Dynamic Graphic Visual Communication Effect Analysis Based on User Mental Model

Zhiwei Jiang
Zhejiang A & F University, Hangzhou Zhejiang, 311300, China

Abstract
In order to analyze the dynamic graphics with clear and smooth edges, this paper proposes a dynamic graphic visual communication effect analysis method based on user mental model, which uses the user mental model with a fixed size to extract the visual communication information effectively. The user mental model with 6 layers of design depth could analyze the graphics with clearer edges and suppress the ringing effect of the edges to some extent; using a larger sample library for analysis could avoid over-fitting. Experimental results show that: although the method in this paper has no obvious advantages in visual communication effect of user mental model on the Dong small analysis library; however, in the large analysis library such as ImageNet, the dynamic graphics analyzed by this method have better performance in subjective visual feeling and objective graphics quality evaluation.

Keywords: Dynamic Graphics, Visual Communication, Effect Analysis, User Mental Model.

1. INTRODUCTION
Graphic visual communication (abbreviated as GVC) analysis method uses low resolution (LR) graphics to restore the corresponding high resolution (HR) graphics, designed to increase the high-frequency components in the LR graphics, is a classic hot topic in the field of computer vision (Cazetta, Galetti, Rezende, and Schaefer, 2015). GVC analysis method can be approximately classified as three methods including the interpolation method, the method based on reconstruction, and the method based on user mental model. As the input of GVC analysis method of dynamic graphics only has one LR graphics, it usually does HR Analysis by using the method based on the user mental model (Descout, Decottignies, Vaconsin, Fortineau, Barbault-Foucher, and Rieutord, 2015). The used sample library could come from the similar structures with different scales of LR graphics itself, also from the external LR-HR sample library. In recent years, with the improvement of hardware computing power, the user mental model is widely used in the field of computer vision, and has made outstanding achievements; inspired by this, GVC field began to introduce the user mental model thought (Zhang and Zheng, 2016). Compared with the traditional graphics model which can only manually extract the characteristics and do relatively simple model fit, the user mental model can be automatically made user mental model get hierarchical feature representation (Mannay, 2015).

This method uses a smaller user mental model, splitting the large user mental model into multiple models could better extract the graphic edge information and deepen the dynamic graphic visual communication effect. As the number of models increases, in order to avoid over-fitting, the method uses a larger sample library to analyze the depth model. The experimental results show that compared with the traditional graphic model method, this method has better visual communication effect on dynamic graphics.

2. RELATED WORK

2.1. Dynamic Graphic GVC Analysis Method
The GVC analysis method of dynamic graphics usually adopts the method based on the user mental model, and approximately classifies the different frameworks as four types including forecasting model method, edge-based method, graphical statistical model method, user mental model method based on graphic block, among which the user mental model method based on graphic block has the most prominent effects, and emerged a large number of classical method framework. An example based user mental model method is used to predict the analysis process of LR graphic block to HR graphic block by the belief propagation Markov random field. This method needs a large number of analysis samples and is very sensitive to noise (Hatton, Zhao, Gorantla, Chae, Ahlbrand, and Xu, 2015). Based on the locally linear embedding GVC method, it is assumed that there are similar manifold vectors in the LR/HR graphic block space. The LR graphic block can be expressed as a linear weighted fusion of the neighborhood LR samples, the weights of the LR samples are passed to the HR samples, integrated and generated the HR graphics block (Palevicius and Ragulskis, 2015). The HR graphics is reconstructed using the graphic model coefficients of the LR graphic block. The method can adaptively select the neighborhood to avoid over-fitting or under-fitting, but the process of analyzing library doing graphic model encoding is very time-consuming (Feng, Zhang, Chen, Zhao, and Chen, 2015). As the user mental model is applied
The objective of the GVC method is to increase the visual communication effect of the graph analysis through the gradient descent method to minimize the loss model. It consists of a large number of HR graphics \( \{X_i\} \) and its corresponding LR graphics \( \{Y_i\} \). The third layer model can be expressed as an analogy of GVC analysis process expressed based on graphic model. The three-layer model respectively corresponds to the three functions of graphic block extraction and representation, non-linear mapping, and dynamic graphic analysis. The model structure of the method is shown in Figure 1.

