

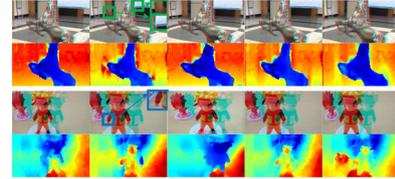


DOI: 10.12086/oe.2021.200186

基于网格形变的立体变焦视觉优化

周 莘, 柴雄力, 邵 枫*

宁波大学信息科学与工程学院, 浙江 宁波 315211



摘要: 立体图像的视觉变焦优化是近年来图像处理和计算机视觉领域热门的基础研究问题。3D图像的变焦视觉增强技术越来越受到关注。为此,本文从相机变焦拍摄的模型出发,提出一种基于网格形变的立体变焦视觉优化方法,力求提高3D立体视觉的体验。首先利用数字变焦方法模拟相机模型对目标区域进行适当的视觉放大,然后根据相机变焦距离建立起参考图像与目标图像之间的映射关系,并提取前景目标对象,接着使用改进型深度恰可察觉深度模型来引导目标对象的自适应深度调整。最后结合本文所设计的七个网格优化能量项,对图像网格进行优化,以提高该目标对象的视觉感,并确保整幅立体图像的良好视觉体验效果。与现有的数字变焦方法相比,所提出的方法在图像目标对象的尺寸控制方面和目标对象的深度调整方面都具有更好的效果。

关键词: 立体图像; 相机变焦; 改进型深度恰可察觉深度模型; 网格优化能量项

中图分类号: TP391

文献标志码: A

周莘, 柴雄力, 邵枫. 基于网格形变的立体变焦视觉优化[J]. 光电工程, 2021, 48(4): 200186

Zhou S, Chai X L, Shao F. Stereoscopic zoom for visual optimization based on grid deformation[J]. *Opto-Electron Eng*, 2021, 48(4): 200186

Stereoscopic zoom for visual optimization based on grid deformation

Zhou Shen, Chai Xiongli, Shao Feng*

College of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract: Stereoscopic image zoom optimization is a popular basic research problem in the field of image processing and computer vision in recent years. The zoom visual enhancement technology of 3D images has attracted more and more attention. To this end, this paper proposes a method of stereoscopic zoom vision optimization based on grid deformation from the model of camera zoom shooting, and strives to improve the experience of 3D stereoscopic vision. Firstly, use the digital zoom method to simulate the camera model to properly zoom in on the target area, and then establish the mapping relationship between the reference image and the target image according to the camera zoom distance. Secondly, extract the foreground target object and use the modified just noticeable depth difference (JNDiD) model to guide the adaptive depth adjustment of the target object. Finally, combined with the seven grid-optimized energy terms designed in this paper, the image grid is optimized to improve the visual perception of the target object and ensure a good visual experience for the entire stereoscopic image. Compared with the

收稿日期: 2020-05-27; 收到修改稿日期: 2020-07-20

基金项目: 国家自然科学基金资助项目(61622109); 宁波市自然科学基金项目(2017A610112)

作者简介: 周莘(1994-), 男, 硕士研究生, 主要从事三维立体图像视觉优化的研究。E-mail: 1134482159@qq.com

通信作者: 邵枫(1980-), 男, 教授, 博导, 主要从事三维视频信号编码与质量评价方面的研究。E-mail: shaofeng@126.com

版权所有©2021 中国科学院光电技术研究所

existing digital zoom method, the proposed method has better effects on the size control of the image target object and the depth adjustment of the target object.

Keywords: stereoscopic image; camera focal length; modified just noticeable depth difference model; grid-optimized energy term

1 引言

立体图像目标对象的视觉变焦增强技术越来越受到关注, 并且该技术在数字成像等领域有着重要的贡献。图像的视觉变焦是根据相机焦距的变化将目标对象放大的一个过程。但是对于立体图像而言, 目标对象的尺寸会随着变焦操作的变化而变化, 与此同时也会影响其深度和 3D 立体视觉体验(即视觉舒适度)。为此, 有许多专家学者在立体图像重定向方面对此进行了研究。一般的重定向操作是通过匹配目标设备的纵横比和分辨率^[1]将原图像进行重新定位。虽然重新定位或者缩放的目标与选择显示原始内容部分的目标对象理论上要保持相同, 但是重新定位未能有效地减少目标对象的信息丢失。Devernavy 等^[2]提出了基于形变的立体图像重组的方法, 该方法未能有效对目标对象进行保持和缩放, 并且未考虑立体视差的优化。Basha 等^[3]提出了基于立体图像 Seam Caving 的几何一致法, 其通过对接缝雕刻进行一般化处理, 以便在两个图像中能够同时雕刻一对接缝, 使外观和深度的失真最小化。接缝选择和视差图修改受到几何约束, 这些约束考虑了图像中遮挡像素和遮挡像素之间的可见性关系。Luo 等^[4]提出了基于补丁的几何一致性立体图像编辑模型, 其首先引入了与深度相关的补丁对的相似性度量, 以区分具有不同深度结构的图像内容。其次, 使用联合补丁对搜索法来处理两个视图之间的相关性。该方法可生成具有良好立体感和视觉质量的立体图像。Lee 等^[5]提出了基于图层的立体图像调整法, 它将输入的立体图像对分解为一组图层, 然后利用视差图创建相应图层, 接着采用基于 Warping 方法利用能量函数来指导每一层的图像变形, 最后形成调整后的图像。Yan 等^[6]提出了一种内容感知的立体网格变形模型, 该方法通过视差缩放来确定显著区域的缩放因子以调整显著对象。Niu 等^[7]利用美学规则裁剪和缩放立体图像来实现立体图像内容重组, 然而当对象目标比较分散并且背景有重要信息时, 裁剪使得重组后的图像失去了部分重要信息。

对于 3D 立体图像, 通常需要调整深度以适应不同显示设备的不同舒适区域或深度范围。因此, 有学

者在其工作中使用了立体品质因数对目标对象进行优化以调整视觉舒适度和深度感知^[8-12]。视差或深度调整最简单方法是在常规视差调整方法中增加或减少对象或整个场景的水平视差^[13], 缺点是由于深度不连续而不可避免地发生形变和能量损失。视差移动法旨在将深度或视差范围调整至所需范围内。Lei 等^[14]通过调整场景的最佳零视差平面(zero disparity plane, ZDP)以满足视觉舒适性约束。Park 等^[15]通过利用 3D 空间中的平面逼近来调整立体图像中对象的深度。Park 等^[16]直接调整了选定对象的坐标。在该方法中, 目标对象中的所有像素都投影到 3D 空间中的虚拟正面平行的平面上, 并对垂直于 3D 立体显示屏的移动投影像素进行调整。但是, 这些方法将不可避免地最终的立体图像中造成孔洞或遮挡。Yan 等^[17]提出了一种线性深度映射法, 其采用深度相干性能量项和运动相干性能量项来调整立体视频的深度范围。Sohn 等^[18]提出了一种串扰减少法, 其在不可校正区域内通过使用视差调整操作来替换可见串扰。Guan 等^[19]提出了一种基于几何感知的深度百分比度量标准, 其通过遮蔽场景而不是视差来描述不同观看条件下的深度感知。Oh 等^[20]提出了一种通过视差移位和散焦过滤的视觉舒适度提高方法, 以在不适感和临场感之间达到最佳平衡。但是, Yan 等^[17]、Sohn 等^[18]和 Guan 等^[19]的方法虽然考虑到了深度自适应的问题, 但未解决所生成的图像的视觉舒适度问题; Oh 等^[20]的方法虽然考虑了视觉舒适度与深度之间的问题, 但未能解决深度自适应的问题。

为了更好地提高 3D 视觉体验, 本文提出一种基于网格形变的立体变焦视觉优化的方法, 主要创新包括: 1) 利用相机焦距引导前景目标对象尺寸的调整, 并且在图像变焦后将目标对象的坐标点映射至原图相应位置的过程中, 结合坐标映射能量项、线保持能量项和边缘能量项来防止产生空洞等图像失真问题; 2) 使用改进型深度恰可察觉模型来自适应地调整图像目标对象与全局图像之间的深度, 从而提高图像视觉舒适度; 3) 综合利用七个能量项来保证前景目标对象与图像背景之间的良好过渡以提高能量优化后立体图像的视觉质量。

2 实验预处理

图 1 为优化的框架, 本文先将变焦后的图像进行目标对象显著性检测, 同时根据显著信息生成二值掩膜图像并结合变焦后的图像信息来提取该图像的前景目标对象。为了能更好地适应网格优化, 实验中建立起图像特征与网格的对应关系, 随后将变焦前后的目标对象的特征网格点进行映射。在特征网格坐标点映射的过程中使用七个能量项来优化目标对象与原图背景之间边缘网格、深度等重要信息, 最后绘制出结果输出图。

