

基于 YCbCr 光场数据的视差重建算法

史立根, 邱钧, 刘畅*, 邓小娟

北京信息科技大学理学院应用数学研究所, 北京 100101

摘要 在光场计算成像中, 场景深度重建问题被转化为视差重建问题。通过引入 YCbCr 颜色空间光场数据实现了基于光场数据单体化的高效视差重建。在 Y 通道进行区域匹配可避免 RGB 三通道匹配的冗余计算, 进而提高了匹配效率。在 Cb、Cr 通道进行单体化可实现单体边缘去遮挡和内部视差一致性约束, 解决了遮挡区域和平滑区域的误匹配问题。Cb、Cr 色度信息为单体化提供了有效聚类信息, 将区域生长和二分法相结合实现了单体的精准分割。在单体边缘处, 根据单体边缘选取匹配窗口的形状和可见视点, 避免了边缘遮挡误匹配的出现。在单体内部, 基于视差一致性先验实现了视差图的优化。模拟数据实验和实际数据实验的结果表明, 所提方法在纹理、遮挡和平滑区域均能实现精确重建, 对遮挡区域的视差重建具有很强的鲁棒性。

关键词 成像系统; 光场; 视差重建; YCbCr 颜色空间; 单体化; 高效匹配

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202259.0211002

Disparity Reconstruction Algorithm Based on YCbCr Light Field Data

Shi Ligen, Qiu Jun, Liu Chang*, Deng Xiaojuan

Institute of Applied Mathematics, School of Applied Science, Beijing Information Science and Technology University, Beijing 100101, China

Abstract In the light field computational imaging, scene depth reconstruction is transformed into a problem of disparity reconstruction. Efficient disparity reconstruction based on monomer light field data is realized by introducing YCbCr color space light field data. Region matching in Y channel can avoid the redundant calculation of RGB three-channel matching, and improve the matching efficiency. Monomer edge occlusion and internal disparity consistency constraint can be realized by Cb, Cr channel monomer to solve the problem of mismatching between occlusion region and smooth region. Cb, Cr chromaticity information provides effective clustering information for monomer. The accurate segmentation of monomer is realized by combining the region growth and dichotomy. At the edge of the monomer, the matching window shape and the visible viewpoint are selected according to the edge of the monomer to avoid the appearance of the edge occlusion mismatching. In the monomer, the disparity map is optimized based on disparity consistency prior. Experimental results of simulated data and real data show that the proposed method can achieve accurate reconstruction in texture, occlusion, and smooth regions, and has strong robustness for disparity reconstruction in occlusion regions.

Key words imaging systems; light field; disparity reconstruction; YCbCr color space; monomer; efficient matching

收稿日期: 2021-03-08; 修回日期: 2021-03-14; 录用日期: 2021-03-17

基金项目: 国家自然科学基金(61931003, 61871042, 61801031)、北京市教委科研计划一般项目(KM201911232009)

通信作者: *liuchang@bistu.edu.cn

1 引言

光场包含光线的空间和角度信息^[1-2],被广泛应用于场景渲染、深度重建及三维重构中^[3-4]。光场数据的采集方式主要有微透镜阵列^[5-6]、相机阵列^[7-8]、编码掩膜^[9-10]和聚焦堆栈^[11]等。视差重建的方法主要包括:1)基于亚光圈图像的匹配方法,通过逐点匹配的方式获取场景视差信息^[12-21]。虽然该类算法可以降低光场频谱混叠和视角串扰的干扰,但是在平滑区域和遮挡区域的重建精度较低。2)基于EPI(epipolar plane images)^[22-23]的方法,通过求得EPI图像中直线的斜率来获取场景的深度信息。虽然该类方法在遮挡区域能得到较好的结果,但是该类方法存在计算量大、对噪声敏感的缺点。3)基于聚焦堆栈的方法,通过计算像点的聚焦程度来获取物点的深度信息^[24-26]。由于像点的聚焦程度很难准确刻画,故该类算法很难重建出场景中的精细结构,进而通常采用多线索融合的方法对精细度进行补偿。4)基于深度学习的方法^[27-29],其可通过设计神经网络来代替复杂的数学建模,但该类方法需要大量的标签数据且网络泛化能力不足。

在亚光圈图像获取视差信息方面,Liu等^[11]通过分析遮挡存在时多视角图像的像素特点,对场景空间进行离散化,利用平面扫描获取深度信息。Heber等^[15]将参考图像的颜色作为双向滤波器的平均值,利用形成的双边一致性度量进行深度重建。虽然该算法对遮挡区域重建具有一定的鲁棒性,但是存在对噪声敏感的缺点。Jeon等^[18]采用立体匹配的方法估计中心视图与同一行和同一列中其他视图间的视差值,并采用中值运算符对所有估计值进行最终组合。该算法只能处理同一行或同一列的遮挡,且会给视差图组合带来较大误差。

由双平面光场的几何关系可知,视差与深度存在反比例关系,故深度重建问题被转化为视差重建

问题^[14, 17-18, 21-24, 27-29]。光场数据中包含多个等间隔视角下的透视投影图像,视差信息蕴含在光场数据空角域的耦合关系中,这是场景视差重建的基础。因此,基于亚光圈图像匹配的方法是场景视差重建的一类基本方法,这类方法面临的问题:1)光场数据占用大量的计算资源;2)平滑区域和遮挡区域匹配困难^[13, 17, 25-26]。

本文将光场数据在RGB下的场景纹理信息映射到YCbCr中的Y通道中、颜色信息映射到Cb、Cr通道中,形成YCbCr颜色空间下的光场数据(简称YCbCr光场数据)。利用Y通道充分表达图像纹理信息的优势,可避免RGB空间出现的多通道计算问题,从而提高了区域匹配效率。利用Cb、Cr通道充分表达图像色度信息的优势,可分割识别单体区域,从而降低了平滑区域和遮挡区域的匹配误差。本文基于YCbCr颜色空间光场数据的单体化,实现了利用光场数据高效、高精度地重建视差信息。

2 基于YCbCr光场数据的视差重建算法

基于YCbCr光场数据在YCbCr颜色空间下建立视差的距离测度函数,采用线搜索方式重建视差信息。

2.1 基于YCbCr光场数据的视差重建模型

对于自然场景,物体纹理信息主要由辐亮度描述,且物体呈色度一致现象。场景的纹理信息主要蕴含于Y通道、色度信息主要蕴含于Cb、Cr通道。因此,YCbCr颜色空间可表示亮度参量和色度参量^[30]。物体纹理信息可以实现光场数据的区域匹配,色度信息可以实现光场数据的单体分割。

以中心子孔径图像 $L(u_0, v_0, x, y)$ 为参照,利用区域匹配算法获取 $L(u_0, v_0, x, y)$ 对应的视差图。在四维光场中,基于视差 $d(x, y)$ 与视点距离的等比关系,建立关于视差的目标函数,即

$$E[d(x, y)] = \sum_{u_i, v_j \in \Phi} \sum_{c \in C} \left| L[u_i, v_j, x + \Delta u_i \cdot d(x, y), y + \Delta v_j \cdot d(x, y)] - L(u_0, v_0, x, y) \right|, \quad (1)$$

