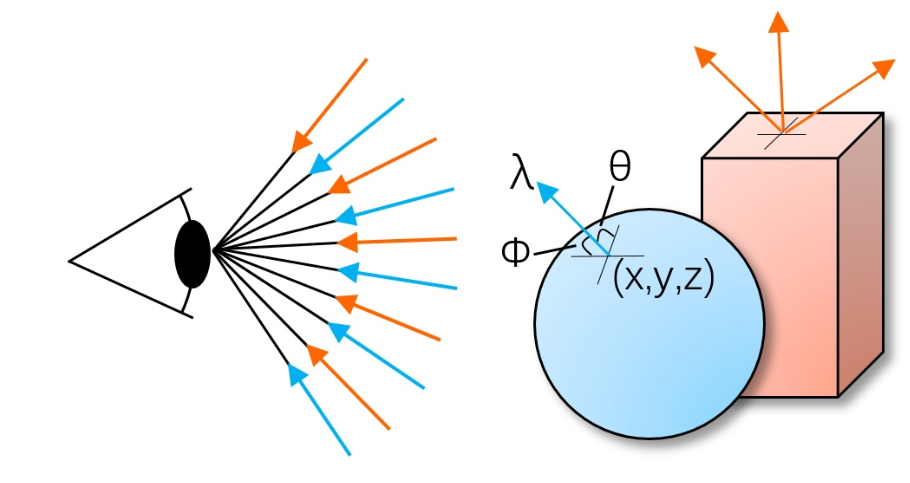
2020年03月18日

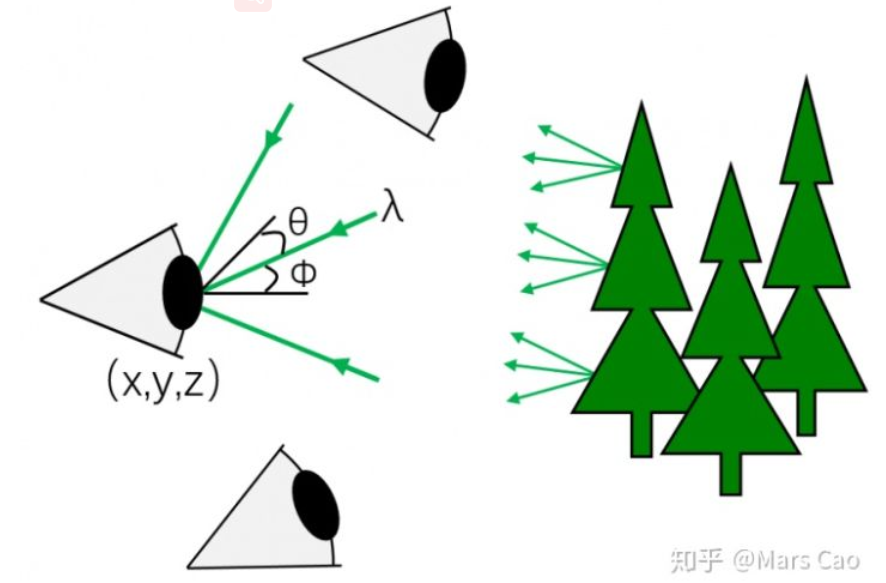
Mars说光场（1）— 光场技术综述

**摘要：**

全光函数（Plenoptic Function）包含7个维度，是表示光场的数学模型。由于光路是可逆的，以发光表面为中心来描述光线集合衍生出与光场类似的概念——反射场（Reflectance Field）。反射场也具有7个维度的信息，但每个维度的定义与光场不尽相同。但实际情况一般只采用4D光场模型，从而降低7个维度导致的采集、处理、传输的负担。

