

基于光线空间插值的任意视点绘制

蒋刚毅^{1,3}, 范良忠², 郁梅^{1,3}, 陈 愚¹

(1. 宁波大学电路与系统研究所, 浙江宁波 315211; 2. 浙江大学宁波理工学院, 浙江宁波 315100;
3. 南京大学计算机软件新技术国家重点实验室, 江苏南京 210093)

摘 要: 本文提出了一种基于改进动态规划的光线空间插值方法, 区别于传统的动态规划算法, 本文先利用扫描线间的平滑约束来消除条纹效应, 并采用增量计算方案来加速动态规划算法中最耗时的匹配代价计算. 然后, 采用以视点为中心的视差估计方法插值出稠密的光线空间数据. 最后, 从光线空间数据快速生成高质量的任意视点图像. 该方案采用 Middlebury 和 3DTV 数据库作为测试数据, 实验结果表明, 新方法在计算效率和生成视点图像的 PSNR 上优于块匹配插值方法和视差域滤波方法.

关键词: 任意视点绘制; 动态规划; 光线空间插值

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2009) 08-1799-05

Arbitrary Viewpoint Rendering Based on Ray-Space Interpolation

JIANG Gang-yi^{1,3}, FAN Liang-zhong², YU Mei^{1,3}, CHEN Ken¹

(1. Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China;
2. Laboratory of Information and Optimization, Ningbo Institute of Technology, Zhejiang University, Ningbo, Zhejiang 315100, China;
3. National Key Laboratory of Software New Technology, Nanjing University, Nanjing, Jiangsu 210093, China)

Abstract: This paper presents an improved dynamic programming (DP) based ray-space interpolation for arbitrary view rendering. Different from the conventional DP method, a smoothing constraint of inter-scanlines is applied to eliminate the "streaking" artifact, and the most time-consuming step of computing matching cost is carried out by an incremental computation scheme. Then, dense ray-space data are interpolated by a view-centered disparity estimation method. Finally, arbitrary viewpoint image with high quality can be rendered from this dense ray-space easily. The proposed scheme is evaluated on the Middlebury data set and 3DTV test data. The experimental results show that the PSNR and computational efficiency of the proposed rendering method are much higher than that of the other two schemes.

Key words: arbitrary view rendering; dynamic programming; ray-space interpolation

1 引言

自由视点视频 (FVV, free viewpoint video) 是一种新的交互式多媒体系统, 它能提供观察场景的深度信息并允许用户在一定范围内任意选择欣赏角度^[1,2]. 目前, FVV 系统的场景表示方法主要可分为两类^[3]: 一类是基于几何模型的方法, 这种方法将场景表示成 3D 网格, 采用 3D 表面重构和纹理映射来再现真实世界中的物体. 使用这类方法对一些自然场景进行建模将会比较复杂并且耗时; 另一类是基于图像的方法, 这种方法对已有的视点图像进行重采样来获得新图像, 不需要对场景进行 3D 建模, 然而, 为了获得高质量的虚拟视点图像, 必须对场景进行稠密采样. 已有的研究表明, 合成的新

视点图像质量与原始视点数量成正比.

基于光线空间的 FVV 系统通过将稠密的多视点图像转换到光线空间中, 无需依赖任何几何信息就能绘制出任意视点图像^[4]. 但动态地捕获光线空间数据比较困难. 为了降低系统中相机的数量, 需要采用插值技术生成相关中间视点图像. Fujii 等提出了自适应滤波插值和基于视差的插值等方法. MPEG 中的三维音视频组证实了基于光线空间的方法适合于 FVV 系统^[5].

本文主要研究基于光线空间插值法的任意视点绘制技术, 提出了一种基于改进动态规划的光线空间插值方法, 以利用少量的实际视点图像生成稠密的光线空间数据, 进而由稠密的光线空间数据绘制出非相机平面的任意视点图像.

收稿日期: 2007-04-24; 修回日期: 2009-03-16

基金项目: 国家自然科学基金 (No. 60872094, No. 60832003); 教育部新世纪人才支持计划项目 (No. NCET-06-0537); 教育部博士点基金 (No. 20081646003); 宁波市自然科学基金 (No. 2008A610024)

2 基于光线空间插值的任意视点绘制方法

对于图 1(a)所示的多视点视频采集系统,其对应的光线空间可以用 4D 函数 $\{f(x, y, \theta, \phi)\}$ 表示,这里 (θ, ϕ) 表示光线的角度, (x, y) 表示光线和平面 $z=0$ 的交点, $f(x, y, \theta, \phi)$ 表示光线的强度.若忽略垂直方向的视角(即 $\phi=0$),则光线空间可用 3D 函数 $\{f(x, y, u =$