式中: Φ 表示物点 (x, y) 的可视视点的集合; c 表示颜色通道; C 表示光场数据颜色通道的集合,一般包含R、G、B通道; $\Delta u_i = u_i - u_0$; $\Delta v_j = v_j - v_0$; (u_i, v_j) 表示视点位置; (u_0, v_0) 表示中心视点位置。

由于图像中局部区域的像素值相近且存在噪

声等因素的影响,故将单个像素点作为匹配基元的匹配方法的鲁棒性较差。利用 $w(x, y)$ 匹配窗口代替单个像素点作为匹配基元,以提高匹配方法的鲁棒性。为了消除匹配中的颜色计算冗余,在纹理通道(Y通道)中进行区域匹配,此时关于视差 $d(x, y)$ 的目标函数为

$$\hat{E}[d(x, y)] = \sum_{(u, v) \in \Phi} \sum_{(x', y') \in \omega(x, y)} h(x', y') |L[u, v, x' + \Delta u, d(x, y), y' + \Delta v, d(x, y)] - L(u, v, x', y')|, \quad (2)$$

式中: (x', y') 为窗口 $\omega(x, y)$ 中的像素; $h(x', y')$ 为像素 (x', y') 的权重系数。

利用区域匹配方法实现线搜索, 极小化 $\hat{E}[d(x, y)]$ 可得到像素 (x, y) 的视差 $d^*(x, y)$, 其表达式为

$$d^*(x, y) = \arg \min_d \{ \hat{E}[d(x, y)] \}. \quad (3)$$

2.2 视差重建算法

在 YCbCr 颜色空间下, 首先对 YCbCr 光场数据进行单体分割, 然后利用分割好的单体在 Y 通道实现区域匹配, 最后对单体内部的视差信息进行一致性优化。

2.2.1 YCbCr 光场数据单体化

单体是场景中具有视差一致性的单一物体, 是场景中视差信息具有分片光滑性质的区域。场景的单体化是在中心视图 $L(u_0, v_0, x, y)$ 中划分出单体在 (x, y) 空间上的支集。设 D 表示场景在 (x, y) 空间上的支集, D_k 表示第 k 个单体在 (x, y) 空间上的支集, 则

$$\begin{cases} \bigcup_{k=1}^K D_k = D \\ \bigcap_{k=1}^K D_k = \emptyset \end{cases}, \quad (4)$$

式中: K 为单体总数。

Cb、Cr 通道可更完备地表达色度信息。当单体

内部具有明显的色度一致性时, 单体在 Cb、Cr 通道中的聚类状态明显优于 RGB 颜色空间, 进而可利用单体内的色度聚类状态对单体进行准确分割。

利用区域生长和二分法实现单体的分割。设第 k 个单体的种子点坐标为 (x_k, y_k) , 如图 1(a) 中星号标记所示。在单体边界外选取 n 个背景点, 如图 1(b) 中圆圈标记所示。设 (x_k, y_k) 中心的单体集合为 $\phi_{p,k} = \{(x_k, y_k)\}$, $k=1, \dots, K$, 对应的背景点集合设为 $\phi_{b,k} = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ 。为了防止 CbCr 颜色空间单体化在边界发生越界, 利用 Y 通道的梯度进行约束, 建立关于颜色距离和空间结构的单体化能量函数, 其表达式为

$$E_k(x, y) = [M_{Cb}(x, y) - \overline{M}_{Cb,k}]^2 + [M_{Cr}(x, y) - \overline{M}_{Cr,k}]^2 + \gamma |\nabla Y(x, y)|, \quad (5)$$

式中: $[M_{Cb}(x, y) - \overline{M}_{Cb,k}]^2 + [M_{Cr}(x, y) - \overline{M}_{Cr,k}]^2$ 为颜色距离项; $\overline{M}_{Cb,k}, \overline{M}_{Cr,k}$ 表示单体集合 $\phi_{p,k} = \{(x_k, y_k)\}$ 内所有像素的 Cb、Cr 均值, $M_{Cb}(x, y)$ 表示像素点 (x, y) 处的 Cb 值, $M_{Cr}(x, y)$ 表示像素点 (x, y) 处的 Cr 值; $\gamma |\nabla Y(x, y)|$ 为空间结构项, γ 为权重因子, $\nabla Y(x, y)$ 为像素点 (x, y) 在 Y 通道的梯度, 即 $|\nabla Y(x, y)| = |Y(x+1, y) - Y(x-1, y)| + |Y(x, y+1) - Y(x, y-1)|$ 。

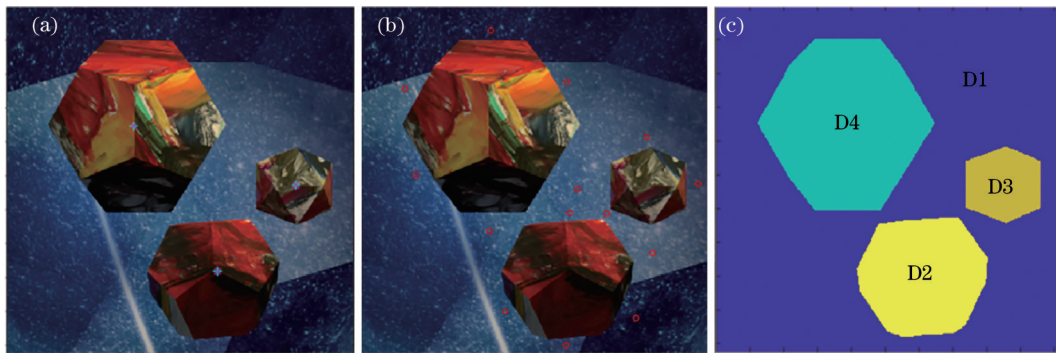


图 1 单体化数据。(a) 种子点选取; (b) 背景点选取; (c) 单体化数据结果

Fig. 1 Monomer data. (a) Selection of seed points; (b) selection of background points; (c) results of monomer data

定义像素点 (x, y) 的阈值分割函数 $\tau(x, y)$ 为

$$\tau(x, y) = [M_{Cb}(x, y) - M_{Cb}(x_i, y_i)]^2 + [M_{Cr}(x, y) - M_{Cr}(x_i, y_i)]^2, \quad (6)$$

式中: (x_i, y_i) 表示给定的背景点集合 $\phi_{b,k}$ 中距离当

前像素点 (x, y) 最近的像素点, 即 $(x_i, y_i) = \arg \min_{(x_i, y_i) \in \phi_{b,k}} [\sqrt{(x-x_i)^2 + (y-y_i)^2}]$, 其中 (x_i, y_i) 为给定的背景点集合 $\phi_{b,k}$ 中第 i 个像素点的坐标。

