

# Magic Leap 工作原理 - 从视野到 GPU

2016 年 8 月 1 日 翻译: [曲晓峰](#) 原作: [Jono MacDougall](#)

原文链接: <http://gpuofthebrain.com/blog/2016/7/22/how-magic-leap-will-work>

Magic Leap 一直对他们的工作细节保密。我们现在仅仅知道他们的系统是一个全新的系统, 并且远超一切消费者熟知的现存竞争对手。这也难怪 Magic Leap 想要为自己的革命性的系统保密。众多的企业都在伺机窥探, 想要了解究竟是什么技术, 让人们如此的兴奋。这种技术, 就是 Apple 想要拥有的, 那种潜在的革命性的带有真正创新意味的“新事物”; 也是微软在 Hololens 上想要努力实现, 却又远未实现的; 又是受 Google Glass 启发, 却又明显领先几代的。这个技术到底是什么, 它究竟是如何工作的? 我调查了在 Magic Leap 工作的人所做的演讲、专利、工作申请和他们的背景, 试图寻找这个问题的答案。

## 总体来说



这是常规大小眼镜。但相机又往哪放呢? 注: 这可不是 Magic Leap 的产品。

在谈论细节之前, 我们先总体上谈谈这是个什么技术。简单说, Magic Leap 在做一设备, 这个设备能把物体投影到人的视场, 而且真实得远超我们现在所见的其它类似设备。Magic Leap 的设备由两部分组成: 一副眼镜和一个便携的口袋式投影/计算模块。这模块是差不多手机大小的没屏幕的长方体。这个便携式计算模块通过数据线连接眼镜。眼镜与普通眼镜有着类似的尺寸和结构, 也许稍胖一点点。体积小是该产品的一个重要特征。体积小意味着该产品可以在社交场合里佩戴使用, 也许有可能会像智能手机一样地便携易用, 随处可用。

## 便携式投影计算模块



如图所示, 一个与 Magic Leap 便携模块大小类似的充电宝。

Magic Leap 最吸引突出的是能把一大部分必须的硬件设备从眼镜本体上拆除, 放在另外一个独立的模块中。Hololens, 刚好相反, 其显著地削减了头戴设备中各个组件的尺寸, 但也只能做到现在的程度。那便携模块里面都有些什么呢? 可能有如下部分:

- 电池

这块电池的容量大约相当于现在的智能手机, 也许还得再多点。如果要替代智能手机, 那更得是相当牛的一块大电池。估摸着差不多至少得 5000 毫安时。

- CPU/GPU

肯定要用最新一代的移动 CPU。估计会用高通的。幸运的是, 他们应该用不到高端图形处理, 因为混合现实 (Mixed Reality MR) 只需要渲染局部, 不需要渲染

整个场景。这样就避免了虚拟现实（Virtual Reality VR）所需的高强度图形处理。

- 内存

跟智能手机差不多，估计 3 - 4 GB。

- 定制的 [SLAM](#) 芯片

这个是把虚拟物体摆放到真实世界所必需的。他们可能会自己流片，或者采用 [Movidius](#) 或其它类似的芯片。

- 4G/Wifi/蓝牙
- SIM 卡
- GPS 芯片
- 相机

眼镜上肯定要有一堆相机，但便携模块也不一定就没有。头戴设备上的 **SLAM** 相机跟普通数码相机是不同的。眼镜的体积有限，可能容不下一个高功耗的相机，只能放进便携模块。这样设计的好处是可以减轻他人对隐私侵犯的担忧，没有便携模块只有眼镜的话，没法拍照。

- 激光投影仪

这是该设备最主要的创新。把投影系统从眼镜上拿走，挪到便携模块上，使得产品的体积显著缩小。投影的光由便携模块生成，然后通过光纤传导到头戴设备。后面我们会详细分析一下其工作原理。

## 眼镜

当我们尽可能把所有东西都塞进便携模块之后，眼镜上剩下啥了？眼镜上得安装如下组件：

- [惯性测量单元](#)

