

研究与试制

DOI:10.19453/j.cnki.1005-488x.2023.02.009

一种基于人眼凝视点自适应的中心凹图像传输算法

王 燕¹, 严利民^{1,2}

(1. 上海大学 机电工程与自动化学院, 上海 200444; 2. 上海大学 微电子研究与开发中心, 上海 200444)

摘 要: 从人眼视觉系统的角度出发, 结合凝视点位置的变化和人眼中心凹的采样特性, 提出一种基于人眼凝视点自适应的中心凹图像传输算法, 降低数据传输带宽。实验结果表明, 当视距为 2 且凝视点位置位于图像中心时, 综合压缩比为 12.257, 数据传输时钟频率为 12.532MHz, 传输的数据总量为原先数据总量的 8.158%, 有效地降低了数据传输总量, 极大地缓解了数据传输带宽的压力, 使得同一时钟频率下能够实现更高分辨率和刷新率的视频图像。

关键词: 近眼显示; 凝视点; 自适应; 图像传输

中图分类号: TN911.73 **文献标志码:** A **文章编号:** 1005-488X(2023)02-0160-06

An Adaptive Foveal Image Transmission Algorithm Based on Human Eye Gaze Point

WANG Yan, YAN Limin

(1. School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, CHN;
2. Microelectronics Research and Development Center, Shanghai University, Shanghai 200444, CHN)

Abstract: With the development of display devices towards high resolution and high refresh rate, the amount of video image data was increasing exponentially, which brought challenges to the bandwidth resources of the display system. From the perspective of human visual system, combining the change of the gaze point position and the sampling characteristics of the human eye foveal, an adaptive foveal image transmission algorithm based on human eye gaze point was proposed to reduce the data transmission bandwidth. The experimental results showed when the viewing distance was 2 and the gaze point position was at the center of the image, the comprehensive compression ratio was 12.257, the data transmission clock frequency was 12.532 MHz, and the total amount of data transmitted was 8.158% of the original total. The total amount of data transmission was greatly reduced, and the pressure on data transmission bandwidth was effectively relieved, so that video images with higher resolution and refresh rate could be displayed under the same clock frequency.

Key words: close-eye display; gaze point; self-adaptation; image transmission

收稿日期: 2023-02-01

作者简介: 王 燕(2001—), 女, 研究方向为图像处理; (E-mail: 2789401636@qq.com)

严利民(1971—), 男, 副教授, 从事平板显示驱动技术的研究。 (E-mail: yanlm@shu.edu.cn)

引 言

近眼显示器由于体积小、质量轻和易携带等特性,受到科研界和商业界的青睐。但用户对人机交互机制和沉浸式体验的高要求^[1],产生海量待传输的视频图像数据,若直接传输原始的视频图像数据,则为传输带宽和存储空间带来巨大挑战。

近年来研究人员把人眼视觉特性运用到近眼显示器的数据传输领域,并提出运用人眼的视觉特性对需要数据传输的视频图像进行压缩,从而降低数据传输带宽。较为重要的两种算法是基于人眼视觉特性和基于多尺度金字塔的多分辨率图像传输算法,但两种算法均未考虑人眼凝视图像位置改变的情况。

利用人眼凝视点位置的改变与人眼视网膜上视觉信息采集方式不一致的特性,提出一种基于人眼凝视点自适应的中心凹图像传输算法,在人眼凝视区域传输高分辨率的图像,非凝视区域传输低分辨率的图像,模拟凝视点位置改变的人眼凝视效果,有效地减少了数据传输带宽,大幅地降低了待传输的视频图像数据总量。

1 凝视点自适应模型

1.1 中心凹采样特性

视锥细胞在人眼视网膜上的分布并不均匀,在中心凹区域分布密度最高,具有较高的空间分辨率和较强的辨色能力,距离中心凹区域越远,视锥细胞^[2]的分布密度越低,空间分辨率和辨色能力越弱。图1为人眼视网膜上视锥细胞分布示意图,图中每个小圆圈代表单个视锥细胞。该示意图^[3]形象地反映了视网膜上不同区域对视觉信息采集能力的强弱存在空间上的差异。