若 $E_k(x, y) \leq \tau(x, y)$, 则当前像素点 (x, y) 属

于第 k 个单体, 否则停止增长, 即

$$\psi_{p,k} = \begin{cases} \psi_{p,k} \cup M, E_k(x,y) \leq \tau(x,y) \\ \psi_{p,k}, \text{others} \end{cases} \quad (7)$$

式中: M 为 $\{(x,y)\}$ 。

当场景中的色度信息有明显区分时, 单体的分割精度主要由颜色距离项贡献。当场景中的色度信息区分不明显时, 空间结构项起决定作用。在自然场景中, 物体色度信息呈聚类分布, 单体的划分主要依赖颜色距离项, 故空间结构项 γ 取较小的权重因子。当 γ 取值为 0.02 时, Platonic 场景分割的

单体结果如图 1(c) 所示。

2.2.2 Y 通道实现区域匹配

Y 通道综合了 RGB 三通道的纹理特征, 在 Y 通道做区域匹配可以减少 RGB 三通道匹配的冗余计算。图 2 展示了融合 RGB 三通道、Y 通道和 R、G、B 单通道重建两个典型光场数据的视差。从视差重建的结果来看, Y 通道的匹配结果要优于 R、G、B 单通道, 且能达到融合 RGB 三通道的匹配结果。从计算效率来看, Y 通道匹配运算的时间是 RGB 三通道的 1/3。

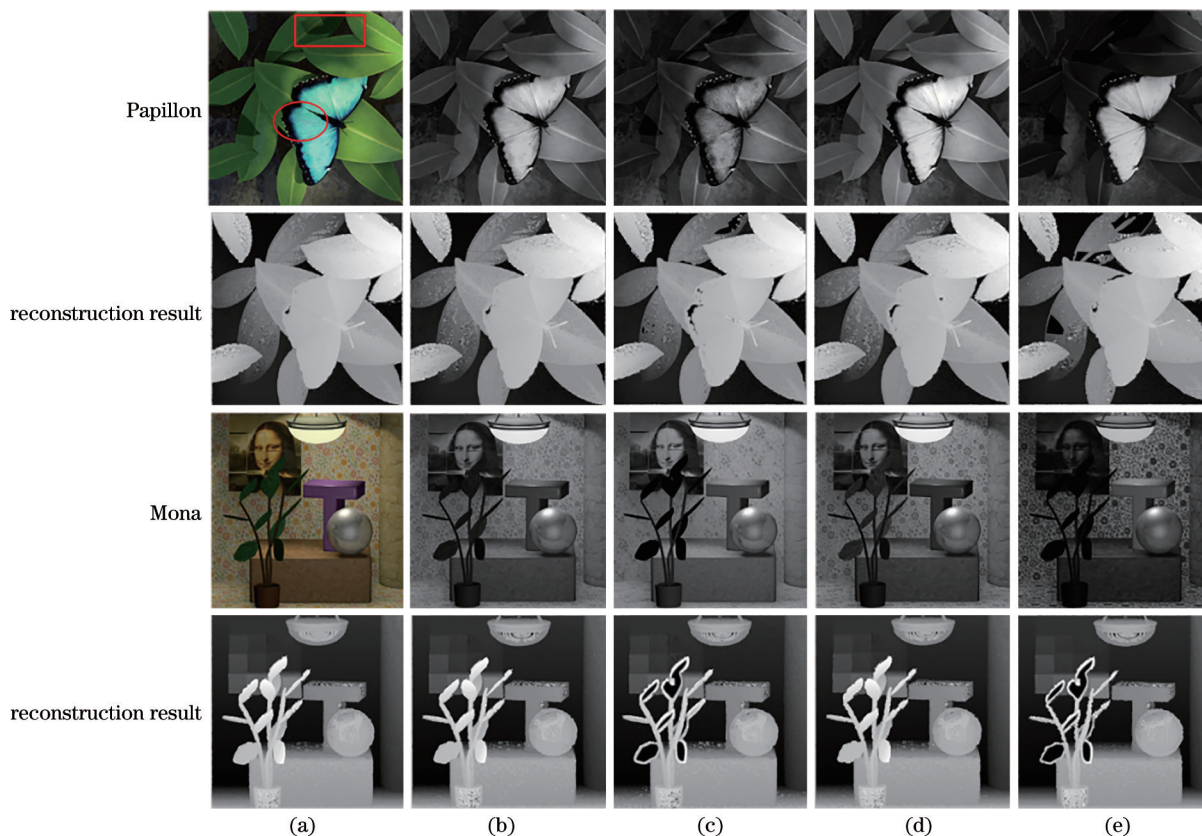


图 2 Papillon 和 Mona 场景不同通道视差重建的结果。(a) RGB^[31]; (b) Y 通道; (c) R 通道; (d) G 通道; (e) B 通道
Fig. 2 Results of disparity reconstruction in different channels of Papillon and Mona scenes. (a) RGB^[31]; (b) Y channel; (c) R channel; (d) G channel; (e) B channel

区域匹配的核心步骤是选取匹配点 (x,y) 的匹配窗口 $w(x,y)$ 。随着视点的变化, 单体的边界处会发生遮挡, 其部分几何结构被覆盖。为了精确地重建, 需要根据单体的边界选取不同形状的匹配窗口并进行视点筛选, 进而保证匹配窗口在筛选出的

视点中均可见。本文给出自适应匹配窗口, 使窗口内只包含同一类型的单体, 如图 3 所示。设匹配点 (x,y) 属于第 k 个单体 D_k , 则匹配窗口 $w(x,y)$ 内的像素点 (x',y') 的权重系数 $h(x',y')$ 为

$$h(x',y') = \begin{cases} \frac{1}{2\pi\sigma^2} e^{-\frac{(x'-x)^2+(y'-y)^2}{2\sigma^2}}, & (x',y') \in D_k \cap w(x,y) \\ 0, & \text{others} \end{cases} \quad (8)$$



图 3 匹配窗口选择示意图

Fig. 3 Schematic diagram of matching window selection

式中: $w(x, y)$ 是边长为 d 的矩形区域; $\sigma = \frac{1}{4}d$ 。

在边界遮挡区域, 根据匹配窗口的形状, 将匹配窗口分为 8 个方向, 每个方向对应一种预设的可视点组合。预设可视点组合如图 4 所示, 以 9×9 的视点为例, 标号 (1)~(8) 对应 8 种不同的视点组合。根据匹配窗口的形状找到匹配窗口的方向, 选择对应的视点组合。图 5 为匹配窗口方向的选择。图 5(a)~(c) 中箭头的指向表示该匹配窗口的方向。图 5(d) 中的待匹配点为单体内部纹理区域, 匹配窗口形状为矩形匹配窗口, 不存在匹配窗口方向。图 5(a) 所示的匹配窗口对应的视点组合为图 4 中标号为 (6) 的图。

