HUMAN POSE ESTIMATION

GROUP 1

Scrum Master-Anupa Philip

Team Leader-Aarathy Sasi

Team Members-Anju Thampi,Ida phiyus,Jiya P.S.,Rosemol Augustine

Introduction

Analysis of human behaviour from image sequences has attracted much attention for decades because it is very useful and helpful to have a system automatically observing human actions and behaviours via CCTV in many places. For example, detecting theft actions in shop store, alarms for intruders at home, detecting falls in hospital and alerting security guards for harmful actions in train station. Two main keys of this technology is to track configuration of human body along the image sequence and to analyse them through knowledge provided by experts for each applications.While the configurations of interest depends on applications concerned, position and orientation are basic configurations necessary for further analysis.

Although there have been many proposed techniques on human detection and pose estimation to extract location and postures, it is noticed that there have been very few concerning orientation. Estimation of human orientation is considered a key to study and understand human behaviours and there have been many works focusing on doing so. However, most of previous works are sensor-based technologies but visual-based. Based on our knowledges, there are not many visual-based techniques for estimating orientation of human. Here we categorise techniques of orientation estimation in computer vision into two groups,feature-based and silhouette-based. While the feature- based techniques focus on estimating orientation degree from given image features, silhouette-based techniques estimate from foreground pixels obtained from background subtraction process.

Abstract

Estimation of human orientation is considered a key configuration to study and understand human behaviours. There have been many works focusing on this. However, most of previous works are sensor-based technologies but visual-based technologies. In this paper, we proposed a method to estimate the orientation of human in a given image patch. We have observed the performance of a feature called Histogram of Oriented Gradients(HOG) on human detection.The results show that the proposed method can effectively estimate the orientation of human pose inside an image with fast and simple additional operations required.

Histogram Of Oriented Gradient

The Histogram of Oriented Gradients(HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms,scale-invariant feature transformdescriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

Calculating the HOG feature descriptor.

Step 1 : Preprocessing
Step 2 : Calculate the Gradient Images
Step 3 : Calculate Histogram of Gradients in 88 cells
Step 4 : 1616 Block Normalization
Step 5 : Calculate the HOG feature vector

The essential thought behind the histogram of oriented gradients descriptor is that local object appearance and shape within an image can be described by the distribution of intensity gradients or edge directions. The image is divided into small connected regions called cells, and for the pixels within each cell, a histogram of gradient directions is compiled. The descriptor is the concatenation of these histograms. For improved accuracy, the local histograms can be contrast-normalized by calculating a measure of the intensity across a larger region of the image, called a block, and then using this value to normalize all cells within the block. This normalization results in better invariance to changes in illumination and shadowing.The HOG descriptor has a few key advantages over other descriptors. Since it operates on local cells, it is invariant to geometric and photometric transformations, except for object orientation. Such changes would only appear in larger spatial regions.

Algorithm Used

SVM

The algorithm used in this project is SVM(support Vector machine).Support Vector Machine (SVM) is a supervised machine learning algorithm which can be used for both classification or regression challenges. However,it is mostly usedinclassification problems.In this algorithm, we ploteach data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, weperform classification by finding the hyper-plane thatdifferentiate the two classes very well.maximizing the distances between nearest data point (either class) and hyper-plane willhelp us to decide the right hyper-plane. This distance is called as Margin. We need to identify the correct hyperplane having highest margin then only classification can be done correctly,If we choose the hyperplan having low margin their there is chance of occuring miss-classification.

Python, scikit-learn is a widely usedlibrary for implementing machine learning algorithms, SVM is also available in scikit-learn library and follow the same structure.In this project For detecting human in an image, Support Vector Machine (SVM) is introduced to handle the task of human/non-human classification in each detection window by training SVM with human and non-human.

Dataset

The dataset used in this project is INRIA person.

1. Contributions

- Images from personal digital image collections taken over a long time period. Usually the original positive images were of very high resolution (approx. 2592x1944 pixels), so we have cropped these images to highlight persons. Many people are bystanders taken from the backgrounds of these input photos, so ideally there is no particular bias in their pose.
- Few of images are taken from the web using google images.

Note

- Only upright persons (with person height greater than 100) are marked in each image

2. Original Images
   - Images Folders 'Train' and 'Test' correspond, respectively, to original training and test images. Both folders have two sub folders:
     (a) 'pos' (positive training or test images)
     (b) 'neg' (negative training or test images.)
3. Normalized
   - Normalized Images Folders 'train-64x128-H96' and 'test-64x128-H96' correspond to normalized dataset as used in above referenced paper. Both folders have two sub folders:
     (a) 'pos' (normalized positive training or test images centered on the person with their left-right reflections)
     (b) 'neg' (containing original negative training or test images).

Note

- images in folder 'train/pos' are of 96x160 pixels (a margin of 16 pixels around each side), and images in folder 'test/pos' are of 70x134 pixels (a margin of 3 pixels around each side). This has been done to avoid boundary conditions (thus to avoid any particular bias in the classifier). In both folders, use the centered 64x128 pixels window for original detection task.

