

# 一种用于视点图像绘制的光线空间插值方法

符冉迪<sup>1),4)</sup> 范良忠<sup>1),2),3)</sup> 郁梅<sup>1),4)</sup>

<sup>1)</sup>(宁波大学电路与系统研究所,宁波 315211) <sup>2)</sup>(中国科学院计算技术研究所,北京 100080)

<sup>3)</sup>(中国科学院研究生院,北京 100080) <sup>4)</sup>(北京大学视听信息处理国家重点实验室,北京 100871)

**摘要** 光线空间表示是实现实时复杂场景自由视点视频的有效方法。提出了基于特征点检测的快速光线空间插值方法,首先采用自适应阈值提取光线空间片或图像域的特征点,然后结合遮挡问题采用相关性准则匹配特征点,根据匹配的特征点将光线空间分割成不同对象区域并进行插值,以实现快速任意视点图像绘制。实验结果表明,该方法能快速插值光线空间数据,生成的光线空间数据信噪比高、绘制的任意视点图像的主客观质量也较高。

**关键词** 图像绘制 光线空间插值 特征点

中图法分类号:TN911.73 文献标识码:A 文章编号:1006-8961(2008)05-0998-05

## A Ray-space Interpolation Method for Viewpoint Image Rendering

FU Ran-di<sup>1),4)</sup>, FAN Liang-zhong<sup>1),2),3)</sup>, YU Mei<sup>1),4)</sup>

<sup>1)</sup>(Institute of Circuits and Systems, Ningbo University, Ningbo 315211)

<sup>2)</sup>(Institute of Computer Technology, Beijing 100080)

<sup>3)</sup>(Graduate School of Chinese Academy of Sciences, Beijing 100080)

<sup>4)</sup>(National Laboratory on Machine Perception, Peking University, Beijing 100871)

**Abstract** Ray-space representation is an effective approach to realizing real-time free viewpoint video system with complicated scenes. A feature point based ray-space interpolation method is proposed for arbitrary viewpoint image rendering in this paper. At first, feature points of ray-space slice or image are extracted, then, considering occluded problem, the feature points are matched based on correlation-based criterion, and each ray-space slice is segmented into different regions. Finally, an interpolation operation is applied to the segmented regions. Experimental results show that the proposed method can interpolate ray-space data fast and render arbitrary viewpoint images, and PSNR of the interpolated ray-space data is high, as well as quality of the rendered arbitrary viewpoint images.

**Keywords** Viewpoint image rendering, ray-space interpolation, feature point

## 1 引言

交互性是下一代媒体研究的主要内容,自由地选择视点位置观察场景的不同侧面是交互性的一种重要体现。实现这一目标主要有两类方法,一类是对场景进行3维建模,通过投影、裁减、消隐、纹理映射和光照

模型加以渲染。虽然3维建模技术能够表示复杂场景,但构造过程非常复杂,在生成复杂场景的真实感图像和实时绘制方面存在明显的困难。另一类方法称为基于图像的绘制,它从不同位置和方向对真实场景采样,然后利用这些采样图像间的几何关系绘制任意视点<sup>[1]</sup>。由于它不需要对场景建模,并且视点绘制利用真实采样的图像,因此具有绘制速度快,真实感强的优

**基金项目:**国家自然科学基金项目(60472100,60672073);教育部科学技术研究重点项目(206059);浙江省自然科学基金项目(Y105577);浙江省教育厅科研计划项目(20070870)

**收稿日期:**2006-12-20; **改回日期:**2007-05-31

**第一作者简介:**符冉迪(1971~),男,讲师。主要从事计算机图像处理和模式识别技术研究。E-mail: furandi@nbu.edu.cn

点。光线空间表示是一种基于图像绘制的方法,是实现复杂场景实时自由视点视频系统的一种有效途径<sup>[2]</sup>。光线空间表示方法需要大量相机从不同方位对场景实时采样,而实际系统中相机不可能安置得十分紧密,这样经变换后只能得到稀疏的光线空间数据,必须通过插值生成稠密的光线空间数据。为了适合实时应用,插值算法应尽可能满足计算量小、易于实现等要求<sup>[3]</sup>。本文讨论分析了光线空间的插值技术,提出了基于遮挡分析与特征点检测的快速光线空间插值新算法。