2.1 模拟变焦和深度与理想坐标的计算

为了建立起相机变焦前后目标对象坐标映射的关系, 实验中将计算变焦后图像的理想坐标, 并根据理想的目标对象坐标信息与原图坐标信息来实现精准匹

配。在典型的立体摄影系统中, 两个摄像机的光学中心之间的距离是基线, 如图 2(a)摄像机的焦距是 f 。实物点 P (位于投影中心的距离 z 处)在左右图像平面上投影, 产生投影点 p_L 和 p_R 。当焦距改变至 f' 时, 投影点相对之前的位置发生改变。此时两点之间的视差具有以下形式:

$$d(z) = p'_L - p'_R = B \cdot \frac{f'}{z'} \quad (1)$$

式中: B 为立体相机的基线距离, 在实验中模拟人的双目距离, z' 为该像素点变焦后的深度值, f' 为镜头变焦后的焦距。

本文用到的相机模型加入了景深信息。由于在景深范围内图像具有良好的清晰度, 因此实验中为了消除模糊因素的影响, 根据景深范围内的深度方向来实现变焦效果。如图 2(b), P_1 是实体处在正常拍摄位置

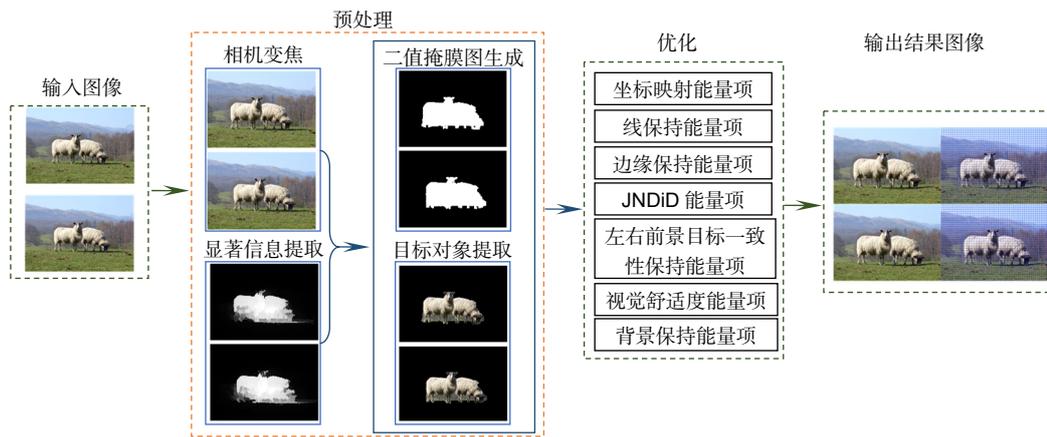


图 1 立体变焦视觉优化框架
Fig. 1 Stereo zoom optimization framework

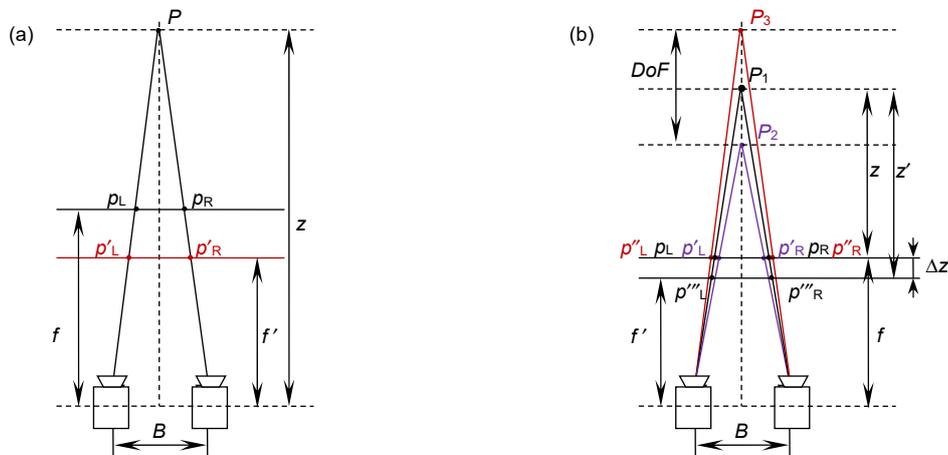


图 2 立体变焦系统。(a) 传统相机模型; (b) 本文中用到的相机模型
Fig. 2 Stereoscopic zoom system. (a) Traditional camera model; (b) Camera model used in this article

时, p_L 与 p_R (黑色)是左右相机聚焦在该位置时, 图像视点在相机屏幕上的投影点; 而当实体 P_2 处在前景深的位置时, p'_L 与 p'_R (紫色)是左右相机聚焦在该位置时, 图像视点在相机屏幕上的投影点; 而当实体 P_3 处在后景深的位置时, p''_L 与 p''_R (红色)是左右相机聚焦在该位置时, 图像视点在相机屏幕上的投影点。当改变焦距后, 形成图像投影点 p'''_L 与 p'''_R , 对应的深度 z 会随之变化。深度与焦距之间的关系:

$$z = \frac{Bf}{d + d_{\text{offs}}}, \quad (2)$$

式中: z 为原始深度, f 为原始焦距, d 为像素视差, d_{offs} 为相机主点差异。

在图 2(b)中相机对聚焦物体进行放大, 但是由于相机镜头缩小而获得的图像不能在同时刻同场景下获得背景先验信息, 因为这会导致图像边缘的信息丢失, 所以本实验只考虑基于相机镜头放大下的图像实验。

根据上述参数以及在景深范围内映射的目标对象深度, 可获取该物体对象的焦距值:

$$f' = \frac{z' \cdot (d + d_{\text{offs}})}{B}, \quad (3)$$

式中 z' 为在景深范围内的目标对象深度。

根据对象映射前后深度的变化, 可求得映射后的视差深度。具体公式:

$$\Delta z = |z - z'|, \quad (4)$$

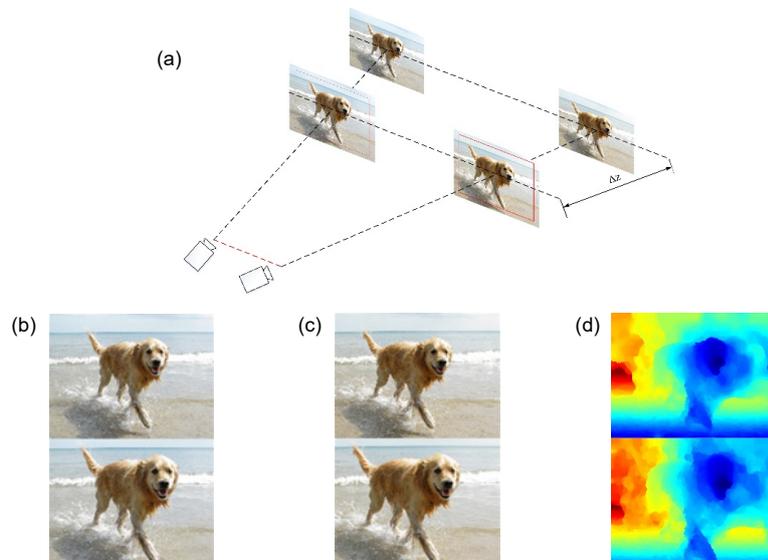


图 3 Δz 变化下图像与视差结果。

(a) Δz 变化的相机模型; (b) 原始左右图像; (c) 变焦后的左右图像; (d) 对应左右图像的视差图

Fig. 3 Image and parallax results under Δz changes.

(a) Camera model under Δz change; (b) Original images; (c) Zoomed images; (d) Corresponding disparity images

$$d'(\Delta z) = p'''_L - p'''_R = e \cdot \frac{f'}{z + \Delta z}, \quad (5)$$

式中 p'''_L 与 p'''_R 为变焦后对应于该焦距下的图像投影点。

本实验在相机镜头放大的过程中会引起目标对象深度的变化。但是由于显示屏幕尺寸保持不变, 与变焦前的图像相比, 焦距的增加通常可以捕捉到更加清晰的细节。根据 Δz 变化的视觉结果如图 3 所示。

实验中在物理深度方向上考虑了景深的因素, 同时实现相机镜头变焦的过程。该过程在保持相同分辨率时使用图像的中心作为主轴并以此作为 Δz 变化的方向。当理想图像平面上的像素 (x, y) 被投影到图像平面上时, 新的水平和垂直坐标 (\bar{x}, \bar{y}) 可被计算为

$$\bar{x} = \frac{W}{2} - \left(\frac{W}{2} - x \right) \cdot \left(\frac{z}{z - \Delta z} \right), \quad (6)$$

$$\bar{y} = \frac{H}{2} - \left(\frac{H}{2} - y \right) \cdot \left(\frac{z}{z - \Delta z} \right), \quad (7)$$

式中 W 和 H 是图像的宽度和高度。

2.2 显著检测和对象分割

本文采用根据立体图像显著信息用以提取目标对象的思路, 在实验中使用立体图像显著计算方法^[21], 该方法利用图像的边缘和立体视差信息来判别图像显著性特征, 同时兼顾左右图像显著特征的一致性, 随后使用目标对象提取算法^[22]以获得变焦后的目标对象, 效果如图 4 所示。