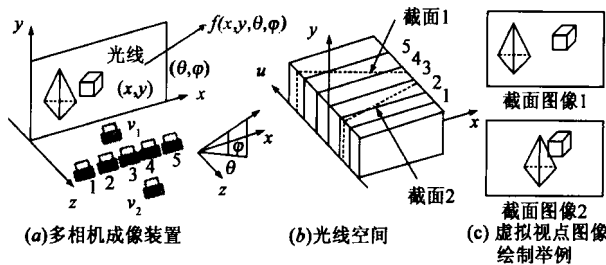


图1 自由视点视频系统的任意视点绘制

$\tan\theta\}$ 表示(如图 1(b)所示).光线空间表示的一个重要特征是任意视点图像可以从光线空间中简单采样获得^[6].当要生成任意视点图像时,仅需要从光线空间数据中垂直截取相应的切面即可,图 1(c)给出了生成的虚拟视点 1 和视点 2 图像.在实际应用中,若采用少量相机,那么转化得到的光线空间数据势必稀疏.因此,需要研究高质量的插值方法生成虚拟的中间视点图像.现有代表性的光线空间插值方法包括块匹配插值(BMI, block-based matching interpolation)^[7]和视差域滤波插值(DDFI, disparity domain filter interpolation)^[8].图 2 阐述了估计立体图像对中间 α 位置视差的原理.其左视点、右视点和待估计视点之间的强度存在以下关系:

$$I_\alpha(x, y) = I_L(x + \alpha \cdot d, y) = I_R(x + (\alpha - 1) \cdot d, y) \quad (1)$$

这里 $I(x, y)$ 表示虚拟视点图像上 (x, y) 处的强度.

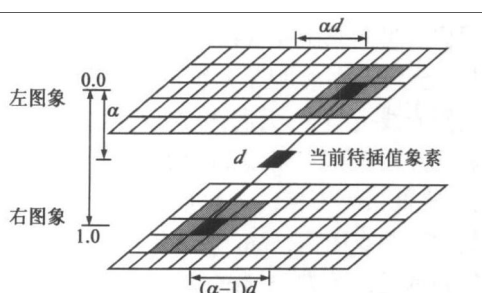


图2 视点为中心的视差估计

BMI 方法采用绝对误差和 SAD 作为相似性度量,从最大视差范围内选择使 SAD 最小的一对像素点来插值当前像素点.当每一个点都计算完后就可以得到稠密的视差图.DDFI 方法在能量函数中增加了平滑视差约束,确保当前点的视差与先前相邻点视差保持连续.这些方法的不足之处在于:

(1)采用固定大小的匹配窗口很难模拟物体的边界部分,因为它假设窗口内的像素点视差都是相同的.

(2)没有考虑遮挡问题,当两个实际相机的间距比较大的时候,这种现象不可避免.

(3)在光线空间插值方法中,实时性要求非常重要.上述两种算法的计算复杂度接近于 $O(mdw^2)$,这里 m 表示图像的像素数, d 表示最大视差, w 为匹配窗口大小.

鉴于以上问题,我们提出了一种基于改进动态规划的快速估计中间视点视差图方法,它能有效克服遮挡问题,并能很好地保持物体的边界,同时,可以明显降低计算复杂度,从而快速实现高质量的任意视点图像绘制.

本文算法包括三部分,首先,采用改进的动态规划算法求取扫描线间的相关性,将匹配像素投影到待插值的中间视点位置,得到稀疏的视差图,并采用空洞填补策略填补遮挡位置的视差.然后,利用左右视点中的对应像素双向插值得到稠密视差图中的匹配像素值;而对于遮挡像素,则根据遮挡类型,分别从左视点或右视点单向插值得到.最后,在得到密集多视点图像后,从光线空间中绘制出非相机平面的任意视点图像.

2.1 改进的动态规划算法

在动态规划方法中,相关性问题的可以看作是搜索一条使能量函数 $E(x, y, d)$ 代价最小的路径^[9],如式(2)所描述,该能量函数包括数据项 E_{data} 和平滑项 E_{smooth} 两部分.其中,数据项为匹配代价,如式(3)所示;平滑项为行间约束,与传统仅考虑当前扫描线的平滑约束不同,本文方法采用的约束如式(4)所示,增加了扫描线间约束.这样得到的视差图将有效地去除条纹效应.其中参数 n 为匹配窗口的一半大小, λ 为一常数,用于控制平滑的程度.

