ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

A Multi-Feature and Machine Learning Graded Pedestrian Detection Method

Ning Gao

Ningbo Institute of Technology, Zhejiang University, Zhejiang, China

n@gao.us

Abstract

Pedestrian detection is a hot topic in computer vision and has a wide application prospect in robotics, unmanned driving, virtual reality technology and some military fields. Pedestrians are in the most vulnerable position in the whole traffic system and are the most vulnerable people in traffic accidents. Data show that the number of casualties among pedestrians in traffic accidents is still high. In the field of computer vision and pattern recognition, a very important task is to make computers analyze images intelligently and understand images like people. If we can make the vehicle more intelligent and use pedestrian detection as a driver's assistant system, we can automatically detect the pedestrians in front of the vehicle. In this paper, based on artificial intelligence technology, multi-feature and machine learning hierarchical pedestrian detection methods are studied. It is expected that useful information can be obtained through uncertainty and machine learning to achieve a high degree of unity of theory and application, so as to better serve the development of artificial intelligence.

Keywords

Pedestrian Detection, Intelligence, Traffic, Artificial Intelligence.

1. Introduction

Pedestrian detection is to segment the pedestrians in video sequences or images from the background and locate them accurately. It is one of the most active research topics in the field of computer vision at present [1]. While enjoying the convenience, comfort and quickness brought about by road development, it also bears the negative impact brought by it. In recent years, with the continuous development of science and technology, a large number of achievements have emerged in the research on biological feature recognition of human body [2]. In the field of computer vision and pattern recognition, a very important task is to make computers analyze images intelligently like people and understand images [3]. Behavior-based feature recognition, such as gait and voice recognition, has the characteristics of easy detection without target cooperation compared with physiological feature recognition [4]. In the field of intelligent transportation, pedestrian detection can reduce the driver's visual blind zone. Through real-time detection, pedestrians in the scene can be predicted, the safety of driving vehicles can be improved, and traffic accidents can be further reduced. Pedestrian detection has broad application prospects in the fields of automatic driving, intelligent security, virtual reality, and motion analysis.

Vision is the main way for humans to obtain outside information. Through vision, humans perceive the brightness, color, movement and size of external things. Traditional monitoring is simply to collect images, and it is necessary to rely on the administrator's real-time monitoring to deal with emergencies in a timely manner [5]. Pedestrian detection can be applied to the field of intelligent security by detecting, identifying and alarming pedestrians in the surveillance scene [6]. Suspicious

ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

people can be discovered in time, effectively preventing crimes and anomalies. If we can make the vehicle more intelligent and use pedestrian detection as the driver's auxiliary system, then pedestrians in front of the vehicle can be automatically detected [7]. The assisted driving system based on pedestrian detection technology can automatically analyze the traffic conditions on the road and detect pedestrians and vehicles. In view of the advantages and disadvantages of various sensors, practical applications often complement each other through information fusion of various sensors, thereby effectively improving the comprehensive performance of the pedestrian detection system [8]. If we want future computers to communicate freely, this requires computers to have a strong ability to perceive the external environment and to extract important information from the environment independently. Based on artificial intelligence technology, this paper studies multi-features and machine learning hierarchical pedestrian detection methods. Achieve a high degree of unity of theory and application to better serve the development of artificial intelligence.

2. Human Posture Recognition

2.1 Attitude and Feature Generation

The main focus of pedestrian detection research is the extraction of time features. By improving the feature extraction algorithm, the extracted features have the strongest representation ability as possible, which is the most important means of the technological progress of pedestrian detection. Dendrites connected to many other neuronal axons receive signals from transmitters, and nerve cells use some method to superimpose signals from all dendrites. The risk caused by uncertainty is undeniable. If we want to comprehensively analyze and deal with the uncertain factors in data, we must correctly understand and grasp it. A weak classifier is trained on the training data, the error rate of the classifier is calculated, and then the weak classifier is trained again on the same data set. Articulated human pose recognition in images is very difficult in computer vision, and the algorithms used must deal with a lot of data. The human brain has unsupervised learning ability. When a neuron cell receives frequent stimulation within a period of time, it is more likely to be excited when stimulated again. The multi-site method decomposes the changes of the human body into the changes of the appearance of each site and the changes of the relative positions between each site, and considers these two factors separately.