2.2.3 单体内部视差信息的一致性优化

场景中的视差信息具有局部一致性, 可被分片光滑函数逼近。单体的视差可被具有紧支集且在支集内连续、高阶可导的函数逼近。在线搜索视差中, 最小视差间隔为 0.01, 重建的初始视差图有锯齿状, 故需要通过单体优化, 消除计算误差。采用全变差正则化^[32]方法对单体数据子集 D_k 重建的结果进行进一步优化

$$\mu_k^*(x, y) = \arg \min_{\mu_k(x, y)} \left\{ \sum_{(x, y) \in D_k} [\mu_k(x, y) - d_k^*(x, y)]^2 + \lambda V_{D_k}[\mu_k(x, y)] \right\}, \quad (9)$$

$$V_{D_k}[\mu_k(x, y)] = \sum_{(x, y) \in D_k} [|\mu_k(x+1, y) - \mu_k(x, y)| + |\mu_k(x, y+1) - \mu_k(x, y)|], \quad (10)$$

式中: (9) 式等号右边第一项是数据保真项, 用来度量优化后的视差图与初始求得的视差图的欧氏距离; (9) 式等号右边第二项 $V_{D_k}[\mu_k(x, y)]$ 是光滑正则项, 用视差图的梯度 1 范数来逼近单体内部视差的一致性先验; $\mu_k^*(x, y)$ 为全变差正则化模型计算出的视差图; $d_k^*(x, y)$ 为初始视差图; λ 为权重系数; $\mu_k(x, y)$ 为优化视差 $d_k^*(x, y)$ 后得到的估计结果。

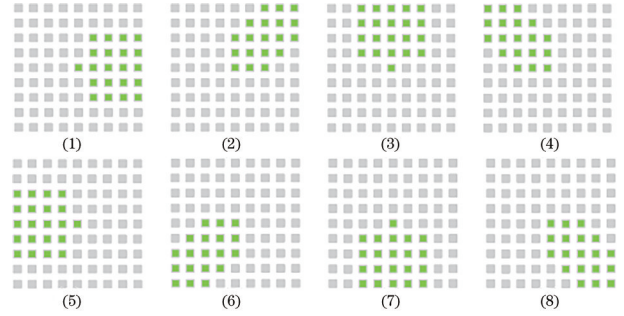


图 4 预设视点组合

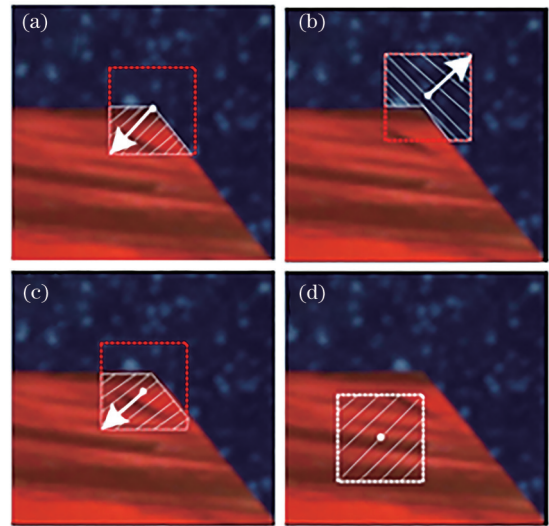


图 5 匹配窗口方向的选择

Fig. 5 Selection of matching window direction

本文对 TV (total variation) 模型的求解采用了文献 [32] 中的方法。

为了避免单体外部视差的影响, 对单体进行边缘扩充处理。对单体外边缘上的每一个像素 (x, y) 都向单体外部扩充 8 邻域, 并赋值为它本身, 如图 6 所示。此时, 对单体数据子集 D_k 重建结果优化的函数为

$$\mu_k^*(x, y) = \arg \min_{\mu_k(x, y)} \left\{ \sum_{(x, y) \in D_k^*} [\mu_k(x, y) - d_k^*(x, y)]^2 + \lambda V_{D_k^*}[\mu_k(x, y)] \right\}, \quad (15)$$

式中: D_k^* 为单体数据子集 D_k 扩充的集合。

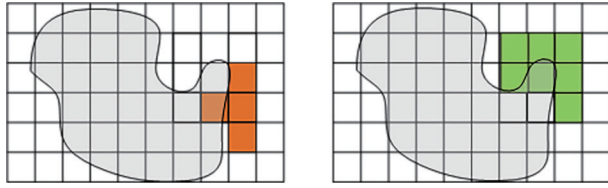


图 6 单体化数据的边缘扩充

Fig. 6 Edge extension of monomer data

3 实验与分析

使用公开的 4D Light Field Benchmark^[33-34] 标准光场数据、HCI 数据集和真实场景的光场数据测试所提算法。对于模拟数据,用重建的视差图与真实视差图进行比较。对于实际光场数据,分析重建视差图的精度。对单体内部和边界进行误差分析,从而验证所提算法在边界遮挡区域、平滑区域的视差重建精度。选取标准场景与经典的算法来比较评估所提算法的性能。计算平台包含 16 GB 的 RAM, NVIDIA GeForce GTX 1650, AMD Ryzen7 4800H @ 2.90 GHz CPU, 并使用 MATLAB 2019b 进行计算。

3.1 光场数据实验

对于模拟数据的实验,采用标准数据库提供的光场数据进行视差重建。对于实际数据的实验,通过搭建成像系统获取真实光场数据,进行单体化和视差重建。

3.1.1 模拟光场数据实验

每个场景提供的光场数据为 9×9 子孔径图像(空间分辨率为 $512 \text{ pixel} \times 512 \text{ pixel}$)。实验结果给出了单体分割后的单体图、初始视差图与优化处理后的视差图,如图 7 所示。其中 Backgammon、Dots、Pyramids 和 Cotton 的场景构造以及算法评估的量化指标在 3.3 节中给出。参数设置: γ 权重因子取 0.02, $w(x, y)$ 的边长 d 取 11 pixel, 搜索视差间隔设为 0.01 pixel, TV 模型的 λ 权重系数设为 0.2。

3.1.2 实测光场数据实验

在实测数据实验中,采用三轴平移台控制相机,以 1 mm 为视点间隔,采集同一平面内 9×9 视点阵列中每个视点位置下的二维图像,从而获取四维光场数据。相机型号为 SONY NEX-5C(分辨率为

$4592 \text{ pixel} \times 3056 \text{ pixel}$), 三轴平移台型号为 GCM-125302AM(精度为 0.1 mm)。

在实验中,利用相机获取单幅子孔径图像数据,并裁剪出包含物体信息的图像中心区域用于成像计算,裁剪后图像的分辨率为 $1024 \text{ pixel} \times 1024 \text{ pixel}$ 。实测光场数据的目的是验证算法对遮挡区域的鲁棒性。实际深度范围为 85~100 cm。参数设置: γ 权重因子取 0.02, $w(x, y)$ 的边长 d 取 15 pixel, 搜索视差间隔设为 0.01 pixel, TV 模型 λ 权重系数设为 0.2。真实数据的视差重建结果如图 8 所示。

3.2 单体边缘与内部的误差分析

遮挡通常存在于单体的边界处,平滑通常存在于单体的内部。为了进一步客观评估所提算法对遮挡、平滑区域的处理效果,对 Platonic 场景进行了分析,如图 9 所示。在真实视差图与计算得到的视差图中,选择某一行像素进行比较,如图 9 中虚线所示,选取的行像素包含了遮挡、平滑区域。