\[
E_j(Y) = \max(0, W_j * Y + B_j)
\]

(1)

Where, \( W_1 \) and \( B_1 \) represent the filter and the deviation, * represents the analysis operation, \( W_1 \) contains \( n_1 \) \( c \times f_1 \times f_1 \) filter, \( c \) is the number of channels included in the graphics, \( f_1 \) is the size of the filter space, that is, \( W_1 \) does \( n_1 \) times of analysis on the graphics, the user mental model is \( c \times f_1 \times f_1 \), the layer output \( n_1 \) feature vector. \( B_1 \) is a \( n_1 \)-dimensional vector, each of its elements corresponding to a filter. Activation model uses ReLU, i.e. \( \max(0, x) \).

The first layer model extracts the \( n_1 \) dimension of the graphic block; the second layer model maps the \( n_1 \)-dimensional eigenvector to the \( n_2 \)-dimensional eigenvector, which is the feature-to-feature mapping. The second layer model can be expressed as

\[
E_j(Y) = \max(0, W_2 * E_j(Y) + B_2)
\]

(2)

Where, \( W_2 \) contains \( n_2 \) \( n_1 \times f_2 \times f_2 \) of the filter, \( B_2 \) is \( n_2 \)-dimensional vector.

Traditional methods usually define the fusion method in advance, for example, the overlapping parts take the mean; while the GVCCNN method uses user mental model method to merge the graphic blocks, and analyzes the complete HR graphics through the third layer model. The third layer model can be expressed as

\[
E_j(Y) = W_3 * E_2(Y) + B_3
\]

(3)

Where, \( W_3 \) contains \( c \) \( n_2 \times f_3 \times f_3 \) filter, \( B_3 \) is a \( c \)-dimensional vector.

According to formula (1) ~ (3), if solving end-to-end model \( F(Y) \), it requires the user mental model obtaining the model parameters \( \Theta \sim \{ W_1, W_2, W_3, B_1, B_2, B_3 \} \). If there exists a dynamic graphics library consisting of a large number of HR graphics \( \{X_i\} \) and its corresponding LR graphics \( \{Y_i\} \), \( \Theta \) can be solved through the interpolation between minimization of \( F(Y_i, \Theta) \) and original HR graphics \( X_i \). GVCCNN uses mean squared error (MSE) as loss model

\[
L(\Theta) = \frac{1}{n} \sum_{i=1}^{n} \| F(Y_i; \Theta) - X_i \|^2 ,
\]

Where \( n \) is the number of analysis samples. GVCCNN uses the standard backward propagation random gradient descent method to minimize the loss model.

It should be noted that the GVCCNN model does not have a pooling layer. Through the results of analysis does downsampling, reduce some parameters, obtain invariance such as rotation, translation, stretch etc. The objective of the GVC method is to increase the visual communication effect of the graphics, which is

\[
F_{\Theta}(Y, X) = \max(0, W * Y + B)
\]

(4)
contrary to the goal of reducing the visual communication effect of the pooling layer. Therefore, both the method in this paper and GVCCNN method do not use pooling layer.

Compared with the traditional user mental model method, the GVCCNN method could express the user mental model in features, extract dynamic graphic visual communication features in more favor of analysis. Experimental results show that the GVCCNN has some advantages over the general dynamic, in the objective evaluation goes beyond a lot of classic methods, obtains analytical graphics with a clear edge and rich visual communication effects, and has good visual subjective feelings, but some graphical edges could be observed with an obvious ringing effect.