In the field of machine vision, there may be several scene changes in a video. For a fixed scene, the images in the scene may evolve continuously. If there is a moving object, that is, there is relative motion between the moving object and the background image. Then the velocity vector formed by the moving target must be different from the velocity vector of the neighboring background. There is a certain gap in intelligence, especially limited by many factors in practical application, and further improvement is needed. Fig. 1 is a frame of a motion detection and tracking system using dense disparity variance technology.

The intensive searching of images by sliding window will also produce a large number of windows, which not only increases the calculation amount, but also makes the detection performance of the system very sensitive to the detection accuracy of the classifier. Feature recognition based on time series is different from general traditional feature extraction. For general feature extraction methods, the order is not taken into account, and the relationship between its variables is independent of each other [9]. In view of the complicated detection results of rectangular sliding windows and square sliding windows at all levels, it is necessary to fuse them, and a window fusion algorithm is proposed. The sample library can be rapidly and effectively expanded according to the actual detection conditions, such as customizing a relatively close sample library according to different backgrounds and even different feature extraction methods. The number of hidden layers can be defined according to the model of the network. The number of hidden nodes is also uncertain, but usually the larger the number, the stronger the neural network is. The reason why machine learning has become a key part of artificial intelligence is that it gives computers more advanced functions, enabling computers to simulate or realize the learning process of human brain.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

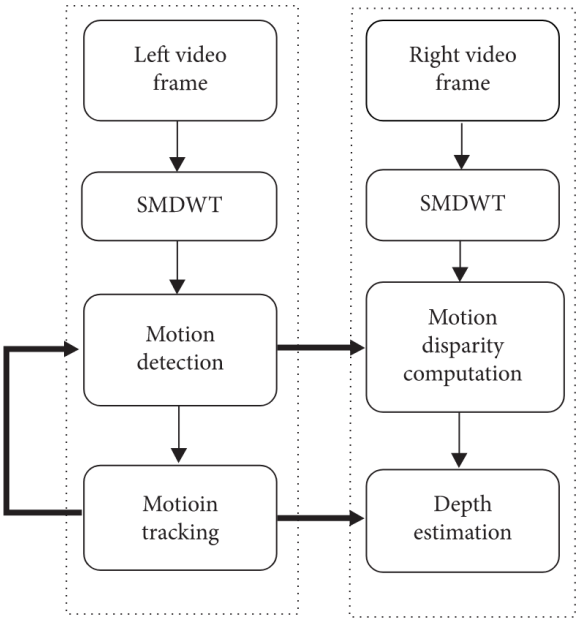

Fig. 1 Motion detection and tracking system framework using dense parallax variance technique

2.2 Data Set and Training

The measurement standard of pedestrian detection is difficult to draw up, because there are too many factors that affect the performance of pedestrian detection. The recognition of human body posture requires a lot of calculation. Before the appearance of deep neural network, most recognition methods were based on graphic structure, which transformed human body model into a set of rigid templates and a set of paired models in the form of tree structure. Tracking the vehicle trajectory is the premise of accurately acquiring the moving target features and realizing the detection of abnormal traffic events, while the detection and tracking of the moving vehicle is the most critical part in the process of acquiring the vehicle trajectory. The difficulty of detecting samples, the advantages and disadvantages of feature extraction algorithms, the excellent classification techniques and the design of scale adaptation algorithms have direct and significant effects on the detection results. Feature extraction technology is the most basic and key technology in the whole pedestrian detection system. Many times we know that some features will have certain value, but each feature alone will not produce very good results. For the technical performance of feature extraction, attention is usually paid to the detection performance, but the dimension of extracted feature vectors and the detection time of the system are also very important performance inspection points.

According to the location of the target centroid, it is necessary to control the tripod head to drive the camera to rotate. In order to change the field of view, the target is always within the field of view. The necessary action points can be quickly located. Table 1 shows the detection rate when pedestrians are not blocked. Table 2 shows the detection rate with occlusion.

