Research Journal of Applied Sciences Engineering and Technology 4(16): 2639-2644, 2012
ISSN: 2040-7467

Submitted: December 02, 2011 Accepted: January 04, 2012 Published: August 15, 2012

# Moving Object Detection Based on the Histograms of Oriented Gradients and Cloud Model

[1]Yuyong Cui, [1]Zhiyuan Zeng, [1]Bitao Fu and [2]Weihong Cui
[1]Digital Engineering and Simulation Research Center, Huazhong University of Science and Technology, Wuhan, China
[2]School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China

**Abstract:** We present a Histograms of Oriented Gradients (HOG) features and the cloud model approach for the moving object detection from video sequences. Our model is based on the HOG features for the moving object and then uses the cloud model to find the moving object. First the HOG features are described from the image and then a HOG map is used by the cloud model to detect the moving object. The experiment shows our method is relative effect and has advantage in the detecting moving object.

**Keywords:** Cloud model, concept rising, HOG feature map, moving object detection

## INTRODUCTION

Moving object detection is one of the fundamental challenging work in computer vision, which are used in the video surveillance (Belongie and Malik, 2002), traffic control, medical diagnosis (Dryden and Mardia, 1998) and so on. In the paper, we concentrate on the HOG feature and the cloud model to represent and detection. We know the HOG feature is good at people or cars in a static image, The HOG descriptors are reminiscent of edge orientation histograms (Dalal and Triggs, 2005), SIFT descriptors (Lowe, 2004) and shape contexts (Belongie and Malik, 2002), but these descriptors use a dense grid of the uniformly spaced cells and compute overlapping local contract normalizations for improved effective result. A study of HOG in image is performed in our works.

The cloud model is proposed by Li and Du (2005), which is a model of uncertainty transition between a linguistic term of a qualitative concept and its numerical representation. Moving object detection and segmentation methods in recent years has been proposed, classical methods like optic flow (Zhang *et al*., 1993), frame difference and now the statistics method such as Markov random field method (Ren and Sun, 2005; Fu *et al*., 2008; Zabih *et al*., 1999) and another is the artificial Intelligence method as graph cut method (Jung *et al*., 2010), Automata Cellular and Binary tree and so on (Boykov and Kolmogorov, 2004; Cucchiara *et al*., 2003; Rother *et al*., 2004), from these method, each need more parameters than the cloud model, for the model only consider three numerical characters. So in our method we use the cloud model to consider the detection of the moving object (Boykov and Funka-Lea, 2006).

This study considers the cloud model based the HOG feature, which can be explained the spatial information and cloud model work well in information expressing. And our method seeks to decrease the detection time than other method based on HOG feature.

**Histogram of gradient feature:** The HOG feature is proposed in Dalal and Triggs (2005), it works well in image explanation we can let $\phi$ (x, y) and r(x, y) be the orientation and magnitude of the intensity gradient at a pixel (x, y) in an image. Like in Dalal and Triggs (2005), we compute gradients using finite difference filters, [-1, 0, 1] and its transpose, it has proved that the simple expression of the gradients can get a good result of the object. The color image is use the largest gradient magnitude to define $\phi$ and r at each pixel.

To get the gradient orientation at each pixel, a method is proposed with either a contrast sensitive (B1) or insensitive (B2) to be discredited into one of p values. Define it as follow:

$$B_1(x,y) = round\left(\frac{P\varphi(x,y)}{2\pi}\right) \bmod p \quad (1)$$

$$B_2(x,y) = round\left(\frac{p\varphi(x,y)}{\pi}\right) \bmod p \quad (2)$$

$B_1$ or $B_2$ always set as B for many works; a fundamental nonlinearity of the HOG based on pixel is to calculate a weighted vote for an edge orientation histogram channel on the orientation of the gradient element centered on it. Then the votes are accumulated

**Corresponding Author:** Yuyong Cui, Digital Engineering and Simulation Research Center, Huazhong University of Science and Technology, Wuhan, China

