A Hierarchical Computational Model of Selective Visual Attention

Qiaorong Zhang
College of Computer and Information Engineering, Henan University of Finance and Economics, Zhengzhou, China
Email: zqrzqh@126.com

Huimin Xiao
Department of Mathematics and Information, Henan University of Finance and Economics, Zhengzhou, China
Email: huiminxiao@126.com

Haibo Liu
College of Computer Science and Technology, Harbin Engineering University, Harbin, China

***Abstract*—Computational model of visual attention has got more and more attention in machine vision and image processing. A hierarchical computational model for selective visual attention is proposed in this paper. This model simulates the attention mechanism from far (coarse) to near (fine) of human visual system. Firstly, the input image is analyzed at the coarsest resolution and visual saliency of each part is computed at this level. Regions of attention are selected according to the saliency. Then the sub-regions of the selected region compete for attention at a finer resolution. This process is done iteratively until every salient region and its sub-regions have been processed at different levels respectively. The proposed model has been tested on many natural images. Experiment results show that the proposed model is valid and the attention results are consistent with human visual system.**

***Index Terms*—visual attention, image processing, visual saliency, saliency map, hierarchical selection**

I. INTRODUCTION

Visual selective attention is the mechanism that makes human and other primates find what they want in a clustered vision scene rapidly. With visual attention mechanism, they can concentrate on the important things for further processing and ignore other unimportant things. With visual attention mechanism, the important or interested information has the priority to get processed firstly and the computation and reaction speed can be improved. In recent years, visual attention has attracted more and more researchers in active vision and image processing field.

Some researchers have attempted to develop a computational model of visual attention for image processing and machine vision by simulating human visual attention mechanism. And some computational models of visual attention have been proposed in recent years. Most of them are based on visual saliency map. The most influential model is the computational model proposed by Itti *et al.* in 1998 [1] [2] [3]. Firstly, they extracted early visual features like intensity, color and orientation. Then saliency maps of these features were computed by using center-surround operation at different scales. An integrated saliency map was obtained by integrating these feature saliency maps. Finally the focus of attention is got and shifted according to the integrated saliency map. But this model didn’t analyze the effects of different features and just summed these normalized feature maps together. At the stage of attention, only a fixed image scale (resolution) was considered and the size of the attention region is also fixed. Dirk Walther improved this model by selecting the most effective feature and the size and shape of attention regions were arbitrary [4] [5]. But the second problem was not solved yet.

Ma *et al.* proposed a computational method based on color contrast of every pixel in 2003 [6]. Only color feature was used to compute the visual saliency. It is appropriate when the color feature is the most useful feature for the image. It can’t give the right saliency result if the actual contrast resides in other features. The attention regions were extracted by using fuzzy growing and often included the backgrounds or some wrong locations. Zhang *et al.* used the similar method to compute the visual saliency in 2005 [7]. They used multiple features like intensity, color and texture. But they didn’t analyzing the different effect of each feature and just summed them up. These two models didn’t consider the effect of different image scales.

Hou *et al.* presented a spectra residual method to compute visual saliency in 2008 [8]. This method is simpler and more efficient than other existing method. But only intensity feature was used in the method. It could not get the right result if intensity was not the useful feature. This method only used a fixed scale to analyze the visual saliency and the method how the focus of attention was obtained and shifted was not given in their model.

Corresponding author: Qiaorong Zhang

doi:10.4304/jsw.6.9.1705-1712

Therefore, most of the existing models have some shortcomings. Firstly, the different feature maps with different scales competed together to get the focus of attention. The different effects of feature and scale are neglected. Actually the feature and scale have different effect on visual saliency [8] [9] [10]. Secondly, the focus of attention was shifted from one region to another at a fixed image scale. In fact, this is not consistent with human visual attention mechanism. It is known that human visual system has resolution-varying sampling sensing so that it can not only quickly explore a scene but also flexibly acquire interesting information through far to near and coarse to fine by the limited visual processing resources [11][12][13]. When a region is attended in human visual field, people often select sub-regions inside to get more information and shift to another region until the details of the current region are understood.

In this paper, a hierarchical computational model of selective visual attention is proposed. The input image is analyzed at the coarsest resolution. When an area is attended, the attention area is processed as the input image at a finer resolution and the regions inside the area are attended. This process is done iteratively until every region at every resolution has got attention.

