目前，AR光学所广泛采用的三个技术路径分别有：传统同轴光路、波导和自由曲面。尽管著名的 Magic Leap 公司提出他们实现了“光场显示”，但其基本技术途径仍然是波导显示，因此还是难以避免波导显示固有的弊端（关于“光场显示”，我们会在以后的文章里详细说明）。由于其在成像效果、成本和量产性等方面的优势，自由曲面AR光学已经被国内外研发AR眼镜的企业，如联想、悉见、爱普生、亮风台、Meta、ODG、骁龙等公司所采用，并成为目前能量产销售的AR产品的核心部件。可见，自由曲面AR光学对AR行业的发展起到了至关重要的作用。

其实，自由曲面是目前光学最前沿的设计理论和方法，不仅在AR方面，在其他光学设计方面也发挥了巨大的作用。

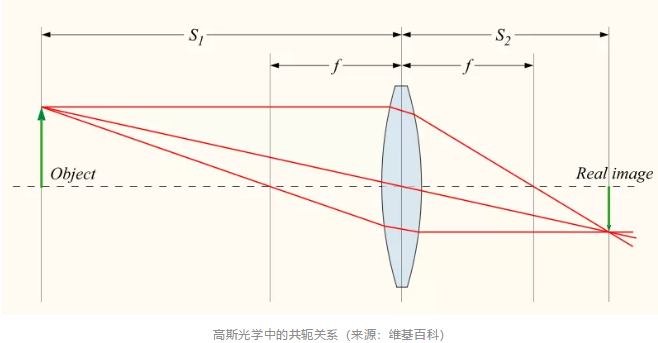
**第一章 自由曲面的发展**

- 1 -

球面光学系统

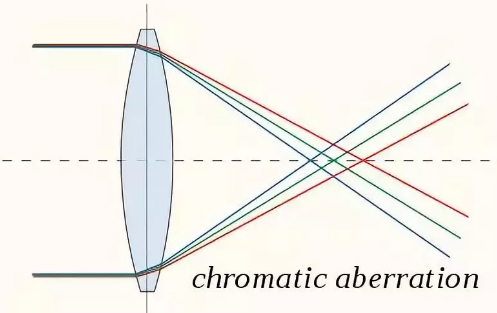
球面是最被广泛采用的光学表面，在球面光学系统中，光学表面可以看作从某个固定半径的球面上截取的一部分，表面上各点处曲率相等。

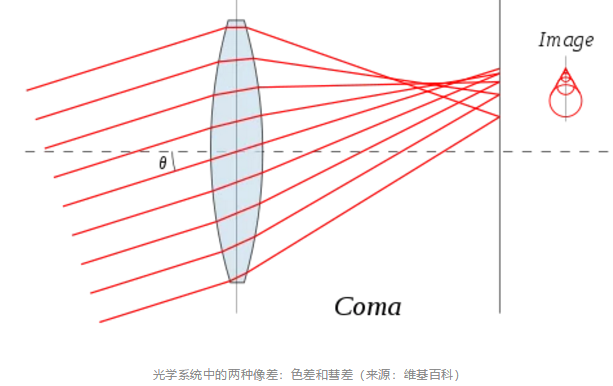
球面光学系统被广泛应用的原因之一是其有着最简单明确的几何构造。只需用到三角函数的有关知识，经过几何推导和傍轴近似，我们便不难推得理想光学系统的一般理论，使得对于点、线、面而言，光学系统的物像之间都有明确的共轭关系。这项工作由高斯在1841年完成，因此理想光学系统理论又被称作“高斯光学”。



球面被广泛采用的另一个原因是它的易加工性。传统的光学元件都采用不同折射率和阿贝数的玻璃作为材料，其加工工序通常包括切割、研磨、抛光以及之后的镀膜。其中研磨和抛光都需要特定曲率的磨具与光学表面匹配，过程中磨具和光学元件绕不同的转轴旋转并保持贴合。显然，只有球面才能适应这样的加工方式。