## 2 光线空间插值算法

### 2.1 光线空间表示与插值问题

光线空间方法将3D空间的光线用4维函数  $\{f(x, y, \theta, \phi)\}$  来表示,如图1(a)所示,其中,  $(x, y)$  表示光线与参考平面  $z=0$  的交点坐标,  $(\theta, \phi)$  表示光线的方向,  $f(x, y, \theta, \phi)$  表示光线的强度。为简化问题描述,在垂直视差  $\phi$  恒定时,考虑如图1(b)所示3D光线空间数据  $\{f(x, y, u = \tan\theta)\}$ 。图1(b)中,定义平面  $\{f(x, y_k, u)\}$  (当  $y_k$  为一常数) 为一个片。在基于光线空间的自由视点视频系统中,某一视点图像的绘制仅仅是从光线空间中截取与该视点相对应的特定平面的数据。因此,光线空间数据插值效果决定了绘制的任意视点图像的质量。光线空间插值方法主要有基于像素匹配的光线空间插值 (pixel-based matching interpolation PMI) 和基于块匹配的光线空间插值 (block-based matching interpolation BMI) 等<sup>[4]</sup>; 当视差范围较大或纹理信息较少时, BMI 方法优于 PMI 方法。光线空间插值的难点在于寻找最佳的像素对来插值当前待插像素,它需要解决噪声干扰、无纹理区域或周期性纹理区域的误匹配、遮挡区域的检测等问题。

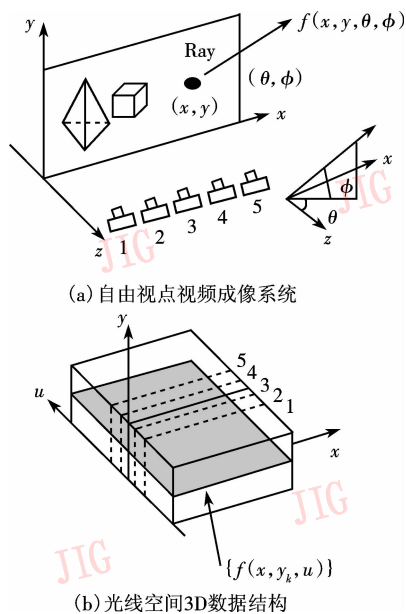


图1 自由视点视频系统与光线空间数据结构  
Fig.1 Free viewpoint video system and ray-space data

### 2.2 遮挡现象与特征点检测

图2给出了对3个具有不同深度的简单对象进行多视点图像采集,然后转化得到光线空间片。从场景来看,对象A距离相机最近,对象A的斜线斜率最大,它将对象B和对象C部分遮挡,反映在光线空间片中,对象B和C所对应的斜线被对象A对应的斜线部分覆盖;如对象B在相机0和相机1的成像中是可见的、而相机2成像中则是被部分遮挡的,这种遮挡现象会影响虚拟视点图像的绘制质量。有效的插值算法必须能快速地检测遮挡区域,这里首先提取光线空间片(或图像域)的特征点,对特征点进行匹配,并考虑遮挡和噪声问题,然后采用基于区域的插值方法进行插值以实现视点绘制。

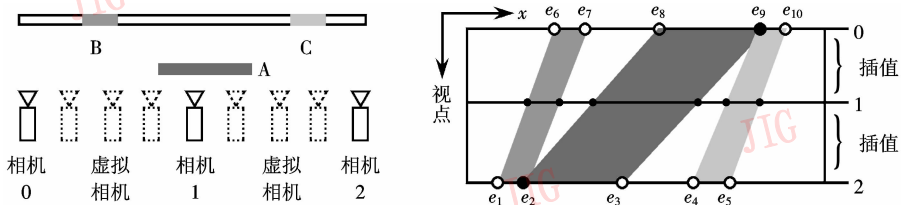


图2 多视点成像中的遮挡现象  
Fig.2 Occlusion in multiview imaging

自由视点视频成像系统中的  $L$  个平行相机同时成像将形成尺寸为  $M \times N$  的  $L$  个视点图像  $\{X_1(i, j)\}, \{X_2(i, j)\}, \dots, \{X_L(i, j)\}$ , 则可构成  $N$  个大小为  $M \times L$  的稀疏光线空间片  $\{I_1(x, u)\}, \dots, \{I_L(x,$

$u)\}, \dots, \{I_N(x, u)\}$ 。

#### (1) 图像域的特征提取方法

对于最左边、最右边视点图像采用 SUSAN (small univalue segment assimilating nucleus) 算子进

行图像域的边缘和角点的检测。其阈值的选取等可参照已有相关方法<sup>[5]</sup>。在图像域提取的特征,再对应到光线空间域中,进行光线空间片中的特征点多行匹配,基于块匹配,以确定准确的特征点集,并在匹配过程中将光线空间片分割为不同的对象区域,包括遮挡区与非遮挡区。