This paper is organized as follows. Section II describes the outline and details of the proposed hierarchical selective model. Section III gives some experimental results and discussion. Section IV presents our conclusions and prospects.

II. HIERARCHIAL SELECTIVE COMPUTATIONAL MODEL

Fig. 1 shows the diagram of our proposed computational model of hierarchical selective visual attention.

A. Hierarchical Selection

Different image scales have different influences on the visual system. The scale in our model is equal to the resolution of the input image. When the resolution of the image is coarse, detail features are omitted. When the image is in a fine resolution, detail features are paid more attention to than large features. For example, when we observe a scene from far away, we can see a house. But when we approach it little by little, we pay more attention to the roof, windows of the house. That is to say, human visual system has a flexible attention process from far to near and from coarse to fine.

There are s+1 levels in our hierarchical selective model. Level s is the coarsest resolution and level 0 is the finest resolution which is the original resolution of the input image. Assume that the resolution of the original image is W*H, the resolution of the image at level s is $(W/2^{s})*(H/2^{s})$. The detail of hierarchical selectivity is described as follows:

1) Selective attention begins at the coarsest resolution (level s) and the attention area is the whole input image.
2) Every part of the input image competes for the focus of the attention. The visual saliency of each part is computed to get a saliency map according to their features.

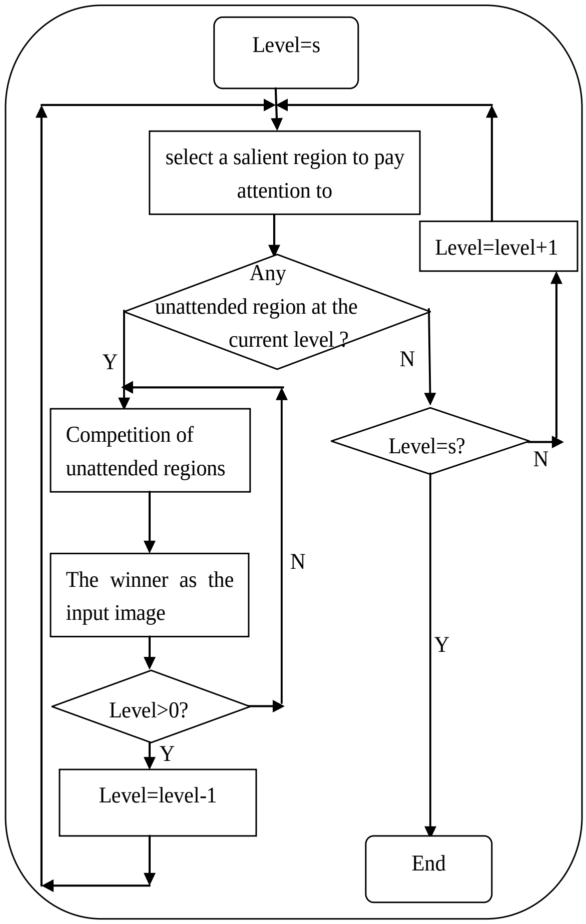

Figure 1. Block diagram of the proposed model

3) Extract salient regions based on the saliency map.
4) According to the saliency map, choose a salient region which has not been focused to be the next attention area.
5) If all the salient regions have obtained attention and the current level is s, then go to step 8).
6) If there is no unattended salient region and the current level is less than s, then return the last level (level=level+1), then go to 4).
7) If the current level is the finest resolution, then go to step 4); else choose the attention area as the input image and analyze it at the finer level (level=level-1), then go to step 2).
8) Stop.

B. Competition of Feature

If a region is salient and remarkable in the image and can get the focus of attention, at least one kind of feature of it is different with its surroundings. That is to say, the saliency of each part in the image is determined by the attributes (features) of it and its surroundings. The most used visual features are intensity, color and orientation. But in our experiments, we found that orientation feature is not very useful in natural images. Therefore in our model, we only use two kinds of feature which are color feature and intensity feature.

Because HSI (Hue, Saturation and Intensity) color space is consistent with human color perception system and is better than RGB color space, the input image is transformed from RGB space to HSI space using (1).

$$\begin{cases} H = \dfrac{1}{360}[90 - Arc\tan(\dfrac{F}{\sqrt{3}}) + \{0, G > B; 180, G < B\}] \\ S = 1 - [\dfrac{\min(R,G,B)}{I}] \\ I = \dfrac{(R+G+B)}{3} \\ F = \dfrac{2R - G - B}{G - B} \end{cases} . \quad (1)$$

Then we use I channel in (1) to represent intensity feature of the input image. H (hue) channel and S (saturation) channel are used to describe the color feature of the image.