人类观看视频图像时,人眼视网膜的中心凹区域(凝视区域)对视觉信息的采样能力^[4]较强,若此处的图像存在失真,则将被人眼所感知;中心凹外围区域对视觉信息的采样能力较弱,即图像无论是清晰还是模糊,人眼均感知模糊。

1.2 凝视点自适应模型

定义 E 为装载视角的容器,其包含所有位置像素点对应的视角,如公式(1)所示。

$$E = \{e_i, i \in \mathbb{N}^+\} \quad (1)$$

式中, e_i 表示第 i 个视角, i 属于正整数。

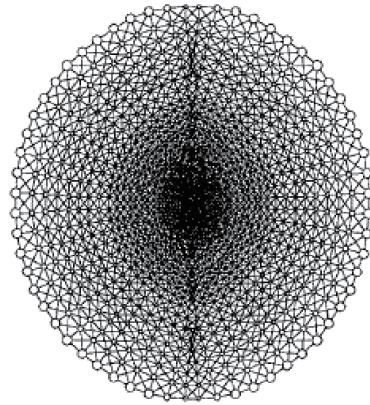


图1 人眼视网膜上视锥细胞分布示意图

Fig.1 Schematic of cone cell distribution on human retina

视角容器 E 包含了图像所有像素点对应的视角。由于近眼显示器通常显示高分辨率的视频图像,因此观看图像所需要的视角数量巨大。

假设凝视点位置位于图像的右下角, X_m 和 X_n 为图像中的两个像素点, X_m 和 X_n 到图像几何中心的距离相等, X_m 对应的视角为 e_m , X_n 对应的视角为 e_n ,线段 OX_r 为 $\angle X_m O X_n$ 的平分线,如图2所示。由图可知,视角 e_m 等于视角 e_n 。因此,公式(1)的视角容器中存在大量数值相同的视角。

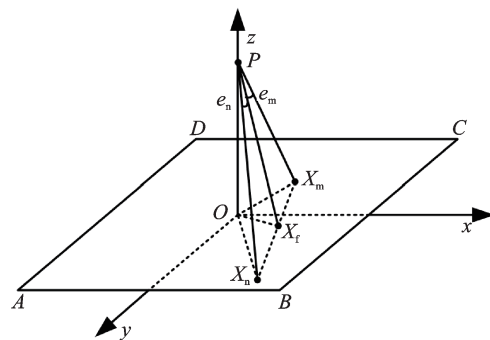


图2 凝视点位置位于右下角的凝视光路几何示意图

Fig.2 Geometric diagram of gaze optical path with staring point at the lower right corner

针对上述问题,提出一种星形距离判别法,将求解视角时的二维运算,降为一维运算,有效地减少了求解视角的运算量。显然,观看图像时存在一个最大视角和一个最小视角,最大视角存在于图像四个角的像素点上,凝视点位置所对应的视角为最小视角。通过求解出不同凝视点位置所对应的最大视角和最小视角,并在最大视角所处像素点和最小视角所处像素点之间间隔单位像素距离,选取虚拟像素点,将这些虚拟像素点作为求解视角的输入,便可获得一组数量最少且数值互异的视角。

图 3 为星形距离判别法的示意图,图中 A、B、C、D 为图像四个角上的像素点,即平面 ABCD 为图像坐标平面, X_f 为凝视点, e_A 、 e_B 、 e_C 、 e_D 分别为 A、B、C、D 所对应的视角,P 到 A、B、C、D 的距离分别为

PA 、 PB 、 PC 、 PD ,如公式(2)~(5)所示。 X_f 到图像四角的距离为最大距离,依次定义为 XA 、 XB 、 XC 、 XD ,如公式(6)所示。