然而球面光学系统的缺陷也显而易见。由于理想光学系统理论是建立在傍轴近似的基础上的，即不同视场和孔径的光线的入射角都应该足够小，否则，光线就会偏离理想情况而形成几何像差。同时，不同波长的光受到不同程度的折射，又会产生色差。这些像差无法用单个透镜来消除，因此，球面光学系统往往需要较多数量和不同材质的镜片来联合校正这些像差。

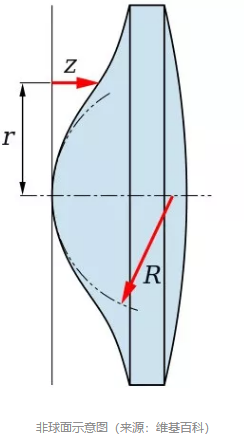




- 2 -

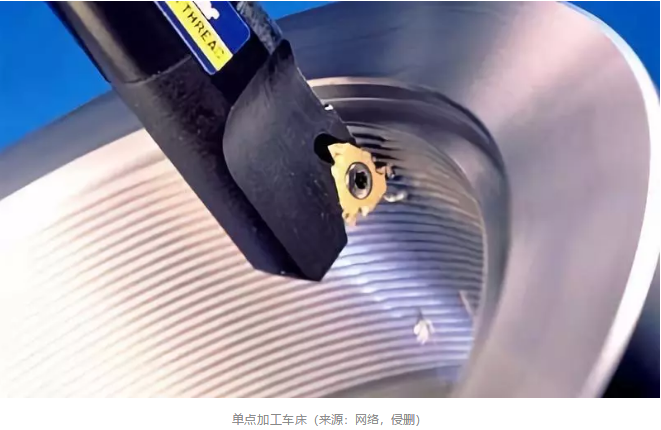
非球面光学系统

广义的非球面指不能用球面定义描述的面型，即各处曲率不一致。我们通常所说的非球面主要指用含有非球面系数的高次多项式来表示的面型，其中心到边缘的曲率连续发生变化，且一般情况下为旋转对称结构。常见的抛物面、双曲面和椭球面都属于非球面的范畴。



随着计算机技术的发展，传统的查对数表和手工计算被更高效的计算机辅助设计所取代，大计算量和高迭代次数的光学优化方式成为可能。相比于传统的球面光学系统，非球面光学系统在面型上具有更高的自由度，因此可以在一定程度上校正更多高级像差，从而减少球面的数量，优化系统结构。比如，我们可以在光阑附近使用非球面校正高级球差等与口径有关的像差，在远离光阑的位置使用非球面校正像散和畸变等与视场有关的像差。世界上第一款低成本可量产的非球面摄影镜头是柯达在1982年推出的 12.5mm f/2.8 镜头，它采用模压玻璃制作。如今，非球面表面已广泛用于各种相机镜头和投影镜头中。

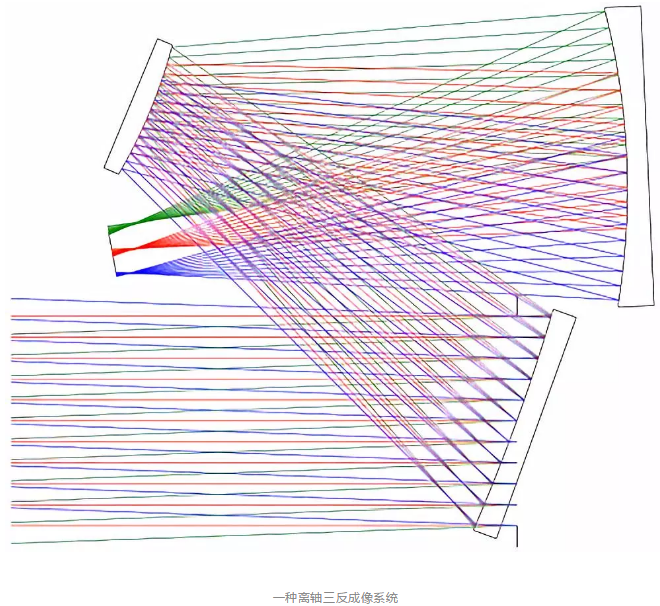
非球面光学系统的加工与球面光学系统有着明显的不同，加工球面所使用的工艺并不能完全适用于非球面光学系统。但由于非球面一般都具有旋转对称的几何结构，因此随着单点金刚石车削工艺的成熟，非球面表面也可以通过车削的方式进行批量生产。