$$E(x, y, d) = E_{data}(x, y, d) + E_{smooth}(x, y, d) \quad (2)$$

$$E_{data}(x, y, d) = C(x, y, d) = \sum_{i=-n}^n \sum_{j=-n}^n |I_L(x+j, y+i) - I_R(x+d+j, y+i)| \quad (3)$$

$$E_{smooth}(x, y, d) = \lambda \cdot (|d(x-1, y) - d(x, y)| + |d(x, y-1) - d(x, y)|) \quad (4)$$

$$C(x, y+1, d) = C(x, y, d) + \Psi(x, y+1, d) \quad (5)$$

$$\Psi(x, y+1, d) = \sum_{j=-n}^n |I_L(x+j, y+n+1) - I_R(x+d+j, y+n+1)| - \sum_{j=-n}^n |I_L(x+j, y-n) - I_R(x+d+j, y-n)| \quad (6)$$

$$\Psi(x, y+1, d) = \Psi(x-1, y+1, d) + |I_L(x+n, y+n+1) - I_R(x+d+n, y+n+1)| - |I_L(x-n-1, y+n+1) - I_R(x+d-n-1, y+n+1)| - |I_L(x+n, y-n) - I_R(x+d+n, y-n)|$$

$$+ |I_L(x-n-1, y-n) - I_R(x+d-n-1, y-n)| \quad (7)$$

在上面公式中,计算 $C(x, y, d)$ 是最为耗时的. 为了提高算法的速度,新算法采用增量计算的方法来减少冗余计算.图 3 给出了增量计算方法的示意图.当得到 $C(x, y, d)$ 之后,可以按照式(5)得到 $C(x, y+1, d)$,其中 $\Psi(x, y+1, d)$ 表示匹配窗口中最上边行和最下边行的绝对误差和,如式(6)所示.进一步简化, $\Psi(x, y+1, d)$ 可以从 $\Psi(x-1, y+1, d)$ 获得,如式(7)所示,这样仅需要计算图中四个黑色标记的点之间的误差即可.通过增量计算方法,使得匹配代价的计算独立于所选择的匹配窗口大小,因此,新方法在基本不增加计算量的同时,获得更加精确的匹配代价.

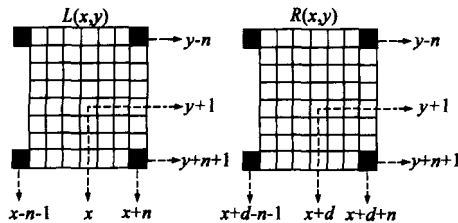


图3 增量计算方法示意图

2.2 中间视点图像插值与任意视点绘制

当采用改进动态规划方法估计得到扫描线之间的相关性后,将匹配的像素标记为匹配区,然后将这些像素点投影到虚拟视点位置,得到稀疏的视差图,并将空洞的像素标记为遮挡区,针对这些空洞,从左右两个方向分别查找它最邻近的标记为匹配区的像素点,相应的视差表示为 d_L 和 d_R ,用式(8)计算该点视差.由于在实际场景中,背景物体被前景物体所遮挡,因此,遮挡的像素点往往对应于背景物体.考虑到场景深度反映到视差图中具有视差越大离相机越近,视差越小离相机越远的特征,从左右两个匹配点中选取视差小的那个作为当前遮挡像素的视差并标记为左遮挡或右遮挡.这样就得到中间视点位置的稠密视差图.

$$d = \begin{cases} d_L, & \text{if } d_L \leq d_R \\ d_R, & \text{if } d_L > d_R \end{cases} \quad (8)$$

为了得到中间视点图像,采用图 4 所示方法插值,对于匹配区像素,从左右两个视点图像中线性插值出一对像素点,再结合当前位置 α ,线性插值出当前像素点.对于标记为左遮挡或右遮挡的像素点,分别从左图像或右图像线性插值出一个像素点,简单拷贝到当前

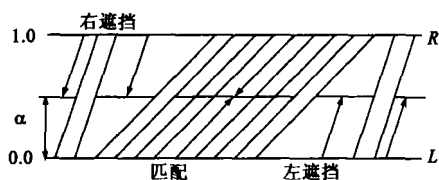


图4 中间视点图像插值

待插值像素位置即可.