就是常见的加速度传感器、陀螺仪和指南针。

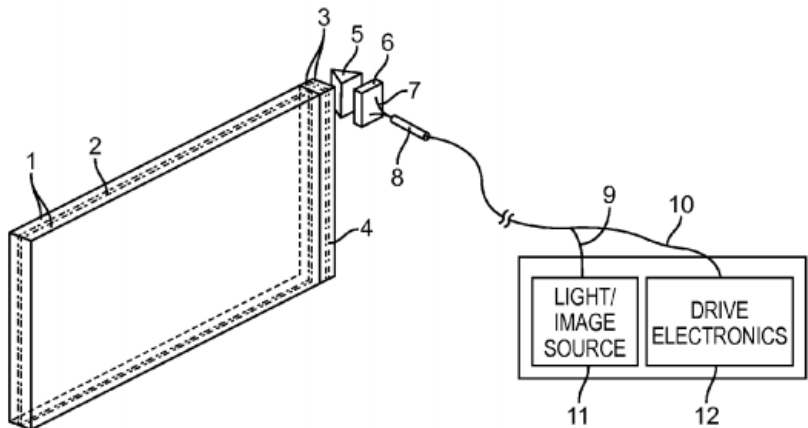
- 耳机

也许会用 Google Glass 上的那种骨传导耳机。这要看他们的设计思路是不是要与身体相协调。骨传导的优势是你既能听见耳机播放的音乐，同时听到其他的声音。

- 麦克风
- 光学部分
- 相机

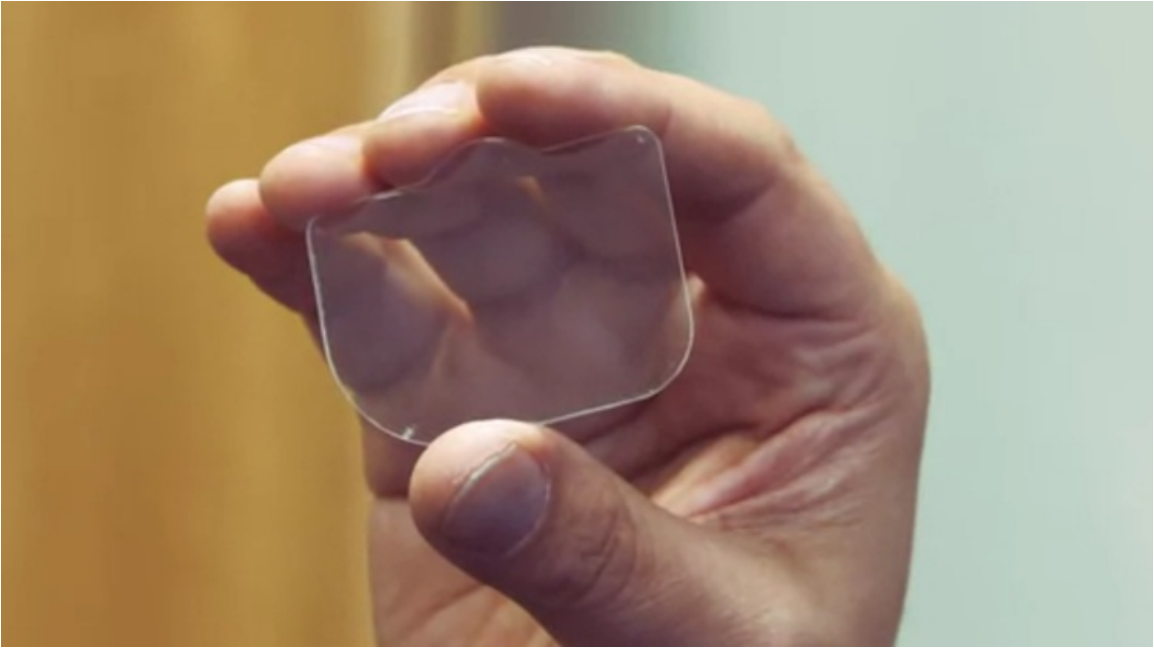
光学部分和相机是最有趣的组件，我们详细分析一下。

## 光学部分



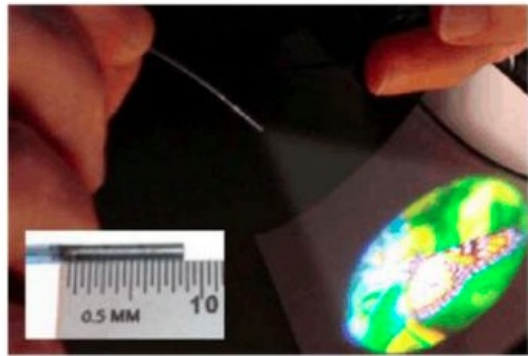
根据专利文献，Magic Leap 使用的光学设备比 Hololens 和 Google Glass 使用的传统投