- 3 -

自由曲面光学

从20世纪50年代起，随着国防和民用光电技术的不断发展，各种复杂的光学和机械结构开始出现，这对现代成像系统的性能、像质、体积和重量等指标都提出了更高的要求。如航空航天领域所采用的离轴折反式光路中，系统的非对称性带来了更多的非对称和更高级像差，光学系统需要向超薄、超简的方向发展，这些仅靠传统的球面和对称非球面已难以满足。此时，非球面系统的更高级阶段——自由曲面开始登上历史舞台。



我们通常所说的自由曲面一般指没有旋转对称轴的复杂非常规连续曲面，或者说，可以是任何形状的表面。显然，不论是在初始结构计算、面型描述、系统优化还是制作加工上，自由曲面都面临着相当大的难以把握性。

3.1 自由曲面的初始结构获得

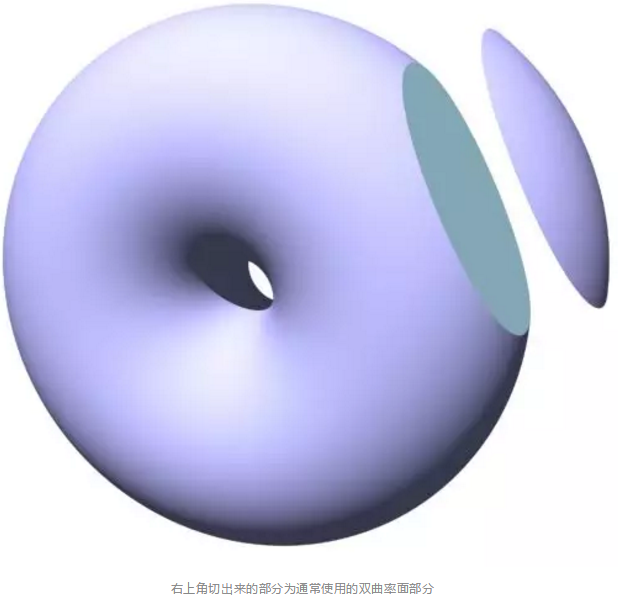
自由曲面初始结构的获得可以从球面或非球面逐步逼近而来，或通过光线追迹得到点云再进行曲面拟合。对于前者，通常先设计低阶曲面，以满足系统的光学特性参数和结构要求，如焦距和数值孔径。在之后的优化过程中，再逐步加入其他参数，使其向着更复杂的面型发展，以满足像质要求。对于后者而言，点云通常基于费马原理直接获得，即对于共轭的一对物像点，沿不同路径传播的光线具有相同的光程。得到点云之后，再选择合适的数学方式对其进行拟合和描述。

3.2 自由曲面的面型描述方法

自由曲面的的面型描述方法非常丰富，其遵循的原则是灵活多变、能描述多种复杂面型、像差校正能力强、光线追迹和优化收敛速度快。按照面型控制方式，这些描述方法大致可以分为两类：全局控制曲面和局部控制曲面。对于前者而言，每个参数都会对表面的全局形状产生影响，因此我们只改变其中一个参数，曲面各处的矢高和斜率都会发生改变。此类描述方法包括各种多项式定义的方法。而对于后者而言，每个参数对曲面面型变化的作用范围有限，因此可以通过改变某个或某几个参数来调整局部面型。此类描述方法包括三次样条曲面、非均匀有理样条曲面以及高斯基函数组合曲面等等。

这里我们例举一些常用的自由曲面描述方法：

双曲率面——又叫镯面或马鞍面，由在 x-z 平面（或 y-z 平面，z 轴与光轴方向平行，下同）内的一段曲线（圆弧或高阶曲线）绕与 x 轴（或 y 轴）平行的直线旋转一定角度形成，其特点是具有旋转对称轴，但该对称轴并不与光轴平行。双曲率面的 x 和 y 方向通常具有不同的曲率。特殊情况下，当这两个曲率相同时，双曲率面就退化为球面，而当其中一个方向的曲率为0时，双曲率面则退化为柱面。