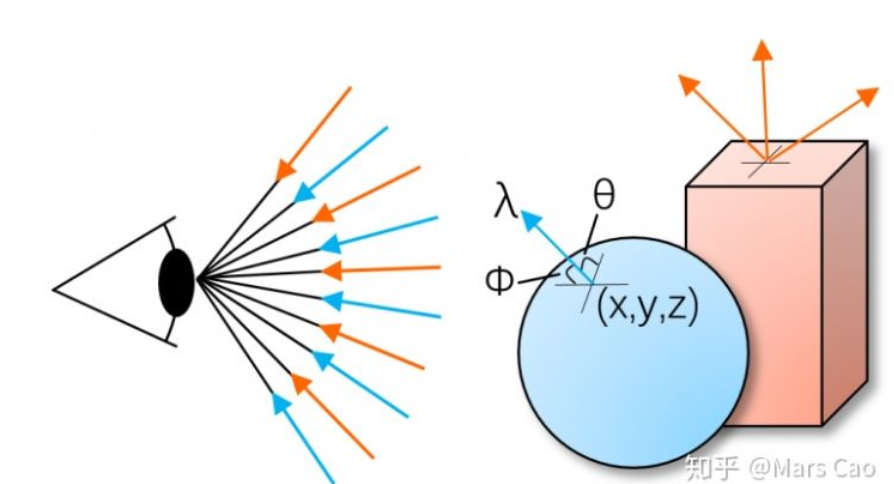
**一、历史进程：**

Light Field”最早出现在Alexander Gershun于1936年在莫斯科发表的一篇经典文章中，由美国MIT的Parry Moon和Gregory Timoshenko在1939年翻译为英文[2]，Parry Moon在1981年提出的“Photic Field”才是当前学术界所研究的“光场”[3]。光场的采集[4]和3D显示[5,6]两个方面，现在成为主要的研究方向，开设的课程——计算摄像学（Computational Photography）



1、什么是光场：

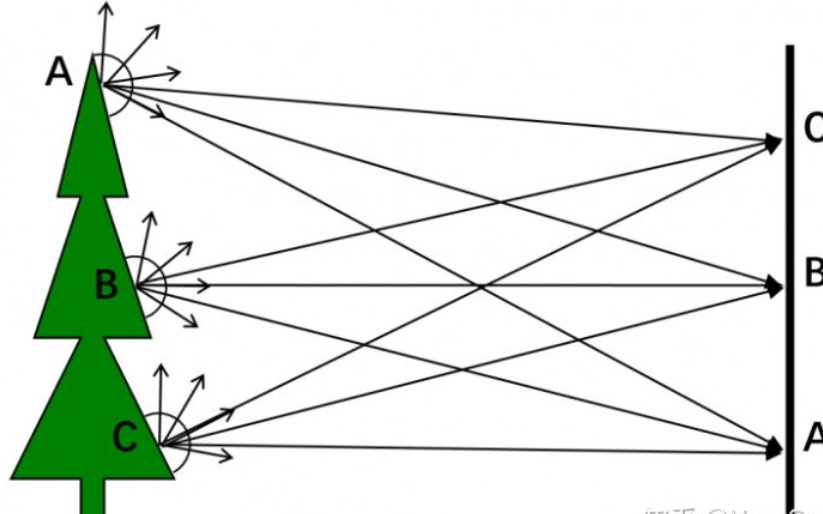
光场（Light field）：用（x, y, z）表示人眼在三维空间中的位置坐标，光线可以从不同的角度进入人眼，用（θ, Φ）表示进入人眼光线的水平夹角和垂直夹角，每条光线具有不同的颜色和亮度，可以用光线的波长（λ）来统一表示，进入人眼的光线随着时间（t）的推移会发生变化。所以三维世界中的光线可以表示为7个维度的全光函数（Plenoptic Function,）**P(x, y, z, θ, Φ, λ, t)。**