Researchers also carried out experiments of increasing the number of layers and changing the size of the user mental model and concluded from the experimental results that the deeper was not absolutely the better. GVCCNN method first layer analysis uses the 9×9 relatively large user mental model; while some people put forward different ideas, believing that from the first layer analysis avoiding using large user mental model but using equivalent user mental model stack to increase the model depth could effectively increase the determination of the judgment model, and reduce the number of weighting parameters. Enlightened by the above, this paper refers to the GVCCNN model structure, and aiming at the problem of graphic edge analysis uses the 3×3 user mental model to suppress the ringing effect to a certain extent, deepen the number of model layers, and obtain clearer edges.

3. USER MENTAL DYNAMIC GRAPHIC VISUAL COMMUNICATION ANALYSIS

A deeper model could be used to better extract the visual communication features in the dynamic graphics. When analyzing a larger model is equivalent to several smaller model continuous analysis, for example, consecutive analysis of two 3x3 models can be equivalent to one analysis of the 5×5 model. Change the analysis graphics through one analysis layer with the user mental model size of 5×5 to consecutive through two analysis layers with the user mental model size of 3×3, the process of which is equivalent in the angle of analytical calculation, but because of the introduction of a non-linear layer, it can extract more complex and effective graphic features; for the analytical calculation, suppose c represents the number of graphical channels, then $2 \times 3 \times 3 \times c = 18c$ has less calculation than $5 \times 5 \times c = 25c$.

In the process of GVC analysis, in order to obtain a clear edge, the graphic visual communication information is more important than the brightness information. The method in this paper increases the model to 6 layers and double the model number, and uses a fixed 3×3 model, which is the smallest model that represents the graphic features and could effectively extract the visual communication information. In the analysis process, it uses zero filling retain the original graphic information to the maximum, avoids loss of graphic information after analysis, makes the analysis process input and output equal, and analyzes the whole HR graphics including the edge.

Considering that larger and deeper models will be fitted when using small data volume analysis, this paper uses a larger dynamic graphics library to analyze the model. In the later experiment part, it will prove that the use of large data volume analysis could fully play the advantages of 6-layer user mental model and effectively enhance the method performance.

The user mental model network structure designed in this paper is shown in Figure 2.

![Figure 2. Model Structure of Method in This Paper](image)

In this method, the first layer model is shown in formula (1), where $W1$ contains $n1$ 3×3 filters, which could be understood as $W1$ does $n1$ analysis of the graphics, the used user mental model is 3×3, the layer outputs $n1$ feature maps; $B1$ is a $n1$-dimensional vector, $n1=128$. Activation model uses ReLU.

The structure of the second to fifth layers is similar to that of the first layer,

$$F_i(Y) = \max(0, W_i \ast F_i(Y) + B_i)$$

for $i = 2, ..., 5$. 

$$F_1(Y) = \max(0, W_i \ast F_1(Y) + B_i)$$
Where, W2 contains n2 filters, n2=64, which is halved compared to the first layer model; similarly, the number of filters at the third to fifth layer is sequentially n3=32, n4=16, n5=8, the filter size is fixed 3×3. Activation model uses ReLU.

The model for the sixth layer is the HR analysis model, which is the same as the last layer model of GVCCNN

\[ F(Y) = W_s * F_Y (Y) + B_s. \]

In order to obtain the model parameters \( \Theta = \{ W_1, W_2, W_3, W_4, W_5, W_6, B_1, B_2, B_3, B_4, B_5, B_6 \} \), the method in this paper continues to use MSE same as the GVCCNN method as the loss model, and use standard backward propagation random gradient descent method minimizes the loss model.

4. EXPERIMENT AND RESULT ANALYSIS

This method analysis uses the platform as the 3.5GHz Intel i5 CPU, 32 GB memory, 6GB video memory GTX980 Ti graphics card, the analysis data is put on the SSD cache to improve the speed.

