 2 is trained using extracted features from the previous step. In order to show the efficiency of this network in detecting the type of test input image class, we randomly selected half of the images of the databases as train images and the other half as test images. After training, this network is able to distinguish among the images of three classes with an accuracy of 100%. To determine the suitable number of neurons in the hidden layer, 30% of train images are selected randomly, where half of them are considered as train images and the rest as test images. Then FFNN with the number of neurons in the hidden layer varying from 4 to 10 was trained with these train images. Next due to these changes, the sensitivity of classification for test images was assessed 7 times. The mean of these 7 times was calculated for each varied neuron. The best result was obtained for the hidden layer with 6 neurons. The three neurons used in the output layer will determine the atmospheric situations.

Table 2. Specifications of FFNN used to distinguish atmospheric situation class.

Functions	Type
Number of hidden layers	1
Number of neurons in hidden layer	6
Number of neurons in output layer	3
Activation functions for hidden layer	Hyperbolic tangent sigmoid
Activation functions for output layer	Linear
Learning function in back-propagation	Levenberg-marquardt

3.4. ROI detection

ROI detection for classes 1 and 2 is nearly the same but for class 3 is different, which is mentioned later.

3.4.1. ROI detection for classes 1 and 2

In order to detect ROIs for classes 1 and 2, first, the input image I is normalized with Eq. (1), and results in I_n.

$$I_n(x,y) = I(x,y) / \max|I| \quad \text{for} \quad x = 1:i,\ y = 1:j \qquad (1)$$

In Eq. (1), max|I| is the maximum absolute image value, (x,y) is the location of pixel in the image, and i×j is the image dimension (Fig. 7). The two pictures look like the same. The gray value for each pixel in the normalized image is between 0 and 1 in order to simplify the computations in the next steps.

Figure 7. A sample of thermal image (a) and normalized thermal image (b).

In the second step, the image pixels with lower values than the mean value (mean(I_n)) are set to zero with Eq. (2) (Figure. 8).

$$F = \begin{cases} I_n(x,y), & I_n(x,y) \le \text{mean}(I_n) \\ 0, & I_n(x,y) > \text{mean}(I_n) \end{cases} \text{ for } x = 1:i,\ y = 1:j \qquad (2)$$

Figure 8. Adjustment of normalized thermal image.

In the third step, the edge values of F are calculated through canny edge detector with δ=1.4, DT=0.4×mean(I_n) and UT=mean(I_n). UT and DT are the upper and down thresholds, respectively, and the result of edge detection is L(x,y) for x=1:i and y=1:j. It should be mentioned that the image margin with the thickness of two pixels is forced to zero to remove the effect of the edge of image margin before edge detection (Figure. 9).

Figure 9. Edge detection with canny edge detector.

In the fourth step, the convex hull morphological operator [21] is done on L(x,y) and binary image CH(x,y) for x=1:i and y=1:j is made (Figure. 10).

Figure 10. Applying convex hull morphological operator.

In the fifth step, the area of each block of CH is calculated, and the unnecessary blocks are omitted based on their area. The minimum and maximum connected pixels that are acceptable are stated in Table 3. In the sixth step, the mass center of each remained block is marked as ROIs (Figure. 11).

Figure 11. Center detection of ROIs with possibility of presence of one pedestrian (a), several pedestrians (b).

Then if the horizontal and vertical distances between two centers are less than thresholds T_1 and T_2, respectively, the region between two centers is considered as central point of final box (Figure. 12).

Figure 12. Omitting unnecessary centers in Fig. 11-a.

Subsequently, if there are unnecessary centers, they will be omitted. T_1 is the length threshold, and is 0.6×M. T_2 is the width threshold, and is 0.4×N. M and N are the length and width of the box (B), respectively. In the seventh step, boxes with the dimension of (M,N), which are determined by train database, are located in ROIs of F (Figure. 13).

Figure 13. ROIs with possibility of presence of one pedestrian (a) and possibility of presence of several pedestrians (b).

The farther the camera from the ground, the smaller is the box dimension. Finally, these boxes are stored as ROIs. If there are several pedestrians next to one another in the input image, their images will result in a bigger joint connected component. Hence, if the number of pixels in a connected component is more than a threshold T_3, the region relevant to the connected component is considered as the region of several pedestrians (image (b) of Fig. 11). After that, these regions are divided into same (M*N) boxes with 0.3×M overlapping length and 0.4×N overlapping width. Threshold T_3 is determined by the distance between camera and the ground, the quality of the camera, the dimensions of input image, and so on (Figure. 14).

Figure 14. Dividing ROIs with possibility of presence of several pedestrians in same blocks.

Figure.15 shows the steps of detecting ROIs for four thermal images. In columns one and three, the algorithm is stopped in step 7 because there is no possibility of the presence of several pedestrians.

3.4.2. ROI detection for class 3

We can use the method used in section 3.4.1 to detect ROIs in class 3. In case of using this method, the detected ROIs will be increased, which reduces the speed and accuracy of detection. In the meantime, it is possible that some pedestrians are considered as background image since most of pedestrians are darker than the background image.