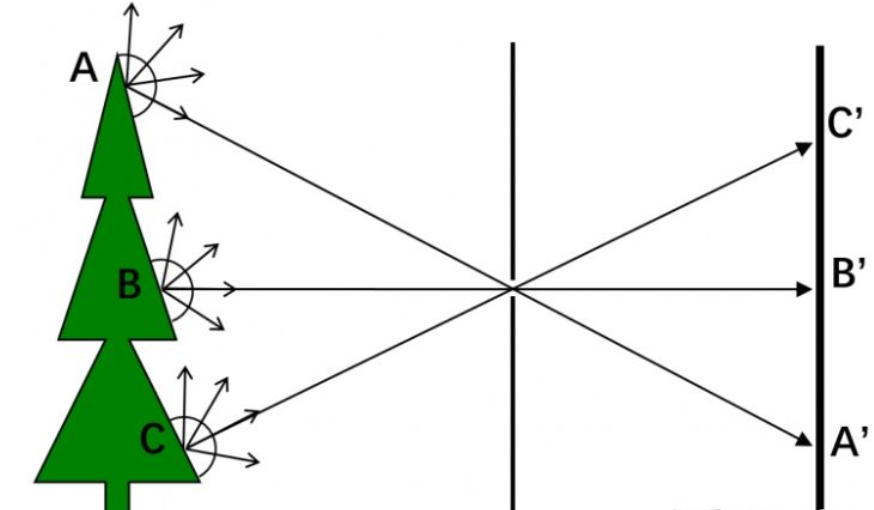
反射场（Reflectance Field）： 物体表面发光点的位置可以用（x, y, z）三个维度来表示；对于物体表面的一个发光点，总是向180度半球范围内发光，其发光方向可以用水平角度和垂直角度（θ, Φ）来表示；发出光线的波长表示为（λ）；物体表面的光线随着时间（t）的推移会发生变化。

两者主要的区别是：光场是以人眼为中心描述空间中所有的光线，反射场是以物体表面发光点为中心描述空间中所有的光线。光场中的任一条光线都可以在反射场中找到。（光场所描述的光线集合与反射场所描述的光线集合是完全一致的。）

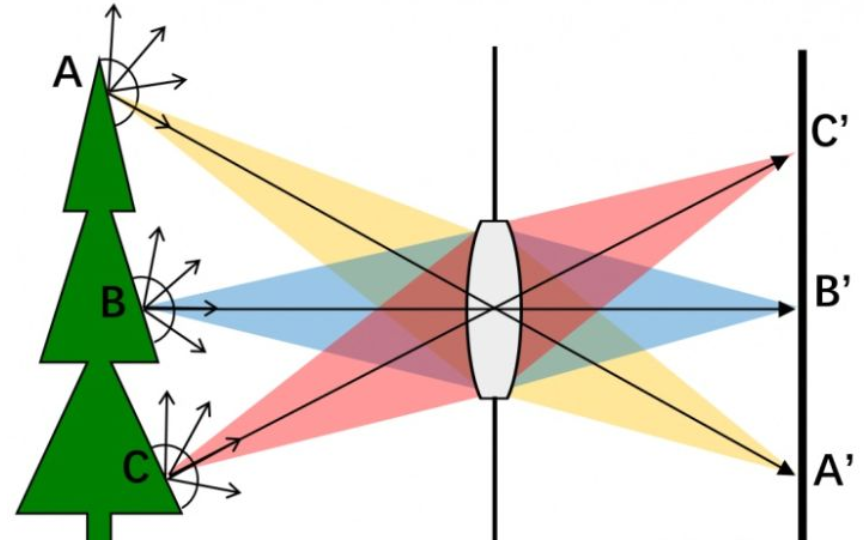
2、为什么研究光场：



物体表面A、B、C三点都在向半球180度范围内发出光线，对于成像传感器（CCD）上的感光像素A'会同时接收到来自A、B、C三点的光线，因此A'点的像素值近似为物体表面上A、B、C三点的平均值。类似的情况也会发生在CCD上的B'和C'点的像素。因此，如果把相机上的镜头去掉，那么拍摄的图片将是噪声图像。



小孔成像的原理：CCD上A'点只接收到来自物体表面A点的光线。类似的，CCD上B'和C'点也相应只接收到物体表面B点和C的点光线。因此，在CCD上可以成倒立的像。