由于单体内部视差信息具有一致性,因此视差图中单体的边界位置会发生跳跃。图 10 为水平方向的视差曲线。可以发现,初始视差(虚线)和优化后的视差(实线)在真实视差(点划线)发生跳跃的位置处均发生跳变,说明所提算法在重建结果中能够很好地保持视差边缘,进而在遮挡区域处能够实现精确重建。此外,初始视差在单体内部产生的锯齿现象经过全变差正则化优化后,会变得更加光滑,进而更接近真实视差信息。

对于实际采集的数据,选取场景中心视图的第 480 行对遮挡区域视差重建结果进行分析,如图 11 所示。图 11(a)、(c) 中类星号标记为该行的边界分隔点。图 11(c) 是边界分隔点与同行的视差曲线的吻合图,可以看出,视差曲线在边界分隔点处间断,表明所提算法在边缘遮挡区域能够实现精确重建。

3.3 算法的评估与比较

选取 4D Light Field Benchmark 数据库中 4 种不同类型的场景,并与 LF^[18]、epil^[23]、LF_OCC^[25]、MV^[34]、mvmv0^[34] 等算法进行比较评估。其中, Backgammon 场景由两个平行倾斜的背景平面和一个反向倾斜的前景平面组成,该场景主要用来评估精细结构和遮挡边界。Dots 场景是由 15 个相同的

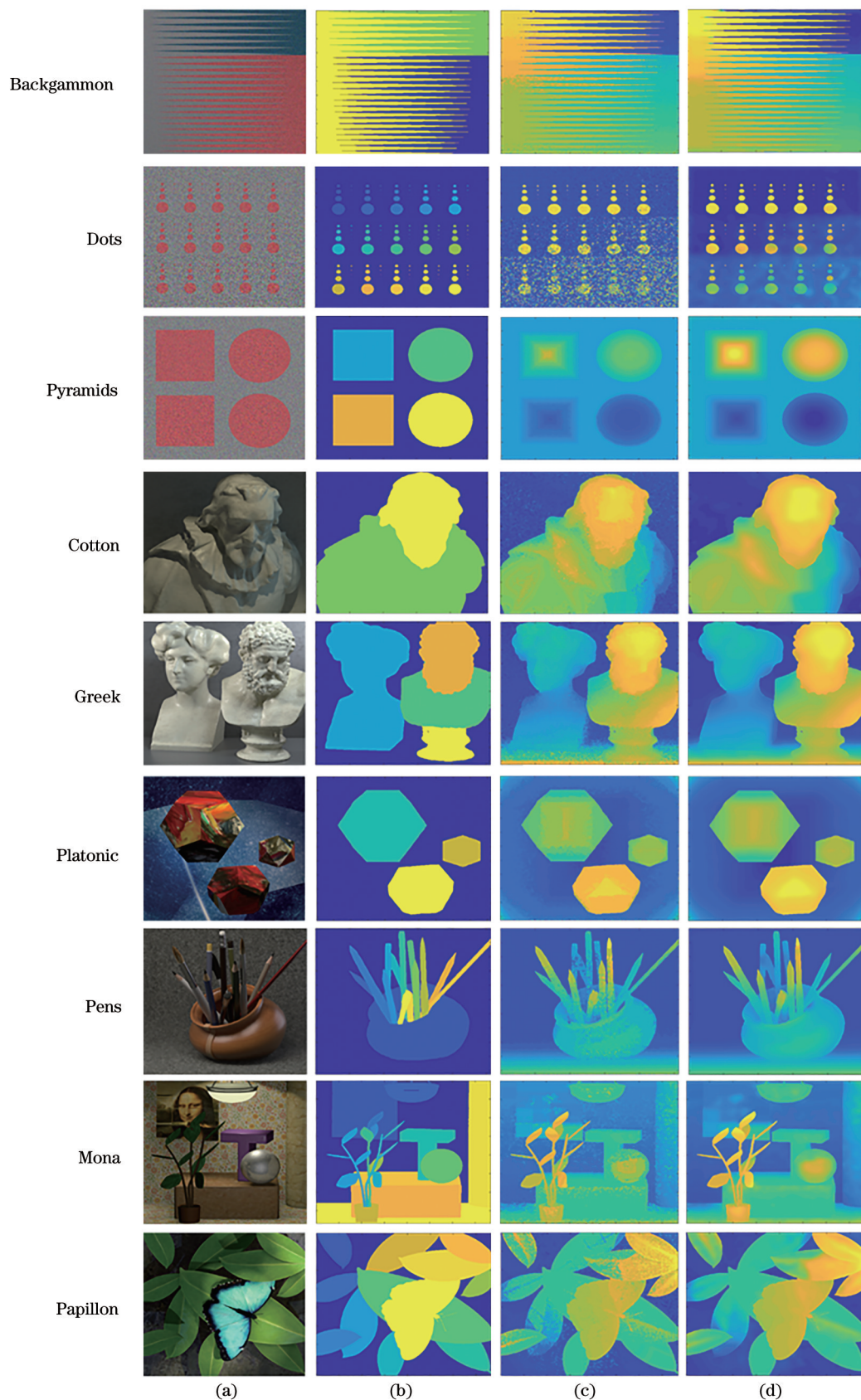


图 7 合成数据视差重建结果。(a)中心视图^[31,34];(b)单体图;(c)初始视差图;(d)优化视差图
 Fig. 7 Disparity reconstruction results of synthetic data. (a) Center view image^[31,34];
 (b) monomer map; (c) initial disparity map; (d) optimized disparity map

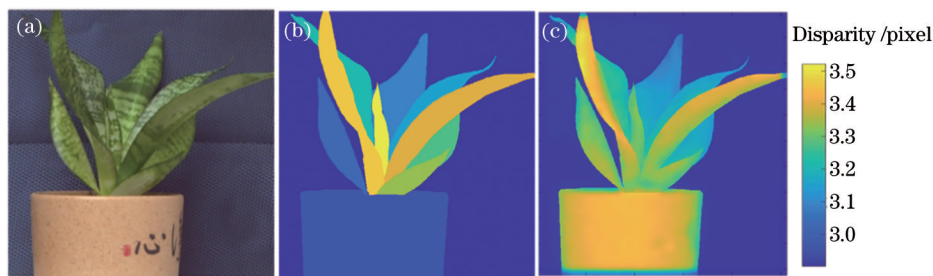


图 8 真实数据的视差重建结果。(a)中心视图;(b)单体图;(c)视差图

Fig. 8 Disparity estimation result of real data. (a) Center view image; (b) monomer map; (c) disparity map