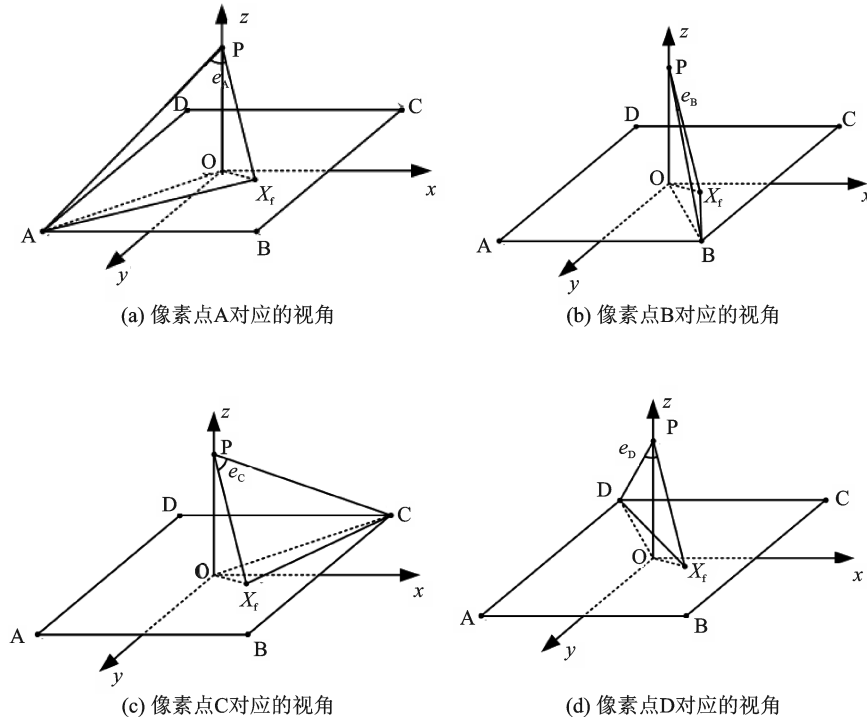


图 3 星形距离判别法示意图

Fig.3 Schematic of star distance discrimination method

$$PA = \sqrt{\left(0 + \frac{N}{2}\right)^2 + \left(0 - \frac{N}{2}\right)^2 + (Nv - 0)^2} \quad (2)$$

$$PB = \sqrt{\left(0 - \frac{N}{2}\right)^2 + \left(0 - \frac{N}{2}\right)^2 + (Nv - 0)^2} \quad (3)$$

$$PC = \sqrt{\left(0 - \frac{N}{2}\right)^2 + \left(0 + \frac{N}{2}\right)^2 + (Nv - 0)^2} \quad (4)$$

$$PD = \sqrt{\left(0 + \frac{N}{2}\right)^2 + \left(0 + \frac{N}{2}\right)^2 + (Nv - 0)^2} \quad (5)$$

由公式(2)~(5)可知, PA 、 PB 、 PC 、 PD 四者大小一致。则新的视角容器 E 可表示为:

$$E = \{e_i, i \in [1, \max\{XA, XB, XC, XD\}], i \in \mathbb{N}^+\} \quad (6)$$

此时,视角容器 E 中装载的视角数量远远小于公式(1)装载的视角数量。同理,可以通过星形距离判别法获得不同凝视点处的视角容器。

如图 4 所示,使用 Matlab 软件绘制出分辨率为 $1\ 600 \times 3 \times 1\ 600$ 、视距为 2 的图像的凝视点自适应的视敏度曲线。图中,凝视点坐标为 $(0, 0)$, xoy 平面表示图像的坐标平面,纵轴为人眼空间临界频率。由图可知,凝视点位置附近人眼空间临界频率较高,人眼感知图像细节能力较强,可感知出图像