图 4 对象分割。(a) 经变焦后得到的图像; (b) 图像的显著目标; (c) 目标分割
Fig. 4 Object segmentation. (a) Image after zooming; (b) Salient target of the image; (c) Object segmentation

2.3 景深计算

本文用改进型深度恰可察觉模型调节图像深度, 其中景深范围为该模型的关键参数。改进型深度恰可察觉模型中的虚拟视距由真实视距与景深距离决定, 同时为了确定该模型中对应的深度阈值, 实验中需要对景深范围值作进一步计算。

根据图 5 中相机模型可知, 景深距离可由近点和远点决定, 即该数值可计算为近点允许弥散圆中心与远点允许弥散圆中心之间的距离。焦点与近点之间的距离称为前景深, 前景深距离公式:

$$\Delta L_1 = \frac{F\delta L^2}{f^2 + F\delta L}, \quad (8)$$

式中: F 为光圈值, δ 为允许弥散圆直径, L 为拍摄物的距离。

根据图 5 中相机的模型也可得知, 焦点与远点之间的距离, 即该数值可计算为焦点与远点弥散圆中心之间的距离称为后景深。后景深公式:

$$\Delta L_2 = \frac{F\delta L^2}{f^2 - F\delta L}, \quad (9)$$

景深距离范围公式:

$$\Delta L = \Delta L_1 + \Delta L_2 = \frac{2F\delta f^2 L^2}{f^4 - F^2 \delta^2 L^2}, \quad (10)$$

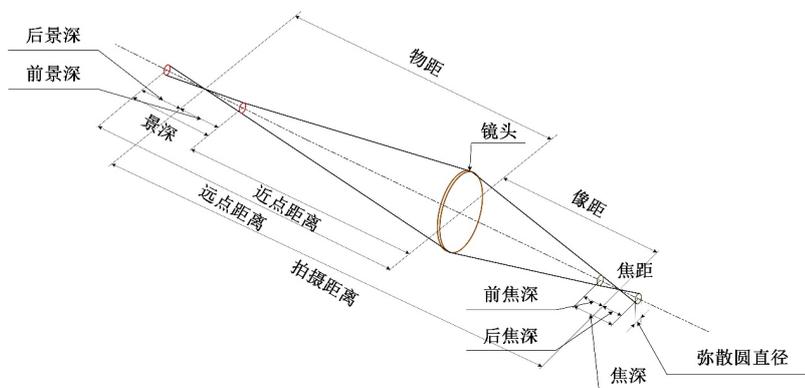


图 5 相机拍摄的景深模型

Fig. 5 Depth of field model

2.4 改进型深度恰可察觉模型

在 Li 等人^[24]的工作中, 改进型深度恰可察觉模型中的恰可察觉阈值是由基础阈值 Δd_{base} 和附加阈值 $\Delta d_{additional}$ 构成的。基础阈值是受现实情况下的视距影响, 而附加阈值则受景深距离影响^[23], 对于景深范围内物体的深度阈值可通过结合虚拟视距计算得到, 虚拟视距计算式:

$$v' = v + \Delta L, \quad (11)$$

式中: v 为视距, ΔL 为景深距离范围。

最终的改进型深度恰可察觉阈值表达式:

$$\Delta d_{NDID}(v') = \Delta d_{base}(v) + \Delta d_{additional}(v'), \quad (12)$$

式中基础阈值和附加阈值分别表示为

$$\Delta d_{base}(v) = 0.056 \cdot v' - 0.031v, \quad (13)$$

$$\Delta d_{additional}(v') = \begin{cases} -0.1 \cdot (v' - d_n) + \xi & v' < d_n \\ 0 & d_n < v' < d_f \\ -0.03 \cdot (v' - d_f) & v' > d_f \end{cases}, \quad (14)$$

式中: d_f 、 d_n 为中心深度区域的上下界值, 而 $\xi = 0.056 \cdot (v - d_n)$ 。

为了适应三个深度区域中的不同深度效果, 将恰可察觉阈值定义为三段线性函数, 从而使其适用于具有特定模型参数的不同类型的显示器。最终可得:

$$\Delta d_{\text{JNDID}} = \begin{cases} 0.056v' - 0.031v - 0.1(v' - d_n) + \xi & v' < d_n \\ 0.056v' - 0.031v & d_n \leq v' \leq d_f \\ 0.056v' - 0.031v - 0.03(v' - d_f) & v' > d_f \end{cases} \quad (15)$$

3 网格能量项的优化

实验参考 Shao 等^[11]的优化工作。首先在左右立体原图像中创建一组尺寸大小为 30×30 的常规网格, 并将各个网格顶点定义为 $\mathbf{u}_{i,L}$ 与 $\mathbf{u}_{i,R}$, 对于每个顶点有 $\mathbf{u}_{i,L} = (x_{i,L}, y_{i,L})$ 与 $\mathbf{u}_{i,R} = (x_{i,R}, y_{i,R})$, 其中 $x_{i,L}$ 与 $y_{i,L}$ 为左图像的水平与垂直坐标, $x_{i,R}$ 和 $y_{i,R}$ 为右图像水平和垂直坐标。根据式(6)、式(7)可以计算获得理想坐标点, 可各自表示为 $\hat{\mathbf{u}}_{i,L}$ 与 $\hat{\mathbf{u}}_{i,R}$ 。

3.1 坐标映射能量项

根据左右目标坐标与理想坐标可得:

$$E_{\text{coor}_L} = \sum_{\mathbf{u}'_{i,L} \in N_L} \|(\mathbf{u}'_{i,L} - \hat{\mathbf{u}}_{i,L})\|^2, \quad (16)$$

$$E_{\text{coor}_R} = \sum_{\mathbf{u}'_{i,R} \in N_R} \|(\mathbf{u}'_{i,R} - \hat{\mathbf{u}}_{i,R})\|^2, \quad (17)$$

式中: \mathbf{u}'_i 为数字变焦后的目标网格点, $\hat{\mathbf{u}}_i$ 为目标对象理想坐标上的网格点, N_L 与 N_R 分别为左右图像目标对象上所有网格点。左、右图像坐标映射总能量为

$$E_{\text{coor}} = \sum_{\mathbf{u}'_{i,L} \in N_L} \|(\mathbf{u}'_{i,L} - \hat{\mathbf{u}}_{i,L})\|^2 + \sum_{\mathbf{u}'_{i,R} \in N_R} \|(\mathbf{u}'_{i,R} - \hat{\mathbf{u}}_{i,R})\|^2. \quad (18)$$

3.2 线保持能量项

在映射图像目标对象网格顶点的坐标过程中, 有时其网格顶点所在的网格线会发生变形, 从而导致图像目标对象发生形变。为了防止网格线过度形变, 需要进行网格线的保持。能量项如下所示:

$$E_{\text{line}}(\hat{\mathbf{u}}'_i, \hat{\mathbf{u}}'_j) = \|(\hat{\mathbf{u}}'_i - \hat{\mathbf{u}}'_j) - S_e(\hat{\mathbf{u}}_i - \hat{\mathbf{u}}_j)\|^2, \quad (19)$$

式中: $S_e = (\mathbf{e}_{ij}^T \mathbf{e}_{ij})^{-1} \mathbf{e}_{ij}^T \hat{\mathbf{e}}_{ij}$, 原始边 $\mathbf{e}_{ij} = \hat{\mathbf{u}}_i - \hat{\mathbf{u}}_j$ 为两个顶点的集合之间的差值, 变形边缘 $\hat{\mathbf{e}}'_{ij} = \hat{\mathbf{u}}'_i - \hat{\mathbf{u}}'_j$ 为顶点的集合之间的差值。 $\hat{\mathbf{u}}_i$ 和 $\hat{\mathbf{u}}'_i$ 分别表示左右图像整体映射前后的网格点。左图像和右图像的线保持总能量为

$$E_{\text{line}} = \sum_{\langle \mathbf{u}'_{i,L}, \mathbf{u}'_{j,L} \rangle \in N'_L} E_{\text{line}}(\mathbf{u}'_{i,L}, \mathbf{u}'_{j,L}) + \sum_{\langle \mathbf{u}'_{i,R}, \mathbf{u}'_{j,R} \rangle \in N'_R} E_{\text{line}}(\mathbf{u}'_{i,R}, \mathbf{u}'_{j,R}), \quad (20)$$

式中 N'_L 与 N'_R 分别为左右图像上所有映射后的网格点。

3.3 改进型深度恰可察觉能量项

为了对图像中不同深度的对象进行处理, 实验采用改进型深度恰可察觉深度模型(即式(15))对处于不

同深度的对象进行恰可察觉深度化处理。该模型通过设置深度阈值 d_f 和 d_n 将图像的视差深度域进行划分, 从而起到恰可察觉深度的目的。

根据景深范围, 用 Δd_{JNDID} 阈值来区分景深范围内目标对象之间的深度, 达到恰可察觉深度的目的, 阈值可由式(15)得到。基于改进型深度恰可察觉深度模型和图像目标对象与其周围区域的深度用来构造能量项 E_{JNDID} , 如下所示:

$$E_{\text{JNDID}} = \sum_{O_i} (\Delta d_{\text{JNDID}} - |\tilde{D}_O - D_{O_i}|)^2, \quad (21)$$