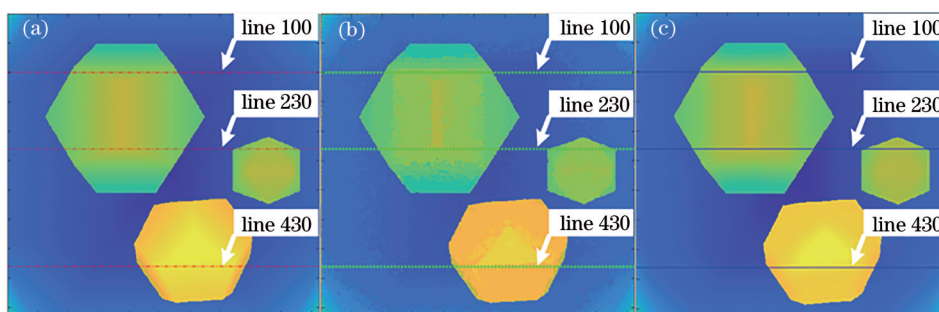


图 9 Platonic 场景分析。(a)真实视差图;(b)初始视差图;(c)优化视差图

Fig. 9 Analysis of Platonic scene. (a) Real disparity map; (b) initial disparity map; (c) optimized disparity map

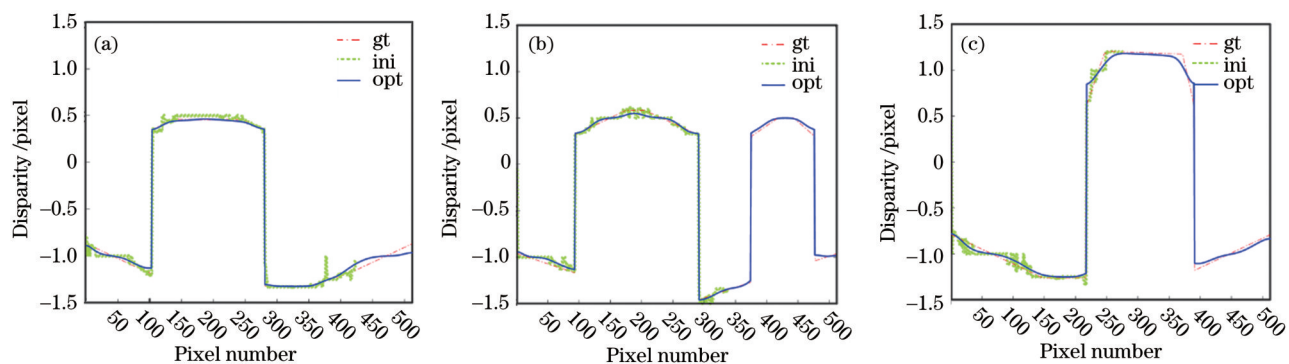


图 10 水平方向的视差曲线。(a)第 100 行;(b)第 230 行;(c)第 430 行

Fig. 10 Disparity curves in the horizontal direction. (a) 100th line; (b) 230th line; (c) 430th line

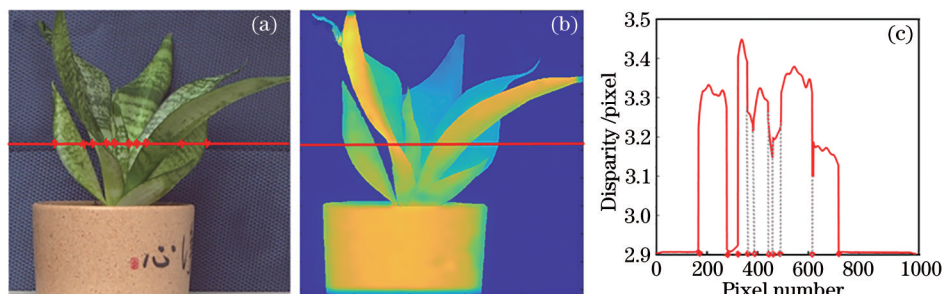


图 11 真实数据遮挡的评估。(a)中心视图;(b)视差图;(c)视差曲线与边界点的吻合图

Fig. 11 Evaluation of occlusion of real data. (a) Center view image; (b) disparity map; (c) fitting diagram of disparity curve to boundary point

网格单元组成,每个单元由 9 个越来越小的共面圆组成。在场景中添加了方差在 0~0.2 范围内线性

增长的高斯噪声,该场景用来评估算法的抗噪能力。Pyramids 场景是由上半球和棱锥突出平面以

及下半球和棱锥嵌入平面组成,该场景用来评估算法重建凹凸几何球形状的性能。Cotton 场景主要评估算法对平滑纹理区域的重建精度。

图 12 为不同算法的重建结果,从 Backgammon 场景可以看出,LF_OCC、mvcmv0 边缘处的噪声点很多,epil 丢失了边缘信息,LF 在前景的间隙处有一定程度的模糊。相比之下,所提算法能够很好地保持边缘结构,在前景的间隙处能重建出更多的细节信息。所提算法对噪声也具有很好的鲁棒性,这

一点可以从 Dots 场景中看到。相比 LF、LF_OCC、mvcmv0,所提算法在噪声水平很高的时候能够重建出更多的前景。对于 Pyramids 场景,所提算法能在凹凸的倾斜面上重建的视差曲面更加光滑。对于 Cotton 场景,所提算法在前景的边界处能够保持更多的细节。

以 Backgammon 场景为代表,将细节部分放大,如图 13 所示。可以看出,所提算法相比其他算法在边缘等精细结构处能得到较精确的重建结果。

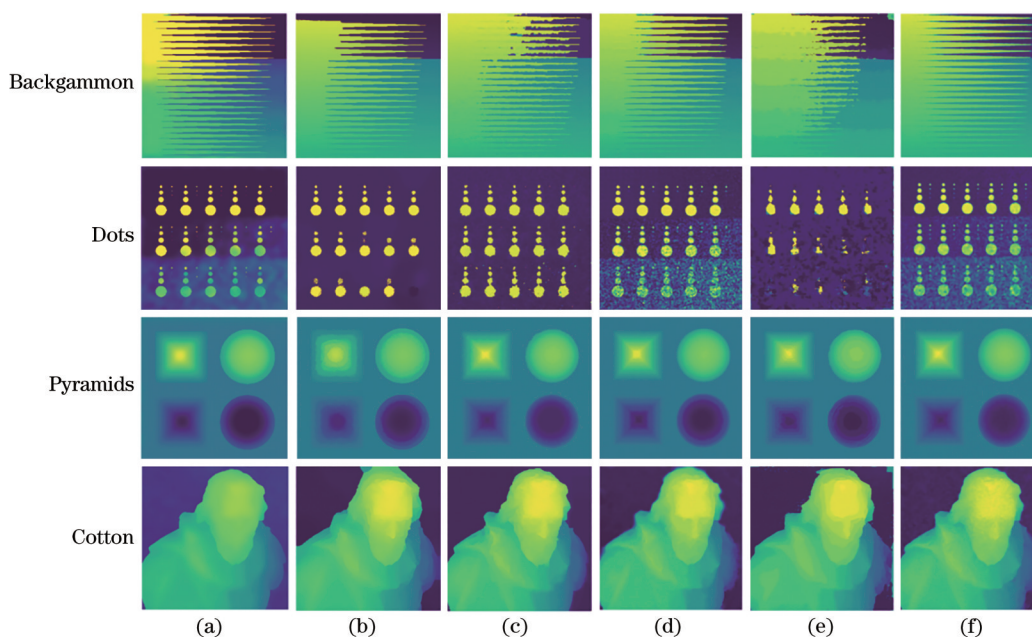


图 12 不同算法的重建结果,图 12(b)–(f)数据来自文献[34]。(a)所提算法;(b) LF;(c) LF_OCC;(d) MV;(e) mvcmv0;(f) epil