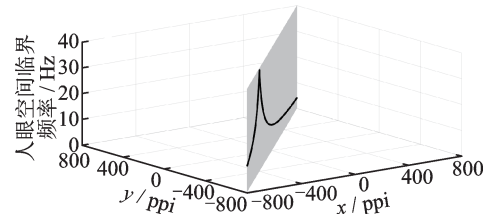


图 4 凝视点自适应的视敏度曲线

Fig.4 Adaptive visual acuity curve of staring point

的轻微失真;随着与凝视点位置的距离逐渐增加,人眼空间临界频率逐渐下降,人眼感知图像细节能力减弱。

2 凝视点自适应的中心凹图像传输算法

2.1 算法流程

基于人眼凝视点自适应的中心凹图像传输算法重点在于计算人眼空间临界频率、构造模糊图像以及图像融合三个步骤。该算法的流程图如图 5

所示。

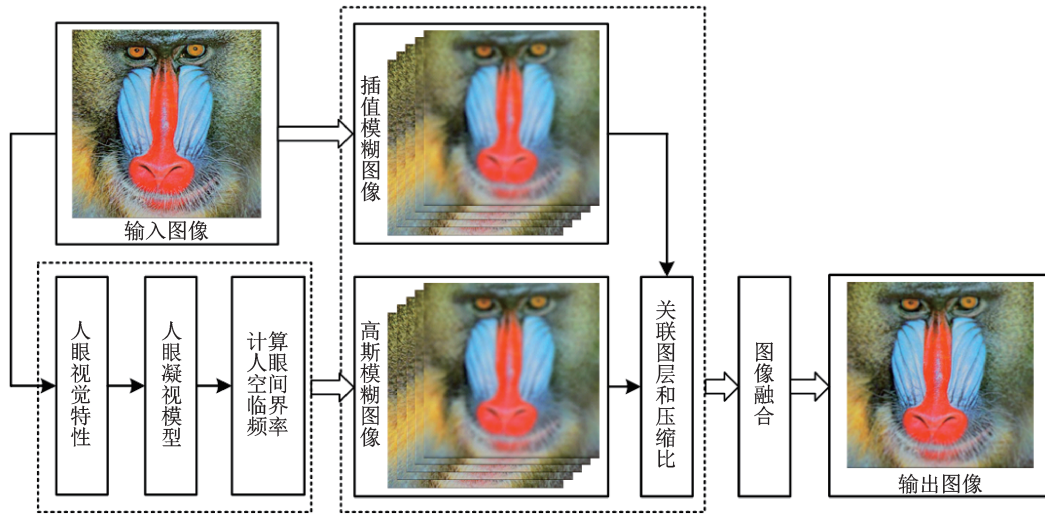


图5 基于人眼凝视点自适应的中心凹图像传输算法流程图

Fig.5 Flow chart of concave image transmission algorithm based on human eye's viewpoint adaptation

该算法的主要步骤如下:

(1)结合人眼视觉特性对传统的凝视光路几何模型进行优化,使得人眼凝视模型适用于凝视视频图像的位置发生改变的情况,并在此基础上计算出不同视角所处的人眼空间临界频率。

(2)根据人眼空间临界频率和高斯核半径的数学表达式,获得不同视角所对应的高斯核半径,并对输入图像进行高斯模糊处理,得到高斯模糊图像集合。同时使用双线性插值算法处理输入图像,获得不同缩放系数下的插值模糊图像集合。

(3)根据实际像素距离对图层进行融合,形成不同图层的交界处平滑且符合人眼主观视觉感知的图像。

2.2 模糊图像构造

在空间频域中,可以把人眼看作一个低通线性系统,而高斯滤波是一种线性平滑滤波^[5],结合人眼在凝视区分辨细节能力强,非凝视区分辨细节能力弱的视觉特性可知人眼对某些高频信息不敏感,可以使用高斯滤波去除。