影系统小很多。如上图所示，光源是与头戴设备的主体分离的，这就是为什么我们可以推测光源在便携模块中。其次，该设备的镜片系统也非常小。示意图所以虽然明显不是等比例绘制，但也一定表示相关元件的大致尺寸。我们唯一真正看到的元件是镜片。比较图中上侧的 5、6、7、8 元件和镜片的宽度，我们不难看出相对大小。

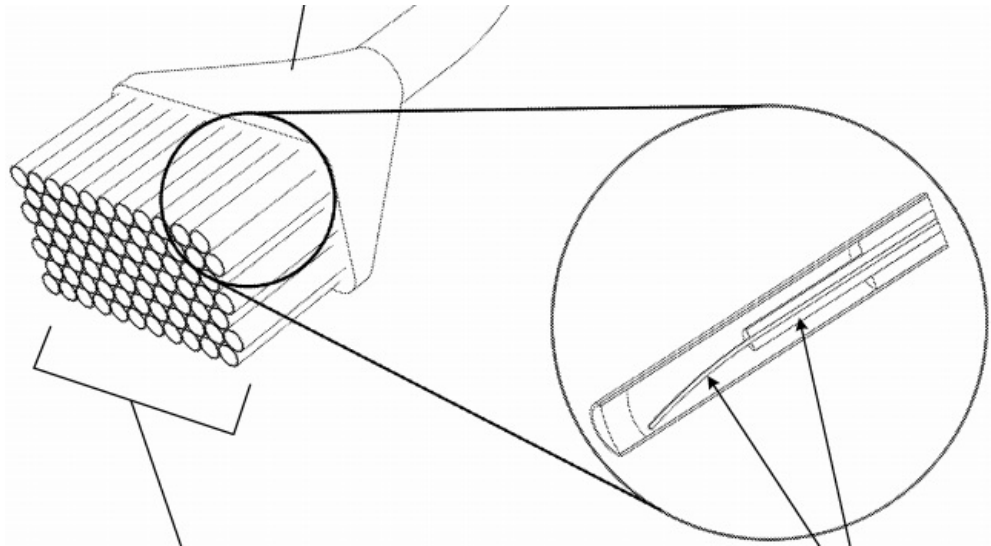


这说明什么？他们怎么能把光学部分缩小这么多，同时还号称能实现光场显示、高分辨率和惊人的视野？答案包括两部分：光纤扫描显示与光子光场芯片。

## 光纤扫描显示



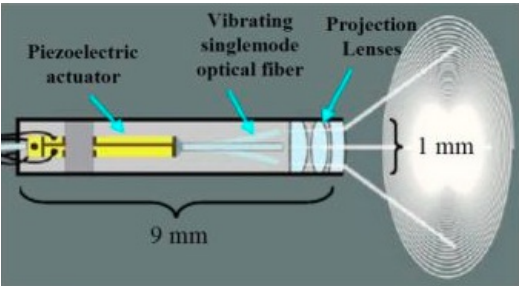
光纤扫描显示是从未在消费电子产品中使用过的全新的显示系统。我们只能通过其 [2013 年的专利申请文件](#) 进行了解。这份专利申请文件有些日子了，所以关于系统性能的一些细节可能已经不准确了，但基本概念应该还是一样的。该系统使用一组致动光纤来扫描输出远大于其数值孔径的图像。这就像老式的显像管电视。只不过不是扫描输出电子到荧幕来激发荧光粉发光，而是直接扫描输出光。扫描是通过压电致动器来实现的。扫描频率保持在大约几十 kHz。但实际图像刷新率并没有那么高。因为需要多次扫描（专利里面举例如 250 次）才生成一整幅图像。这完全改变了我们对于分辨率的概念。这个技术的图像分辨率取决于光纤的扫描频率，光纤可以汇聚的最小光斑尺寸（这决定了像素大小），生成一幅图像所需的扫描次数和刷新帧率。考虑到专利申请之后，他们对于该技术的进一步优化，其分辨率应该远超现有的消费电子产品。



一组光纤扫描单元紧密排列为一束来增加显示的尺寸。每根光纤扫描单元的宽度为1 mm。

除了分辨率和帧率，一个宽广的视场（Field of View FOV）同样也是显示逼真的全息图像的关键。关于这一块，专利的背景信息部分有一段很有趣的描述。

头戴显示器（Head Mounted Display HMD）的视场可以由微型显示器的图像尺寸和观察光路共同决定。人类的视觉系统的视场，水平大约 200°，垂直大约 130°。但大多数头戴显示器仅仅提供了 40° 左右的视场。..... 大约 50 - 60 弧秒的弧度分辨率代表着 20/20 的视力表分辨率（译者注：视力测试标准中正常人眼较高的成绩，参见 [WIKI](#)）。而弧度分辨率，是由微型显示器的像素密度决定的。为了匹配一般人的视觉系统，头戴显示器应在水平 40°、垂直 40° 的视场内提供 20/20 视力表分辨率。以 50 弧秒计算，相当于八百万像素（8 Mpx）。如果把视场拓展到水平 120°、垂直 80°，则需要五千万像素（50 Mpx）。