1) Visual saliency: Different feature makes different contribution to the visual saliency. Firstly, we need to compute the feature saliency maps of these features. Then according to the generated feature saliency maps, select the feature which has the most contribution to the saliency. The method to compute the saliency has been proposed in our other paper [14] in detail. Here we just describe the main idea of the method in brief.

To obtain a more correct and robust measure of visual saliency, when we compute the visual saliency of each part, we consider three kinds of saliency which are local saliency, global saliency and rarity saliency.

Local saliency means the distinctness between a region and its environment. In frequency domain, an image can be decomposed into magnitude spectrum and phase spectrum. It has been discussed in [15] [16] that phase spectrum is very important in image reconstruction. If we reconstruct the image with phase spectrum only or with a random changed magnitude spectrum, the reconstructed image can reserve the structure information and less distort the original image [15]. But if we reconstruct the image with a random changed phase spectrum, the reconstructed image severely distort the original image. It is indicated that phase spectrum represents the information of value changing at each position whereas magnitude spectrum represents the particular value at each position. Because we only care the change of feature value, we reconstruct the image with phase only to eliminate the influence of magnitude spectrum and get the local saliency using (2).

$$\begin{cases} F(u,v) = \sum_{x=1}^{M}\sum_{y=1}^{m} f(x,y) e^{\frac{-j2\pi ux}{M}} e^{\frac{-j2\pi vy}{N}} \\ \qquad\quad = R(u,v) + jI(u,v) \\ P(u,v) = \arctan(\dfrac{I(u,v)}{R(u,v)}) \\ S_{local}(x,y) = \dfrac{1}{M*N}\sum_{u=1}^{M}\sum_{v=1}^{N} P(u,v) e^{\frac{-j2\pi ux}{M}} e^{\frac{-j2\pi vy}{N}} \end{cases} \quad (2)$$

Where *f(x,y)* means the feature map with dimension *M*N*. The values *F(u,v)* are the DFT coefficients of *f(x,y)*. $S_{local}(x,y)$ means the local saliency value of pixel *(x,y)*.

Only considering local saliency is not enough because high local saliency values often lie in boundaries between salient areas and the background. So we use global saliency as well. Global saliency map of a feature map can be computed using (3).

$$\begin{cases} S_{Global}(x,y) = e^{\frac{\left| f(x,y) - f_{avg}(x,y) \right|}{f_{avg}(x,y)}} \\ f_{avg}(x,y) = \dfrac{1}{M*N}\sum_{x=1}^{M}\sum_{y=1}^{N} f(x,y) \end{cases} . \quad (3)$$

Rarity saliency means the less a feature value occurs the more possible it belongs to a salient area. The rarity saliency can be computed using (4).

$$S_{Rarity}(x,y) = \frac{1}{hist(f(x,y))} . \quad (4)$$

Where *f(x,y)* is the feature value of pixel *(x,y)* in the feature map and *hist(•)* is the histogram of the feature map.

Finally, local saliency map, global saliency map and rarity saliency map need to be combined into a feature conspicuity map. Different weights need to be applied to each saliency result. The feature conspicuity map can be generated using (5).

$$\begin{cases} V = \dfrac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left| f(x,y - \dfrac{1}{M*N}\sum_{i=1}^{M}\sum_{j=1}^{N} f(x,y) \right| \\ w_i = \dfrac{V_i}{\sum_{i=1}^{3} V_i} \\ C_F = w_1 * S_{Local} + w_2 * S_{Global} + w_3 * S_{Rarity} \end{cases} .(5)$$

The process of generating intensity feature saliency map is shown in Fig. 2. Firstly, the intensity feature of the input image is extracted. Then local saliency, global saliency and rarity saliency of the intensity map are calculated. Then these three kinds of saliency results are combined into an intensity conspicuity map. The images in the top row are original images and their intensity feature maps are shown in the second row. The local saliency maps, global saliency maps and rarity saliency maps of the intensity feature maps are shown in the third, fourth and fifth row respectively. The final intensity feature conspicuity maps are shown in the bottom row.