二维高斯函数的数学表达式如公式(7)所示。

$$G(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \quad (7)$$

式中, $G(x,y)$ 为权值, σ 为高斯核半径,高斯核半径的大小和图像模糊程度呈正相关。

根据文献可知^[6], σ 的数学表达式如公式(8)所示。

$$\sigma=\frac{F}{1.86f_m} \quad (F \text{ 为常数 } 45) \quad (8)$$

为了获得不同模糊程度且符合人眼主观视觉感知的图像,文章采用不同半径的高斯核对原始图像进行高斯滤波处理,如公式(9)所示。

$$G_i(x,y)=S(x,y)*G(\sigma_i) \quad (9)$$

式中, $G_i(x,y)$ 为经过处理的高斯模糊图像, $S(x,y)$ 表示原始图像, $G(\sigma_i)$ 为不同窗口大小的高斯卷积模板。其中,高斯模糊图像的数量等于视角容器 E 中视角的数量。将所有的高斯模糊图像称为高斯模糊图像集合。如公式(10)所示。

$$GS(x,y)=\{G_i(x,y),i\in N^+\} \quad (10)$$

同时使用双线性插值对输入图像进行处理获得插值模糊图像。双线性插值算法的核函数是一个低通滤波器,借助其核函数的特性,可去除图像中人眼不敏感的高频信息。采用双线性插值算法对原始图像先进行 R_j 倍的缩小,再进行 R_j 倍的放大,获得不同模糊程度的插值模糊图像 $B_j(x,y)$,如公式(11)所示。

$$B_j(x,y)=S(x,y)\cdot BLN(R_j)\cdot BLN\left(\frac{1}{R_j}\right) \quad (11)$$

式中, $S(x,y)$ 为原始图像, \cdot 表示插值运算符, $BLN(R_j)$ 表示进行 R_j 倍的缩小, $BLN(1/R_j)$ 表示进行 R_j 倍的放大。分析可知,缩放系数越大,高频分量损失越多。

将所有插值模糊图像称为插值模糊图像集合,如公式(12)所示。

$$BS(x, y) = \{B_j(x, y), j \in \mathbb{N}^+\} \quad (12)$$

2.3 人眼凝视点自适应的图像融合

在主观因素和客观因素的影响下,人类凝视图像的位置发生变化。显然,将各个图层融合成符合人眼视觉特性的图像时,凝视点的位置不同将导致所需的图层数量不同。因此,需要构建出适用于任意凝视点位置的像素和压缩比关系图,即像素和压缩比的极限关系图,如图6所示。其中,视距为2,横轴为压缩比,纵轴为像素,曲线为凝视点自适应的视敏度曲线,灰黑色部分表示压缩比。

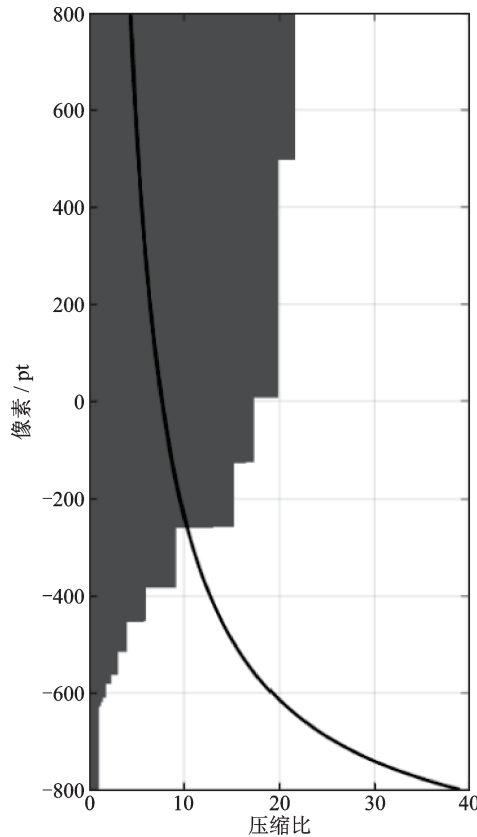


图6 像素和压缩比的极限关系图

Fig.6 Limit graph of pixel and compression ratio

为了表示原始图像与图层之间是否存在重叠,定义 AF_k 、 BF_k 、 CF_k 、 DF_k ,它们分别表示第 k 层图层与原始图像四个边界的标识量。当图层与原始图像无重叠部分,则标识量为1,反之,则为0。它们的数学表达式如公式(13)~(16)所示。