这里谈到了两件事。第一，消费级显示器的分辨率远远小于大视场所需。这就是为什么 Hololens 想要增大视场的话如此的艰难。第二，这显示出 Magic Leap 的野心。他们想提供一个水平 120°、垂直 80° 的视场。这个视场比 Oculus Rift 的视场还大，同时分辨率也远超。他们有没有实现呢？现在还很难说，但至少专利里面已经提到了一些技术参数，同时不要忘了这些还是三年前的数据。他们很有可能已经改进提高了这个技术。

像素间距是从一个像素中心到相邻像素中心的距离。它限制了图像的分辨率。传统的微型显示器，例如 Hololens 所用的，其像素间距在 4 - 5 微米。像素间距限制了这些显示器的分辨率，也因此限制了生成的视场。专利申请文件表明，扫描式光纤显示器能够生成 0.6 微米的像素间距，提升了一个数量级。

那究竟能达到什么分辨率？专利里有一段提到一个 4375 x 2300 分辨率，但我觉得还不止。这是在描述基本方法时举的一个例子，后面还讨论了多核光纤对于性能的提升。我认为其分辨率会远远高于该分辨率。这对于宽广的视场是至关重要的。

最后专利提到 120° 视场的这句特别值得注意：

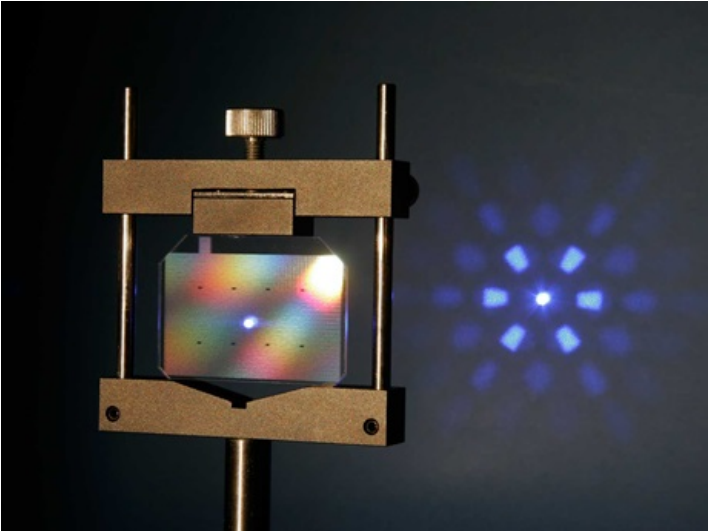
以上所述技术可以用来制造具有宽广视野的头戴式或其他近眼显示方式的超高分辨率显示器。

我认为这充分验证了，其视场将至少大于 40°，接近其宣称的 120° 也并不是不可思议的。要我下注的话，我赌 90°。

## 光子光场芯片

第一次听说 Rony Abovitz 给他的镜片起名叫“光子光场芯片”，我都郁闷了。别总给这些早就有的东西起一些不着调的名字了。就叫它 Rony 镜片就好了。但随着我逐渐加深理

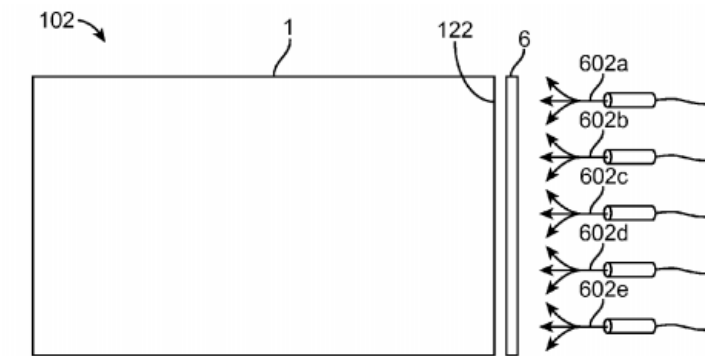
解，它还真不是一个简单的镜片。它究竟是干什么的，为什么它比其他镜片更有意思？我们先来了解一下衍射光学元件。