复曲面——表面在正交的两个方向上分别具有独立的曲率和各阶系数，其特点是具有互相垂直的两个对称面，即 x-z 平面和 y-z 平面。与双曲率面不同的是，复曲面不一定具有旋转对称轴。

XY 多项式曲面——在非球面的基础上增加了各阶单项式得到的曲面，打破了非球面原有的旋转对称性。可以看作是对非球面的更高级修正。进一步地，我们还可以设计复曲面基底 XY 多项式曲面，即以复曲面为基底，增加各阶单项式，从而结合复曲面和 XY 多项式曲面各自的优势，也为光学设计提供更多的自由度。

梯形畸变校正曲面——由美国 ORA 公司（现属 Sysnopsys 公司）的 J. Rogers 提出的一种自由曲面。它在形式上与传统对称非球面类似，但在代入非球面公式之前，首先对 x 和 y 坐标做不同程度的变换（倾斜和缩放），从而校正由带有光焦度的离轴反射镜所带来的梯形畸变。由于坐标变换的存在，这类表面也不再具有传统非球面的旋转对称性。

Forbes 曲面——由美国 QED 公司的 G. Forbes 提出的一种正交曲面。它通过经过优选的标准雅可比多项式正交基函数系的方式来定义偏离球面的非球面系数项，使各项系数都有十分明确的物理含义，并且具有唯一性。

Zernike 多项式曲面——由诺奖得主F. Zernike 提出的一种曲面，它由一系列在圆域内正交的基函数组成。这意味着定义在该圆域内的函数如果用 Zernike 多项式来拟合，无论使用的项数有多少，其各项系数始终保持不变。此外，它还容易与经典的塞德尔像差建立联系，也是它得到普遍应用的主要原因。

高斯基函数复合曲面——由美国中佛罗里达大学的 O. Cakmakci 等提出的一种局部面型可控的自由曲面，它可以是在二次曲面的基础上叠加一组线性拓扑形状分布的高斯曲面，也可以抛离基底项而直接由一系列高斯函数组合而成。该方法对于像差的控制力更强，与 Zernike 圆域正交的描述方式相比，对矩形或其它形状的非球面描述能力更强，很容易实现面型的局部控制。

非均匀有理 B 样条曲面（NURBS 曲面）——一种非常优秀的曲面描述方法，广泛应用于现有三维 CAD 软件中。该方法使用一系列带有权重的顶点来控制面型，各顶点呈拓扑矩形排列。这是一种典型的局部控制曲面，即每个顶点仅影响周围局部区域的面型，因此 NURBS 曲面可以表示出非常复杂的面型。1991 年国际标准化组织（ISO）颁布的关于工业产品数据交换的 STEP 国际标准，把 NURBS 作为定义工业产品形状的唯一数学方法。

分段环形面和拼接非球面——顾名思义，是由各段曲面拼接而成。对于分段环形面而言，每段曲面由一个三次多项式定义，曲面整体为旋转对称。而拼接非球面以非球面和环形面为基础。显然，各段表面在接线处需保证边界点相接且一阶导数连续，才能使曲面整体光滑。

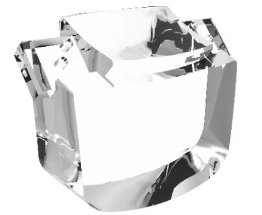
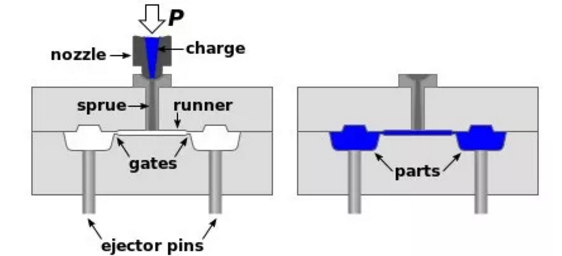
3.3 自由曲面的光学设计优化