Table 1 Detection rate of pedestrians without occlusion

	Detection rate (%)	Discrimination rate (%)
Video 1	94.7	82.8
Video 2	93.2	77.4

Table 2 Detection rate of pedestrian occlusion

	Detection rate (%)	Discrimination rate (%)
Video 1	80.5	77.8
Video 2	78.6	73.3

ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

When a moving object moves, the difference in gray value of the pixel between adjacent frames increases. Randomly select input and output data and submit it to the network. Calculate the output of each neuron in the hidden layer:

$$f(t)=\sum_{j=1}^{N}\sum_{k\in Z}d_k^j\phi_{jk}(t)+\sum_{k\in Z}c_k^N\phi_{Nk}(t) \tag{1}$$

Calculate the response of the output layer neurons:

$$E_{mi}=\sum_{i=1}^{k}(i\Delta t)\cdot\left|S_{mi}\right|^2 \tag{2}$$

Calculate the error of the output layer neurons using the given output data:

$$\vec{E}=\frac{E_{mi}}{\sqrt{\sum_{i=1}^{k}E_{mi}}} \tag{3}$$

For the same training set, the algorithm can train different weak classifiers, which can form a strong classifier. Deep learning ingeniously constructs a machine learning model with multiple hidden layers. Through massive training data, the machine can automatically learn good features without the process of manually selecting features, thus obtaining the essential features of the data set. Under the condition that the edge of pedestrians is not so obvious, missed detection may occur. Moreover, the algorithm only uses the intensity information of the edge and ignores the angle information. For the number of nodes in the hidden layer, if the number of nodes is too small, the network model can only extract a small amount of information, and the mapping ability is poor. The network consists of visual layer and hidden layer. Layers and layers are interconnected, but there is no connection between units in the layer. The advantage of regularization in each batch of training is that it does not need to deal with the initialization of the network in a special way, and it can use a higher learning rate to achieve the purpose of fast convergence. In order to deal with the pedestrians with different sizes in the image, a pyramid method is used to detect the pedestrians with different sizes of sliding windows, and finally the windows at all levels are coincided.

3. Performance Analysis of Pedestrian Detection Model

In the training process, as the training difficulty increases, the error of the weak classifier selected in each round will gradually increase, while the weight of the weak classifier will correspondingly decrease. After determining the matching criteria, it is necessary to find the correct matching search method to estimate the moving target and continuously determine its trajectory. In pedestrian detection systems based on confusing moving windows, sliding windows extract many feature vectors, and the dimension of feature vectors has a great influence on detection speed. Generally speaking, the first few classifiers are only responsible for excluding simple negative samples, and often only a small number of features are needed to participate in the calculation [10]. In order to remove noise, gray-scale images are smoothed by discrete Walsh smoothing templates at different smoothing scales before image processing. After multiple depths of features are superimposed, their feature information will be integrated together, and chromaticity information can be integrated into one dimension. The accuracy in the training set is very high, but the accuracy in the test set is not high. When the aspect ratio is greater than a certain threshold, or the width value is smaller than a preset threshold, or the area is within a threshold range, pedestrians are considered to be present.

When recommending sample markers, the imbalance of positive abnormal sample distribution is fully considered, abnormal uncertain samples are selected by active learning, and the detection model is continuously updated iteratively through expert interaction. If the pixel features or pixel areas at corresponding positions have certain differences, it is considered that the pixel points or pixel areas at these positions in the video image of the frame have changed. Then these change regions are

ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

extracted according to different judgment criteria to form the foreground target motion region. Fig. 2 is a human body height estimation model using a reference height

