Quality Assessment Based on Edge and Saliency for Synthesized Virtual View

Yan ZHANG*

School of Computer and Remote Sensing Information Technology, North China Institute of Aerospace Engineering, Langfang, China 065000

*Corresponding author

Keywords: Quality assessment, Virtual view, Edge information, Visual saliency.

Abstract. In multi-view video technology based on DIBR, virtual view quality assessment is more difficult question due to the virtual view's special distortion. Virtual view synthesized by depth estimation reference software (DERS) and view synthesis reference software (VSRS) has relationship with adjacent views. The main distortion in virtual view occurs on the edge, so we extract edges synthesized by adjacent views edge information and the corresponding depth map. In this paper we compared the edge information of virtual view and combined with stereoscopic visual saliency prediction to assess virtual view quality. Experiments proved that the results of this metric are advantageous and consistent with human visual perception.

Introduction

For multi-view video with different number of views, additional virtual views can be generated using Depth Image Based Rendering (DIBR) [1], where the depth map is used to shift the texture image to the correct position in the intermediate view. The quality assessment of the synthesized virtual view is critical for the 3DTV (3D television) technology. The synthesized virtual view induces new types of artifacts such as edge distortion and deformation due to the synthesis process, compression methods or display technologies [2, 3]. For the special distortion of the synthesized virtual view, the general 2D assessment method can not reflect the influence of distortion, so that the assessment results can't be accordance with the human visual perception. Bosc et al. analyses the synthesized virtual view from the edge extraction, but these methods are based on the image edge extraction operator. The edge extraction method needs to be improved, and the extraction results have a great relationship with the edge extraction factor and texture complexity, so the results have great subjectivity [4, 5]. The salient region of the human visual system is also called the region of interest (ROI), and the attention of the human eye to the region of interest is usually higher than that of the non interested region [6]. In this paper, we proposed a novel image quality assessment metric based on edge pixel differences and ROI. Experiments are performed by test sequences, and the results are consistent with human visual perception.

Edge Pixel Differences Method

Fig. 1 shows the depth map distortion due to inaccurate depth estimation error. Wrong pixel position is wrong, thus causes deformation and position change of objects in virtual view. In Fig. 1(a), \( m \) should be transformed to virtual point \( m' \) by \( M \) in the depth map of the virtual point \( m' \), if the depth estimation error occurs, \( M \) is transferred to the \( M_d \)position. comparatively, \( m \) will be transformed to virtual point \( m_d' \) by \( M_d \). Then shown in Fig. 1(b), the virtual view is deformed. The distortion is mainly reflected in the objects on the edge. Therefore, we analyze the edge pixels and calculate their changes.
The main procedures of the proposed metric are as follows:
1) Extract the difference map between the original image and the distorted image;
2) If the corresponding pixel brightness in difference map is greater than the threshold $T$, the corresponding distorted image pixel is denoted as $Z(i, j) = 1$, otherwise it is denoted as $Z(i, j) = 0$, $i = 1 \ldots M; j = 1 \ldots N$; where $i$ and $j$ are the pixel position and $M, N$ is the size of the image;
3) Extract the number of pixels which $Z(i, j) = 1$, the sum is denoted as $n$;
4) The quality of distorted images is $\omega = \frac{n}{M \times N}, \omega \in [0,1]$.

Experiments are performed by test sequence, and the results are shown in Fig. 2. Subjectively, (b) is better than (c), where the main distortion occurs in the red circle in (c), while (b) has no obvious distortion. While the result using SSIM [7] method is opposite. So subjective perception can not be reflected by the SSIM method. Our proposed edge difference method in this paper is consistent with the subjective results.

**ROI of Synthesized Virtual View**
When people look at stereo images, they often pay close attention to the regions of interest. So we assess the stereo images according to the region of interest. The main procedures of the proposed metric are as follows:
1) Extract the region of interest;  
2) Classify the image as follows: region of interest and non region of interest;  
3) Assign the weight factor for each region of the image;  
4) Assess every regions and then sum them.