式中: O_i 为目标对象区域, D_O 为图像中目标对象的深度值, D_{O_i} 为图像目标对象周围的区域的深度值。

3.4 边界保持能量项

在目标对象网格顶点映射过程中, 由于网格线的形变会导致图像中的目标对象发生扭曲。为了防止这种现象发生, 实验采用了过渡平滑的方法使相邻网格形变一致, 故引入边界保持能量项, 该能量项:

$$E_{\text{edge}}(\mathbf{u}''_i) = \|(\mathbf{u}''_i - \tilde{\mathbf{u}}'_i)\|^2, \quad (22)$$

式中: \mathbf{u}'' 表示图像中前景目标对象映射后的图像目标对象坐标顶点集合, $\tilde{\mathbf{u}}'$ 表示目标映射前图像目标对象顶点集合。边界保持总能量为

$$E_{\text{edge}} = \sum_{\mathbf{u}''_{i,L} \in C_L} E(\mathbf{u}''_{i,L}) + \sum_{\mathbf{u}''_{i,R} \in C_R} E(\mathbf{u}''_{i,R}), \quad (23)$$

式中 C_L 与 C_R 分别表示立体左右图中目标对象的边界区域。

3.5 左右前景目标一致性保持能量项

将左图像和右图像同时进行优化时, 可能会出现左图像和右图像的目标对象区域不一致的情况, 从而引起观看者的不适。因此需要把左右图像的目标对象的位置确定下来, 即需要保持左右视图的目标对象区域一致性变化。本实验中采用 SIFT 匹配的方法, 将左图像和右图像中的目标对象建立起对应的关系。首先用 SIFT 选取出关键点 f , 可以表示为围绕它的四个顶点的线性组合:

$$f = \sum_{i=1}^4 \beta_i \mathbf{u}_i, \quad (24)$$

式中: β_i 表示坐标网格四个顶点的距离权重。实验中考虑两个维度的目标对象网格差异, 即水平差异与垂直差异。优化过程中, 我们使用水平差异一致性与垂直对准这两个能量项并从两个维度来保持左右前景目标对象一致性。水平差异一致性能量定义为

$$E_D = \sum_{f_{i,L}, f_{i,R} \in P} ((f_{i,R}[x] - f_{i,L}[x]) - \hat{d}_i)^2, \quad (25)$$

式中: P 为左图像与右图像的目标对象中的所有关键顶点的集合, $f_{i,L}$ 与 $f_{i,R}$ 分别为左右图像中的关键点。符号 $[x]$ 代表提取 2D 矢量的 x 分量, \hat{d} 为计算的理想视差, 当相机深度变化 Δz 时, 计算式:

$$\hat{d} = \frac{Bf}{Z - \Delta z} - d_{\text{offs}}, \quad (26)$$

而垂直对准能量则定义为

$$E_V = \sum_{f_{i,L}, f_{i,R} \in P} (f_{i,R}[y] - f_{i,L}[y])^2, \quad (27)$$

式中: 符号 $[y]$ 代表提取 2D 矢量的 y 分量。左图像与右图像前景一致性保持的总能量如下:

$$E_{\text{cons}} = E_D + E_V, \quad (28)$$

式中 $f_{i,L}$ 与 $f_{i,R}$ 为左图像与右图像匹配的特征点所在的网格面内的关键网格点。

3.6 视觉舒适度能量项

实验过程中由于难以直接控制网格顶点以控制目标对象的深度。在本步骤中深度控制的目的是调整关键点的深度, 再根据线性深度函数调整关键点周围顶点的深度。根据对象中的所有关键点的平均值, 计算对象的平均深度:

$$\bar{Z} = \frac{1}{n} \sum_{i \in \mathcal{Z}} \tilde{Z}_i, \quad (29)$$

式中: \mathcal{Z} 是所选对象中所有关键点的集合, \tilde{Z}_i 是所选对象中每个像素的目标深度值, n 是对象中关键点数量。设 Z_N 和 Z_F 是所选对象中最近和最远的深度^[23], Z_{min} 和 Z_{max} 是临界 -1° 和 1° 的感知深度^[10], 对于平均深度低于 -1° 的情况, 线性深度函数定义为

$$f(\tilde{Z}) = q^- \cdot (\tilde{Z} - Z_N) + Z_N, \quad (30)$$

式中: \tilde{Z}_i 代表目标深度值, Z_F 与 Z_N 分别为目标对象的最近和最远深度, q^- 表示为

$$q^- = \frac{L_D - Z_{\text{min}}}{L_D - Z_N}, \quad (31)$$

式中: L_D 为视觉距离(即视距)。而对于平均深度大于 1° 的情况, 线性深度函数定义为

$$f(\tilde{Z}_i) = Z_F - q^+ \cdot (Z_F - \tilde{Z}_i), \quad (32)$$

$$q^+ = \frac{Z_{\text{max}} - L_D}{Z_F - L_D}. \quad (33)$$

线性深度函数可自适应地调整目标对象及其周围背景点的深度, 从而保持目标对象舒适度。虽然通过上述的线性深度函数可将目标对象深度调至舒适度范围, 但与理想的深度值还存在一定的差异。为了减少这种微小的深度差异, 在舒适度能量项设计时, 实验在实现过程中赋予目标像素点较高的权重值, 同时结

合线性深度函数所获得的深度值来增强深度感。舒适度能量项设计如下:

$$E_{\text{comf}} = \sum_{i \in \Omega} \omega(\tilde{Z}_i) \cdot (\hat{Z}_i - f(\tilde{Z}_i))^2. \quad (34)$$

式中: Ω 为所有目标区域中的关键点, \hat{Z}_i 为目标区域的预期深度值, $\omega(\tilde{Z}_i)$ 如下:

$$\omega(\tilde{Z}_i) = \exp\left(\frac{|\tilde{Z}_i - L_D|}{Z_F - Z_N}\right). \quad (35)$$

3.7 背景保持能量项

在前景目标对象映射并移动的过程中, 为了减小对背景区域的影响, 需要保持背景坐标不变。左图像与右图像能量项如下所示:

$$E_{\text{back}_L} = \sum_{\mathbf{u}_{B_i,L} \in B_L} \|(\bar{\mathbf{u}}_{B_i,L}'' - \mathbf{u}_{B_i,L})\|^2, \quad (36)$$

$$E_{\text{back}_R} = \sum_{\mathbf{u}_{B_i,R} \in B_R} \|(\bar{\mathbf{u}}_{B_i,R}'' - \mathbf{u}_{B_i,R})\|^2, \quad (37)$$

式中: B_L 与 B_R 为左右立体图像的背景区域, $\bar{\mathbf{u}}_{B_i,L}''$ 与 $\bar{\mathbf{u}}_{B_i,R}''$ 为左右图像目标对象优化后的背景坐标集, $\mathbf{u}_{B_i,L}$ 与 $\mathbf{u}_{B_i,R}$ 为左右图像背景的原始坐标集, 其总能量为

$$E_{\text{back}} = \sum_{\mathbf{u}_{B_i,L} \in B_L} \|(\bar{\mathbf{u}}_{B_i,L}'' - \mathbf{u}_{B_i,L})\|^2 + \sum_{\mathbf{u}_{B_i,R} \in B_R} \|(\bar{\mathbf{u}}_{B_i,R}'' - \mathbf{u}_{B_i,R})\|^2. \quad (38)$$

3.8 总能量

根据前面所述, 实验中总的优化能量项由坐标映射能量项、线保持能量项、改进型深度恰可察觉深度能量项、边界保持能量项、左右前景目标一致性保持能量项、视觉舒适度能量项和背景保持能量项分别乘以对应的权重系数而得到。总的能量优化公式:

$$E_{\text{total}} = \lambda_1 E_{\text{coor}} + \lambda_2 E_{\text{line}} + \lambda_3 E_{\text{NDDiD}} + \lambda_4 E_{\text{edge}} + \lambda_5 E_{\text{cons}} + \lambda_6 E_{\text{back}} + \lambda_7 E_{\text{comf}}. \quad (39)$$

式中: λ_1 、 λ_2 、 λ_3 、 λ_4 、 λ_5 、 λ_6 和 λ_7 为对应能量项的权重。最后, 通过求解最小二乘线性方程计算出最佳的网格点坐标。

4 实验结果与分析

4.1 实验结果分析

为了验证本文提出的方法, 进行了相关实验。本实验用到的立体图像来源于 NBU_VCA 图库、CVPR2012 图库和 S3DImage 图库。本次实验中摄像机镜头的参数值均为普通相机正常拍摄下的相机参数。拍摄光圈值设置为 5.6 mm, 允许弥散圆的直径设置为 0.035 mm, 视距 2500 mm, 相机初始焦距设为 250 mm,