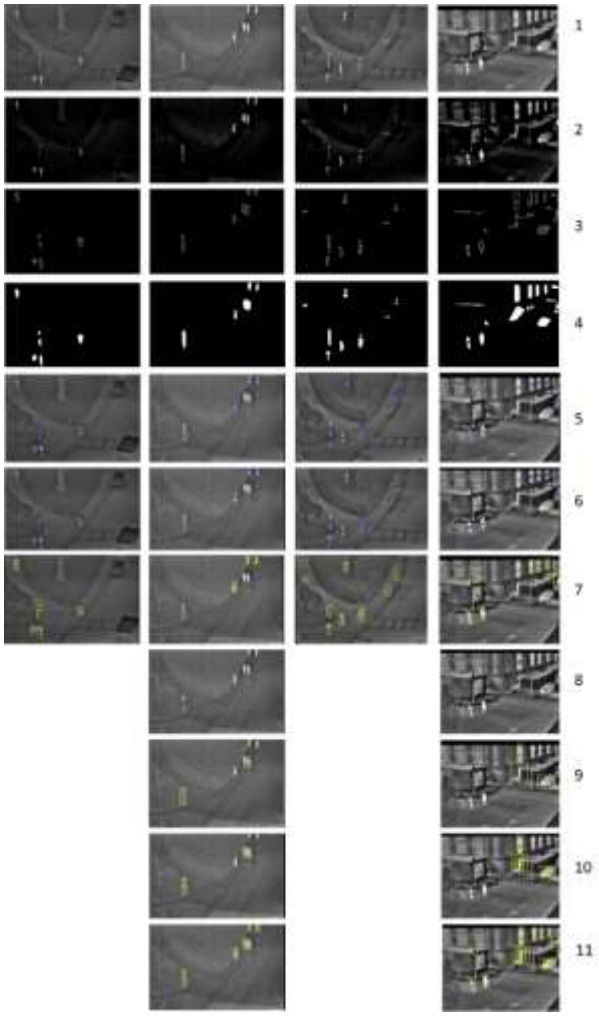

Figure 15. Classes 1 & 2 ROI detection (row 1: normalized thermal images, row 2: adjustment, row 3: edge detection of row 2 with canny edge detector, row 4: applying convex hull morphological operator, row 5: ROI center detection with possibility of presence of one pedestrian, row 6: omitting unnecessary centers, row 7: ROI block division with possibility of presence of one pedestrian, row 8: ROI center detection with possibility of presence of several pedestrians, row 9: ROI block division with possibility of presence of several pedestrians, row 10: dividing row 9 images into blocks, row 11: ROI detection with possibility of presence of pedestrians).

Therefore, three images of train dataset are selected randomly and considered as the main images in step 1 (Figure 16).

Figure 16. Selecting three images randomly as main images.

In step 2, each of the three images is normalized with maximum absolute value using Eq. (1). In step 3, K_p is obtained through subtracting normalized images (I_{np}) from an image with the same dimension of main images with the value of 1 using Eq. (3) (Figure 17)

$$K_p = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}_{i\times j} - \begin{pmatrix} I_{np}(1,1) & \cdots & I_{np}(1,j) \\ \vdots & \ddots & \vdots \\ I_{np}(i,1) & \cdots & I_{np}(i,j) \end{pmatrix}_{i\times j} \quad (3)$$

where $i \times j$ is the image dimension, and I_{np} are the normalized images, where p = 1, 2, and 3.

Figure 17. Subtracting main normalized images from matrix with same dimension with value of 1.

Then steps 2 and 3 are repeated for the input test image so that K_0 is resulted (Figure 18).

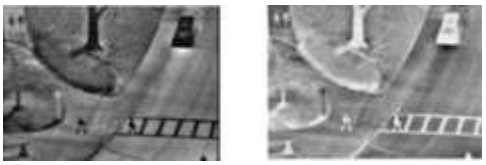

Figure. 18. Repeating steps 2 and 3 for a sample of test images.

In step 4, K_0 is subtracted from K_p using Eq. (4) to obtain S_p for p = 1, 2, and 3 (Figure 19).

$$S_p = K_p - K_0 \quad \text{for} \quad p = 1,2,3 \quad (4)$$

Figure 19. Subtracting K_0 from K_p for p = 1, 2, and 3.

In step 5, the edge detection is applied on S_p with UT = mean I_{np}, DT = 0.4 × mean I_{np} and $\delta = 1.4$ using canny edge detector where p = 1, 2, and 3 (Figure 20).

Figure 20. Edge detection of S_p with canny edge detector, where p = 1, 2, and 3.

Then In step 6, the convex hull of them is calculated so that SH_p are resulted, where p = 1, 2, and 3 (Figure 21).

Figure 21. Applying convex hull morphological operator on Fig. 20

In step 7, common regions between SH_p are considered as ROIs (Figure 22).

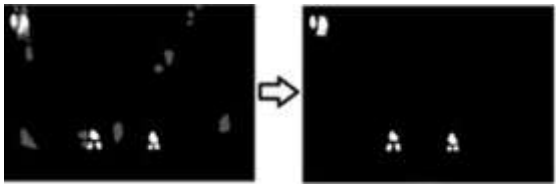

Figure 22. Finding common regions on images of Fig. 21.

In step 8, the center of box with the dimension of (M,N), which are determined by train database, is located on the center of ROIs (Fig. 23).

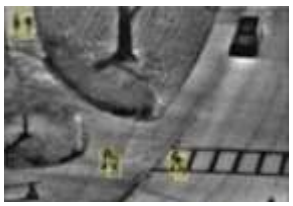

Figure 23. Detection of ROIs with possibility of pedestrians' presence.

3.5. Feature extraction

In this section, ROIs are divided into blocks, and the necessary features are extracted from each one of them. This method is the same for extracting features in the atmospheric situations of classes 1 and 2, and is different in the atmospheric situations of class 3 in some steps.

3.5.1. Feature extraction in atmospheric situations of classes 1 and 2

In this section, the necessary features for each box in the atmospheric situations of classes 1 and 2 are extracted. The boxes are regions, which are detected as ROIs in the previous sections. First, the mean value of box (β) is calculated. In the second step, each box is adjusted by Eq. (5) to obtain $Ba_{i,j}$ in order to omit the unnecessary pixels and to highlight important pixels.

$$Ba_{i,j}=\begin{cases} B_{i,j}, & \alpha\times\beta \le B_{i,j} \le 2\alpha\times\beta \\ 0, & B_{i,j} < \alpha\times\beta \\ 1, & B_{i,j} > 2\alpha\times\beta \end{cases} \quad \beta=\text{mean B} \qquad (5)$$

In the third step, a median filter with the dimension of q*p is applied on $Ba_{i,j}$ to obtain $Bf_{i,j}$.