为了生成图 5(a)中 $P(x, z)$ 处的虚拟图像,需要将前面插值得到的多视点图像转换到光线空间域(如图 5(b))中进行处理.虚拟视点与光线空间数据满足如下关系^[10]:

$$X = x + z \tan \theta \quad (9)$$

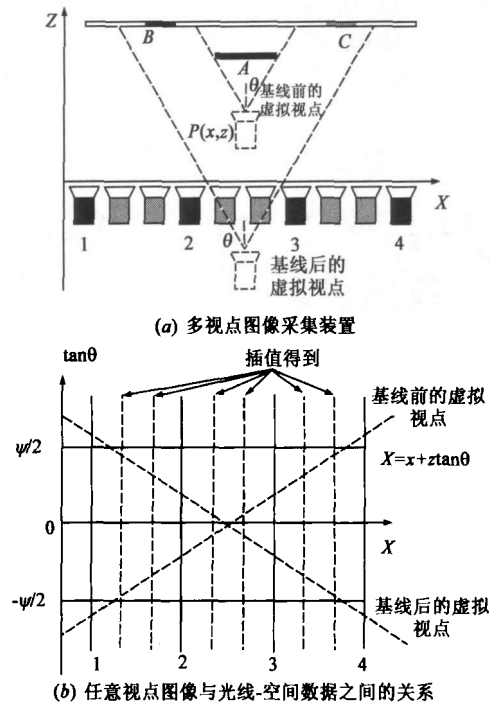


图5 任意视点绘制

图中参数 ψ 为视角范围.显然,相机平面前后的虚拟视点图绘制只要从光线空间中截取不同的切面即可实现.

3 实验结果

为了衡量虚拟视点图像绘制方法,将本文方法与 BMI、DDFI 方法进行比较.绘制虚拟视点图像采用 Middlebury 数据集 (Tsukuba, Sawtooth, Cones, Venus 和 Teddy); 绘制任意视点图像采用 Nagoya 大学提供的 Akko&Kayo 序列,该序列由 5 个水平相机采集,相邻两个相机间距为 5cm,最大视差为 20 个像素.所有实验都运行在 CPU 为 1.8G 的 PC 上,实验中,匹配窗口大小为 7×7 ,平滑参数 λ 为 7.

表 1 给出了各种方法插值中间视点图像的 PSNR 和运行时间.从 PSNR 度量上看,本文方法与 DDFI 和 BMI 方法相比至少提高了 0.8dB.而从运行时间上看,本文方法比 DDFI 方法减少了至少 10 倍.

图 6 给出了三种方法插值中间视点图像的结果比较.从效果来看,尽管主观质量的区不特别明显,但本文方法在任意视点图像绘制的遮挡区域要比 BMI 方法和 DDFI 方法更好.在图 6 中,本文方法的 PSNR 比

DDFI方法高出3.8dB,并且本文方法运算时间比其它两种方法降低了12倍多.实验结果表明本文方法所得到新视点图像效果是最好的,并明显提高了任意视点图像绘制速度.

表1 各种方法插值中间视点图像的PSNR和运行时间 单位:dB/秒

测试图像	BMI	DDFI	本文方法
Venus	36.67/3.25	36.77/3.24	38.02/0.42
Teddy	31.11/8.00	31.60/8.05	35.47/0.58
Cones	29.64/7.98	30.56/8.00	32.14/0.72
Tsukuba	28.81/6.77	30.30/6.72	31.12/0.56
Sawtooth	35.86/3.27	36.01/3.23	36.84/0.34



(a) BMI (PSNR=31.11dB) (b) DDFI (PSNR=31.60dB)
(c) 本文方法(PSNR=35.47dB) (d) BMI方法对应的视差图
(e) DDFI方法对应的视差图 (f) 本文方法对应的视差图
图6 各种方法绘制得到的中间视点图像(teddy)

在绘制任意视点图像的实验中,采用 Akko&Kayo 序列,每两个相机中间插值出 19 个中间视点图像.然后,将这些稠密的多视点图像投影到光线空间.当给定一个虚拟视点位置,相应的视点图像可以从稠密的光线空间中提取出来.图 7 给出了本文方法绘制的部分任意视点图像结果,图 7(a)是与实际相机同一水平线上的虚拟视点图像,图 7(b)和图 7(c)则分别是向前和向后的虚拟视点图像.向前的虚拟视点图像视角变窄,但更容易看清楚细节;向后的虚拟视点图像视角变宽,可以看清场景的全貌.由于没有考虑垂直方向上的插值,因此在图像的上下两侧有黑边出现,距离基线越远,黑边范围越大.这可以通过采用 2 维相机阵列采集系统的方式加以解决.