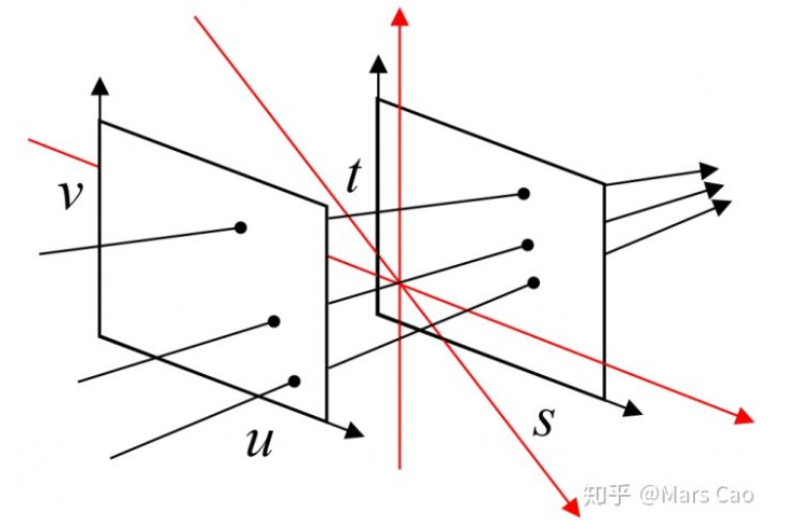


实际相机成像原理：（因为小孔直径较小会导致通光亮非常小，信噪比非常低，成像传感器无法采集到有效的信号；如果小孔直径足够小，当与光波长相当时还会产生**衍射**现象。而小孔直径过大会导致**成像模糊**。）现代的成像设备用透镜来替代小孔，从而既能保证足够的通光量，又避免了成像模糊。物体表面A点在一定角度范围内发出的光线经过透镜聚焦在成像传感器A’点，并对该角度范围内所有光线进行积分，积分结果作为A点像素值。这大大增加了成像的信噪比，但同时也将A点在该角度范围内各方向的光线耦合在一起。

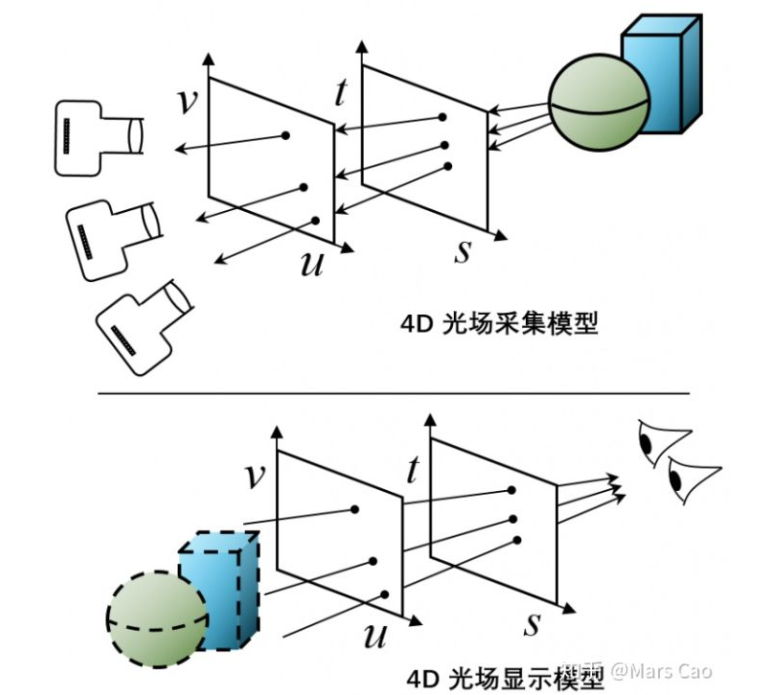
注：小孔成像模型是光场成像的一种降维形式，只采集了（x, y, λ, t）四个维度的信息。RGB-D相机多了一个维度信息（x, y, z, λ, t）。相比于全光函数，主要丢失的是光线的方向信息（θ, Φ），这就带来了一定的问题：在图像采集方面，可以通过调节焦距来选择聚焦平面，然而无论如何调节都只能确保一个平面清晰成像，而太近或太远的物体都会成像模糊。