变焦后焦距为 260 mm, 根据相机参数可计算变焦前的景深范围大小为 39.202 mm。本次实验中改进型深度恰可察觉深度模型在深度上的阈值为 19, 27, 36。本文为体现各个能量项的通用性, 实验中通过增加实验次数来获得对应权重值, 最后将其设置为 2、1、4、2、30、3、2, 如表 1 所示。

表 1 不同能量项的权重

Table 1 Weights of different energy terms

权重参数	λ_1	λ_2	λ_3	λ_4	λ_5	λ_6	λ_7
权重值	2	1	4	2	30	3	2

图 6 是本次实验结果, 可以看出图像中的前景目标对象在优化过后得到了放大, 目标对象的深度范围也得到了调整, 保持了较好的视觉效果。

本文方法中各项能量项参数是通过多次实验数据

获得的, 而实验中的每对立体图像尺寸不一且视差深度有差异, 为了避免网格顶点优化过程中的图像出现空洞, 本文将参数值调至最优。为了反映出本文方法在视觉上的优势, 本文增设了在不同焦距下优化目标对象图像的实验, 效果如图 7 所示。

4.2 目标显著性检测对实验结果的影响

目标显著性检测在本算法中是关键性的步骤。目标显著性检测的好坏直接影响目标二值图像的生成, 而二值图像可区分图像中的像素点是否在目标区域内。在优化处理过程中, 特别是在目标边缘区域, 如果显著性检测方法未检测到该区域的显著信息, 从而使未检测到的显著区域与检测到的显著区域非等比例放大, 进而影响视觉效果。

图 8 中的结果可以看出, 两种图像目标显著性检测的结果差距过大。结果较差的方法检测的显著

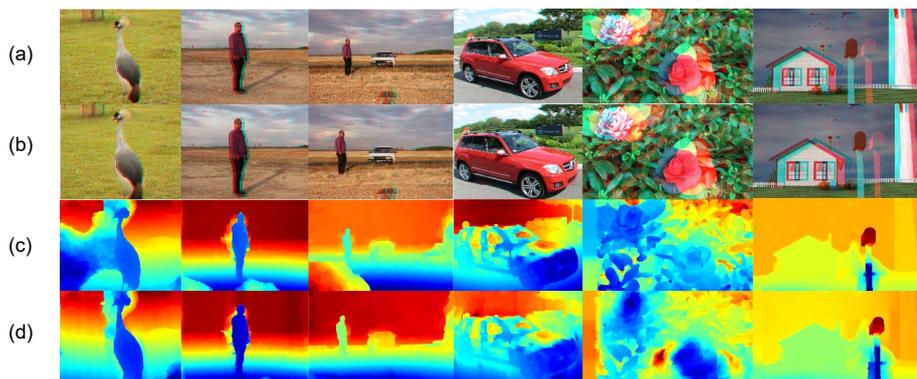


图 6 实验结果。(a) 原图像红蓝图; (b) 本文实验红蓝图; (c) 原图像视差图; (d) 本文实验视差图
Fig. 6 Result of experiment. (a) Original red-cyan anaglyph; (b) Red-cyan anaglyph after optimization; (c) Original disparity; (d) Proposed disparity

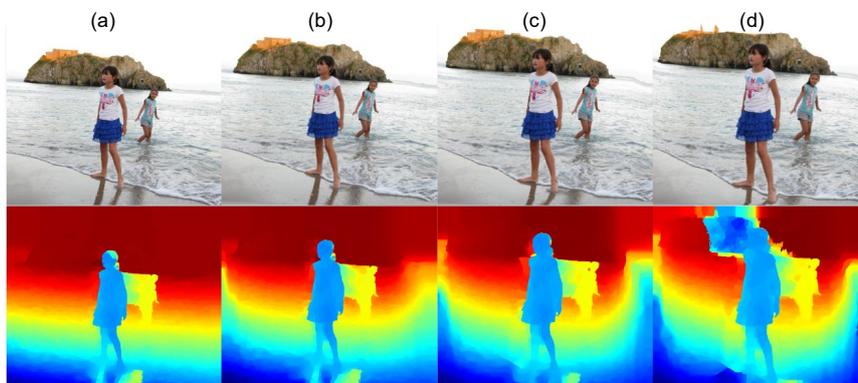


图 7 不同焦距下的目标对象。(a) 原图像及其视差图; (b) 焦距为 260 mm 的图像及其视差图; (c) 焦距为 270 mm 的图像及其视差图; (d) 焦距为 280 mm 的图像及其视差图
Fig. 7 The target object at different focal distances. (a) Original image and its disparity map; (b) The image with focal distance at 260 mm and its disparity map; (c) The image with focal distance at 270 mm and its disparity map; (d) The image focal distance at 280 mm and disparity map

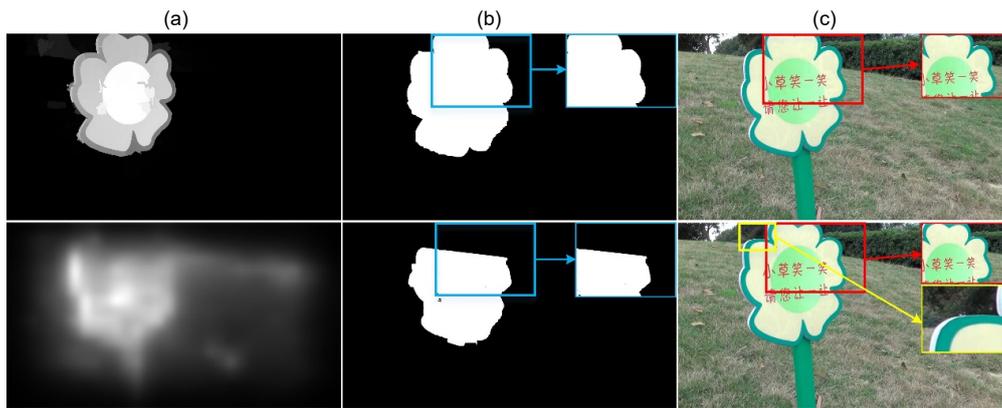


图 8 目标显著性检测对比实验。(a) 显著信息检测结果; (b) 二值图像; (c) 对应的优化结果

Fig. 8 Contrast experiment of saliency detection.

(a) Image of saliency detection result; (b) Binary image; (c) Corresponding optimization result

信息较为分散,无法检测出显著性目标对象,从而使得生成的二值图像结构不完整。从图 8(b)第二行结果可以看出,由于显著性检测效果较差导致其生成的目标二值图像有残缺,从而导致其优化后的结果上半部分过小(红色区域)(图 8(c))。黄色区域的信息失真也是由二值图像信息残缺的结果引起的,因为未检测到的区域无法与检测到的区域同比例放大。

4.3 对比实验

为了验证本文方法的效果,加入三组对比实验进行测试。图 9 展示了 3D 渲染^[24]、普通伸缩变换方法、控制深度方法^[15]和本文方法。本小节主要对图像的内容尺寸控制和深度控制两方面进行分析。

1) 内容尺寸控制: 3D 渲染方法的虚拟视点绘制中虽然对图像对象大小有调整,但是根据实验效果可

以看出,第一次实验中对背景保护得不够好,如图中背景的门发生弯曲;第二次实验中图像目标对象的手臂上出现了失真。其主要原因是该方法不需要提取目标对象,在对目标对象的优化过程中强行将前景与背景结合从而导致了图像的失真。伸缩变换方法优化过程中未考虑到保持与原图有相同的尺寸,仅仅通过伸缩来放大图像中的目标对象。该方法仅在图像整体的尺寸上进行了线性调整,而目标对象很明显也是跟着图像整体尺寸的变化而变化。而 Lei 等人^[15]所提出的方法从实验结果上可以看出,背景信息约束性较强,导致其目标放大的方法不具有一定的优势。本文方法在控制目标对象大小并兼顾图像目标对象与背景之间过渡的方面有一定的优势。

2) 深度控制: 3D 渲染方法与 Lei 等人所提出的方

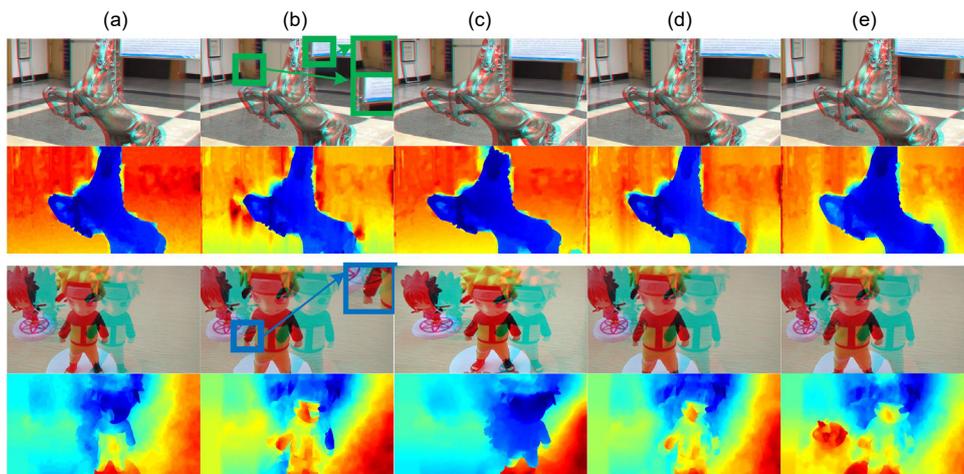


图 9 两组立体图像序列的对比实验。

(a) 源图像; (b) 3D 渲染方法; (c) 伸缩变换方法; (d) 深度控制方法; (e) 本文方法

Fig. 9 Comparative experiment of two sets of stereoscopic image sequences.