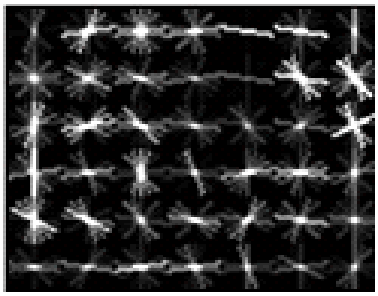


Fig. 1: A car of HOG feature

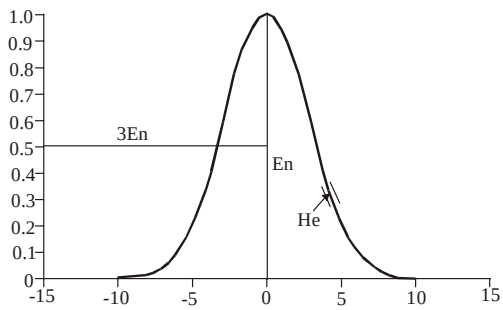



Fig. 2: Digital characteristic of cloud model

into orientation bins over local spatial regions that is called cells, cells is a rectangular or radial and the orientation bins is $B_1$ or $B_2$, vote function can be set the magnitude of the pixel, or its square, or a clipped form of the magnitude. Experiments show the gratitude itself can get a good result. So in our paper it will use this as the vote function. A block is set as 2×2 blocks and the block (Andriluka *et al*., 2009) is normalized for each of the above HOG geometries (Duetscher *et al*., 2000). Let v be the unnormalized vector, then L2-norm or L2-Hys, L1-norm, for L1-norm compute easily then it will used in our method. a car of the HOG feature is given in Fig. 1.

**The cloud model:** The cloud model is a uncertainty and linguistic contact with the quality method. Let U be the set U={u}, as the universe of discourse and T a linguistic term associated with U, $C_T(u)$ stand a tendency with the random number and it is from 0 to 1, it can be set as follow:

$$C_T(u): U \rightarrow [0,1], \forall u \in U, u \rightarrow C_T(u) \quad (3)$$

The cloud model is a one-point to multi-point transition, producing a membership cloud, rather than a membership curve such as the traditional fuzzy subordinate function which can be seen as one point to one point. The total shape of the cloud is visible, elastic and boundless; this is the name of cloud original reason.

A cloud always can be expressed by three digital characteristics: expected value Ex, entropy En and hyper entropy He.

Ex is the position at U corresponding to the gravity center of the cloud and it can most effectively show the qualitative concept. If the elements with the Ex value ,it is compatible with the qualitative concept, in other words, the element with Ex in the universe of discourse fully belong to the object represented by the cloud model.

En is the measure of the concept uncertainty and it is decided by the concept of fuzziness and randomness. This is a measure of the qualitative concept randomness, reflecting the bias of the concept. Another is shown the qualitative of other concept, for it is a domain in the space. The entropy is bigger, the concept is a lager numerical scale and the concept is fuzzier.

He, the entropy of Entropy (En) is a measure of the entropy uncertainty, it reflects the discrete degree of cloud drops. It can be decided by the fuzziness and randomness of En. He stands for the discrete degree of the cloud model and the random of the concept.

Ex, En, He is given and the number of the droplets, then a cloud model is given, A cloud model is shown in Fig. 2 and the cloud model with Ex = 0, En = 3 He = 0.1 and n = 5000.

**Our detection method:** A method based on the HOG feature and cloud model is proposed in the paper, in this section, the method can be descried as follow:

Given an image, the image can be used HOG describer change it as the HOG feature map. From the feature map, we can give the cloud model to classifier the feature map. The cloud model is based on the HOG feature map.

An image always we set a16×16 pixels as a block and 2×2 as a cells, the cell is based on the 9-bin orientation voted with the weight. Then a HOG feature map can be constructed. Then from the HOG feature map to consider the cloud model,

First, the feature map is seen as an image and from which we set a histogram of the feature:

**Step 1:** Calculate the histogram h(x) of the HOG feature.

**Step 2:** Find the maximum value of h(x) take its corresponding value as the $Ex_i$, of the current cloud model.

**Step 3:** Use the $h(Ex_i)$ as the current model's magnitude value.

**Step 4:** Suppose the $En_i$ of the cloud model and allowed error $\epsilon$.

**Step 5:** From the $Ex_i$, $En_i$ and the allowed error $\epsilon$ to calculate the $He_i$ .

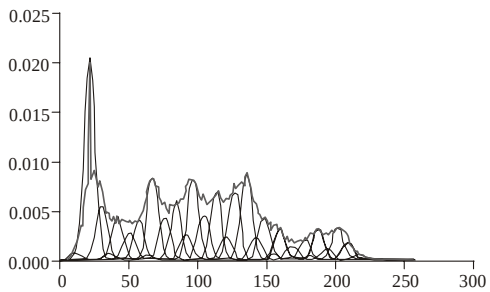
**Step 6:** Construct a cloud set and add the cloud ($Ex_i$, $En_i$, $He_i$).