3、光场4D参数化

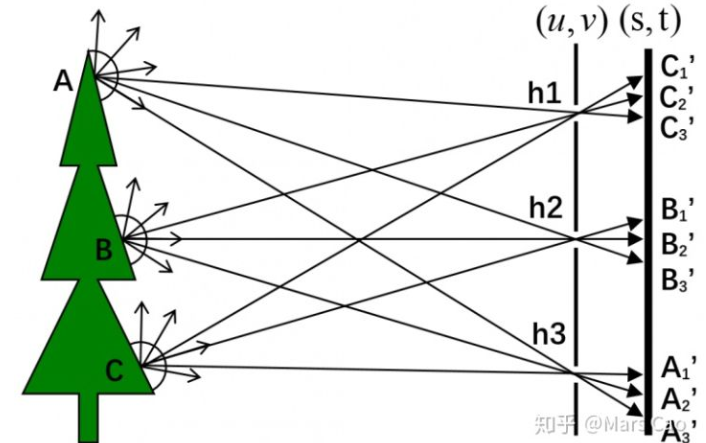
由于7D全光函数的描述，难以对7D光场这么庞大的数据进行实时处理和传输，因此需要对7D光场进行简化降维



（u,v,s,t）4D光场模型：两个不共面的平面（u,v）和（s,t），如果一条光线与这两个平面各有一个交点，则该光线可以用这两个交点唯一表示。（4D模型并不能完备地描述三维空间中所有的光线，与（u,v）或（s,t）平面所平行的光线就不能被该4D模型所表示，但是其可以完备的表示人眼成像所需要的全部光线，因为当光线与人眼前视方向垂直时，该光线不会进入人眼。）



4D光场模型具有可逆性，既能表示光场采集，又能表示光场显示。对于光场采集模型，右侧物体发出的光线经过（s,t）和（u,v）平面的4D参数化表示，被记录成4D光场。对于光场显示模型，经过（u,v）和（s,t）平面的调制可以模拟出左侧物体表面的光线，从而使人眼“看见”并不存在的物体。



物体表面A、B、C三点发出的光线首先到达（u,v）平面，假设（u,v）平面上有三个小孔h1、h2、h3，则A、B、C三点发出的光线经三个小孔分别到达（s,t）平面。A、B、C三点在半球范围内三个不同方向的光线被同时记录下来，例如A点三个方向的光线分别被（s,t）平面上A3’、B3’、C3’记录。如果（u,v）平面上小孔数量更多，且（s,t）平面上的像素足够密集，则可以采集到空间中更多方向的光线。图中展示的是（u,v）（s,t）光场采集模型在垂直方向上的切面图，实际上可以采集到A、B、C三点9个不同方向（3x3）的光线。

图像分辨率和FOV（Field Of View）是传统相机成像性能的主要指标。衡量4D光场的指标不仅有图像分辨率和FOV，还有角度分辨率和FOP（Field Of Parallax）。

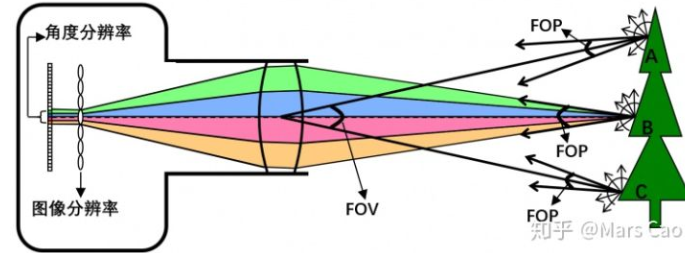
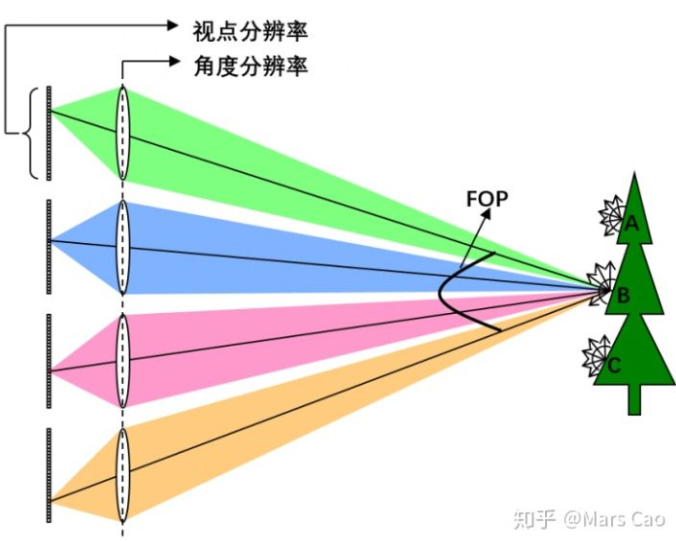