(a) Source image; (b) 3D rendering; (c) Telescopic transformation; (d) Depth control; (e) Proposed method

法在目标对象深度控制方面未考虑到目标对象之间的深度关系,无法做到目标对象之间的深度自适应。由于伸缩变换是将图像内容按其线性比例进行缩放,从而其深度也是线性地调整,同样也无法实现目标对象之间的深度适应。本文所提出的方法运用了改进型深度恰可察觉深度模型,该模型根据深度阈值可自适应地控制目标对象之间的深度。

目前现有其他的图像质量评价指标难以对基于相机变焦操作下的目标尺寸优化图像进行评价,所以为了客观地评价本文方法的优越性,实验将本文方法和三种其他方法所得结果进行逆序操作,使得逆序结果目标尺寸与原图目标尺寸一致。本文方法的逆序操作是依据变焦前后的参数并利用网格顶点信息再对坐标进行一次映射以还原目标对象,3D渲染方法和深度控制方法则以建立前后视差深度关系的方式还原目标对象,而伸缩变换方法则以尺度变换的方式还原目标对象。从理论上讲,实验重建结果图像应该与原图参考图像的信息一致,因此在实验中引入PSNR与SSIM指标用以评价实验结果图像的质量。实验将图像库分成NBU_VCA 图库、CVPR2012 图库和 S3DImage 图库三类,每一个类选 20 对图像进行测试,最后将每个类的指标求平均,如表 2 所示。本文方法与对比方法的

PSNR 和 SSIM^[25]得分不高,这是因为本文方法在逆序过程中目标对象的网格又进行一次映射,在调整网格坐标信息的过程中可能出现偏差,从而导致图像的失真;3D渲染方法与深度控制方法则在视察深度建立关系时出现误差从而导致失真;伸缩变换法则是在尺度变化时就已产生失真。

4.4 不同能量项组合

为了验证能量项的有效性,本小节对文中所述相关能量项进行组合并用实验验证效果。实验过程中各网格约束能量项将网格顶点转化为对应矩阵,然后将各矩阵级联成最小二乘法的输入参数,最后求解网格优化顶点。本节对不同的能量项组合进行了实验,如表 3 所示。实验中用某一个能量项同等维度下的零矩阵来代替该移除的能量项,而对应的权重系数值不变。具体能量项组合实验设计如下:

1) 方案二中的坐标映射能量项是根据变焦前后目标对象坐标点之间的映射关系来进行优化的。缺少了该能量项后就是缺少了坐标映射关系,其结果会出现目标对象坐标匹配错误,效果如图 10 所示。

2) 方案三中的线保持能量项是由相邻网格之间坐标点的连线来表示。设置该能量项是防止目标对象

表 2 通过客观指标 PSNR 与 SSIM 比较四种实验方法

Table 2 Comparison of four experimental methods through objective indicators PSNR and SSIM

测试图像库	3D 渲染		伸缩变换		深度控制		本文方法	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
NBU_VCA 图库	12.62	0.3987	12.97	0.4786	12.61	0.4103	14.94	0.5062
CVPR2012 图库	19.77	0.4453	19.68	0.5237	19.73	0.4523	20.32	0.5519
S3DImage 图库	19.09	0.4227	18.75	0.4913	21.34	0.4937	22.66	0.5732
Average	17.16	0.4222	17.13	0.4979	17.89	0.4521	19.31	0.5438

表 3 不同能量项组合

Table 3 Combination of different energy terms

方案	能量项组合
方案一	坐标映射能量项+线保持能量项+改进型恰可察觉深度能量项+边界保持能量项+左右前景目标一致性保持能量项+视觉舒适度能量项+背景保持能量项
方案二	方案一中除去坐标映射能量项
方案三	方案一中除去线保持能量项
方案四	方案一中除去改进型恰可察觉深度能量项
方案五	方案一中除去边界保持能量项
方案六	方案一中除去左右前景目标一致性保持能量项
方案七	方案一中除去视觉舒适度能量项
方案八	方案一中除去背景保持能量项

与背景之间的区域过渡形变。缺少该能量项会使目标对象内容随网格坐标点发生偏移, 从效果上看会有空洞出现, 如图 11 所示。

3) 方案四中由于缺少改进型恰可察觉深度能量项, 所以无法使用改进型恰可察觉深度阈值来调节深度感知的目标深度。改进型恰可察觉深度阈值是由视距(即变焦距离)决定, 根据式(15), 现实物体的景深范围距离落入改进型恰可察觉深度模型中的不同深度范围, 该模型会根据视差深度值选取对应阈值。从图 12 中可以看出, 缺少改进型恰可察觉深度能量项在前景的感知深度范围上没有起到调整作用。

4) 方案五中的边界保持能量项是用来保护目标对象的边界信息, 该能量项通过对应变焦前后目标对象的边界网格点坐标来进行优化。依据原图像中目标对象边缘的网格点信息的参考和约束, 可保证在变焦后目标对象的形状, 其效果如图 13 所示。

5) 方案六中左右前景目标一致性能量项是基于 SIFT 算法选出左右图像的关键点并结合理想视差而得到的, 而这过程中保持左右图中目标对象内容一致性是关键。该能量项的缺失会导致左右视图中目标对象内容无法得到很好的匹配。从图 14 中可以看出, 缺失了该能量项后左右图像的目标对象形状大小不一, 而加上该能量项后可有效地保持左右目标对象形状。

6) 方案七中视觉的舒适度是通过视差深度信息来呈现。实验中赋予了目标对象较高的权重值, 同时用深度线性函数来调整目标深度, 而在未加视觉舒适度能量项的结果中可以看出目标对象与背景没有较好的区分度, 深度方面也没有很好的层次感, 效果如图 15 所示。

7) 方案八中由于前景目标对象根据焦距发生了尺寸上的变化, 进而与背景区域发生相对移动, 从而使得边缘像素点发生重叠而未能有效地保留。背景保

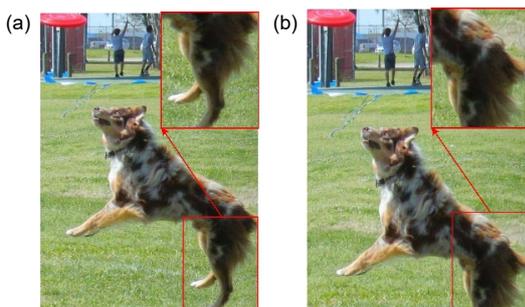


图 10 是否包含坐标映射能量项对比。

- (a) 包含坐标映射能量项的结果;
- (b) 未包含该能量项的结果

Fig. 10 Comparison of whether the coordinate mapping energy term is included. (a) Contain the coordinate mapping energy term; (b) This energy term is not included

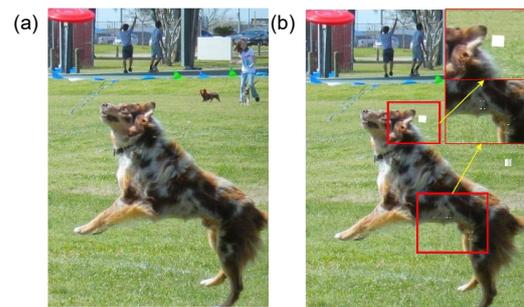


图 11 是否包含线保持能量项对比。

- (a) 包含线保持能量项的结果;
- (b) 未包含该能量项的结果

Fig. 11 Comparison of whether the line retention energy term is included. (a) Contain the line retention energy term; (b) This energy term is not included



图 12 是否包含改进型恰可察觉深度能量项对比。

- (a) 包含改进型恰可察觉深度能量项的结果;
- (b) 未包含该能量项的结果

Fig. 12 Comparison of whether the modified just noticeable depth difference energy term is included. (a) Contain the modified just noticeable depth difference energy term; (b) This energy term is not included

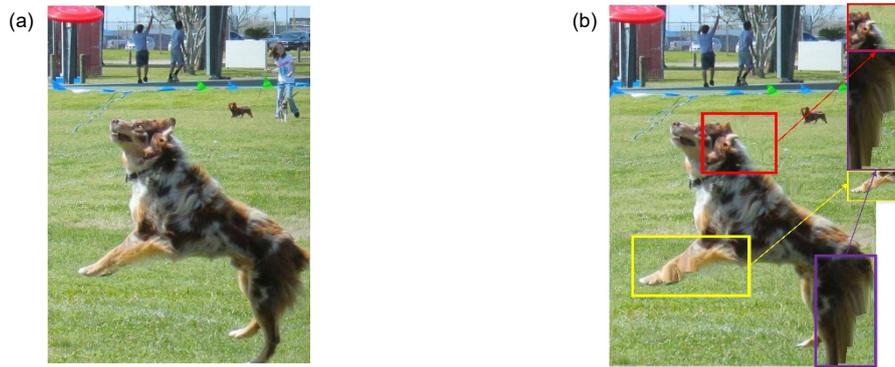


图 13 是否包含边界保持能量项对比。(a) 包含边界保持能量项的结果; (b) 未包含该能量项的结果
Fig. 13 Comparison of whether the boundary retention energy term is included.
(a) Contain the boundary retention energy term; (b) This energy term is not included