(2) 光线空间域上的特征提取方式

令光线空间片  $\{I_k(x, u)\}$  的梯度图像为  $\{G_k(x, u)\}$ , 可由  $G_k(x, u) = |I_k(x + 1, u) - I_k(x, u)|$  计算得到。为降低计算复杂度,在确定特征点时的阈值  $T$  仅由每个光线空间片的第 1 行计算得到, 设第 1 行梯度图像的均值和方差为  $\mu$  和  $\sigma$ , 则阈值为

$$T = \begin{cases} \mu + \sigma & \mu + \sigma \leq T_0 \\ T_0 & \mu + \sigma > T_0 \end{cases} \quad (1)$$

这里,当  $\mu + \sigma$  较大时,将阈值  $T$  设成  $T_0$ , 以防止约束过强而漏检深度不连续点,  $T_0$  是由实验确定的经验常数。对于每一  $(x, u)$ , 若  $G_k(x, u) \geq T$ , 则该点为特征点; 否则为非特征点。由于左遮挡和右遮挡现象主要发生在最左边和最右边的相机对应的扫描线上。因此,采用较小的阈值提取光线空间片第 1 行和最后 1 行的特征点, 它们对应于多视点图像中最左边和最右边的视点图像。若有多个连续特征点, 则选取强度变化最大点作为候选特征点, 然后将这两行的特征点在其他行上进行匹配。

(3) 光线空间片中的特征点筛选

PMI 与 BMI 方法在进行像素点匹配时仅参考最邻近的两行图像数据, 没有充分利用多视点图像空间强度一致性的特征。针对第 1 行和最后一行被提取出来的特征点, 在匹配这些特征点的方向时考虑空间一致性约束。图 3 中  $(x, u)$  位置处的黑色方块为当前待匹配的特征点, 估计该点所在的光线方向时, 采用多行原始数据来计算相似性度量

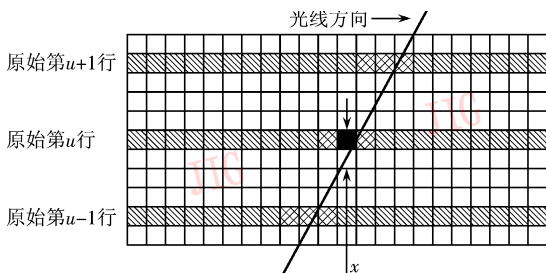


图 3 光线空间片中的特征点多行匹配方法

Fig. 3 Multiple lines matching method in ray space slice

$$\rho(d) = \sum_{k=-M}^M \sum_{w=-W}^W |I(x+w+d_{u+k}, u+k) - I(x+w, u)| \quad (4)$$

这里,  $M$  为靠近  $u$  的行数,  $W$  为块匹配窗口大小,  $d$  表示第  $u$  行和第  $u+1$  行间的视差,  $d_{u+k}$  表示第  $u$  行和  $u+k$  行间的视差,  $d_{u+k} = kd$ , 则当前特征点所在的光线方向为  $d_{best} = \arg\{\min \rho(d)\}$ 。多行匹配方法能较准确地计算出第 1 行和最后一行中的特征点的光线方向。然后针对每个特征点, 基于块匹配, 分别对最下面行和最上面行中的特征点在相邻行中寻找最佳匹配点, 当匹配点的视差  $d_{match}$  满足  $|d_{match} - d_{best}| \leq 2$  时, 则将该匹配点添加到  $f(x, u)$  中, 并且以该点为新的特征点继续去下一行中匹配。如果不满足上面的条件, 则认为该初始特征点为遮挡点或噪声点, 将该点重新置为非特征点。通过将匹配准确的特征点继续延伸的方法进行特征点筛选, 从而得到比较准确的特征点集合, 这些提取出来的特征点将稀疏的光线空间片分割为不同的对象区域。

2.3 基于特征点检测与遮挡的光线空间插值方法

特征点图像将光线空间片分割成不同斜率方向的区域, 在区域内部, 光线方向基本一致。因此, 采用基于区域的插值算法能快速有效地插值稀疏的光线空间片。图 4 给出了基于区域的光线空间插值示意图, 其插值过程包括:

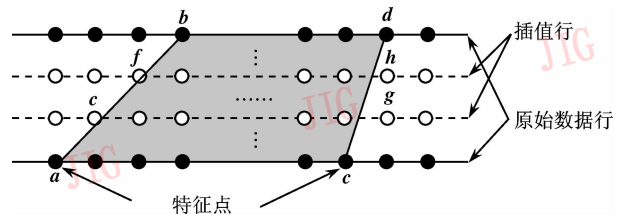


图 4 基于区域的特征点检测光线空间插值

Fig. 4 Region-based ray-space interpolation method with detection of feature points