In the fourth step, $Bb_{i,j}$ is made binary with Eq. (6).

$$Bb_{i,j}=\begin{cases} 1, & Bf_{i,j} > \frac{\alpha}{2} \\ 0, & \text{otherwise} \end{cases} \qquad (6)$$

Figure 24 shows steps 1 to 4 for three samples of pedestrian and non-pedestrian blocks (α, p, and q are stated in Table 3).

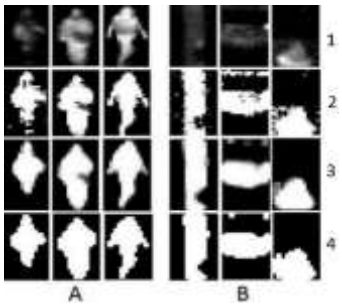

Figure 24. Creating binary image for three samples of pedestrian (A) and non-pedestrian (B) (row 1: samples of pedestrians and non-pedestrians, row 2: adjusting contrast of samples of row 1, row 3: applying median filter, row 4: creating binary image).

Table 3. Parameters of proposed method in databases.

Database	Database 1			Database 2
Atmospheric situation	class1 (fine)	class2 (bad)	class3 (hot)	class 1 (fine)
M*N (box dimension)	28*16	28*16	28*16	36*16
Minimum connected pixels for one pedestrian	30	20	25	80
Maximum connected pixels for one pedestrian	350	350	350	460
Maximum connected pixels for several pedestrian	1500	1500	1500	1700
p*q (median filter dimension)	4*3	4*3	-	5*3
Value of α for adjusting in feature extraction	0.7	0.6	-	0.7

In the fifth step, a weight center point of binary image is determined and the box center is transferred to this determined weight center points. This step can be repeated for several times in order to increase the accuracy in the next steps, if necessary. It should be mentioned that connected pedestrians should be separated before starting step 5. To do so, a segmentation method is utilized to separate the close pedestrians that are fused incorrectly, and then we can differentiate the overlapped pedestrians. The minimum value in the center region of vertical projection curve and the maximum value in the middle region of horizontal projection curve usually represent the junction of two pedestrians together. Finally, all length and width pixels proportionate to these values are made zero, and then step 5 starts. As it

can be seen in Figure 25, the pedestrian in the upper part of ROI transfers to the middle of box 1 (the upper box) after two iterations of step 5.

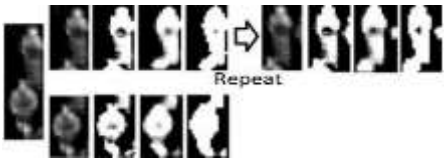

Figure 25. Locating pedestrian in middle of box and transforming box center into weight center points.

Figure 26 shows the separation of connected pedestrians in separate boxes.

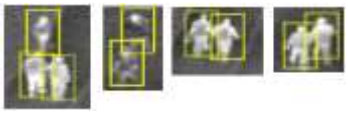

Figure 26. Separation of connected pedestrians.

In the sixth step, this box is divided into some blocks. Then the ratio of mean and entropy of each block to the whole box is calculated, and this calculation is used as feature to train FFNN with 20 neurons in the hidden layer. Empirically, the best size for each sub-block is obtained 4*4. Entropy (Eq. (7)) is a statistical measure that can be used to characterize the texture of the input thermal image.

$$\text{Entropy} = H = -\sum_{I=0}^{N_g-1} P(I).\log_2\left[P(I)\right] \qquad (7)$$

where P is the probability of number of occurrences of I in thermal image and N_g is the maximum gray level of pixels in the image. The image and graph in Figure 27 show a thermal image and the difference of extracted features among pedestrians (blue lines) and non-pedestrian (red lines).

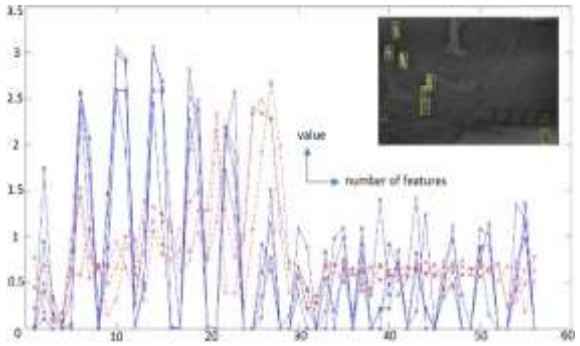

Figure 27. A thermal image and graph of extracted features from samples of pedestrians and non-pedestrians.

The dimension of selected box for this image is 28*16, which is blocked into 28 sub-blocks. The first 28 features are related to the mean features and the second 28 features are related to the entropy.

3.5.2. Feature extraction in atmospheric situations of classes 1 and 2

In order to extract features in the atmospheric situations of class 3, each box of ROI is divided into some blocks. Then, like classes 1 and 2, ratio of mean and entropy of each block to the whole box is calculated, and this calculation is used as feature to train FFNN with 20 neurons in the hidden layer.

3.6. Training FFNN classifier to detect pedestrians from non-pedestrians

In this section, the FFNN classifier with 20 neurons in the hidden layer is trained with the features obtained from Section (3.5). Parameters of this FFNN are the same as Table 2. ROC (receiver operating characteristic) curve is used to determine the number of neurons in the hidden layer. As a matter of fact, in an ROC curve, true positive rate (TPR) or sensitivity is plotted as a function of the false positive rate (FPR) or specificity for different cut-off points of a parameter. After identifying ROIs in the training step, 200 boxes of pedestrians and 200 boxes of non-pedestrians are selected from each three classes of atmospheric situations randomly. Then entropy and mean features of these ROIs are extracted. Next, FFNN is trained with half of these features, the number of neurons in the hidden layer varying from 12 to 24. In each training step, the value of threshold that distinguishes the pedestrians from non-pedestrians is identified. After each training step, the other half of the features is exerted to FFNN and the value is determined. If this value is bigger than the value of threshold in the training step, that ROI is considered as pedestrian, and if it is smaller, ROI is considered as non-pedestrian. If the area under the ROC curve (AUC) is bigger, the accuracy of classifiers is higher. Table 4 shows AUC for FFNN; the number of neurons in the hidden layer vary from 12 to 24 for the two different experiments.