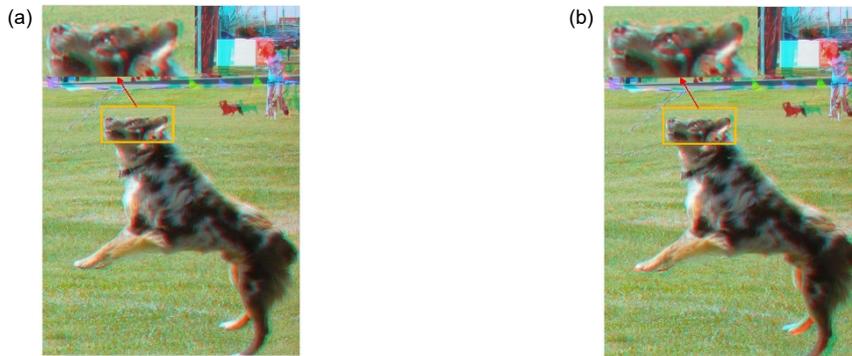


图 14 是否包含前景目标一致性能量项对比。
(a) 包含前景目标一致性能量项的结果; (b) 未包含该能量项的结果
Fig. 14 Comparison of whether the foreground consistent energy term is included.
(a) Contain the foreground consistent energy term; (b) This energy term is not included

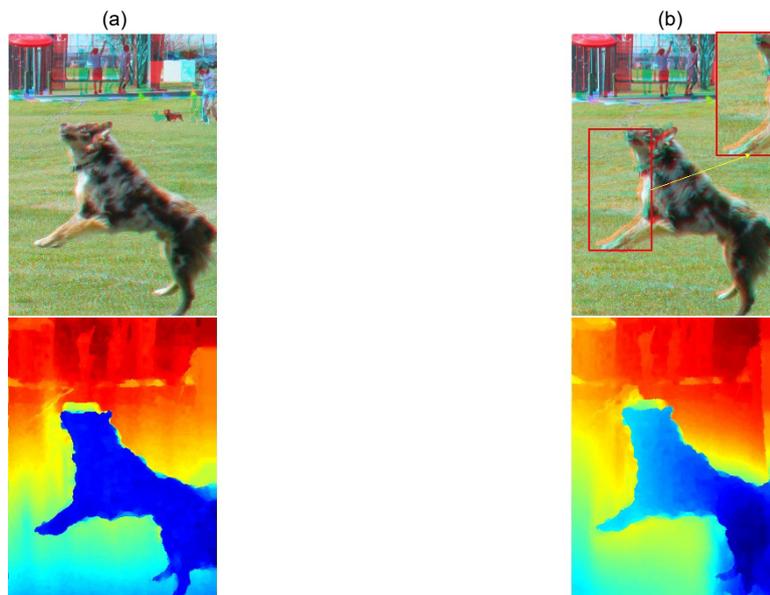


图 15 是否包含视觉舒适度能量项对比。
(a) 包含了视觉舒适度能量项的红蓝图以及视差图的结果; (b) 未包含该能量项的红蓝图以及视差图的结果
Fig. 15 Comparison of whether the visual comfort energy term is included. (a) Included the red-cyan anaglyph and the disparity images of the visual comfort energy term; (b) The red-cyan anaglyph and the disparity images without the energy term

持能量项考虑目标对象尺寸变化前的背景像素点信息, 并建立起同一目标对象区域周围的背景像素点在尺寸变化前后的关系, 通过最小二乘法的优化有效避免了该问题。从图 16 中可以看出, 加入该能量项时, 前景目标对象与背景区域之间网格形变相对较为缓和; 未加入该能量项的网格形变较为紧凑。

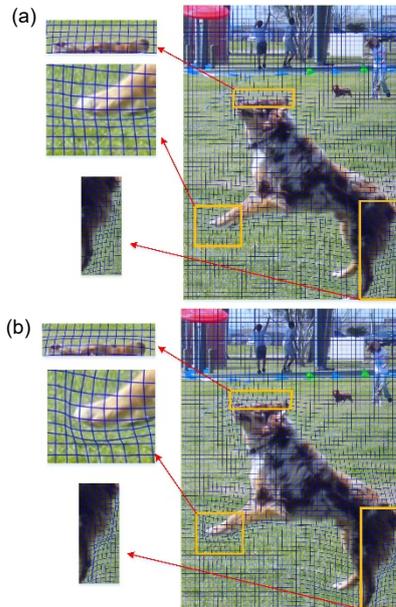


图 16 是否包含背景保持能量项对比。

- (a) 包含背景保持能量项的结果;
- (b) 未包含该能量项的结果

Fig. 16 Comparison of whether the background preservation energy term is included.
(a) Contain the background preservation energy term;
(b) This energy term is not included

4.5 实验复杂度分析

本文实验和其他三种对比实验都在 Intel i5-8400 CPU, 2.80 GHz 和 8 GB 内存的硬件条件下进行。本文方法在上述硬件条件下处理一对 PNG 格式的立体图像优化, 平均耗时 61.2 s, 前期的显著检测与二值图像生成平均耗时 41.9 s, 总共耗时 103.1 s。本文方法在优化过程中的复杂度主要集中在最小二乘法计算优化网格顶点的过程, 该过程耗时 58.9 s, 占优化过程总时间的 96.2%。四种实验方法的时间效率对比如表 4 所示。

4.6 局限性

本文方法的局限性在于目标对象背景区域的复杂度。对于复杂度较高的背景图像, 在优化过程中会破坏其背景的几何结构, 而在背景区域复杂度较低且目标对象之间的重叠度相对较低的情况下, 可以得到很不错的视觉效果。例如图 17 所示目标对象得到放大, 但是其形状有部分扭曲。主要原因是本文方法在对目标对象放大的同时对背景中的物体形状进行了保持, 导致目标对象被压缩, 无法保持良好的形状。

表 4 四种实验方法的时间效率对比

Table 4 Comparison of time efficiency of the four experimental methods

实验方法	预处理时间/s	优化处理时间/s	总时间/s
伸缩变换方法	0	107.2	107.2
3D 渲染方法 ^[24]	0	112.7	112.7
深度控制方法 ^[15]	42.3	65.9	108.2
本文方法	41.9	61.2	103.1

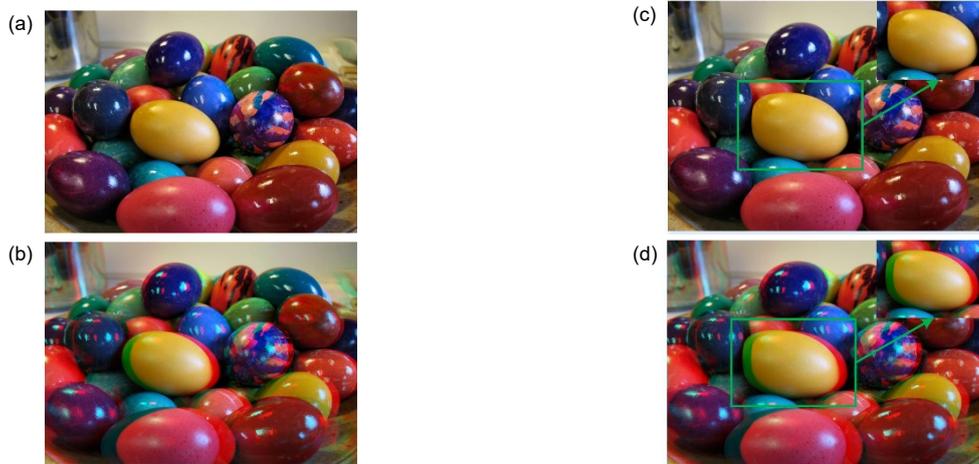


图 17 本文方法的局限性分析。(a) 原图像; (b) 原图像红蓝图; (c) 优化图像; (d) 优化红蓝图

Fig. 17 Limitations of the proposed method.