(1)  $a$  和  $c$  为特征图像上最邻近的两个特征点, 它们在上一行中匹配的特征点分别为  $b$  和  $d$ 。

(2) 根据  $a$  与  $b$  及  $c$  与  $d$  所在的光线在同一直线上的特点, 代入直线方程计算出中间待插值行上的点  $e$ 、 $f$  和  $g$ 、 $h$  4 个点。

(3)  $e$ 、 $f$  和  $g$ 、 $h$  的强度值分别通过  $a$ 、 $b$  和  $c$ 、 $d$  双线性插值得到。

(4)  $e$  与  $g$  之间或  $f$  与  $h$  之间的像素根据四边形的比例关系线性插值得到。

综上所述, 整个基于遮挡与特征点检测的光线

空间插值算法表述以下:

- (1) 对稀疏光线空间片或视点图像,检测特征点候选点;
- (2) 对特征点候选点,采用多行匹配进行拟合,将光线空间片分解为多个区域;
- (3) 对不同区域,采用基于区域的快速插值。

### 3 实验与分析

实验采用“Xmas”、“Cup”2个多视点彩色图像集作为测试集,图5为其中最左和最右两个视点图像。在CPU为P4 1.7G Hz,内存为512M的PC上实验。“Xmas”来自于Tanimoto实验室,共有101个视点图像,相机间距是3mm,图像分辨率为 $640 \times 480$ pixel;“Cup”通过一个相机在云台上平移采集图像得到,每次移动距离1mm,分辨率及采样数量和“Xmas”序列一样。“Cup”中场景物体离相机距离比较近,其最大水平视差较“Xmas”序列大,遮挡现象明显。这样转换得到480个稠密的光线空间片的分辨率为 $640 \times 101$ pixel。实验根据不同的相机数量插值得到稠密的光线空间数据,从101个视点图像中等间距地选取5个、6个、11个或21个图像来验证插值算法,转换得到的光线空间片的分辨率分别为 $640 \times 5$ pixel、 $640 \times 6$ pixel、 $640 \times 11$ pixel或 $640 \times 21$ pixel,对应需要插值25倍、20倍、10倍或5倍,将插值得到稠密的光线空间片与实际采集得到稠密的光线空间片进行比较。



(a) Xmas的最左最右视点



(b) Cup的最左最右视点

图5 测试彩色图像集

Fig. 5 Test color images

表1为480个光线空间片在不同相机间距下采用BMI方法和本文方法的平均插值PSNR和平均插

表1 光线空间片插值平均PSNR和运行时间

Tab. 1 Average PSNR and computational time of interpolated ray-space slice

测试序列	相机间距(mm)	BMI	方案1	方案2	
				单位: dB/ms	
				$T_0 = 5$	$T_0 = 8$
Cup (彩色)	5	33.84/165	34.88/33	34.89/39	34.89/37
	10	33.11/335	34.19/37	34.23/41	34.23/39
	20	31.58/672	33.34/38	33.50/44	33.50/41
Xmas (彩色)	25	30.35/839	31.87/38	31.99/48	31.98/43
	15	39.31/97	41.03/35	41.03/37	41.03/35
	30	38.81/185	40.18/38	40.28/39	40.27/38
Xmas (彩色)	60	37.72/354	39.47/40	39.74/40	39.71/37
	75	36.60/438	39.18/39	39.53/39	39.50/39

值时间的比较,本文方法有两种方案:方案1对应于图像域特征提取的插值新方法;方案2对应于光线空间域特征提取的插值新方法。表1显示在一定范围内不同阈值 $T_0$ 的选取对本文方法插值后的PSNR和运算时间影响不明显,方案2要略微优于方案1,这是由于方案1中特征点是采用全图像提取的,造成某些像素位置上的特征点丢失。整体上,本文方法的2个方案插值效果明显优于BMI方法,并且插值时间在不同相机间距下变化不太明显,而BMI方法的插值时间随相机间距增大而增加。本文方法在插值时间上的优势得益于需要进行匹配的像素数量相对于BMI方法大大降低了。BMI方法需要对光线空间片中所有待插值像素进行方向搜索,即实施匹配过程,而本文方法仅对特征点进行匹配,其余非特征点的光线方向由特征点的方向插值得到,由此降低了整个插值算法的运行时间。此外,由于特征点位于纹理丰富区域,相对于其他非纹理丰富区域的点能得到准确度更高的匹配。因此,本文方法插值得到的光线空间片具有更高的信噪比,相应绘制的虚拟视点图像质量也较BMI方法有明显提高。图6给出了采用方案2选取 $T_0$ 为8时对Cup中第135个稀疏的光线空间片插值结果。