图9展示了基于微透镜阵列的光场相机的光路示意图，物体表面发出的光线进入相机光圈，然后被解耦和并分别被记录下来。以B点为例，发光点B在半球范围内发出各向异性的光线，但并不是所有的光线都进入相机光圈，只有一定角度内的光线被成功采集，被光场相机采集到的光线的角度范围决定了能够观察的最大视差范围，我们记这个角度为FOP。换句话说，图9中光场相机只能采集到B点FOP角度范围内的光线。但FOP的大小随着发光点与光场相机的距离远近而不同，因此通常采用基线的长度来衡量FOP的大小，图中**主镜头的光圈直径等效为基线长度。（角分辨率：当两件物件相距太近，影像重叠太多，两个影像便无法分办。角分辨度 θmin，就是当两件物件刚好能够分辨时，它们与孔隙所形成的夹角。）**

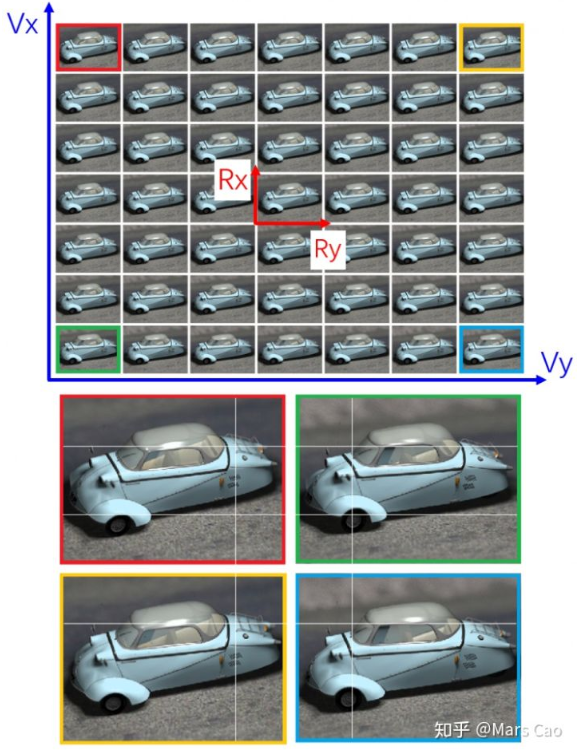
B点在FOP角度范围内的光线被微透镜分成4x4束光线，光场相机的角度分辨率即为4x4，光场相机的角度分辨率表征了一个发光点在FOP角度范围内的光线被离散化的程度。而基于小孔成像模型相机的角度分辨率始终为1x1。光场的视点图像分辨率同样表征了被采集场景表面离散化程度，成像传感器分辨率除以角度分辨率即为视点图像分辨率。