2) Feature Competition: After generating the feature conspicuity maps, we need to analyze the effects of different feature according to the feature conspicuity maps [17]. Itti *et al.* gave four combination strategies in [3]. The weights of different features were equal in the naive linear combination method. So the results were not satisfying. The method of linear combination with learned weights was better but it required a prior knowledge of the salient regions. The global non-linear normalization method and the iterative non-linear method both used a local competition strategy in one feature map.

Figure 2. Example of feature conspicuity map

In this paper, a novel and reasonable feature competition strategy is used to combine these feature conspicuity maps into a final saliency map. The strategy and process of feature integration are described as follows. We use salient area, salient point location and salient point distribution to measure the importance of the feature conspicuity maps and compute their weights [14].

Firstly we should extract salient points. Simply threshold these feature saliency maps using a threshold T. A binary version of the feature conspicuity map is obtained using (6).

$$B(x,) = \begin{cases} 1 & if\ C_F(x,y) \geq T \\ 0 & otherwise \end{cases} \quad (6)$$

Where, T is the threshold which can be computed by using the Matlab *graylevel* function. C_F is the feature conspicuity map.

After the binary maps are obtained, the pixels whose values are 1 in the binary map are considered as the salient points. Based on the rarity principle, the more the salient points in a feature saliency map are, the less useful the feature saliency map is. We compute the number of salient points as the measure of salient point area using (7).

$$W_{area} = N \quad (7)$$

Where, W_{area} means the weight of salient point area and N represents the number of salient points. If the area of salient points is larger than 70% of the area of the whole image, the weight of the feature saliency map is set to zero. This means it is not included when in feature integration.

Secondly, People often pay more attention to the region near image center that means the region near the image center are more likely to be a salient region. So we consider the salient point location as a criterion. Compute the average distance between the salient points and the image center as the location criterion using (8).

$$W_{location} = \frac{1}{N}\sum_{i=1}^{N} Dist(sp_i, center). \quad (8)$$

Where, $W_{location}$ is the weight of salient point location. N is the number of salient points. sp_i means each salient point and *center* means the center of the image. *Dist* means the distance between two points.

In addition, if the salient points don't cluster together but distribute separately in the feature saliency map, the feature saliency map is not very useful. So we compute the spatial distribution of salient points using (9) as another criterion.

$$W_{distribution} = \frac{1}{N}\sum_{i=1}^{N} Dist(sp_i, centroid). \quad (9)$$

Where, $W_{distributio}n$ is the weight of salient point spatial distribution. *centroid* means the center of the salient points.

Finally, these feature conspicuity maps are integrated together using (10).

$$\begin{cases} SM = \sum_{i=1}^{m} W_i * C_F^i \\ W_i = \dfrac{\frac{1}{W_{fi}}}{\sum_{i=1}^{m} \frac{1}{W_{fi}}} \\ W_{fi} = W_{area}^i + W_{location}^i + W_{distribution}^i \end{cases} . \quad (10)$$

Where, SM is the integration saliency map and C_F^i means each feature saliency map.

The example of feature competition is shown in Fig. 3.

C. Attention Shift

After the saliency map is obtained, the attention stage begins. The most used attention shift method is that the focus of attention shifts among the salient regions based on the saliency of each region. But as we have pointed out above, human visual system has resolution-varying sampling sensing so that it can not only quickly explore a scene but also flexibly acquire interesting information through far to near and coarse to fine. Therefore, we propose a hierarchical selectivity method for attention shift.

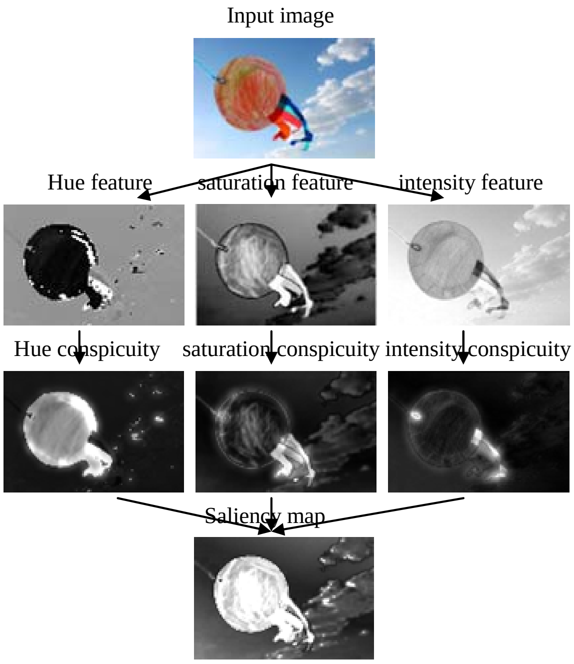

Figure 3. Example of feature competition

Firstly, the binary version of the saliency map at the coarsest level is obtained through the similar method with (6). Then all the salient regions are extracted using the Matlab *bwlabel* function.