$$AF_k = \begin{cases} 1, & \text{if } (\frac{N}{2} + x_f) - \sum_{m=1}^k l_m > 0 \\ 0, & \text{if } (\frac{N}{2} + x_f) - \sum_{m=1}^k l_m \leq 0 \end{cases} \quad (13)$$

$$BF_k = \begin{cases} 1, & \text{if } (\frac{N}{2} - x_f) - \sum_{m=1}^k l_m > 0 \\ 0, & \text{if } (\frac{N}{2} - x_f) - \sum_{m=1}^k l_m \leq 0 \end{cases} \quad (14)$$

$$CF_k = \begin{cases} 1, & \text{if } (\frac{N}{2} - y_f) - \sum_{m=1}^k l_m > 0 \\ 0, & \text{if } (\frac{N}{2} - y_f) - \sum_{m=1}^k l_m \leq 0 \end{cases} \quad (15)$$

$$DF_k = \begin{cases} 1, & \text{if } (\frac{N}{2} + y_f) - \sum_{m=1}^k l_m > 0 \\ 0, & \text{if } (\frac{N}{2} + y_f) - \sum_{m=1}^k l_m \leq 0 \end{cases} \quad (16)$$

为了量化凝视点与第 k 层或原始图像边界之间的距离,定义 L_{Lk} 、 L_{Rk} 、 H_{Uk} 、 H_{Dk} ,其中, k 表示图层序号; H_{Uk} 、 H_{Dk} 、 L_{Lk} 、 L_{Rk} 分别表示第 k 层图层上、下、左、右边界与凝视点间距,数学表达式如公式(17)~(20)所示,单位为像素。

$$L_{Lk} = (\sum_{m=1}^k l_m) \times AF_k + (1 - AF_k) \times (\frac{N}{2} + x_f) \quad (17)$$

$$L_{Rk} = (\sum_{m=1}^k l_m) \times BF_k + (1 - BF_k) \times (\frac{N}{2} - x_f) \quad (18)$$

$$H_{Uk} = (\sum_{m=1}^k l_m) \times CF_k + (1 - CF_k) \times (\frac{N}{2} - y_f) \quad (19)$$

$$H_{Dk} = (\sum_{m=1}^k l_m) \times DF_k + (1 - DF_k) \times (\frac{N}{2} + y_f) \quad (20)$$

通过公式(13)~(20)便可完成凝视点自适应的图像融合,形成凝视区域图像分辨率高,非凝视区域图像分辨率低的人眼凝视点位置改变的凝视效果图,融合示意图如图7所示。其中,红色小圆圈表示凝视点,颜色深的区域表示图像分辨率高,颜色浅的区域表示图像分辨率低,图层左上角的编号为图层序号, ΔH_U 、 ΔH_D 、 ΔL_L 、 ΔL_R 分别表示图层之间上、下、左、右边界最小间距。

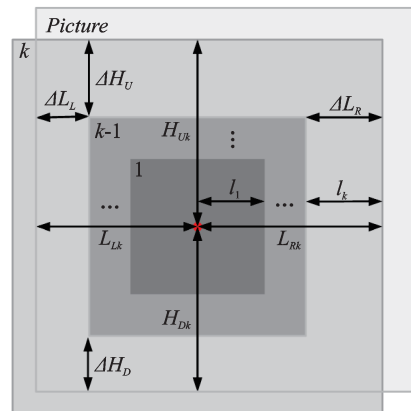


图7 凝视点自适应的图像融合示意图

Fig.7 Schematic of image fusion with adaptive gaze point

3 实验分析

当凝视点位置发生改变时,采用文中所提算法绘制的人眼凝视效果图,如图8所示。其中,红色小圆圈表示人眼凝视点,图(a)为原始图像,图(b)的凝视点位置位于图像中心,图(c)的凝视点位置位于图像左下角,图(d)的凝视点位置位于图像左上角。结合文中算法和人眼视觉系统可知,红色小圆圈附近区域(凝视区域)所处的人眼视角小,人眼空间临界频率高,人眼对高频信息的采样能力强,视网膜上视锥细胞分布密度高;离红色小圆圈距离越远,视角越大,人眼空间临界频率越低,人眼对高频信息的采样能力越弱,视网膜上视锥细胞分布密度低;符合人眼主观视觉感知和中心凹采样特性。