**Step 7:** A residual error set as:

$$h_{res} = h(x) - h(Ex_i) * C(Ex_i, En_i) \quad (4)$$

**Step 8:** Repeat step 2 to 7 until the maximum of h(x) is less than the allowed error $\epsilon$.

(a)




(b)

Fig. 3: A is original image of Lina, b is the extracted concepts with $\epsilon = 0.005$ (red curve represents the original image histogram, the blue curves is the extracted concepts)

In step 7, C ($Ex_i$, $En_i$) is a function of calculating expected value of the cloud model, the allowed error $\epsilon$ is used to control the En and iteration time, by the user defined. And it can be seen that the more the cloud model construct, the more accurate the model is. A example of the concepts from the image is shown in Fig. 3.

**Concept rising:** From the series of atomic cloud, because from our method, we only consider two things, one is the background and the other is the foreground (car, person, animals and so on). cloud model is used to upgrade the concept gradually and generate higher level concepts concept rising is a hierarchical clustering process based the information. And multi-scale is reflected in different level of concepts and it can be realized as follows:

First, find two closed cloud models and the closed can be defined as follow:

$$\text{Min D } (Ex_i, Ex_{i+1}) = |Ex_{i+1} - Ex_i| \tag{5}$$

Then, synthesize the two similar concepts to one high level concept. The merge criteria is calculated from the three formulas:

$$Ex = \frac{Ex_1 En_1' + Ex_2 En_2'}{En_1' + En_2'} \tag{6}$$

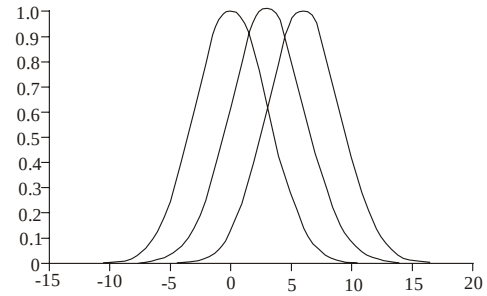



Fig. 4: Three concept result of rising.

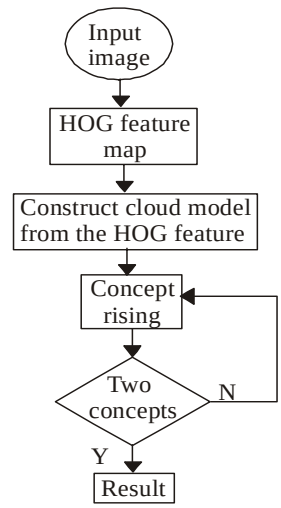



Fig. 5: The flow of the method

$$En = En_1' + En_2' \tag{7}$$

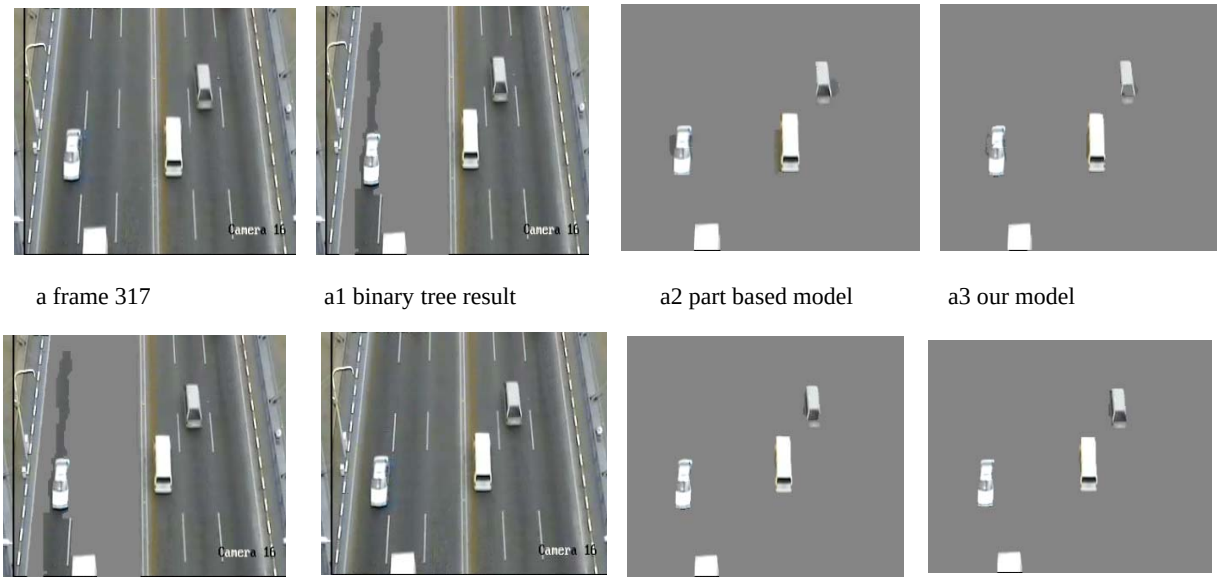
$$He = \frac{He_1 En_1' + He_2 En_2'}{En_1' + En_2'} \tag{8}$$

A simple example is shown in Fig. 4. When we get the new concept and the other non-synthesized concepts to a new layer, from the concepts which represent object, the highest degree of membership is defined as the analysis result of its relevant object.