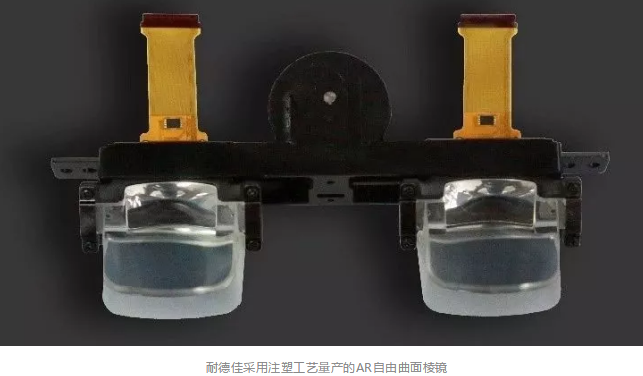
优化是光学设计的重要步骤之一，通过优化我们可以提升系统的成像质量以及控制结构参数。对于几何光学设计，优化过程基于光线追迹。我们通过对不同视场和孔径位置的光线进行采样，逐表面追迹其路径，并分析光线在各表面上的位置，从而计算系统的像差并对其进行控制。同时，在优化的过程中，系统的各优化变量需要收敛于某一组值，使得系统的像差向着局部最小的方向发展。这个过程需要根据光线追迹的结果对面型进行迭代。每一次迭代都需要计算各变量对像质或约束条件的微分，并重新计算优化评价函数。因此，追迹的光线数量越多，描述面型的参数越复杂，迭代所需的计算量越大，系统收敛的速度越慢。而自由曲面由于面型自由度高，需要更加密集地对不同视场和孔径位置的光线进行采样，以防止表面在小范围内产生剧烈变形。所以，自由曲面光学系统的优化要远远难于普通球面或非球面系统的优化。尽管计算机性能日益强大，这些工作都可以由计算机自动完成，但一个较为复杂的自由曲面光学系统往往也需要几天的时间才能优化完成。

由于自由曲面的优化过程中需要对视场进行密集采样，因此像面整体成像质量也更加难以控制。如果采用手动方法平衡系统的像质则会极其复杂耗时，并且很大程度上依赖设计者的经验知识。此时，基于像面整体成像质量的自动平衡算法可以有效减小这部分的工作量。这种方法的主要思想是在迭代过程中采用特定算法对各视场的评价函数分配不同的权重。这种方法使系统在全视场范围内达到均衡的成像质量，甚至能够提高系统整体的成像性能。

3.4 自由曲面光学元件的加工

目前尽管单点金刚石切削技术已经能够实现各种复杂表面的加工，但如果对每个光学表面都采用这种方式进行加工的话，势必付出相当大的成本。目前来看模压玻璃的压注工艺尚不成熟，加工效果不稳定，因此注塑加工是最适合自由曲面光学元件的量产方式。其过程主要是通过金刚石车床加工出自由曲面模芯，闭合后注入树脂材料。这个过程涉及一系列材料本身的特性，如材料的热胀冷缩可能会形成内应力，需要我们不断调整注塑各段阶段的时间和速度。此外，我们还需要对光学表面进行镀膜，以使得系统的反射率和透射率能够符合我们的预期要求。





- 4 -

本章小结

可以看出，光学元件从球面到非球面再到自由曲面的进化过程，体现了光学设计和光学加工越来越高的自由度和复杂性。作为近年来 CAD 技术和生产制造技术蓬勃发展的典型时代产物，自由曲面通过灵活多变的形状和简单易量产的结构实现了各种传统元件无法实现的复杂光学功能。自由曲面光学元件的设计和加工能力，直接体现了一家光学企业的研发实力。尤其是在 VR 和 AR 技术正在崛起的今天，耐德佳和爱普生等公司用自由曲面光学元件配合新型显示器件，为行业提供了多套无与伦比的AR显示模组，对传统的球面或非球面系统实现了降维打击。我们可以预见，现代的光学系统将越来越离不开光学自由曲面，而自由曲面也将在 VR 和 AR 以外的其他应用中大放异彩。