Number of neurons	AUC (experiment 1)	AUC (experiment 2)
12	0.9623	0.9684
14	0.9954	0.9684
16	0.9907	0.9885
18	0.9939	0.9866
20	**0.9955**	**0.9952**
22	0.9870	0.9946
24	0.9954	0.9871

Table 4. AUC for FFNN in which number of neurons in hidden layer vary from 12 to 24.

Figure 28 shows the ROC curve for experiment 2. The biggest value for AUC is obtained from the number of 20 neurons in the hidden layer. The biggest value for AUC is obtained from the number of 20 neurons in the hidden layer.

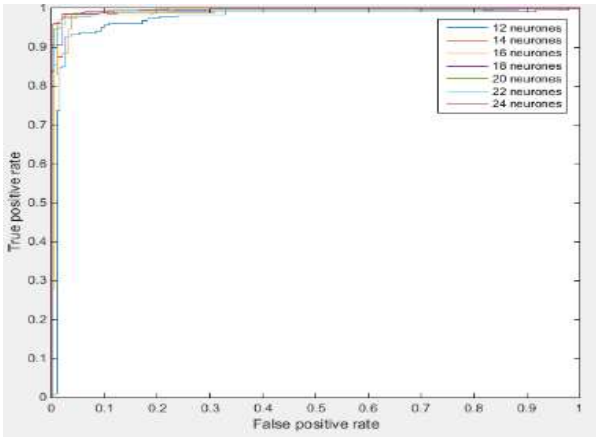

Figure 28. ROC curve for experiment 2.

4. Database

Two databases of OTCBVS benchmark database collection are used to evaluate and show the high efficiency of the proposed method. The OSU thermal pedestrian database (database 1) includes all the atmospheric situations (classes 1, 2, and 3), which cover a variety of environmental conditions such as rainy, cloudy, and sunny days [29]. Database 1 includes thermal images with the dimensions of 360*240 pixels, which are recorded with Raytheon 300D thermal sensor core and 75 mm lens, where camera was mounted on rooftop of an 8-story building. This database includes 10 sub-databases of thermal images, which consist of 284 images. Every image contains an average of 3 to 4 pedestrians. Half of the images of the sub-databases are considered as the training images and the other half are considered as the testing images. The OSU Color-Thermal database (database 2) includes thermal images with the dimensions of 320*240 pixel, which are recorded with Raytheon PalmIR 250D thermal sensor core and 25 mm lens, where camera was mounted on rooftop of a 3-story building [30]. This database includes 3 sub-databases of thermal images, which consist of 5189 images. Since pedestrians are lighter than the background image and the noise is negligible, the images of these 3 sub-databases are considered as class 1. 30% of the images of the sub-database are considered as the training images and the other 70% are considered as the testing images.

5. Parameters of proposed method, experimental results, and discussion

Table 3 shows the parameter values of the proposed method, which are empirically determined considering the dimension of images, dimension of pedestrians in images, and location of the camera. The length and width of the box are a multiple of 4. The dimensions of median filter depend on the dimension of box, and if the box is bigger, the dimension of filter is bigger. According to the dimensions of the pedestrians in training images, dimension of the box, and atmospheric situations, the maximum and minimum numbers of connected pixels for one or several pedestrians are determined. The nearer the camera to the ground, the more is the number of pixels. For example, the maximum number of the connected pixels for one pedestrian is directly related to the dimension of the box, and is 0.8×M×N. The two criteria sensitivity and positive predictive value (PPV) are used to evaluate the accuracy and efficiency of the proposed method in Eqs. (8) and (9), respectively.

$$\text{Sensitivity} = \text{true positive} / \text{total pedestrians} \quad (8)$$

$$\text{PPV} = 1 - (\text{false positive} / \text{total pedestrians}) \quad (9)$$

Sensitivity shows the percent of correct detection of pedestrians. PPV shows the regions that are detected as pedestrians incorrectly. If the number of these regions is fewer, the percentage of these criteria is higher.

5.1. Result of proposed method on databases 1 and 2

The results of atmospheric situation detection for the trained and test images are shown in Table 5, which show that all train and test images are classified correctly.

Table 5. Confusion matrix of atmospheric situations class detection for train and test images

	Train images			Test images		
Class	**1**	**2**	**3**	**1**	**2**	**3**
1	44	0	0	4	0	0
2	0	88	0	0	86	0
3	0	0	12	0	0	11

After determining the atmospheric situations, the related algorithm will be applied. Figure 29 shows pedestrians' detection in some images of database 1 for related works. Rows 1-10 in Fig. 30 show pedestrians' detection in some images of database 1, where each row consists of 7 images. The images of each row are related to the same sub-database. In the images of row 1, pedestrians carrying umbrellas or pedestrians whose intensity is lower are detected very easily. All 91 pedestrians are detected in this sub-database. No wrong detection is observed. The sensitivity and PPV for this sub-database are 100%. The pedestrians who are close to each other are detected very easily and separately in images of

sub-database 2. All 100 pedestrians are detected in this sub-database, and no wrong detection is seen. The sensitivity and PPV for this sub-database are 100%.