Secondly, the salient region which includes the most salient location is selected to be further processed at the next level with a finer resolution. Visual saliency of each part in the selected salient region and the most salient region is selected to be analyzed with the finer resolution. This process is iteratively done until the current level is with the finest resolution. The first salient region is thought to have been completely focused. Then the second salient region is selected and processed in this way.

Finally, when all the salient regions have been processed, the attention stage is finished.

Inhibition of return mechanism is used in attention shift. That means when one salient region has been focused on, the saliency of every pixel in the salient region is set to 0 using (11). Thus when we select another salient region, the salient regions which have been focused on won't be selected again.

$$SM(x,y)=\begin{cases}0 \;\; if \;\; (x,y) \; has \; been \; focused \\ SM(x,y) \;\; otherwise\end{cases} \quad .(11)$$

The example of attention shift will be shown in section III.

III. Experiment Results

To evaluate the performance of the proposed hierarchical computation model of selective visual attention, we have tested it in many natural images. These images are downloaded from the internet or taken with a digital camera. The experiment results and analysis are described in detail.

A. Results

The proposed model has been tested on the computer with Intel Pentium 1.8 GHz and 512M memory using more than 100 natural scene images. Fig. 4 shows the example of hierarchical attention shift.

In Fig. 4, the input image is processed at 4 levels. The resolution at each level is ranging from 100*75 to 800*600. Attention shift begins at the coarsest resolution 100*75. Competition for attention starts in the highest level (level 3). Saliency of each pixel is computed firstly and salient regions of the image are extracted. According to the saliency map, attention is firstly deployed to the winner salient region (here is the region with several cows in front of the house). Then attention is shifted to this region and it is further analyzed at the finer resolution (level 2). Sub-regions of the winner compete for attention at level 2. At this level, the winner is the salient region with three cows at the corner of the house. Again it is further checked at the finer resolution (level 1). Three salient regions are extracted at this level. They compete for attention and are further processed at level 0.

When the winner has no salient region or all the salient regions have attended at level 0, attention shifts again to the coarser level. When attention is shifting, the inhibition of return mechanism is used. That means the saliency of the attended region is set 0 and it will be not included at the next competition. When all the salient regions at the coarsest level (level 4) have been processed, attention shift stops.

The detailed process of attention shift and the salient regions are shown in Fig. 4. The process of attention shift is illustrated by the black arrow. Salient regions of the image at each level are labeled by the white rectangles.

B. Comparision

To evaluate the validation of our proposed model, we also implemented the most popular attention shift method through which focus of attention changes only at a fixed image scale. The detailed implementation is shown in Fig. 5. The pyramids three kinds of feature map with three scales are generated firstly. Feature conspicuity maps are computed at different scales. Each kind of feature conspicuity maps are summed up after resized to one scale. Then these integrated feature conspicuity maps compete to generate the final saliency map. Attention shifts from one location to another according to the saliency map at a fixed scale.

Comparing these two kinds of attention shift method (shown in Fig. 4 and Fig. 5 respectively), we can find that the hierarchical attention shift method is more appropriate to image processing and machine vision because it is more consistent with the resolution-varying sampling mechanism of human visual system. Fox example, it can be used for actual machine vision. When the robot observes a scene using its camera, it can find the most salient area of the scene and then adjust the focal length to observe this area at a finer resolution.