In our method, the HOG feature as the input of our method, from the HOG feature, we know that a cell contains 4×´9 dimension vector, a feature map of HOG can be used as the concept of the cloud model. From the feature map and then use the concept rising the result is given. The flow is shown in Fig. 5

## EXPERIMENTS AND RESULTS

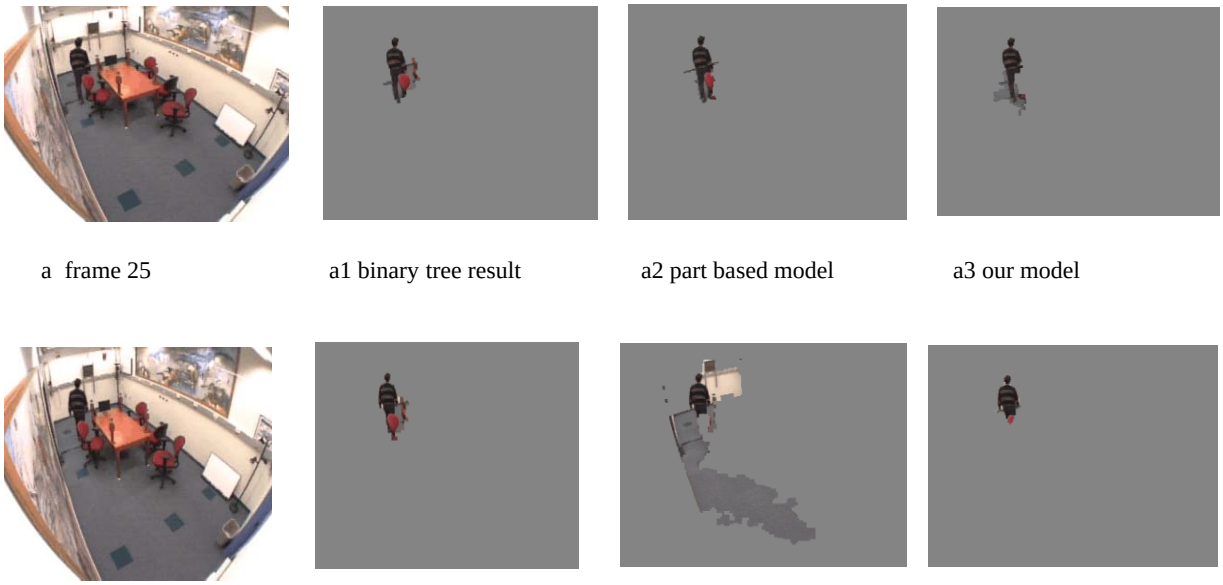
In this section, some result is given based the car and person, from (Dalal and Triggs, 2005), the person of HOG feature perform well (Andriluka *et al*., 2009; Dollar *et al*.,



a frame 317 a1 binary tree result a2 part based model a3 our model

b frame 318 b1 binary tree result b2 part based model b3 our model

Fig. 6: Outdoor scene



a frame 25 a1 binary tree result a2 part based model a3 our model

b frame 32 b1 binary tree result b2 part based model result b3 our result

Fig. 7: Indoor scene

2009) and many researcher has used this method to do the work. But for it need to show a cascade model of the image, it doesn't work so fast, A HOG feature is simply given from the image, the HOG feature is used by the cloud model to detect the moving object. Two examples are given as follow. The algorithm is tested on a 3.06G Intel PC with 2-G RAM.

**Outdoor scene:** The frames 317 and 318 come from Project Function of Monitoring System of the Second Changjiang River Bridge at Wuhan, China. In Fig. 6, there are some results are given, binary tree method and part based model, our method. from the results, the Binary tree method need to choose some pixels as the sample and then to find the moving object. The part based model

works well in the locating of the moving object but it needs to training the samples which we choose from the series of 400 frames as the training set, in our method, it doesn't need to do this work and it work well as the two method. and some results is better than the binary tree method and part based model. Indoor Scene Experiments are chosen two frame of a video from web other is from the web http://www.cs.cmu.edu/~cil/vision.html, the person detection (Eichner and Ferrari, 2010) has be shown that the HOG feature and part based model is well in Fig. 7. As the indoor scene, three methods are tested and the part based model we use some frames as the training samples.