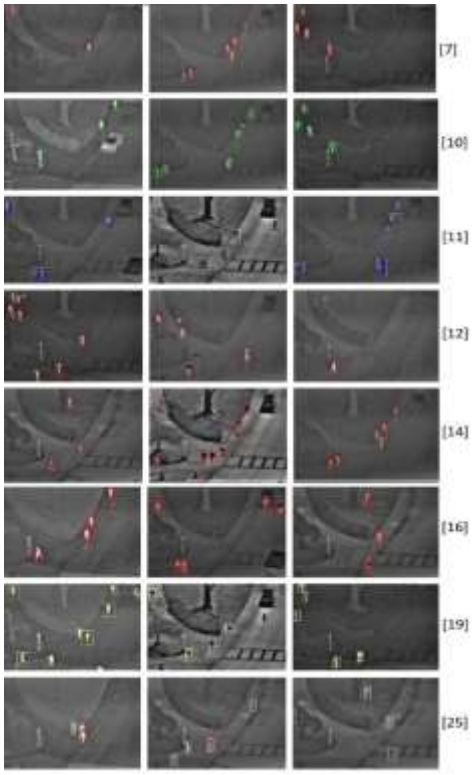

Figure 29. Samples of pedestrians' detection in database 1 for related works.

In recorded images of row 3 in hot days, pedestrians whose intensity difference with background image is very low, are detected easily. Pedestrians who are darker than the background image are detected very easily as well. 100 pedestrians of 101 are detected, and no wrong detection exists. The sensitivity is 99%, and PPV is 100% for this sub-database. Electric poles and cars, which have the intensity value like pedestrians, are not detected as pedestrians. Pedestrian detection in sub-database 4 is perfect, and the sensitivity and PPV for this sub-database are 100%. All 109 pedestrians are detected in sub-database 5 without any wrong detection. The sensitivity and PPV for this sub-database are 100%. In sub-database 6, 96 pedestrians of 97 are detected, where there is no wrong detection. The sensitivity is 99%, and PPV is 100% for this sub-database. The pedestrians who are close to each other are detected very easily and separately in images of sub-database 6. Pedestrian detection in the images of sub-database 7 is perfect, like sub-databases 1, 2, 4, and 5, and all 94 pedestrians are detected. All 99 pedestrians are detected correctly, and only one wrong detection exists in sub-database 8. In this sub-database, pedestrians who do not move in successive images are detected easily and separately. The sensitivity is 100%, and PPV is 99% for this sub-database. Pedestrian detection in the images of sub-database 9 is like sub-database 6, and 94 pedestrians of 95 are detected. The sensitivity is 99%, and PPV is 100% for this sub-database. 93 pedestrians of 97 are detected in sub-database 10, where 3 wrong detections exist. Lack of detection of 4 pedestrians in this sub-database is due to very noisy images due to foggy weather. The sensitivity is 96%, and PPV is 97% for this sub-database.

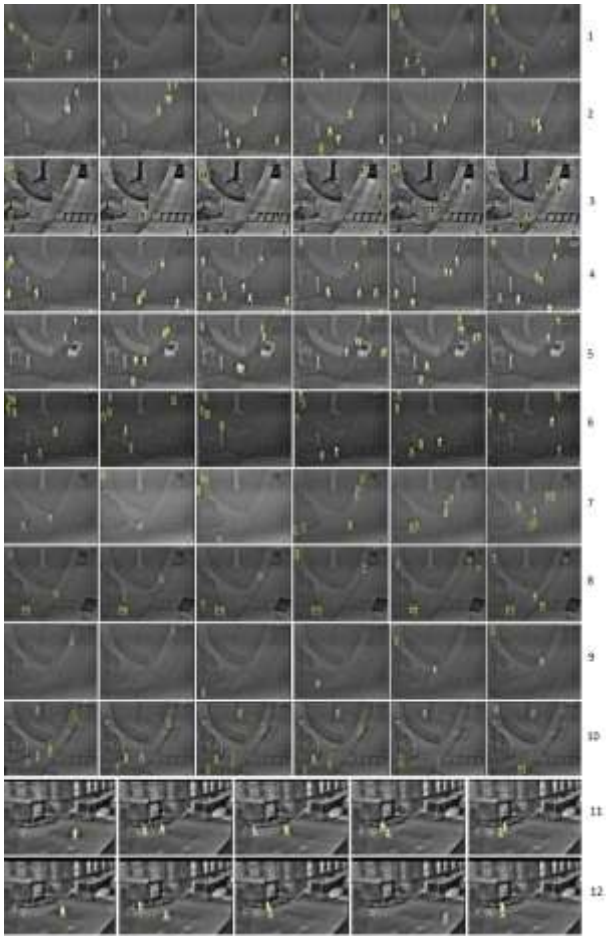

Figure 30. Samples of pedestrians' detection in database 1 (rows 1-10) and database 2 (rows 11-12).

The overall sensitivity and PPV for this database are 99.3% and 99.6%, respectively. Since the number of images in database 2 was high, we selected 200 images of each sub-database randomly. For each sub-database, 60 images were considered as train images, and the other 140 images were considered as test images. The train and test images were selected randomly including 682 pedestrians, where 677 pedestrians were detected and 6 wrong detections occurred with the proposed method. Since two pedestrians were close to each other in some images, 6 pedestrians were not detected. Rows 11 and 12 of Figure 30

show 10 samples of images of this database, where pedestrians were detected correctly. The sensitivity and PPV for this database were 99.3% and 99.1%, respectively. The detection results with the proposed method in databases 1 and 2 are shown briefly in Table 6.

5.2. Noise performance

In this section, the proposed method was applied to noisy images to evaluate the noise performance. Gaussian noise with mean = 0.1 and variance = 0.005 was added to all images of classes 1 and 3, and Gaussian noise with mean = 0.1 and variance = 0.001 was added to all images of class 2. Images A, B, and C in Fig. 31 show the input images, noisy images, and detected pedestrians, respectively. Applying noise to images does not influence sensitivity very much but PPV is somewhat reduced. Since sensitivity is more important than PPV in detection systems, this reduction can be neglected. In detection systems, the detection of all pedestrians is very important, and if several regions are incorrectly detected as pedestrian, it is less challenging. Table 7 shows the detection rate of pedestrians in noisy images of databases 1 and 2 using the proposed method.