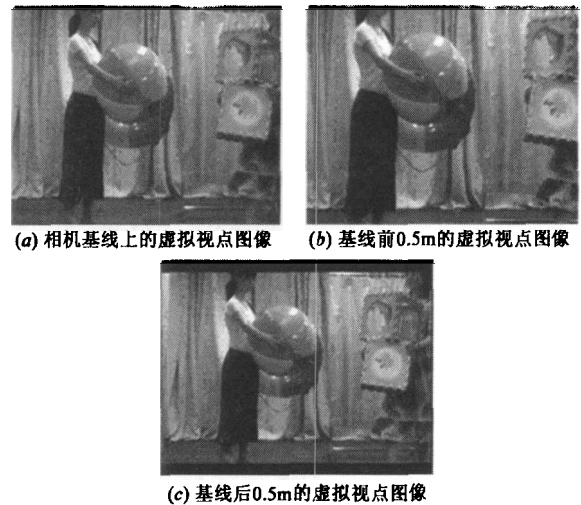


图7 本文方法绘制的各个位置的虚拟视点图像(Akko&Kayo)

4 结论

本文提出了基于改进动态规划算法的光线空间插值与任意视点绘制方法,与传统的块匹配插值方法和视差域滤波方法相比,新方法插值得到的中间视点图像质量和插值速度明显提高,由插值得到的光线空间数据可以进一步快速绘制出任意视点图像.

参考文献:

- [1] Smolic A, Kimata H, Vetro A. Development of MPEG standards for 3D and free viewpoint video [A]. Proc SPIE 6016 [C]. Boston: SPIE Press, Boston, Nov. 2005. 60160R.
- [2] Morvan Y, Farin D, De With P. System architecture for free-viewpoint video and 3D-TV [J]. IEEE Trans on Consumer Electronics, 2008, 54(2): 925 - 932.
- [3] Shum H-Y, Kang S, Chan S-C. Survey of image-based representations and compression techniques [J]. IEEE Trans on CSVT, 2003, 13(11): 1020 - 1037.
- [4] Fujii T, Tanimoto M. Free-Viewpoint TV (FTV) system [A]. Advances in Multimedia Information Processing-PCM 2004 [C]. Berlin/Heidelberg Springer, 2004. 497 - 504.
- [5] Smolic A, Kimata H. Applications and requirements for 3DAV [DB/OL]. ISO/IEC JTC1/SC29/WG11 N5877, July, 2003. <http://iphome.hhi.de/smolic/3DAV.html>, 2003 - 07.
- [6] Manoh K, Yendo T, Fujii T, Tanimoto M. Ray-space acquisition system of all-around convergent views using a rotation mirror [A]. Proc SPIE 6778 [C]. Boston: SPIE Press, Boston, Sept. 2007. 67780C.
- [7] Jiang G, Yu M, Ye X, Fan L, Fu R. New method of ray-space interpolation for free viewpoint video [A]. Proc of IEEE International Conference on Image Processing [C]. Piscataway: IEEE Press, Genova, 2005. 1138 - 1141.
- [8] Droese M, Fujii T, Tanimoto M. Ray-space interpolation based on filtering in disparity domain [A]. Proc of 3D Image Confer-

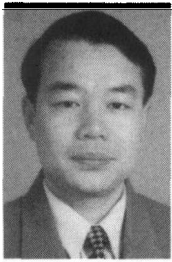
ence[C]. Tokyo, 2004. 213 - 216.

- [9] Wang L, Liao M, Gong M, et al. High-quality real-time stereo using adaptive cost aggregation and dynamic programming[A]. Proc of Int. Symp. on 3D Data Processing, Visualization, and Transmission[C]. Piscataway: IEEE Press, Chapel Hill, June

2006. 798 - 805.

- [10] Fujii T, Kimoto T, Tanimoto M. A new flexible acquisition system of ray-space data for arbitrary objects[J]. IEEE Trans on CSVT, 2000, 10(2): 218 - 224.

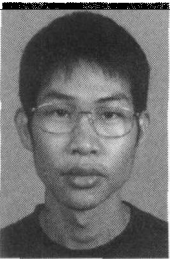
作者简介:



蒋刚毅 男, 1964 年生于浙江绍兴, 博士, 宁波大学信息科学与工程学院教授、博士生导师。主要研究领域为多媒体信息传输与信息安全、图像处理与视频信号编码。
E-mail: jianggangyi@nbu.edu.cn



郁梅 女, 1968 年生于江苏无锡, 博士, 宁波大学信息科学与工程学院教授, 硕士生导师。研究领域包括多媒体信号处理、编码与通信, 计算机视觉及其应用。



范良忠 男, 1964 年生于浙江庆元, 博士, 浙江大学宁波理工学院讲师。主要研究领域为图像处理、多媒体技术、基于视觉的智能控制。



陈恩 男, 1962 年生于四川成都, 博士, 宁波大学信息科学与工程学院副教授, 硕士生导师。主要研究领域为机器视觉图像处理、视频图像处理等。