If the quality of region of interest is $Q_1$, weight factor is $\lambda_1$, the quality of region of interest is $Q_2$, weight factor is $\lambda_2$, the quality of image is $\phi$.

$$\phi = \lambda_1 \cdot Q_1 + \lambda_2 \cdot Q_2$$  
$$\lambda_1 = 1 - \frac{P_1}{M \times N}, \quad \lambda_2 = 1 - \frac{P_2}{M \times N}$$  
where $P_1 + P_2 = M \times N, \quad \lambda_1 + \lambda_2 = 1$

**Proposed Quality Assessment Metric Based on Edge and Saliency for Synthesized Virtual View**

In this paper, the final assessing results are combined with the former two methods shown in Fig. 2 and reflected in Eq. 3.

$$Q = (\omega + \phi) / 2$$  
where Q is the quality of synthesized virtual view.

**Experimental Results**

Four multi-view video sequences released by MPEG, i.e. Pantomime, Book_arrival, Lovebird1, Newspaper are used in the experiments. Table I shows the detailed information of eight sequences. In order to get the virtual view analysis, depth maps used by synthesized virtual view have been processed such as blurred et al.. VSRS3.5 is used in virtual view synthesis algorithm.

<table>
<thead>
<tr>
<th>Test sequences</th>
<th>Original pair (left-right)</th>
<th>Synthesized views</th>
<th>Picture resolution /pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pantomime</td>
<td>37-41</td>
<td>38,39,40</td>
<td>1280×960</td>
</tr>
<tr>
<td>Book_arrival</td>
<td>6-10</td>
<td>7,8,9</td>
<td>1024×768</td>
</tr>
<tr>
<td>Lovebird1</td>
<td>4-8</td>
<td>5,6,7</td>
<td>1024×768</td>
</tr>
<tr>
<td>Newspaper</td>
<td>2-6</td>
<td>3,4,5</td>
<td>1024×768</td>
</tr>
</tbody>
</table>

Three indexes recommended by Video Quality Expert Group (VQEG) are used to compare the advantages and disadvantages for various methods: (1) Correlation Coefficient (CC), (2) Spearman Rank Order Correlation Coefficient (SROCC), (3) Outlier Ratio (OR) [8]. The larger the CC and
SROCC values, the smaller the OR value, the prediction of the model is the better. Shown in Table 2, we can see that CC and SROCC of the proposed method are bigger and OR value is smaller than others methods. So the proposed method is accuracy, monotonic, good consistency, good predictors with subjective perception of human vision.

Table 2. Performance comparison.

<table>
<thead>
<tr>
<th>Method</th>
<th>CC</th>
<th>SROCC</th>
<th>OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSNR</td>
<td>0.8965</td>
<td>0.9368</td>
<td>0.3542</td>
</tr>
<tr>
<td>SSIM</td>
<td>0.9142</td>
<td>0.9566</td>
<td>0.3125</td>
</tr>
<tr>
<td>VQM</td>
<td>0.8979</td>
<td>0.9458</td>
<td>0.3542</td>
</tr>
<tr>
<td>Proposed method</td>
<td>0.9259</td>
<td>0.9891</td>
<td>0.1589</td>
</tr>
</tbody>
</table>

Summary

This paper proposed an efficient quality metric for synthesized virtual view in 3D video. Edge Pixel Differences Method and ROI method are combined in the synthesized virtual view quality assessment. Experimental results over four multi-view test sequences verified the effectiveness of the proposed metric and proved that the metric had higher correlation with human visual perception than other commonly used assessing methods.

Acknowledgement

This research was financially supported by the Science and Technology Project of Hebei Province (15210504), the Youth Foundation of Hebei Educational Committee (QN2015113), the Doctoral Scientific Research Foundation (BKJ-2014-04) and the Project of Langfang science and Technology Bureau (2016011089).

Reference


