People counting based on video processing using histogram of oriented gradient method and support vector machine

Rizqy Bukhari Aziz[1], Primawan Satrio Bindono[2],
Bedy Purnama[3]
rbukharia@gmail.com[1],primawan.satrio@gmail.com[2],
bdp@ittelkom.ac.id[3]
Informatics Faculty, Telkom Institute of Technology

Abstract

Technological developments lately more rapid impact on its application in everyday life. One technology that is currently being one of the rapidly growing field of computer vision. Computer vision is a work computer that aims to model and replicate the human view using software or hardware.

One of the applications being developed, among others, the introduction of biometric hand gesture pose with the goal of identifying what is done by hand in the image are detected according to research by Takashi Abe of Tohoku University, and face recognition to recognize human faces by calculating the weights on each each sub-regionnya (Thanh-Toan DO and Ewa KIJAK, Universit'e de Rennes).

In this paper the author will discuss human detection to counter IT Telkom number of library visitors using the histogram of oriented gradients and a support vector machine.

By knowing the number of visitors and the level density is expected to record a daily visitor to accurately without human error and set the number of visitors so as not to exceed the capacity because it will reduce the comfort while in the library.

Kata Kunci: people counting, HOG, SVM, video processing

1. Pendahuluan

All human work was originally done by hand, began to be changed automatically to facilitate work and reduce the factors that influence both technical and non-technical.

One technology that is currently being one of the rapidly growing field of computer vision. Computer vision [1] is a work computer that aims to model and replicate the human view using software or hardware. Computer vision (image) is the study of reconstruct, interpret, and understand 3D objects from 2D images obtained in terms of the nature of the visible landscape. Computer vision is a combination of image processing and pattern recognition. Image processing is a field that relates to the transformation of the image / images, while the introduction of a pattern associated with the identification of objects or image interpretation.

One of the applications being developed, among others, the introduction of biometric hand gesture pose with the goal of identifying what is done by hand in the image are detected according to research by Takashi Abe of Tohoku University [2], and face recognition to recognize human faces by counting weight to each sub-regionnya (Thanh-Toan DO and Ewa KIJAK, Universit'e de Rennes) [3].

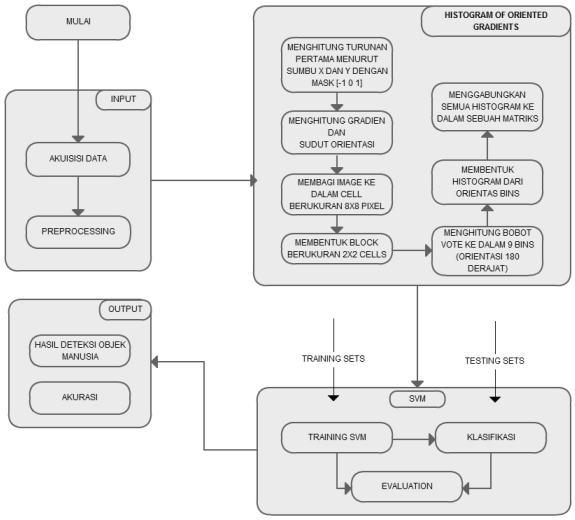
In this paper the author will discuss human detection to counter IT Telkom number of library visitors using the histogram of oriented gradients and a support vector machine.

The reason the authors take the case of calculating the number of visitors to the library is used as the system is still vulnerable manually by human error, calculation error is often a problem in merekap number of visitors each specified period. The method used is a pretty good method to detect and quantify the crowd even in running condition. Shape and appearance of the object being detected by the histograms of oriented gradient method and support vector machine can be well characterized by the distribution of local intensity gradients or edge directions, even without the knowledge of the apparent location of the gradient and edges.

By knowing the number of visitors and the level density is expected to record a daily visitor to accurately without human error and set the number of visitors so as not to exceed the capacity because it will reduce the comfort while in the library.

2. Pengenalan Citra

In general, the activities carried on image recognition is done by stages of digital imagepreprocessing, learning, and detection recognition.



Gambar 1 General Scheme

2.1 Design and analysis

2.1.1 Pre-processing

At the beginning of this process will be a change preprocessing data types using the grayscale image that there are only two colors: black and white without the need to ignore the elements of other colors (RGB) and using gamma compression to reduce light intensity. This process aims to make it easier to implement the method in this case in the process of learning and detection.

2.1.2 Learning

The following are the steps in applying this method:

1) The image is used for training data is obtained with the image that captures video recordings. Training data are divided into 2 types, namely positive and negative training data, training data is a positive image of the results captured video footage of the library IT Telkom with good quality. While the negative training data is capture the video image of the crop with random quality.