Considering that the 6-layer model has more parameters, the method should use a larger dynamic graphics library. The experiment randomly selected about 50,000 graphics (length and width of not more than 512 pixels) from ImageNet to form a dynamic graphics library. The generation mode is in line with the GVCCNN, each of the original graphics is taken as HR graphics X, for which do downsampling with parameter s; and then use the Bicubic method to magnify s times to obtain the graphics Y, where s=2,3,4. Y and X are cut into 32×32 graphic blocks, paired with the model for analysis. In order to facilitate the comparison, this paper only analyzes the gray graphics, uses the peak signal to noise ratio (PSNR) and the structural similarity (SSIM) as the numerical standard for measuring the generated HR graphics quality.

Table 1 shows the comparison of PSNR and SSIM of HR graphics analyzed by GVCCNN and the method in this paper, among which setting magnification times s=2, 3, the GVCCNN and the method in this paper are made 10,000,000 times of iteration analysis, and setting magnification times s=4, both are made 5,000,000 times of iteration analysis, test graphics obtained randomly from ImageNet. It can be seen that the method in this paper is superior to GVCCNN in most cases.

<table>
<thead>
<tr>
<th>Graphic name</th>
<th>Zoom in digits</th>
<th>Bicubic</th>
<th>GVCCNN</th>
<th>Method in this paper</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PSNR/dB</td>
<td>SSIM</td>
<td>PSNR/dB</td>
<td>SSIM</td>
</tr>
<tr>
<td>01119242.JPEG</td>
<td>2</td>
<td>23.28</td>
<td>0.90</td>
<td>25.45</td>
</tr>
<tr>
<td>01119253.JPEG</td>
<td>2</td>
<td>26.77</td>
<td>0.85</td>
<td>29.80</td>
</tr>
<tr>
<td>01119334.JPEG</td>
<td>2</td>
<td>35.53</td>
<td>0.90</td>
<td>37.26</td>
</tr>
<tr>
<td>01119371.JPEG</td>
<td>2</td>
<td>22.40</td>
<td>0.88</td>
<td>25.06</td>
</tr>
<tr>
<td>01119443.JPEG</td>
<td>2</td>
<td>28.33</td>
<td>0.91</td>
<td>31.43</td>
</tr>
<tr>
<td>01119484.JPEG</td>
<td>2</td>
<td>25.65</td>
<td>0.87</td>
<td>26.73</td>
</tr>
<tr>
<td>01119688.JPEG</td>
<td>2</td>
<td>30.34</td>
<td>0.95</td>
<td>35.13</td>
</tr>
<tr>
<td>01119978.JPEG</td>
<td>2</td>
<td>35.89</td>
<td>0.97</td>
<td>39.62</td>
</tr>
</tbody>
</table>

In the subjective visual perspective, the method in this paper could achieve a clearer edge, and suppress the ringing effect to a certain extent. Figure 3 shows the visual communication comparison chart of s=2. It could be seen from the visual communication results of 01119242.JPEG that the edge of the HR graphic letter generated by this method is clearer and smoother, while the periphery of the HR graphic letter generated by GVCCNN method produces an obvious ringing effect. The visual communication results of 01119688.JPEG
and 1119978.JPEG show that the HR graphics of the GVCCNN analysis also has a ringing effect, while the method in this paper suppresses the ringing effect to a certain extent.

![Image](https://example.com/image1.png)

**Figure 3. s=2 Visual Communication Results Comparison Chart**

Figure 4 and 5 are respectively the comparison of visual communication results of s=3 and s=4. It could be seen that in the case of high magnification, the method in this paper could suppress the ringing effect to a certain extent, and the edge of the analysis graphics is smoother. It needs to be point out that, 01119242.JPEG contains letters with thinner strokes, after down sampling using Bicubic zoom, the visual communication effect is greatly lost, the strokes are incomplete; while GVCCNN and the method in this paper take Bicubic zoomed graphics as input, so in the generated HR graphics some of the visual communication effects is difficult to restore.