一个衍射光学元件的例子

可以把衍射光学元件（Diffractive Optical Elements DOEs）想成是一组非常细的镜片。他们可以用来整形，分光，匀化、扩散。**Magic Leap** 使用带有圆形镜片的线性衍射光栅来分光，并生成特定焦距的光束。就是说，它把光导入你的眼中，并让这些光就好像是从一个正确的焦平面发射出来的一样。一般人说，说起来容易做起来难，这个说起来都难。至少我找到的[专利文件](#)是这么写的。

为了生成光场，**Magic Leap** 使用两个分立的元件配置了一个光子芯片。一个元件（原理图中的 6）提取投影光，并将其插入第二个元件（原理图中的 1），第二个元件将光导入人眼。



这两个元件都是使用 DOEs 来完成工作。DOEs 的主要问题在于，他们经过精心调试只能用来进行一个特定工作。他们不能工作于不同的波长，不能实时改变到不同的焦点。为了解决这个问题，**Magic Leap** 堆叠了一组针对不同波长和焦平面优化的 DOEs，作为一个大镜片组来用。这些 DOEs 都非常的薄，跟光的波长在一个尺度，所以加在一起也不会让设备变得太厚。这就是为什么这个光学系统被称为芯片。**Magic Leap** 能够开关 DOEs 的不同层。这样，他们可以改变光到达人眼所使用的路径。这就是他们如何改变图像的焦点，来形成一个真实的光场。就像专利里面说的：

例如，当打开一组 DOEs 里面的第一个的时候，对于一个从正面看进来的观察者，可以产生一个光学观察距离 1 米的像。一组里面的第二个 DOE 打开时，可以生成一个光学观察距离 1.25 米的像。

你可能觉得这个技术很局限，尤其是当你需要一大堆层来产生各种不同焦点的像的时候，但还真不是那么回事。不同的 DOEs 的组合可以产生不同的输出。并不是一个 DOE 对应一个焦平面，而是一个 DOEs 组合对应一个焦平面。

改变 DOEs 组中的激活层，就会改变从光子光场芯片射出的光。他们很有可能比图中多得多的层，但具体多少层就天知道了。

最终，我们明白了 **Magic Leap** 怎么实现其过去宣称的用光来制造暗了。我们分别使用一个外侧的 DOE 和一个内侧的 DOE 就可以像主动降噪耳机一样抵消外界的光。专利里面是这样说的：



这可以用来抵消类似于背景光或真实世界光的平面光波导，在某种程度上类似于主动降噪耳机。

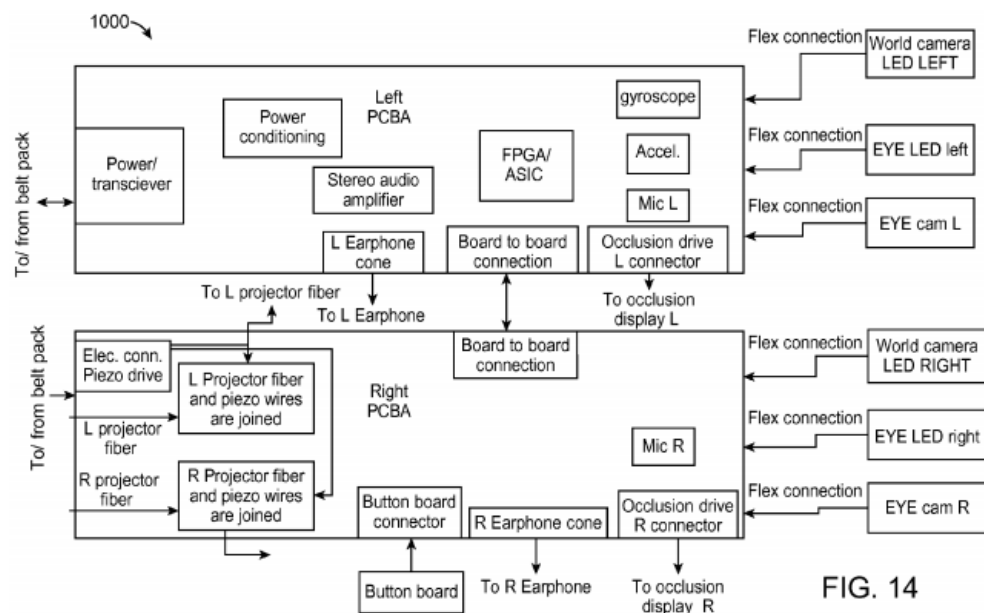
那，为什么这是一个芯片？呃，一个典型的电子芯片基于某种条件改变电子的流向。**Magic Leap** 的光子光场芯片基于某些参数改变光子的通路。我觉得也算是一种芯片了。