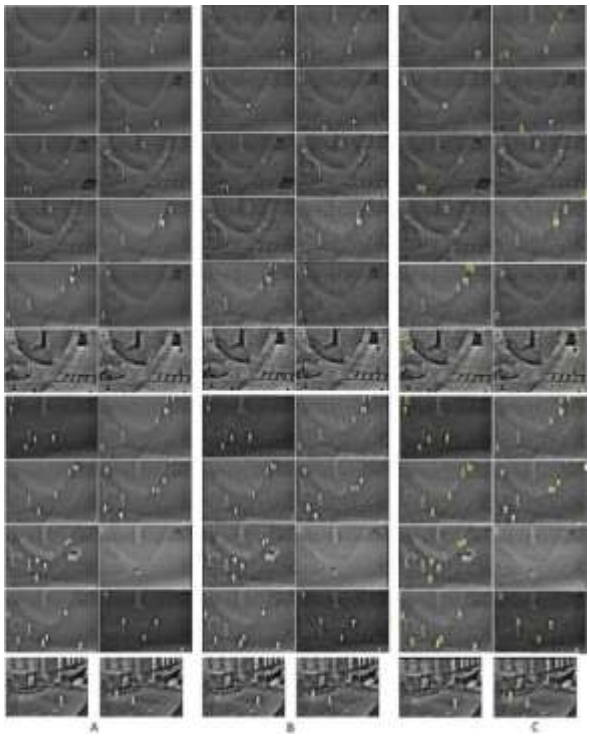

Figure 31. Samples of pedestrians' detection in noisy images (A: input images, B: noisy images, C: pedestrian detection).

5.3. Comparison of proposed method with related works

The proposed method can be generalized on different thermal databases, and is estimated to show good results by adjusting its parameters correctly. Table 8 shows the comparison between pedestrian detection rate (average of sensitivity and PPV) in the proposed method and related works. Sensitivity and PPV were not addressed in some papers such as [2] and [7] so comparison with related works was done by the average of sensitivity and PPV. Table 9 shows the comparison between sensitivity, PPV, and detection rate in the proposed method and method [15]. Method [7] has a low accuracy in sub-databases 1 (88.9%), 3 (68.2%), 6 (88.5%), 7 (87.2%), and 10 (78.8%). This method has a low detection rate in hot days and bad atmospheric situations. Method [10] has a low detection rate in hot days and bad atmospheric situations such as sub-databases 3 (76.5%), 5 (88.5%), 7 (81%), and 10 (89.5%). Method [11] in sub-database 10 has a lower detection rate (96%) than other sub-databases. A large number of negative training samples for each sub-database have been used to raise the detection rate in this method. Method [2] has a low detection rate in hot days (92%), and the detection rate in sub-databases 8, 9, and 10 is not specified. Method [14] has a low detection rate in rainy days (sub-database 7) and a low accuracy in motionless pedestrian detection (sub-database 8). Method [16] has a low detection rate in hot days and bad atmospheric situations such as sub-databases 3 (77%), 5 (80%), 8 (85%), and 10 (84%). This method has a low accuracy in motionless pedestrian detection. Method [25] is more accurate than the other works but this method uses other sub-databases for background extraction. The detection rate of the proposed method on the sub-databases 1, 2, 4, 5, and 7 of database 1 is 100%. The detection rate of the proposed method on the sub-databases 3, 6, 8, and 9 of database 1 is 99.5%. The proposed method in sub-database 10 has a lower detection rate (96.5%) than the other sub-databases. In general, pedestrian detection rate of the proposed method for all images of database 1 is 99.45%, which is higher than the pedestrian detection rate in the related works. Method [15] has a low accuracy in hot days and foggy weather. The proposed method and method [15] on database 2 have the same sensitivity (99.1%) but PPV in the proposed method (99%) is higher than PPV in method [15] (91%). In general, pedestrian detection rate of the proposed method for database 2 is 99.05%

6. Conclusion

In general, the proposed method includes 4 parts: a) atmospheric situations detection, b) detection of ROIs, c) feature extraction, and d) detection of

pedestrians. In the proposed method, initially, the class of atmospheric situation was determined. To do so, 2D DD DT DWT was extracted from train images in 3 levels, and the mean of low frequency sub-band energies was calculated in the third level. These features were used to train FFNN with 6 neurons in the hidden layer in order to determine class of atmospheric situations. In the second part, ROIs were detected using canny edge detector and convex hull morphological operator with adaptive criteria to increase the speed of detection. In the third part, pedestrians and non-pedestrians in ROIs were divided into sub-blocks of 4*4 dimensions, and the ratio of entropy and mean for each sub-block to the whole block was calculated. These features were used to train an FFNN with 20 neurons in the hidden layer in order to distinguish pedestrians from non-pedestrians. Finally, in the last part, atmospheric situation for input test images were determined, and then pedestrians were detected using the relevant algorithm of each class. The ability of performance on different thermal databases, motionless pedestrian detection, high speed detection, independence from atmospheric situations, ability to detect pedestrians with camera movement, detection of pedestrians who are close to one another separately, lack of sensitivity to noise, and independence from background image are some of the advantages of the proposed method. The accuracy of the proposed method for pedestrian detection on two databases is more than 99%.

Table 6. Result of proposed method on databases 1 and 2.

	Database 1											**Database2**
Database	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1-10**	
# Images	31	28	23	18	23	18	22	24	73	24	284	600
# Pedestrians	91	100	101	109	101	97	94	99	95	97	984	682
# Detected pedestrians	91	100	100	109	101	96	94	99	94	93	977	677
#Undetected pedestrians	0	0	1	0	0	1	0	0	1	4	7	5
# Wrong detections	0	0	0	0	0	0	0	1	0	3	4	6
Sensitivity	100%	100%	99%	100%	100%	99%	100%	100%	99%	96%	**99.3%**	**99.3%**
PPV	100%	100%	100%	100%	100%	100%	100%	99%	100%	97%	**99.6%**	**99.1%**

Table 7. Result of proposed method on databases 1 and 2 after applying Gaussian noise.