Tasks

Task-1

Our first task was Human Pose Estimation in which detection of humans in images have been long-standing problems in computer vision. Most successful approaches heavily rely on discriminative models to build appearance detectors for body joints and generative models to constrain possible body configurations. We address the problem of estimating 2d human pose from still real world images. The performance of a feature called Histogram of Oriented Gradients (HOG)on human detection was studied and used for this Estimation. We mostly used Modified HOG-based human detector to calculate the score of confidence of the particular images. After detecting the position there were occurrence of a straight line along the body of image.

Task-2

In our second task of Human Pose Estimation we needed to detection of humans in web-cam by capture of humans in widely. It will be real time detection through web-cam to understand the human pose by using feature caption. While detection of human, we must avoid the background by blurred images.In this we also apply modified HOG-based human detector to calculate the score of confidence. After detecting the position there were occurrence of a straight line along the body of image.

Role of Members

1. Anupa philp : Scrum Master
   - Organises Daily scrum meetings
   - Worked behind the program code.
2. Aarathy: Team Leader
   - Assigned different tasks to the team members after dividing it into sub tasks,
   - Worked behind the program code to detect the humans.
3. Anju Thambi
   - Worked behind the code to drawn a line in human sternum.

4. Rosemole Augustine
   - Worked behind the latex report.
   - Worked behind the program code to detect the humans.
5. Jiya P.S
   - Worked behind the latex and beamer reports.
   - Worked behind the code to drawn a line in human sternum.
6. Ida
   - Worked behind the program code and searching more about different algorithms .

Proof of Concept

14-11 MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

Estimation of Human Body Orientation using Histogram of Oriented Gradients

Kittipanya-ngam Panachit
Institute for Infocomm Research
1 Fusionpolis Way
Singapore, 138632
pngam@i2r.a-star.edu.sg

Ong Soh Guat
Institute for Infocomm Research
1 Fusionpolis Way
Singapore, 138632
sgong@i2r.a-star.edu.sg

Eng How-Lung
Institute for Infocomm Research
1 Fusionpolis Way
Singapore, 138632
hleng@i2r.a-star.edu.sg

Abstract

Estimation of human orientation is considered a key configuration to study and understand human behaviours. There have been many works focusing on this. However, most of previous works are sensor-based technologies but visual-based technologies. In this paper, we proposed a method to estimate the orientation of human in a given image patch. We have observed the performance of a feature called Histogram of Oriented Gradients (HOG) on human detection and found the sensitivity of this feature in in-plane rotation. It is so sensitive that it could be used in classifying images of rotated and non-rotated human body. Therefore we have drawn an idea to exploit a modified HOG descriptor with 'bin shifting' technique to estimate the orientation of human. The results show that the proposed method can effectively estimate the orientation of human pose inside an image with fast and simple additional operations required.

1 Introduction

Analysis of human behaviour from image sequences has attracted much attention for decades because it is very useful and helpful to have a system automatically observing human actions and behaviours via CCTV in many places. For example, detecting theft actions in shop store, alarms for intruders at home, detecting falls in hospital and alerting security guards for harmful actions in train station. Two main keys of this technology is to track configuration of human body along the image sequence and to analyse them through knowledge provided by experts for each applications. While the configurations of interest depends on applications concerned, position and orientation are basic configurations necessary for further analysis. Although there have been many proposed techniques on human detection and pose estimation to extract location and

Estimation of human orientation is considered a key to study and understand human behaviours and there have been many works focusing on doing so. However, most of previous works are sensor-based technologies[1] but visual-based. Based on our knowledges, there are not many visual-based techniques for estimating orientation of human. Here we categorise techniques of orientation estimation in computer vision into two groups, feature-based and silhouette-based. While the feature-based techniques focus on estimating orientation degree from given image features, silhouette-based techniques estimate from foreground pixels obtained from background subtraction process[2, 3]. Iwasawa *et al.*, [2] track the change of the principal axis of the foreground area obtained from background subtraction in Thermal images. Lee *et al.*, [3] estimate ellipse parameters covering foreground image which contains orientation information. Bay *et al.*,[4] proposed to estimate the main orientation of the target from wavelet features. Chen *et al.*, [5] to estimate the orientation from Weber Local Descriptor.

Here we observed the performance of a feature called Histogram of Oriented Gradients (HOG) [6] on human detection and found the sensitivity of this feature in in-plane rotation. It is so sensitive that it could classify rotated and non-rotated human body. Therefore this paper suggests an idea to exploit a modified HOG descriptor to estimate the orientation of human in the given patch assumed there is a human inside. HOG is briefly described in section 2 and the proposed method is explained in section 3. The datasets used in this paper are introduced in section 4. Next, The study of Rotation Sensitivity and the evaluation of the proposed method are illustrated in section 5. Finally, the performance will be discussed and paper will be concluded in the section 6

2 Histogram of Oriented Gradients

Conclusion

The need for better image understanding software is increasing since it can be used in different fields such as medical etc. According to the application of this system, we are intending to apply images for the detection process, checking the position of sitting. Alert to the persons about their position if the position is not correct.

Reference

- IEEE PAPER-Estimation of Human Body Orientation using Histogram of Oriented Gradients