2) the positive training data will be prepared with the same size, while negative training data in-crop

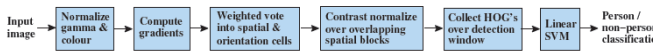
of the original image so that the same size as the positive training data

3) After preparing all the training data, the initial process will perform the function of HOG descriptors with the output generated in the form of vectors of descriptor

4) Using a support vector machine for the classifier. Support vector machine learning is a method that can resolve the problem of multiclass classification very much data quickly

5) negative training data will be processed in the learning phase, if the results obtained false positives were detected wrong but the reality is it will be in retraining

6) The results obtained on vector learning process will be saved and used for the detection process.
This process can be seen in the figure below,



Gambar 2 Learning process

2.1.3 Deteksi

Phase detection aims to detect human objects contained in the image into data testing. Before entering the detection phase, going through the training data preprocessing stage to produce the data in the form of a histogram. From this histogram will be matched with all the results of the training data, will determine the degree of similarity object is an object that is expected or not. The system will give output as the number of objects detected in the human trial data.

2.2 Feature Extraction

2.2.1 R-HOG

Input:

- Input image I
- HOG descriptor parameters
  1) Cell = n x n Pixels
  2) Block = ς x ς cells
  3) Orientetation bins (β)
  4) Overlap (o)

Output: feature vector F = $\{f_1, f_2, \dots, f_k\}$,

Where, k = number of block, size(f_i) = ςςβ

Step-by-step calculation:

- Calculate the first derivative of the image I according to the x axis and y axis
  1) $I_x = h_x * I$ $\quad h_x = [-1\ 0\ 1]$
  2) $I_y = h_y * I$ $\quad h_y = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$
- Calculate the value of the gradient of the image Im of Ix to Iy

$$I_m = \sqrt{(I_x)^2 + (I_y)^2}$$

- Calcualte the value of gradient orientation (angle) I_θ from $I_x$till I_y

$$I_\theta = tan^{-1}\left(I_y / I_x\right) \text{ on range } [0,\pi]$$

- For all *Block* $b_1, b_2, \dots, b_k$:
  - Calculate the weight of the orientation histogram
  - Define the Gaussian window g (bi) with σ = 0.5 * ςn
  - Apply Gaussian window for each block to determine the weight

$$W_i = I_m^{b_i} * g(b_i)$$

  - For each pixel, use weights to calculate β Wi-bin orientation histogram using trilinear interpolation
  - Calculate vi in the above step with the size of ς * ς * β

- Apply L2-norm normalization in vi

  - L2-norm, $f_i = v_i/\sqrt{\|v_i\|^2}$

F = {f_1,f_2, … , f_k} was a HOG feature vector

2.2.2 Support Vector Machine (SVM)

Algoritma training SVM:

Input: x and y as data training, α ← 0 or α ← SVM that has been trained a few

1: C ← some value (10 for example)

2: **repeat**

3: **for all** {xi,yi},{ xj,yj } do

4: Optimize αi and αj

5: **end for**

6: **until** no changes in α or other resource constraint criteria met

Output: support vector (αi>0)

3. Result

In the testing process, the data is an image we made as many as 20 pieces, and 10 .avi formatted video. From the test results obtained the following results.

4. Conclusion

SVM and R-HOG is be able to detect human that included in images and video. As long as that video is in *avi format, and the frame could be read.

Daftar Pustaka

[1] Fisher, Robert., Dawson-Howe, Ken., Fitzgibbon, Andrew., Robertson, Craig., Trucco, Emanuele. 2005. Dictionary of Computer Vision and Image Processing. Wiley.

[2] Takashi, Abe., Okatani, Takayuki., Deguchi, Koichiro., Misra, Arindam. 2011. Hand Gesture Recognition using Histogram of Oriented Gradients and Partial Least Squares Regression. Tohoku University Japan.

[3] DO, Thanh-Toan., KIJAK, Ewa. FACE RECOGNITION USING CO-OCCURRENCE HISTOGRAMS OF ORIENTED GRADIENTS. Universit´e de Rennes 1 IRISA. Rennes, France.

[4] N. Dalal and B. Triggs. 2005. Histograms of Oriented Gradients for Human Detection. In CVPR, pages 886-893.

[5] Wuysang, Felix. 2006. Pembuatan aplikasi pendeteksian fingerprint dengan menggunakan metode gradient histogram. Universitas Kristen petra. jiunkpe-ns-s1-2006-26401190-9202-radient_histogram-abstract_toc.pdf

[6] Tsai, Grace. 2010. Histogram of oriented gradients. University of Michigan. 09_28_Grace.pdf

[7] Cecilia, Janice. 2012. Perancangan program penghitungan kepadatan pengunjung pada area pasar swalayan menggunakan metode histogram of oriented gradients. 2012-1-00547-mtif abstrk.pdf

[8] Changrampadi, Mohamed Hashim. 2011. A Fusion-based Multiclass AdaBoost for Classifying Object Poses Using Visual and IR Images. Chalmers Universit of Technology. Sweden.

[9] Rasmus Ulslev Pedersen. 2008. Module Support Vector Machine. nxtgcc.sourceforge.net/nxtgcc.pdf