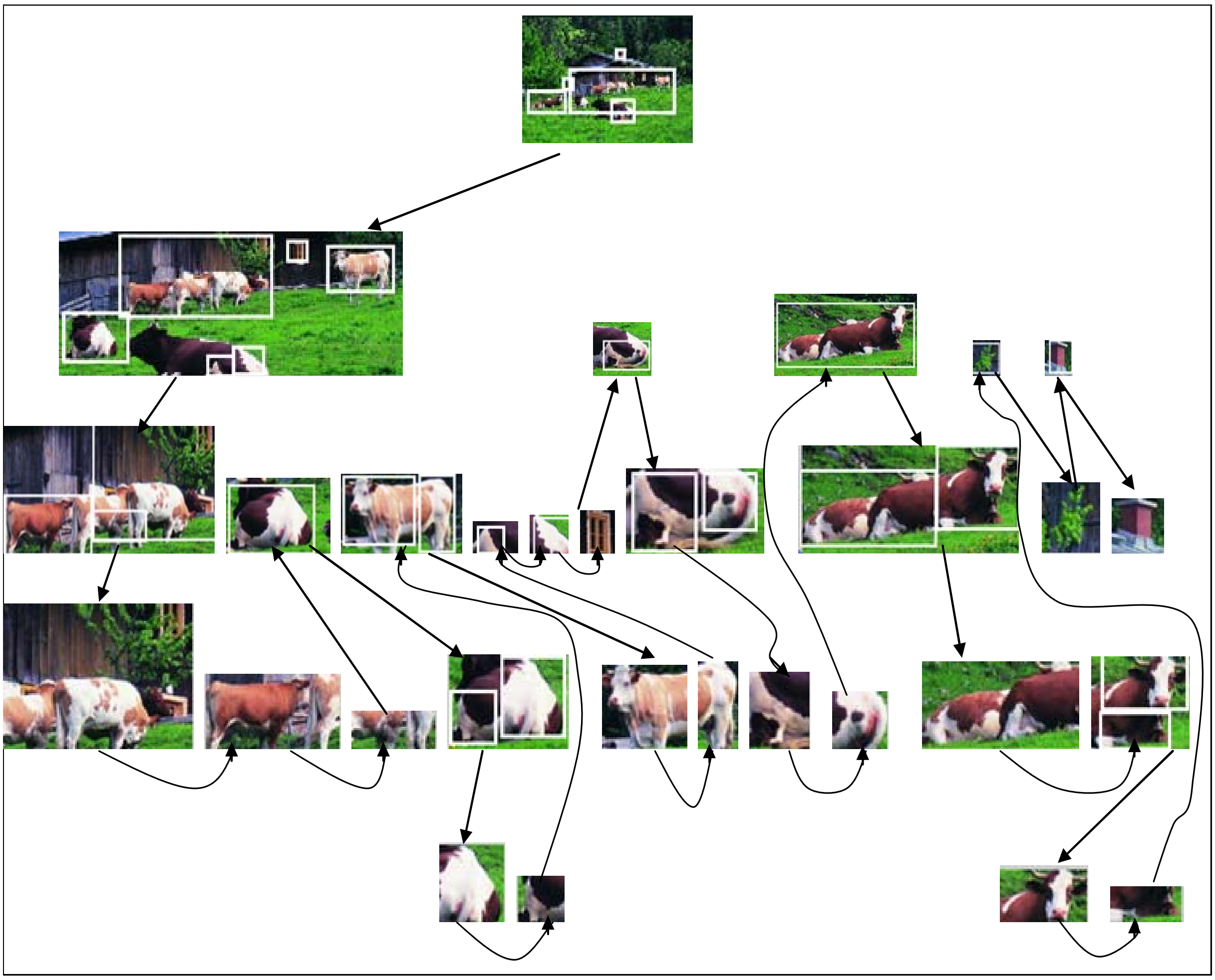

Figure 4. Example of hierarchical attention shift

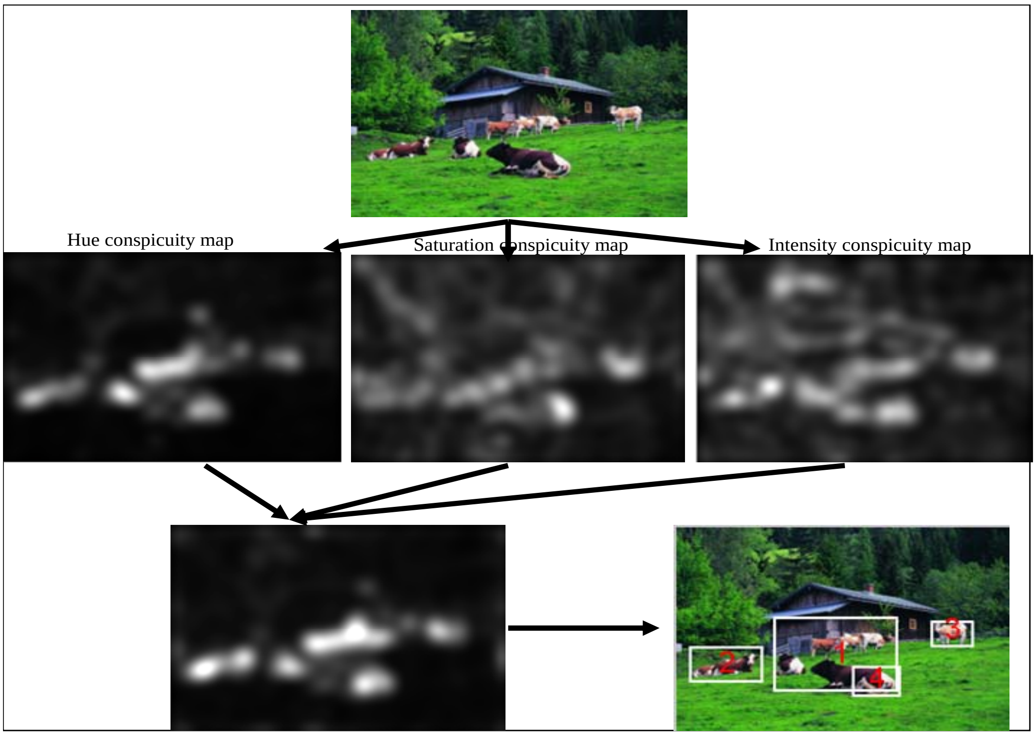

Figure 5. Example of attention shift without hierarchical sekection

C. Discussion

The proposed computational model for visual attention is space-based that means the region that has the maximal saliency value is selected as the focus of attention. But from Fig. 4 we can see that some selected regions don't contain a whole object and some selected regions include some background area. Sometimes attention will shift to a region which has no meanings. This is the defect of space-based attention model.

There are two reasons that are likely to lead to above defects. One reason is the generated saliency map. Although our method has solved some problems of current other methods, the generated saliency maps are not very well in some images which have very clustered background or low-contrast foreground. This need to be further researched. The other reason is the threshold to binary saliency maps. The threshold directly affects the segmentation results. If the threshold is too high not all the pixels of salient regions will be extracted. This will result in low hit rate. Contrarily, a low threshold will lead to too many pixels be extracted. This will result in high false alarm rate. We can make a balance between high hit rate and low false alarm rate to select the threshold.

Some researchers have proposed another attention model which is object-based model. This model can avoid the above problem of space-based model to a certain extent. But the object-based attention model needs perceptual clustering to segment the image into some regions before attention stage. And the same problem may exist in segmentation results.

IV. CONCLUSION

A hierarchical computational model for selective visual attention is proposed in this paper. The focus of attention changes at different layers. Competition begins at the layer the resolution of which is the coarsest. The visual saliency of each pixel in the image at this layer is computed firstly. The winner is selected to be further processed at the next layer with a finer resolution according to the saliency of each area. When a region has been analyzed at the finest layer, attention shift returns to the last layer with a coarser resolution. This computational model can be used in image processing and actual machine vision.