(a) Source image; (b) Original red-cyan anaglyph; (c) Image after optimization; (d) Red-cyan anaglyph after optimization

5 结论

本文提出了一种基于网格形变的立体变焦视觉优化的方法, 利用变焦方法来模拟视觉放大, 并建立起原图像与目标图像之间对象坐标的映射关系, 同时用二值掩膜图像提取前景目标对象, 随后将这些数据信息作为输入, 通过七个能量项的优化来达到良好的视觉效果。在视差深度方面, 用相机变焦参数来调整目标对象的视差深度并结合改进型深度恰可察觉深度模型用以调整目标对象的深度值。在视觉效果方面, 通过线保持、边界保持能量项和其他网格能量项的约束, 使得前景目标对象与背景之间有良好的过渡, 同时也使得目标在视觉上保留了放大的效果。实验结果表明, 本文方法在图像目标对象尺寸控制方面与目标对象的深度调整方面都具有一定的优势。在接下来的工作中, 将考虑图像内容复杂性并结合本文的优化方法进一步提高立体图像的视觉质量。

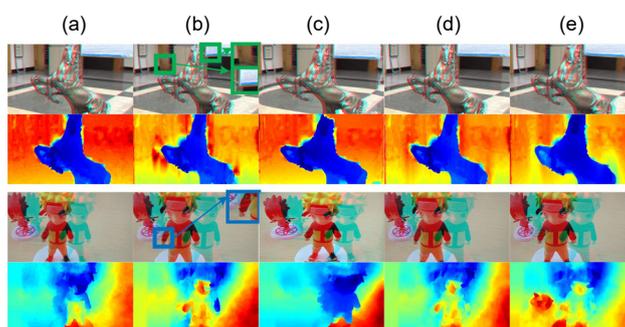
参考文献

- [1] Devernay F, Duchêne S, Ramos-Peon A. Adapting stereoscopic movies to the viewing conditions using depth-preserving and artifact-free novel view synthesis[J]. *Proc SPIE*, 2011, **7863**: 786302.
- [2] Islam M B, Lai-Kuan W, Chee-Onn W, et al. Stereoscopic image warping for enhancing composition aesthetics[C]//*Proceedings of the 3rd IAPR Asian Conference on Pattern Recognition (ACPR)*, Kuala Lumpur, Malaysia, 2015: 645–649.
- [3] Basha D T, Moses Y, Avidan S. Stereo seam carving a geometrically consistent approach[J]. *IEEE Trans Patt Anal Mach Intell*, 2013, **35**(10): 2513–2525.
- [4] Luo S J, Sun Y T, Shen I C, et al. Geometrically consistent stereoscopic image editing using patch-based synthesis[J]. *IEEE Trans Visual Comput Graph*, 2015, **21**(1): 56–67.
- [5] Lee K Y, Chung C D, Chuang Y Y. Scene warping: layer-based stereoscopic image resizing[C]//*2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, 2012: 49–56.
- [6] Yan W Q, Hou C P, Wang B L, et al. Content-aware disparity adjustment for different stereo displays[J]. *Multim Tools Applicat*, 2017, **76**(8): 10465–10479.
- [7] Niu Y Z, Liu F, Feng W C, et al. Aesthetics-based stereoscopic photo cropping for heterogeneous displays[J]. *IEEE Trans Multimed*, 2012, **14**(3): 783–796.
- [8] Tong R F, Zhang Y, Cheng K L. Stereo Pasting: Interactive composition in stereoscopic images[J]. *IEEE Trans Visualizat Comput Graph*, 2013, **19**(8): 1375–1385.
- [9] Chang C H, Liang C K, Chuang Y Y. Content-aware display adaptation and interactive editing for stereoscopic images[J]. *IEEE Trans Multimed*, 2011, **13**(4): 589–601.
- [10] Shao F, Lin W C, Lin W S, et al. QoE-guided warping for stereoscopic image retargeting[J]. *IEEE Trans Image Process*, 2017, **26**(10): 4790–4805.
- [11] Shao F, Shen L B, Jiang Q P, et al. StereoEditor: controllable stereoscopic display by content retargeting[J]. *Opt Exp*, 2017, **25**(26): 33202–33213.
- [12] Lu S P, Florea R M, Cesar P, et al. Efficient depth-aware image deformation adaptation for curved screen displays[C]//*Proceedings of the on Thematic Workshops of ACM Multimedia*, New York, NY, USA, 2017: 442–450.
- [13] Kim D, Sohn K. Depth adjustment for stereoscopic image using visual fatigue prediction and depth-based view synthesis[C]//*Proceedings of the 2010 IEEE International Conference on Multimedia and Expo*, Singapore, 2010: 956–961.
- [14] Lei J J, Peng B, Zhang C Q, et al. Shape-preserving object depth control for stereoscopic images[J]. *IEEE Trans Circuits Syst Video Technol*, 2018, **28**(12): 3333–3344.
- [15] Park H, Lee H, Sull S. Object depth adjustment based on planar approximation in stereo images[C]//*Proceedings of the IEEE International Conference on Consumer Electronics-Berlin (IC-CE-Berlin)*, Berlin, Germany, 2011: 10–14.
- [16] Park H, Lee H, Sull S. Efficient viewer-centric depth adjustment based on virtual fronto-parallel planar projection in stereo 3D images[J]. *IEEE Trans Multimed*, 2014, **16**(2): 326–336.
- [17] Yan T, Lau R W H, Xu Y, et al. Depth mapping for stereoscopic videos[J]. *Int J Comput Vision*, 2013, **102**(1–3): 293–307.
- [18] Sohn H, Jung Y J, Lee S I, et al. Visual comfort amelioration technique for stereoscopic images: disparity remapping to mitigate global and local discomfort causes[J]. *IEEE Trans Circuits Syst Video Technol*, 2014, **24**(5): 745–758.
- [19] Guan S H, Lai Y C, Chen K W, et al. A tool for stereoscopic parameter setting based on geometric perceived depth percentage[J]. *IEEE Trans Circuits Syst Video Technol*, 2016, **26**(2): 290–303.
- [20] Oh H, Kim J, Kim J, et al. Enhancement of visual comfort and sense of presence on stereoscopic 3D images[J]. *IEEE Trans Image Process*, 2017, **26**(8): 3789–3801.
- [21] Wang W G, Shen J B, Yu Y Z, et al. Stereoscopic thumbnail creation via efficient stereo saliency detection[J]. *IEEE Trans Visualizat Comput Graph*, 2017, **23**(8): 2014–2027.
- [22] Gastal E S L, Oliveira M M. Shared sampling for real-time alpha matting[J]. *Comput Graph Forum*, 2010, **29**(2): 575–584.
- [23] Li C H, An P, Shen L Q, et al. A modified just noticeable depth difference model built in perceived depth space[J]. *IEEE Trans Multimed*, 2019, **21**(6): 1464–1475.
- [24] Mao Y, Cheung G, Ji Y S. On constructing z-Dimensional DIBR-synthesized images[J]. *IEEE Trans Multimed*, 2016, **18**(8): 1453–1468.
- [25] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity[J]. *IEEE Trans Image Process*, 2005, **13**(4): 600–612.

Stereoscopic zoom for visual optimization based on grid deformation

Zhou Shen, Chai Xiongli, Shao Feng*

College of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China



Comparative experiment of two sets of stereoscopic image sequences.

(a) Source image; (b) 3D rendering; (c) Telescopic transformation; (d) Depth control; (e) Ours

Overview: With the development of digital imaging technology, more and more attention has been paid to the visual enhancement technology of 3D stereoscopic image objects, which has made important contributions in the fields of AI image modeling, astronomical observation, and digital imaging shooting. For this reason, with the model of camera zoom shooting, a grid deformation-based method is proposed for visual optimization of the stereoscopic zoom. In order to improve the experience of 3D stereoscopic vision, the digital zoom method is used to simulate the camera model to appropriately enlarge the objective area, and then the mapping relationship between the reference image. The target image is established according to the camera zoom distance, and the foreground target object is extracted. Then use the modified just noticeable depth difference model to guide the adaptive depth adjustment of the target object. Finally, the grid of the image is optimized in combination with the designed energy term to improve the visual perception of the target object and ensure good visual experience effect of the entire stereoscopic image. The camera-based zoom operation is used to extract the foreground target object of the stereoscopic image, and then the target object and the transition area between the target object and the image background are optimized. From the perspective of the size of the target object, with the zoomed image and camera focal length as a reference, the foreground target object is enlarged while the background remains unchanged. From the point of view of the image optimization effect, the transition between the foreground target object and the image background is good, to avoid problems such as voids in the transition between the target object and the background and changes in the target depth causing visual discomfort. From the experimental results, the target object is enlarged according to the focal length, and the improved depth can be used to adjust the image depth to improve the visual comfort of the 3D image. Through camera zoom operation and related grid deformation operation, and introduce coordinate mapping energy term, line retention energy term, boundary retention energy term, modified just noticeable depth difference energy term, left and right foreground consistency retention energy term, visual comfort energy term, and background preservation energy term to achieve the visual optimization of stereoscopic zoom. Compared with the existing digital focus adjustment of the enlarged image object, the method proposed in this paper has a certain good effect on the size control of the image target object and the depth adjustment of the target object.

Zhou S, Chai X L, Shao F. Stereoscopic zoom for visual optimization based on grid deformation[J]. *Opto-Electron Eng*, 2021, 48(4): 200186; DOI: 10.12086/oe.2021.200186

Foundation item: National Natural Science Foundation of China (61622109) and National Natural Science Foundation of Ningbo (2017A610112)

* E-mail: shaofeng@126.com