![Image](https://example.com/image2.png)

**Figure 4. s=3 Visual Communication Results Comparison Chart**

**Table 2. PSNR and SSIM Results for Set14 Analysis of HR Graphics**

<table>
<thead>
<tr>
<th>Original graphics</th>
<th>Bicubic</th>
<th>GVCCNN</th>
<th>Algorithm in this paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baboon</td>
<td>23.01</td>
<td>0.64</td>
<td>24.02 0.74 24.10 0.75</td>
</tr>
<tr>
<td>Barbara</td>
<td>25.88</td>
<td>0.81</td>
<td>26.75 0.86 26.65 0.86</td>
</tr>
<tr>
<td>bridge</td>
<td>25.95</td>
<td>0.76</td>
<td>27.48 0.84 27.48 0.84</td>
</tr>
<tr>
<td>coastguard</td>
<td>27.27</td>
<td>0.73</td>
<td>28.84 0.82 28.78 0.82</td>
</tr>
<tr>
<td>Comic</td>
<td>23.14</td>
<td>0.80</td>
<td>25.84 0.89 26.01 0.90</td>
</tr>
<tr>
<td>Face</td>
<td>32.80</td>
<td>0.82</td>
<td>34.13 0.86 34.16 0.87</td>
</tr>
<tr>
<td>Flowers</td>
<td>27.61</td>
<td>0.86</td>
<td>30.85 0.92 30.91 0.92</td>
</tr>
<tr>
<td>Foreman</td>
<td>31.73</td>
<td>0.93</td>
<td>34.53 0.96 35.03 0.96</td>
</tr>
<tr>
<td>Lena</td>
<td>32.10</td>
<td>0.88</td>
<td>34.82 0.91 34.81 0.91</td>
</tr>
<tr>
<td>Man</td>
<td>27.16</td>
<td>0.80</td>
<td>29.16 0.87 29.29 0.87</td>
</tr>
<tr>
<td>Monarh</td>
<td>29.78</td>
<td>0.94</td>
<td>34.99 0.97 35.12 0.97</td>
</tr>
<tr>
<td>Pepper</td>
<td>32.72</td>
<td>0.88</td>
<td>35.37 0.90 35.45 0.90</td>
</tr>
<tr>
<td>ppt3</td>
<td>24.62</td>
<td>0.93</td>
<td>29.16 0.97 29.43 0.98</td>
</tr>
<tr>
<td>Zebra</td>
<td>27.88</td>
<td>0.87</td>
<td>32.18 0.93 31.62 0.93</td>
</tr>
</tbody>
</table>
In order to study the effect of the graphic analysis library on the performance of the method, GVCCNN and the method in this paper use the analysis library of 91 graphs provided by GVCCNN, and do the 1,000,000 times of iteration analysis with a magnification of 2. Table 2 shows the visual communication results of the GVCCNN and the method in this paper in graphic library Set14. It could be seen that the method in this paper is slightly superior to GVCCNN in most cases, but the performance improvement is not obvious, mainly because the number of model of the method in this paper is twice of GVCCNN, and the insufficient data volume in the analysis library causes over-fitting. While Table 1 uses a larger library analysis, this method could effectively play a role so the visual communication results have significant improvement than GVCCNN.

5. CONCLUSIONS

This paper presents a method of analysis of the dynamic graphic visual communication effect based on user mental model, referring to GVCCNN model structure, it uses the user mental model with uniform size, increases the model depth to 6 layers, doubles the number of models, thus could more effectively do GVC analysis for the edge. The method in this paper analyzes that HR graphics has a clearer edge than GVCCNN, and suppresses the ringing effect to a certain extent. Because this method uses larger graphic analysis library, it can effectively avoid model over-fitting. It is necessary to point out that although this method has better visual communication analysis effect on most graphics, however, for graphics with an unclear edge and a large irregular texture, the advantages of this method are not obvious. How to classify and analyze the edge and texture and generate graphics with a clear edge and rich texture is the focus of future work.

REFERENCES