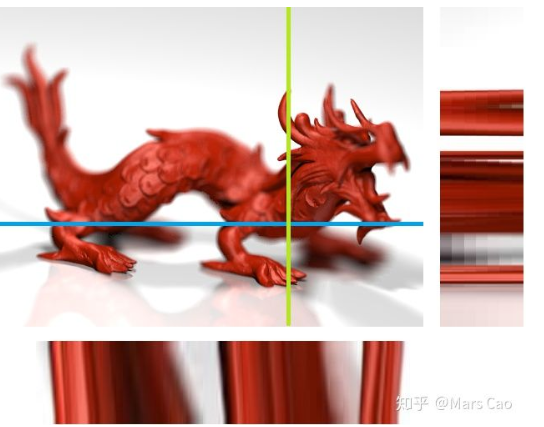
基于相机阵列的光场相机同样可以用视点图像分辨率、角度分辨率、FOV、FOP四个参数来衡量光场相机的各方面性能。B点半球范围内发出的光线中FOP角度范围内的光线被相机阵列分成4x4束并分别被采集。相机的个数4x4即为角度分辨率，单个相机成像传感器的分辨率即为视点图像分辨率。所有相机FOV的交集可以等效为光场相机的FOV。基于相机阵列的光场相机的基线长度为两端相机光心之间的距离。一般而言，基于相机阵列的光场相机比基于微透镜阵列的光场相机具有更长的基线，也就具有更大的FOP角度。

4、光场4D可视化：

4D光场数据可以表示为（Vx, Vy, Rx, Ry），其中（Vx, Vy）表征了角度分辨率，（Rx, Ry）表征视点图像分辨率，表示单个视点图像的分辨率为Rx\*Ry。如图中上侧图展示了7x7光场的可视化，表示共有49个视点图像，每个视点的图像分辨率为384x512。**视点图像分辨率越高，包含的细节越多。角度分辨率越高，单位角度内视点数量越多，视差过度就越平滑。角度分辨率越低，视点就越稀疏，观看光场时视点跳跃感越明显。**假设在10度的可视角度内水平方向上包含了7个视点，相邻视点间的角度为1.67度，在2米的观看距离，相邻两个视点的水平距离为29毫米。当视点个数减少，相邻视点的空间距离变大，观看者移动观看位置时就会感到明显的视点跳跃。



光场中任意两个视点间都存在视差，将光场（Vx, Vy, Rx, Ry）中的部分视点图像放大，如图11中下侧所示；同一行所有视点图像之间只有水平视差，没有垂直视差；同一列所有视点图像之间只有垂直视差，没有水平视差。



光场作为高维数据，不便于可视化分析。为了更好的分析光场中的视差，往往将光场中某一个水平/垂直视点上所有图像的同一行/列像素堆成一幅2D图像，称之为光场切片数据（Light Field Slice）。光场切片图像可以将光场中的水平视差和垂直视差可视化，便于直观分析。如图12中蓝色线条所在的行像素堆叠后就形成了图12中下侧的光场切片图像。类似的，如图11中将光场中同一垂直视点上所有绿色线条所在的列像素堆叠后就形成了图11中右侧的光场切片图像。

**参考文献：**

[1] E. N. Marieb and K. N. Hoehn, Human Anatomy & Physiology (Pearson, 2012).

[2] A. Gershun, “The light field,” Moscow, 1936, P. Moon and G. Timoshenko, translators, J. Math. Phys. XVIII, 51–151 (1939).

[3] Moon P, Spencer D E. The photic field[J]. Cambridge Ma Mit Press P, 1981, 1.

[4] Zhang C, Chen T. Light Field Sampling[J]. Synthesis Lectures on Image Video & Multimedia Processing, 2006(1):102.

[5] Javidi B, Okano F. Three-Dimensional Television, Video, and Display Technology[J]. Materials Today, 2003, 6(2):50.

[6] Ozaktas H M, Onural L. Three-Dimensional Television: Capture, Transmission, Display[J]. Thomas Telford, 2008, 2(1):487 - 488.

[7] E. Adelson and J. Bergen, “The plenoptic function and the elements of early vision,” in Computational Models of Visual Processing (MIT, 1991), pp. 3–20.

[8] Todd Gustavson, George Eastman House. Camera: A history of photography from daguerreotype to digital[M]. Sterling Innovation, 2012.

[9] M. Levoy and P. Hanrahan, Light field rendering[C]. Proceedings of ACM SIGGRAPH, 1996.