Fig. 12 Reconstruction results by different algorithms, data of Figs. 12(b)–(f) is from Ref. [34]. (a) Proposed algorithm; (b) LF; (c) LF_OCC; (d) MV; (e) mvcmv0; (f) epil

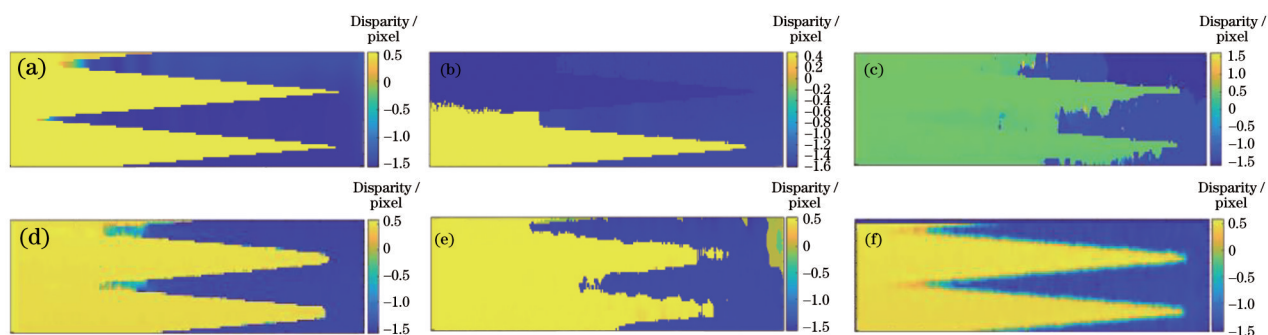


图 13 不同算法得到的 Backgammon 场景重建结果的放大图。(a)所提算法;(b) LF;(c) LF_OCC;(d) MV;(e) mvcmv0;(f) epil

Fig. 13 Enlarged images of reconstruction results of Backgammon scene obtained by different algorithms. (a) Proposed algorithm; (b) LF; (c) LF_OCC; (d) MV; (e) mvcmv0; (f) epil

此外,还计算了不同的定量指标来验证所提算法的有效性。采用雷达图综合呈现不同算法的性能

指标,如图 14 所示,分数越低(交点越接近圆心)表示该项指标性能越高。均方误差(MSE)取 4 个场景

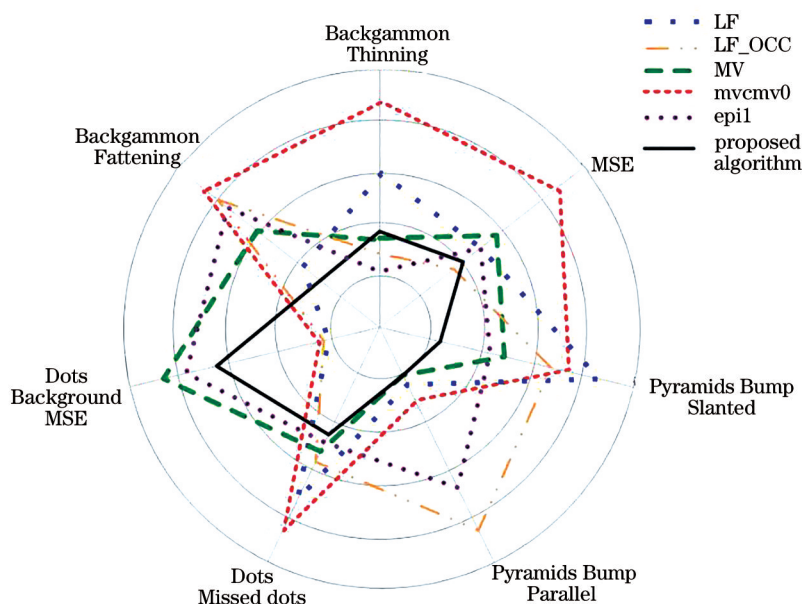


图 14 评估场景的雷达图

Fig. 14 Radar chart of evaluation scenes

MSE 的中值。Pyramids Bump Slanted 与 Pyramids Bump Parallel 是对 Pyramids 场景的倾斜度和平行面进行视差评估。Dots Missed dots 与 Dots Background MSE 是对算法在 Dots 场景中的抗噪性进行评估。Backgammon Thinning 与 Backgammon Fattening 是对算法在 Backgammon 场景中的恢复精细结构的性能进行评估。

所提算法的 Pyramids Bump Slanted 与 Pyramids Bump Parallel 得分较低,表明其对含有倾斜面以及凹凸情况的场景具有更高的重建精度。所提算法的 Dots Missed Dots 得分最低,表明在不同噪声水平下其均能重建出更多场景。然而,所提算法的 Dots Background MSE 不是最低,表明随着噪声水平的增加,虽然其能够重建出更多的场景点,但整体重建精度在降低,尤其是背景部分会受到一定的影响。所提算法的 Backgammon Thinning 得分最低,表明其在 Backgammon 场景中的前景像素被错误计算为背景像素的比例最少。然而,所提算法的 Backgammon Fattening 得分不是最低,表明其在重建出更多细节的同时,也导致部分背景像素被错位计算为前景像素。4 个场景的综合 MSE 评估表明所提算法性能较优。

4 结 论

由于基于光场数据重建视差的区域匹配算法在遮挡区域和平滑区域的计算精度不高,且 RGB 三

通道匹配存在冗余计算,因此提出了基于 YCbCr 光场数据的视差重建算法。所提算法在 Cb、Cr 通道实现单体分割,利用分割的单体在 Y 通道实现高效的区域匹配,并对计算得到的初始视差图进行视差一致优化处理。其中,利用分割的单体边界能提高边缘遮挡区域的匹配精度;对单体进行视差一致性约束能去除平滑区域和噪声带来的影响,进而提高算法的鲁棒性。模拟和实测数据实验表明,所提方法利用四维光场数据能够实现纹理、遮挡和平滑区域的高精度视差重建,且在遮挡区域具有很强的鲁棒性。

所提算法的核心思想是在光场数据的视差重建问题中引入 YCbCr 光场数据以进行单体化的视差重建。精准的单体化为边缘遮挡区域提供了几何先验信息,为平滑区域的精确匹配提供了一致性约束。因此,边缘遮挡区域和平滑区域的重建精度依赖于单体化的精准程度。利用 Cb、Cr 通道蕴含的色度信息实现单体化是一种有效的单体化方式。充分利用单体内部信息的一致性和边界信息的跳跃性,可实现更精准、自适应、自动划分的单体分割,这将使得所提算法具有更广泛的适用性。此外,光场数据单体化将光场数据空间划分成若干个具有一致性的子空间,使得子空间的计算更准确和高效,这一思想可扩展到其他计算成像问题中,形成具有适用性的单体化计算成像技术。