图7给出了由BMI方法和本文方法绘制的Cup序列第40个虚拟视点图像的局部图像比较。BMI方法绘制图像在纸杯上的文字很不清晰,这是由于BMI方法在计算光线方向时没有考虑无纹理区域的误匹配、噪声及遮挡等因素的影响,选择错误的像素点插值所造成的。而本文插值方法绘制出来的文字

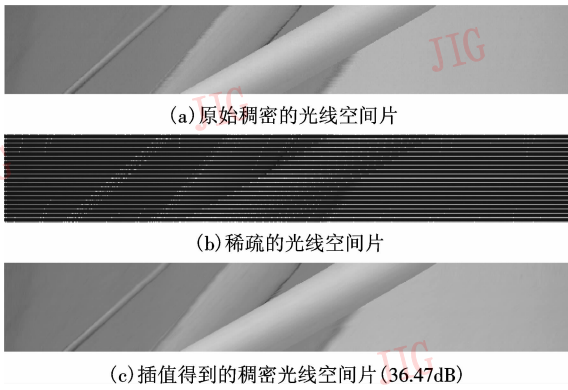


图 6 Cup 的第 135 个光线空间片的插值结果

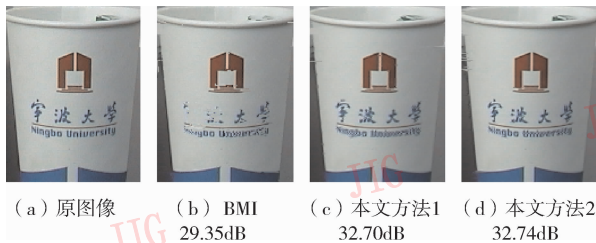
Fig. 6 Interpolating result of 135<sup>th</sup> ray-space slice in Cup

图 7 绘制的第 40 个虚拟视点的局部图像比较

Fig. 7 Comparison of the rendered 40<sup>th</sup> viewpoint images in Cup

则比较清楚,信噪比也明显高得多。此外,利用本文方法能快捷、精确地插值生成稠密的光线空间数据场,进而通过简单地从光线空间数据场中读取相应数据,即可生成任意视点的虚拟立体图像对。图 8 为利用由本文方法在相机间距为 65mm 时得到的光线空间数据绘制的虚拟立体图像对。

图 8 绘制的虚拟立体图像对  
(虚拟立体相机间隔 65mm)Fig. 8 The rendered virtue stereo image pair  
(the interval between two virtue stereo cameras is 65mm)

## 4 结 论

基于光线空间表示的自由视点视频的主要优点在于无需进行复杂场景分析和 3D 建模就可方便地绘制虚拟视点图像,且对场景没有限制,绘制虚拟视点图像与立体图像对的代价低。然而基于光线空间的方法需要大量的相机来生成有效的光线空间,因此,有效的光线空间插值算法是基于光线空间表示的自由视点视频的核心技术之一,对降低系统复杂度、提高合成图像质量有至关重要的作用。本文提出的基于光线空间域特征提取的光线空间插值算法能有效绘制高质量的虚拟视点图像,且算法计算量小。

## 参考文献 (References)

- 1 Shum H Y, Kang S B, Chan S C. Survey of image-based representations and compression techniques[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, **13**(11): 1020 ~ 1037.
- 2 Smolic A, McCutchen D. 3DAV (3D audio video) exploration of videobased rendering technology in MPEG[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, **14**(3): 348 ~ 356.
- 3 Fan Liang-zhong, Jiang Gang-yi, Yu Mei. Realization of FTV (Free-viewpoint TeleVision) based on ray space representation[J]. Journal of Computer-Aided Design & Computer Graphics, 2006, **18**(2): 170 ~ 179. [范良忠, 蒋刚毅, 郁梅. 自由视点电视的光线空间实现方法[J]. 计算机辅助设计与图形学报, 2006, **18**(2): 170 ~ 179.]
- 4 Tehrani MP, Fujii T, Tanimoto M. Offset block matching of multi-view images for ray-space interpolation[J]. The journal of the institute of Image information and Television Engineers, 2004, **58**(4): 540 ~ 548.
- 5 Liu Bo, Zhong Si-dong. A SUSAN corner detector based on adaptive threshold[J]. Infrared Technology, 2006, **28**(6): 331 ~ 333. [刘博, 仲思东. 一种基于自适应阈值的 SUSAN 角点提取方法[J]. 红外技术, 2006, **28**(6): 331 ~ 333.]