Early vision features are also important to construct the saliency map. A simple feature can not entirely represent the character of the salient region. Therefore, multiple features analysis is used in the proposed method. In this paper, we consider colors, intensity as the features of the image. However, it is very likely that there are some other features such as edge and symmetry feature which also should be considered. What feature and how many features should be extracted according to the target will also be included in future work.

Some experiment results and discussion have been presented in this paper. The proposed model is space-based, so the problem of spaced-based model also exists in our model. We can consider object-based attention model and combine object-based and space-based model together to solve the problem. This is the work we should do in future. Only bottom-up visual attention is researched in our model. The research on top-down visual attention to improve the saliency map will be included in future work.

ACKNOWLEDGMENT

This work was supported in part by a grant from National Natural Science Foundation of China (No. 60774041), Science Technology Project of Henan Province of China under Grant (No. 082102320020), Fundamental Research Funds for the Central Universities and Doctoral Fund of Ministry of Education of China (No. 20092304120013), Fundamental Research Funds for the Central Universities under Grant No.HEUCFZ1010 & HEUCF100604.

REFERENCES

[1] Itti L, Kouch C, and Niebur E, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, USA, vol. 20, pp. 1254-1259, November 1998.

[2] Itti L, Kouch C, "Computational modeling of visual attention," *Nature Reviews Neuroscience*, USA, vol. 2, pp.194-230, March 2001.

[3] Itti L, Kouch C, "Feature combination strategies for saliency-based visual attention systems," *Journal of Electronic Imaging*, vol. 10, pp.161-169, January 2003.