	Database 1											**Database 2**
Database	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1-10**	
# Images	31	28	23	18	23	18	22	24	73	24	284	600
# Pedestrians	91	100	101	109	101	97	94	99	95	97	984	682
# Detected pedestrians	88	100	98	109	100	95	94	99	94	92	969	673
# Undetected pedestrians	3	0	3	0	1	2	0	0	1	5	15	9
# Wrong detections	0	1	8	0	2	5	0	3	3	4	26	11
Sensitivity	96.7%	100%	97%	100%	99%	97.9%	100%	100%	99%	94.8%	**98.5%**	**98.7%**
PPV	100%	99%	92%	100%	98%	94.8%	100%	97%	96.8%	95.9%	**97.4%**	**98.4%**

Table 8. Comparison between pedestrians' detection rate (average of sensitivity and PPV) in proposed method and related works on database 1.

	Sub-database number										
Method	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**1-10**
[2]	95%	94%	92%	98%	92%	94%	94%	-	-	-	94.2%
[7]	88.9%	96.3%	68.2%	**100%**	95.5%	88.5%	87.2%	92.5%	98.6%	78.8%	89.4%
[10]	92%	96%	76.5%	96%	88.5%	94.5%	81%	91.5%	93.5%	89.5%	89.9%
[11]	99%	**100%**	95.5%	99.5%	99%	98%	**100%**	98.5%	99.5%	96%	98.5%
[14]	98.5%	**100%**	99.4%	98.9%	**100%**	**100%**	96%	96.5%	**100%**	**100%**	98.9%
[16]	89.5%	90.5%	77%	88.5%	80%	94%	89%	85%	92%	84%	86.9%
[19]	98.5%	97%	99.5%	98.5%	94.5%	98%	99%	87%	**100%**	97.5%	96.9%
[25]	**100%**	99%	**100%**	**100%**	**100%**	99.5%	99.5%	99%	**100%**	97%	99%
Proposed method	**100%**	**100%**	99.5%	**100%**	**100%**	99.5%	**100%**	**99.5%**	99.5%	96.5%	**99.45%**

Table 9. Comparison between pedestrians' detection rate in proposed method and related works on database 2.

Method	**Sensitivity**	**PPV**	**Detection rate**
[15]	99.1%	91%	95/05%
Proposed method	**99.1%**	**99%**	**99.05%**

References

[1] Gade, R., & Moeslund, T. B. (2014). Thermal cameras and applications: a survey. Machine vision and applications, vol. 25, no. 1, pp. 245-262.

[2] Rajkumar, S., & Mouli, P. C. (2015). Target Detection in Infrared Images Using Block-Based Approach. In Informatics and Communication Technologies for Societal Development, Springer, New Delhi, pp. 9-16.

[3] Fang, Y., Yamada, K., Ninomiya, Y., Horn, B. K., & Masaki, I. (2004). A shape-independent method for

pedestrian detection with far-infrared images. IEEE Transactions on Vehicular Technology, vol. 53, no. 6, pp. 1679-1697.

[4] Qi, B., John, V., Liu, Z., & Mita, S. (2016). Pedestrian detection from thermal images: A sparse representation based approach. Infrared Physics & Technology, vol. 76, pp. 157-167.

[5] Negied, N. K., Hemayed, E. E., & Fayek, M. B. (2015). Pedestrians' detection in thermal bands–Critical survey. Journal of Electrical Systems and Information Technology, vol. 2, no. 2, pp. 141-148.

[6] Davis, J. W., & Sharma, V. (2007). Background-subtraction in thermal imagery using contour saliency. International Journal of Computer Vision, vol. 71, no. 2, pp. 161-181.

[7] Soundrapandiyan, R., & Mouli, P. C. (2017). Robust pedestrian detection in infrared images using rotation and scale invariant-based structure element descriptor. International Journal of Signal and Imaging Systems Engineering, vol. 10, no. 3, pp. 157-167.

[8] El Maadi, A., & Maldague, X. (2007). Outdoor infrared video surveillance: A novel dynamic technique for the subtraction of a changing background of IR images. Infrared physics & technology, vol. 49, no. 3, pp. 261-265.

[9] Davis, J. W., & Sharma, V. (2007). Background-subtraction using contour-based fusion of thermal and visible imagery. Computer vision and image understanding, vol. 106, no. 2-3, pp. 162-182.

[10] Li, W., Zheng, D., Zhao, T., & Yang, M. (2012). An effective approach to pedestrian detection in thermal imagery. In Natural Computation (ICNC), 2012 Eighth International Conference on IEEE, pp. 325-329.

[11] Li, Z., Wu, Q., Zhang, J., & Geers, G. (2011). SKRWM based descriptor for pedestrian detection in thermal images. Multimedia Signal Processing (MMSP), 2011 IEEE 13th International Workshop on IEEE. pp. 1-6.

[12] Askari, M., Asadi, M., Asilian, B. A., & Ebrahimpour, H. (2016). Isolated Persian/Arabic handwriting characters: Derivative projection profile features, implemented on GPUs. Journal of AI and Data Mining, vol. 4, pp. 9-17.

[13] Jeon, E. S., Choi, J. S., Lee, J. H., Shin, K. Y., Kim, Y. G., Le, T. T., & Park, K. R. (2015). Human detection based on the generation of a background image by using a far-infrared light camera. Sensors, vol. 15, no. 3, pp. 6763-6788.

[14] Zhao, X., He, Z., Zhang, S., & Liang, D. (2015). Robust pedestrian detection in thermal infrared imagery using a shape distribution histogram feature and modified sparse representation classification. Pattern Recognition, vol. 48, no. 6, pp. 1947-1960.

[15] Li, J., Gong, W., Li, W., & Liu, X. (2010). Robust pedestrian detection in thermal infrared imagery using the wavelet transform. Infrared Physics & Technology, vol. 53, no. 4, pp. 267-273.