我们还缺什么？我们有了光子光场芯片，有了高分辨率投影，但如何构造一幅图像。这个是靠组合。图像是分层绘制的，以便让不同的部件投影到不同的焦距的子图像上。就是说，每一帧都是通过多次扫描来构建的，每一个焦平面都是分别绘制的。

## 相机

**Magic Leap** 试图在相机上实现三个功能。第一个，最明显的，一个生成日常图像的相机。这是他们所使用的最容易理解的相机技术，他们也许就是用一个智能手机市场上最新的类似传感器。这个相机是放在眼镜上还是便携模块上，仍不可知，但总要有个能拍照片的。

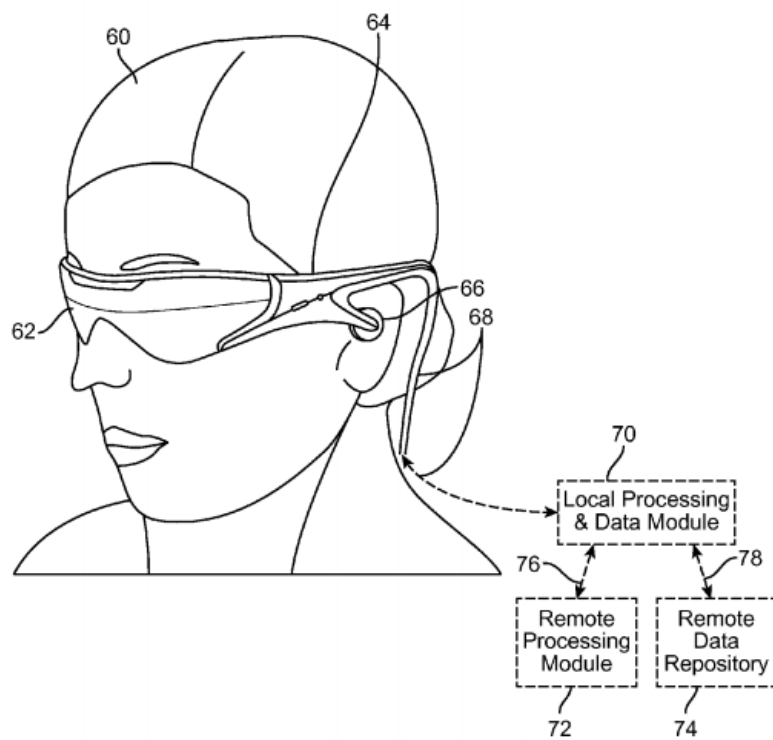
其他两个功能很有意思。**Magic leap** 反复提到其设备具有理解周围世界的能力。在[一个采访](#)里面，提到该设备能识别物体，例如刀和叉子。要想做到这个，他们需要一组相机。我们可以看看这方面做的不错的 **Hololens**。**Hololens** 有一组四个环境感知相机和一个深度相机。从 **Magic Leap** 的专利文档中我们能得到进一步确认。



原理图上显示的是眼镜左右镜腿上的两个元件。上面是左边镜腿的，下面是右边镜腿的。

由以上原理图，我们可以发现两个向外的相机，叫做“world camera”。专利的文字描述暗示可能不止两台相机，原文所述为“一个或多个面向外侧或提供世界视角的相机（每侧）”。暂时，还不知道具体会有多少个相机，也不知道 **Magic Leap** 会把这些元件做得有多小。但我们知道的是，这些要放在眼镜上，而且对于 **SLAM** 非常重要。

最后一个相机的功能也可以从上面的原理图中找到。至少要有两个相机拍摄眼睛。这是用来追踪视线和眼动以便获取焦点和视线方向。同时也会有红外 **LED** 为这两个相机提供照明。眼跟踪对于用户交互很重要。我想“你在看什么”个问题对于你如何与 **Magic Leap** 交互应该是非常重要的。这将是其主要的交互工具，就像是电脑的鼠标。



当然，现在，我没法验证这些信息是真是假，但这些加总在一起，确实感觉上就像 **Magic Leap** 正在做的东西。不管最后这个产品会不会在消费市场上获得成功，这确实是好一阵子没有见到的一个技术产业的真正意义上的创新。我真的无比激动想要看看他们会有什么奇遇，同时也期待该产品在业界会造成什么样的影响。