[4] Dirk Walther, "Interactions of visual attention and object recognition: computational modeling, algorithms, and psychophysics," *PhD thesis*, California Institute of Technology, 2006.

[5] Dirk Walther, Christof Koch, "Modeling attention to salient proto-objects," *Neural Networks*, USA, vol. 19, pp.1395-1407, 2006.

[6] Yufei Ma, Hongjiang Zhang, "Contrast-based Image Attention Analysis by Using Fuzzy Growing," *Proceedings of the 11th ACM International Conference on Multimedia*, March 2003, pp. 374-381.

[7] Zhang Peng, Wang Runsheng, "Detectiong salient regions based on location shift and extent trace," *Journal of Softwarer, China*, vol. 15, pp. 891-898, June 2004.

[8] Xiaodi Hou, Liqing Zhang, "Saliency Detection: A Spectral Residual Approach," *IEEE Conference on Computer Vision and Pattern Recognition*, June 2007, pp.1-8.

[9] Qiaorong Zhang, Guochang Gu, Huimin Xiao, "Image segmentation based on visual attention mechanism," *Journal of Multimedia*, Finland, vol. 4, pp. 363-370, December, 2009.

[10] Zheshen Wang, Baoxin Li, "A Two-Stage Approach to Saliency Detection in Images," *IEEE ICASSP 2008*, April 2008, pp. 965-968

[11] Yaoru Sun, Robert Fisher, "Object-based Visual Attention for Computer Vision," *Artificial Intelligence, USA*, vol. 146, pp. 77-123, January 2003.

[12] Yaoru Sun, Robert Fisher, Fang Wang and Herman Martins Gomes, "A Computer Vision Model for Visual-Object-Based Attention and Eye Movements," *Computer Vision and Image Understanding*, vol. 112, pp. 126-142, February 2008.

[13] Yaoru Sun, "Hierarchical object-based visual attention for machine vision," *PhD thesis*, University of Edinburgh, 2003.

[14] Qiaorong Zhang, Yafeng Zheng, “An improved computational approach for salient region detection,” *Journal of Computers*, Finland, in press.
[15] Xuelei Ni, and Xiaoming Huo, “Statistical Interpretation of the Importance of Phase Information in Signal and Image Reconstruction,” *Elsevier Science*, June 2002.
[16] Peter J. Bex, and Walter Makous, “Spatial frequency, phase, and the contrast of natural images,” *Journal of Optical Society, American*, vol. 19, pp. 1096-1106, June 2002.
[17] Yiqun Hu_, Xing Xie, and Wei-Ying Ma, “Salient Region Detection Using Weighted Feature Maps Based on the Human Visual Attention Model”, *LNCS 3332*, pp. 993-1000, 2004.

Qiaorong Zhang was born in Zhengzhou, China, in Nov 1978. She received her BS degree in 1999 and Master degree in 2002 in computer application technology both from Harbin Engineering University, Harbin China. Currently she is a Ph.D. student with the College of Computer Science and Technology at Harbin Engineering University, Harbin, Heilongjiang, China.

She is also a lecture at College of Computer and Information Engineering at Henan University of Finance and Economics, Zhengzhou, Henan, China.

Ms. Zhang’s research interests and publications have been focused on image processing, artificial intelligence and biologically inspired computing.

Huimin Xiao was born in Henan, China, in August 1963. He received his M.S. degree in control theory from Huazhong Normal University, Wuhan, Hubei, China in 1988, and Ph.D. degree in automation theory and technology from South China University of Technology, Guangzhou, Guangdong, China in 1991.

He was an academy visitor in College of Engineering and Computer Science in Concordia University, Canada. Currently he is a Professor in the College of Information at Henan Finance and Economics University, Zhengzhou, Henan, China.

Professor Xiao’s research interests and publications have been focused on artificial intelligence, intelligent control and automation.

Haibo Liu was born in Harbin, Heilongjiang, China, in 1976. He received his BS degree in 1998, Master degree in 2001 and Ph.D. degree in 2005, in computer application technology from Harbin Engineering University, Harbin China.

He is an Associate Professor in the College of Computer Science and Technology at Harbin Engineering University, Harbin, Heilongjiang, China currently.

Dr. Liu is a member of China Computer Federation. His research interests and publications have been focused on artificial intelligence, intelligent robots and computer vision.