[16] Yang, C., Liu, H., Liao, S., & Wang, S. (2015). Pedestrian Detection in Thermal Infrared Image Using Extreme Learning Machine. In Proceedings of ELM, Springer, Cham. vol. 2, pp. 31-40.

[17] Olmeda, D., de la Escalera, A., & Armingol, J. M. (2012). Contrast invariant features for human detection in far infrared images. In Intelligent Vehicles Symposium (IV), 2012 IEEE, pp. 117-122.

[18] Apatean, A., Rogozan, A., & Bensrhair, A. (2016). Image Features Extraction, Selection and Fusion for Computer Vision. In Image Feature Detectors and Descriptors, Springer, Cham, pp. 75-107.

[19] Davis, J. W., & Keck, M. A. (2005). A two-stage template approach to person detection in thermal imagery. In Application of Computer Vision, 2005. WACV/MOTIONS'05 Volume 1. Seventh IEEE Workshops on IEEE, Vol. 1, pp. 364-369.

[20] Zhang, L., Wu, B., & Nevatia, R. (2007). Pedestrian detection in infrared images based on local shape features. In Computer Vision and Pattern Recognition, CVPR'07, IEEE Conference, pp. 1-8.

[21] Gonzalez, R. C. (2002). Richard E. woods. Digital image processing.

[22] Che, Z. G., Chiang, T. A., & Che, Z. H. (2011). Feed-forward neural networks training: a comparison between genetic algorithm and back-propagation learning algorithm. International Journal of Innovative Computing, Information and Control, Vol. 7, no. 10, pp. 5839-5850.

[23] Hagan, M. T., Demuth, H. B., & Beale, M. H. (2014). Orlando De Jesus. Neural Network Design, 2nd Edition, Cengage Learning.

[24] Selesnick, I. W. (2004). The double-density dual-tree DWT. IEEE Transactions on signal processing, vol. 52, no.5, pp. 1304-1314.

[25] Dai, C., Zheng, Y., & Li, X. (2007). Pedestrian detection and tracking in infrared imagery using shape and appearance. Computer Vision and Image Understanding, vol. 106, no. 2-3, pp. 288-299.

[26] Suard, F., Rakotomamonjy, A., Bensrhair, A., & Broggi, A. (2006). Pedestrian detection using infrared images and histograms of oriented gradients. In Intelligent Vehicles Symposium, 2006 IEEE, pp. 206-212.

[27] Cao, Jiuwen, Kezhi Mao, Jonathan Wu, and Amaury Lendasse. (2015). Proceedings of ELM-2015, Springer, Vol. 2.

[28] Cheng, M. M., Zhang, Z., Lin, W. Y., & Torr, P. (2014). BING: Binarized normed gradients for objectness estimation at 300fps. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3286-3293.

[29] OTCBVS Benchmark Dataset Collection. Receiving date of the database 1: January 16, 2017, <<http://vcipl-okstate.org/pbvs/bench/Data/01/download.html>>.

[30] OTCBVS Benchmark Dataset Collection. Receiving date of the database 3: January 16, 2017, <<http://vcipl-okstate.org/pbvs/bench/Data/03/download.html>>.

شناسایی رهگذران در تصاویر حرارتی فضای باز بر اساس تخمین شرایط آب و هوایی

سید مرتضی غزالی، یاسر بالغی*

[1] دانشگاه صنعتی نوشیروانی بابل، دانشکده مهندسی برق و کامپیوتر، خیابان شریعتی، بابل، ایران

ارسال ۲۰۱۷/۰۶/۱۳؛ بازنگری ۲۰۱۷/۱۱/۰۴؛ پذیرش ۲۰۱۸/۰۱/۰۶

چکیده:

نظارت در تاریکی مطلق و طول روز در تمامی شرایط آب و هوایی یکی از مزیت‌های سیستم‌های تصویربرداری حرارتی است. با وجود گستردگی استفاده از این دوربین‌ها، مشکلات زیادی در خصوص تحلیل تصاویر حرارتی به دلیل ماهیت متغیر ظاهر و نمود رهگذران و متغیر بودن شرایط جوی وجود دارد. در این مقاله الگوریتمی کارا جهت شناسایی رهگذران در تصاویر حرارتی فضای باز، سازگار با شرایط آب و هوایی گوناگون ارائه شده است. در ابتدا نوع شرایط آب و هوایی با استخراج ویژگی مناسب تخمین زده می‌شود. سپس الگوریتمی مرتبط با شرایط آب و هوایی تشخیص داده شده برای شناسایی رهگذران اجرا می‌گردد. برای این منظور تصاویر حرارتی به سه دسته شرایط آب و هوایی، مناسب مانند روزهای آفتابی، نامناسب مانند بارانی یا مه آلود و گرم مانند روزهای تابستان که رهگذران عموماً نسبت به تصویر پس‌زمینه تیره‌ترند، دسته‌بندی می‌شوند. سپس از تصاویر ورودی آموزش در سه سطح تبدیل موجک مختلط دوگانه دودرختی دوبعدی گرفته می‌شود و انرژی ضرایب فرکانس پایین در سطح سوم به عنوان ویژگی‌ای متمایز کننده برای تشخیص نوع شرایط آب و هوایی محاسبه می‌شود. با ویژگی های بدست آمده شبکه عصبی مصنوعی پیش خور برای تعیین دسته شرایط آب و هوایی آموزش داده می‌شود. در پایان الگوریتمی مرتبط با نوع دسته شرایط آب و هوایی برای تشخیص رهگذر به کار گرفته می‌شود. روش پیشنهادی در تشخیص رهگذران از عملکرد بالایی برخوردار بوده به طوری که دقت شناسایی رهگذران بر روی ۲ مجموعه داده معروف حرارتی بیش از ۹۹٪ است.

کلمات کلیدی: تصاویر حرارتی فضای باز، شرایط آب و هوایی، کلاس‌بند شبکه عصبی مصنوعی، تبدیل موجک، شناسایی عابرین.