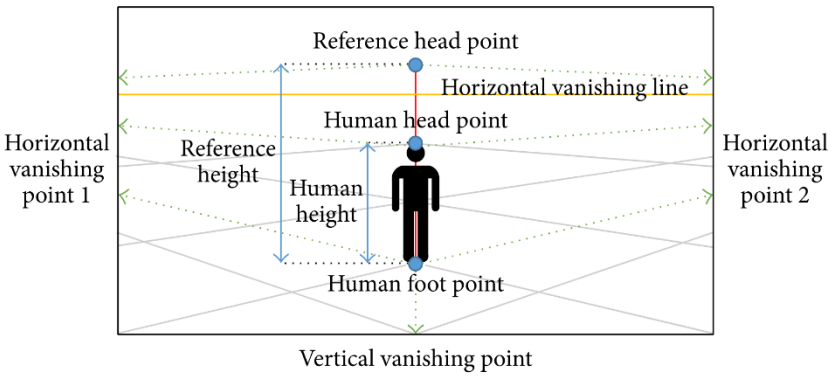

Fig. 2 Human height estimation model using reference height

If the neural network remembers the information of the illumination or chroma of the picture when it trains to recognize the number, then fitting may occur. Sexual reproduction is superior to asexual reproduction in natural selection, because natural selection may not be the healthiest gene. The preprocessing of data set labeling is the process of transforming the information of data set labeling points into heat map. Its purpose is to calculate the similarity between the heat map output by neural network and the standard heat map in the loss node. For multi-feature fusion of pedestrian detection, these three measures for each feature are predetermined. When new features appear, the allocation of basic credibility is continued, and the detection and recognition results and confidence of the current time are generated. In the system detection module, a sliding window is used to slide in the image, and the features of each window of the search are extracted and input into the classifier for detection. The core of machine learning is learning, that is, the system realizes the process from unknown to known through learning algorithms. When it encounters similar problems again, it can achieve more efficient processing. It mainly uses induction, synthesis rather than deduction.

4. Conclusion

With the development of machine learning and computer vision, pedestrian detection technology is bound to mature step by step. The detection of pedestrian targets in images through image processing technology is a key technology in the field of computer vision, and has important application prospects in application fields such as intelligent monitoring and vehicle-assisted driving. As a part of computer vision, pedestrian detection has great social significance and commercial value. At the same time, due to its inherent technical difficulties. Although many detection algorithms have been used to detect pedestrian targets in actual monitoring in different occasions, there are not many pedestrian event detection algorithms that can be used in traffic scenes where pedestrians break in by mistake and interfere with normal traffic operation. Pedestrian detection is an important branch of computer vision, which has great social significance and commercial value and attracts the attention of many researchers. This paper takes pedestrian detection as the background, expecting to obtain useful information through uncertainty and machine learning, to achieve a high degree of unity of theory and application, so as to better serve the development of artificial intelligence. In the future, pedestrian detection technology will be more mature, better applied to intelligent monitoring, intelligent driving and other occasions, for the benefit of all mankind.

References

[1]. Ma Y, Lee E W M, Yuen R K K, et al. An Artificial Intelligence-Based Approach for Simulating Pedestrian Movement[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(11):3159-3170.

[2]. Htike K K, Hogg D. Adapting pedestrian detectors to new domains: A comprehensive review[J]. Engineering Applications of Artificial Intelligence, 2016, 50:142-158.

ISSN: 2414-1895 DOI: 10.6919/ICJE.201911_5(12).0008

[3]. Xia D X, Su S Z, Geng L C, et al. Learning rich features from objectness estimation for human lying-pose detection[J]. Multimedia Systems, 2017, 23(4):515-526.
[4]. Bila C, Sivrikaya F, Khan M A, et al. Vehicles of the Future: A Survey of Research on Safety Issues[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, PP(99):1-20.
[5]. Baltieri D, Vezzani R, Cucchiara R. Mapping Appearance Descriptors on 3D Body Models for People Re-identification[J]. International Journal of Computer Vision, 2015, 111(3):345-364.
[6]. Booth J, Eugenio B D, Cruz I F, et al. Robust Natural Language Processing for Urban Trip Planning[J]. Applied Artificial Intelligence, 2015, 29(9):859-903.
[7]. Glauner P, Meira J A, Valtchev P, et al. The Challenge of Non-Technical Loss Detection Using Artificial Intelligence: A Survey[J]. International Journal of Computational Intelligence Systems, 2017, 10(1):760-775.
[8]. Keating J, Nourbakhsh I. Teaching artificial intelligence and humanity[J]. Communications of the ACM, 2018, 61(2):29-32.
[9]. Liu P, Ni H, Fu B, et al. Detection of partially occluded pedestrians by an enhanced cascade detector[J]. IET Intelligent Transport Systems, 2014, 8(7):621-630.
[10].Fricoteaux L, Thouvenin I, Mestre D. GULLIVER: A decision-making system based on user observation for an adaptive training in informed virtual environments[J]. Engineering Applications of Artificial Intelligence, 2014, 33:47-57.