In Fig. 7, the person is detected by the method of Binary tree and part based model and our method, from the result we can find our method give the effective result. but the leg is lost because we use only the information of the HOG.

From the results of the indoor scene and outdoor scene, our method is good at detecting moving object, but it loses more information to the part based model. And as the binary tree method contains more noise of the background. Our method loses some details based the cloud model. the binary tree needs to find the sample of the foreground manual, the part based model need training process, our method considers only the feature and it is more effective than the other two methods.

## CONCLUSION

In this study, a method based on the HOG and cloud model is presented, first from the image we compute the HOG feature and then the cloud model is used to extract information from the HOG feature, the concept rising is used to refine the moving object. Because it considers the statistics character so it is effective in detecting moving object. Only three number character Ex, En, He are needed to define. From the experiment, it is shown that the result is more coincident with actual facts.

The cloud model always needs to consider concept, but to some condition, the moving object is so similar that make a mistake of the foreground. In HOG feature to cloud model need a more useful method because we only use it as the histogram, it lost the spatial information, next we will consider the spatial information of HOG and to consider the cloud model constructed with feature indeed.

## ACKNOWLEDGMENT

This study is supported in part by National Natural Science Foundation of China No. 41001236 and No. 41101410.

## REFERENCES

Andriluka, M., S. Roth and B. Schiele, 2009. Pictorial structures revisited: People detection and articulated pose estimation. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp: 1014-1021.

Belongie, S. and J. Malik, 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Analysis Machine Intelligence, 24(24).

Boykov, Y. and V. Kolmogorov, 2004. An experimental comparison of min-cut/max-flow algorithms for enrgy minimization in vision. IEEE Trans. PAMI, 26(9): 1124-1137.

Boykov, Y. and G. Funka-Lea, 2006. Graph Cuts and Efficient N-D Image Segmentation. Int. J. Comp. Vision, 70(2): 109-131.

Cucchiara, R., C. Grana, M. Piccardi and A. Prati, 2003. Detecting moving objects, ghosts and shadows in video streams. IEEE T. Patten Anal., 25(10): 1337-1342.

Dalal, N. and B. Triggs, 2005. Histagram of oriented gradients for human detection. IEEE Conference on Computer Vison and Pattern Recognition, pp: 886-893.

Dollar, P., C. Wojek, B. Schiele and P. Perona, 2009. Pedestrian detection: A benchmark. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp: 304-311.

Dryden, I. and K. Mardia, 1998. Statistical Shape Analysis. John Wiley and Sons, chichester.

Duetscher, J., A. Blake and I. Reid, 2000. Articulated body motion capture by annealed particle filtering. IEEE Comp. Soc. Conference Comp. Vision Pattern Recognition, 2: 126-133.

Eichner, M. and V. Ferrari, 2010. We are family: Joint pose estimation of multiple persons. In: European Conference on Computer Vision, pp: 228-242.

Fu, Y., J. Cheng, Z. Li and H. Lu, 2008. Saliency cuts: An automatic approach to object segmentation, in Proc. of IEEE International Conference on Pattern Recognition, pp: 1-4.

Jung, C.H., B. Kim and C. Kim, 2010. Automatic Segmentation of Salient Objects Using Iterative Reversible Graph Cut. ICME2010, pp: 590-595.

Lowe, D., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comp. Vision (IJCV), 60(2): 91-110.

Li, D.Y. and Y. Du, 2005. Uncertainty Artificial Intelligence (in Chinese). National Defense Industrial Press, Beijing.

Ren, M.W. and H. Sun, 2005. A practical method for moving target detection under complex background. Comp. Eng., DOI: CNKI: SUN: JSJC. 0.2005 -20-012, pp: 33-34.

Rother, C., V. Kolmogorov and A. Blake, 2004. Graph cut-interactive foreground extraction using iterated graph cuts, Proceeding. of ACM SIGGRAPH, pp: 309-314.

Zabih, R., J. Miller and K. Mai, 1999. A feature-based algorithm for detecting and classifying production effects. Multimedia Syst., 7(2): 119-128.

Zhang, H.J., A. Kankanhalli, S.W. Smoliar, 1993. Automatic partitioning of full-motion video. Multimedia Syst., 1(1): 10-28.