参 考 文 献

- [1] Gershun A. The light field[J]. *Journal of Mathematics and Physics*, 1939, 18(1/2/3/4): 51-151.
- [2] Adelson E H, Bergen J R. The plenoptic function and the elements of early vision[M]//Landy M, Movshon J A. *Computational models of visual processing*. Cambridge: The MIT Press, 1991.
- [3] Levoy M, Hanrahan P. Light field rendering[C]// *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, August 4-9, 1996, New Orleans, LA, USA. New York: ACM Press, 1996: 31-42. [LinkOut]
- [4] Fang L, Dai Q H. Computational light field imaging [J]. *Acta Optica Sinica*, 2020, 40(1): 0111001.
方璐, 戴琼海. 计算光场成像[J]. *光学学报*, 2020, 40(1): 0111001.
- [5] Ng R, Levoy M, Brédif M, et al. Light field photography with a hand-held plenoptic camera[D]. Stanford: Stanford University, 2005.
- [6] Sun F S, Han X. Multi-objective fast ranging method based on microlens array light field camera [J]. *Acta Optica Sinica*, 2019, 39(8): 0815001.
孙福盛, 韩燮. 基于微透镜阵列型光场相机的多目标快速测距方法[J]. *光学学报*, 2019, 39(8): 0815001.
- [7] Wilburn B, Joshi N, Vaish V, et al. High-speed videography using a dense camera array[C]// *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 27-July 2, 2004, Washington, DC, USA. New York: IEEE Press, 2004: 8168976.
- [8] Wilburn B, Joshi N, Vaish V, et al. High performance imaging using large camera arrays[J]. *ACM Transactions on Graphics*, 2005, 24(3): 765-776.
- [9] Liang C K, Lin T H, Wong B Y, et al. Programmable aperture photography: multiplexed light field acquisition[J]. *ACM Transactions on Graphics*, 2008, 27(3):1-10.
- [10] Nagahara H, Zhou C Y, Watanabe T, et al. Programmable aperture camera using LCoS[M]// Daniilidis K, Maragos P, Paragios N. *Computer vision-ECCV 2010*. Lecture notes in computer science. Heidelberg: Springer, 2010, 6316: 337-350.
- [11] Liu C, Qiu J, Jiang M. Light field reconstruction from projection modeling of focal stack[J]. *Optics Express*, 2017, 25(10): 11377-11388.
- [12] Bishop T E, Favaro P. Plenoptic depth estimation from multiple aliased views[C] //2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, September 27-October 4, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 1622-1629.
- [13] Yu Z, Yu J Y, Lumsdaine A, et al. Plenoptic depth map in the case of occlusions[J]. *Proceedings of SPIE*, 2013, 8667: 86671S.
- [14] Yu Z, Guo X Q, Ling H B, et al. Line assisted light field triangulation and stereo matching[C]//2013 IEEE International Conference on Computer Vision, December 1-8, 2013, Sydney, NSW, Australia. New York: IEEE Press, 2013: 2792-2799.
- [15] Heber S, Ranftl R, Pock T. Variational shape from light field[M]//Heyden A, Kahl F, Olsson C, et al. *Energy minimization methods in computer vision and pattern recognition*. Lecture notes in computer science. Heidelberg: Springer, 2013, 8081: 66-79.
- [16] Fickel G P, Jung C R, Malzbender T, et al. Stereo matching and view interpolation based on image domain triangulation[J]. *IEEE Transactions on Image Processing*, 2013, 22(9): 3353-3365.
- [17] Chen C, Lin H T, Yu Z, et al. Light field stereo matching using bilateral statistics of surface cameras [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 1518-1525.
- [18] Jeon H G, Park J, Choe G, et al. Accurate depth map estimation from a lenslet light field camera[C]// 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1547-1555.
- [19] Buades A, Facciolo G. Reliable multiscale and multiwindow stereo matching[J]. *SIAM Journal on Imaging Sciences*, 2015, 8(2): 888-915.
- [20] Navarro J, Buades A. Reliable light field multiwindow disparity estimation[C]//2016 IEEE International Conference on Image Processing (ICIP), September 25-28, 2016, Phoenix, AZ, USA. New York: IEEE Press, 2016: 1449-1453.
- [21] Mishiba K. Fast depth estimation for light field cameras[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4232-4242.
- [22] Tomic I, Berkner K. Light field scale-depth space transform for dense depth estimation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 441-448.
- [23] Johannsen O, Sulc A, Goldluecke B. What sparse

- light field coding reveals about scene structure[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 3262-3270.
- [24] Tao M W, Srinivasan P P, Malik J, et al. Depth from shading, defocus, and correspondence using light-field angular coherence[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 1940-1948.
- [25] Wang T C, Efros A A, Ramamoorthi R. Occlusion-aware depth estimation using light-field cameras[C]//2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE Press, 2015: 3487-3495.
- [26] Liu X M, Du M Z, Ma Z B, et al. Depth estimation method of light field image based on occlusion scene [J]. *Acta Optica Sinica*, 2020, 40(5): 0510002.
刘晓旻, 杜梦珠, 马治邦, 等. 基于遮挡场景的光场图像深度估计方法[J]. *光学学报*, 2020, 40(5): 0510002.
- [27] Heber S, Yu W, Pock T. Neural EPI-volume networks for shape from light field[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2271-2279.
- [28] Jeon H G, Park J, Choe G, et al. Depth from a light field image with learning-based matching costs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, 41(2): 297-310.
- [29] Shin C, Jeon H G, Yoon Y, et al. EPINET: a fully-convolutional neural network using epipolar geometry for depth from light field images[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 4748-4757.
- [30] Yang Y, Peng Y H, Liu Z G. A fast algorithm for YCbCr to RGB conversion[J]. *IEEE Transactions on Consumer Electronics*, 2007, 53(4): 1490-1493.
- [31] Wanner S, Meister S, Goldluecke B. Datasets and benchmarks for densely sampled 4D light fields[C]//18th International Workshop on Vision, Modeling, and Visualization, September 11-13, 2013, Lugano, Switzerland. Geneva: Eurographics Association, 2013, 13: 225-226.
- [32] Rudin L I, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms[J]. *Physica D*, 1992, 60(1/2/3/4): 259-268.
- [33] Honauer K, Johannsen O, Kondermann D, et al. A dataset and evaluation methodology for depth estimation on 4D light fields[M]//Lai S H, Lepetit V, Nishino K, et al. *Computer vision-ACCV 2016. Lecture notes in computer science*. Cham: Springer, 2017, 10113: 19-34.
- [34] Johannsen O, Honauer K, Goldluecke B, et al. A taxonomy and evaluation of dense light field depth estimation algorithms[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1795-1812.