(a) 原始图像 (b) 凝视点位置位于图像中心

(c) 凝视点位置位于图像左下角 (d) 凝视点位置位于图像左上角

图8 不同凝视点位置的人眼凝视效果图

Fig.8 Effect diagram of human gaze at different gaze points

使用分辨率为 $1\ 600 \times 3 \times 1\ 600$ 的图像验证文章算法,设定凝视点位置位于图像中心,选择4、3、2三个视距,依次采用基于多分辨率融合的纹理金字塔压缩算法、基于多尺度金字塔的压缩算法和文中所提算法计算压缩比,对比结果如表1所示。

表1 凝视点坐标为(0, 0)时,不同算法综合压缩比结果对比

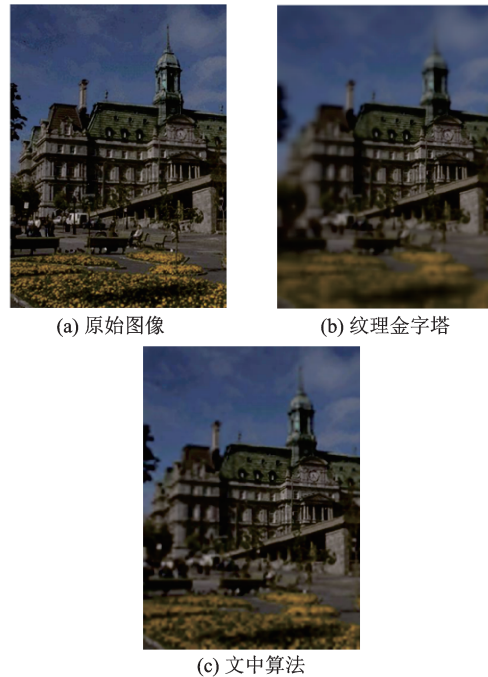
Tab.1 The comparison of the comprehensive compression ratio results of different algorithms when the gaze point coordinate is (0,0)

视距	纹理金字塔	多尺度金字塔	文中算法
4	1.259	1.299	4.283
3	3.600	3.794	7.241
2	8.102	8.489	12.261

纹理金字塔压缩算法和多尺度金字塔压缩算法均采用下采样的方式来降低传输带宽,但都存在相邻图层的过渡区域分界线较明显的问题。为了缓解边界线,纹理金字塔压缩算法在相邻图层之间预留一部分的重叠区域,并根据权值进行融合计算,但这一部分重叠区域降低了数据压缩速率;多尺度金字塔压缩算法选择在图层的交界处采用多尺度滤波图像,但相邻图层间高斯核半径的高差值又无法较好的切合人眼视敏度曲线,文中算法以多图层渐变方式代替下采样图层,既提高了压缩效率又切合了人眼视敏度曲线,更符合人眼主观感知。

综合分析可知,所提算法相较于纹理金字塔压缩算法和多尺度金字塔压缩算法,在相同的视距下具有更高的综合压缩比,可在进一步降低数据传输带宽的同时显示更高分辨率、更高刷新率的视频。

文中所提算法与纹理金字塔压缩算法成像效果对比图如图9所示。可以明显看出,文中所提算法成像效果明显好于纹理金字塔压缩算法。



(a) 原始图像 (b) 纹理金字塔

(c) 文中算法

图9 不同算法的效果对比图

Fig.9 Effects comparison of different algorithms

4 结 论

近年来,随着人机交互机制和沉浸式体验的发展,推动近眼显示设备向高分辨率和高刷新率的方向发展,为近眼显示设备的硬件带宽资源带来了挑战。
(下转第172页)