

# 虚拟现实中的基于图象的绘制技术\*

徐丹 潘志庚 石教英

(浙江大学 CAD&CG 国家重点实验室, 杭州 310027)

**摘要** 对目前出现的各种基于图象的绘制方法进行分类介绍和综述,同时结合作者的研究工作,指出了需要进一步研究的内容。

**关键词** 基于图象的绘制(IBR),实时图形生成,虚拟现实

## 1 引言

随着虚拟现实应用的发展,人们要求实时地生成高质量的画面。例如,一个虚拟环境的完全漫游需要快速模拟6个自由度的镜头运动,包括相机旋转、对象旋转和相机移动。如果用传统的图形学方法,首先要建立场景的三维几何模型,然后对场景物体表面的材料、光照、纹理等进行描述。这样的处理过程非常耗时,往往需要很昂贵的硬件开销,并且还要在场景的复杂程度和图形的真实感两者之间进行折中。

为了解决这一技术问题,人们提出了基于图象的绘制方法(image-based rendering,简称IBR)。IBR是指通过一组预先获得的图象,而不是用三维模型来编码一个场景,并且通过适当组合这些图象来产生场景的新视图,以完成绘制工作。

从已知的图象中合成新视图的IBR方法引起了计算机图形学和计算机视觉界中有关研究人员的普遍兴趣,并且在全世界范围内形成了IBR的研究热潮。与传统的基于模型的绘制方法相比,IBR方法有以下特点:

(1)计算量适中。IBR方法所需的计算量相对较小,因此适合于工作站以及个人计算机上的虚拟现实应用。

(2)独立于场景复杂度。交互显示的开销与场景

的复杂度无关,只与分辨率有关。因此IBR能用于表现非常复杂的场景。

(3)满意的绘制质量。作为已知的源图象不仅可来自于绘制系统,也可从真实环境中捕获,因此可以反映更加丰富的明暗、颜色、纹理等信息,并且计算开销都是一样的。

可以想象,一旦IBR技术走向实用,将会带来计算机图形学的一次飞跃。然而,IBR技术还面临着很多问题,主要有:

(1)表示模式:怎样找到一种简便有效且适合计算机的表示模式,能精确完整地对场景进行编码。

(2)捕获方法:是用手持相机还是数控摄像机、图象采样的数量多少、采样模式及样本均匀性等都会影响问题求解的难度和精度。

(3)走样与空洞:如何解决因采样引起的走样和空洞问题。

(4)信息压缩:基于图象的方法不可避免地面临着大量图象的压缩问题,怎样利用数据间的连贯性,找到合理有效的压缩机制。

(5)完全漫游:如何实现基于表示模式的完全实时漫游。包括模拟相机旋转、对象旋转、相机移动及缩放等相机运动方式。

目前,世界上很多国家的学者在研究和探索这一新方法,并且取得了一批有意义的研究结果。本文结合我们在IBR方面的研究工作,对这一技术进行

综述性介绍。

## 2 IBR 方法概述

IBR 技术的最早尝试源于分支电影<sup>[1~3]</sup>,但作为一种场景表示和绘制的新方法还是近几年的事。尽管如此,已有很多文章和成果发表,提出并讨论了多种实现方法。总结这些方法,我们粗略地把 IBR 方法按表示模式不同归纳为以下 5 类。

### 2.1 基于立体视觉的视图合成方法

基于立体视觉的视图合成方法是指利用立体视觉技术从已知的参考图象中合成相对于新视点的视图。其关键问题是找出每对已知图象之间的对应关系 (correspondence map),即解决立体匹配问题。由对应关系可导出偏差映射,并能进一步估计出场景中可见点的深度信息。利用深度信息可以对已有的图象进行“变形”来合成新视点的图象。也就是说,通过对应关系建立了一个基于图象的场景几何表示。如图 1 所示。这种方法将场景视图及其对应关系组成一个图结构,图中顶点表示不同物理位置的场景视图,边表示邻接视图间的对应关系。

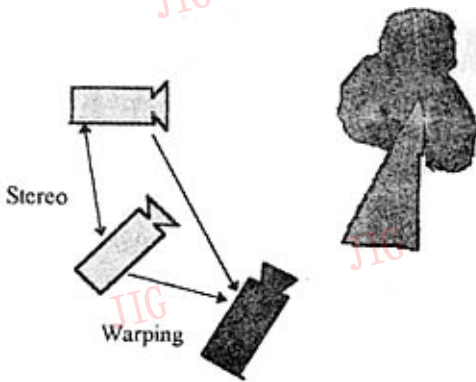


图 1 基于立体和变形的视图合成

Leveau 和 Fangeras 描述了一种从弱校正图象中构造新图象的方法<sup>[4]</sup>,但该方法要靠人工选择 4 对对应点。McMillan 和 Bishop 的 PlenOptic Modeling 系统<sup>[5]</sup>是一个基于柱面全景图象的系统。利用偏振约束 (epipolar constraint) 的柱面偏差计算出相邻两个柱面全景图的偏差映射,然后依据该偏差变形已知的全景图象,从而得到新视点的柱面全景图。另外,也可以采用多基准线的立体算法<sup>[6,7]</sup>。在这种方

法中,真实图象先被映射到已多边形网格化了的深度图上,然后新视图可用标准图形学中的纹理映射方法生成。

还有一类基于立体视觉的方法称为模型与图象混合的方法<sup>[8,9]</sup>。这些方法依据从很多图象对中收集的立体信息恢复出 3D 场景结构。Debevec 等介绍了一种用少数照片构造和绘制真实建筑场景的混合方法<sup>[10]</sup>,利用所建立的粗略模型可以计算出真实场景与模型之间的偏差,而且可以将图形重新投影从而使立体匹配能处理相距较远的图象对。

基于立体视觉的图象合成方法主要有以下优点:(1) 新视图可以由两幅临近的参考图及它们的对应关系合成,整体的几何模型不是必需的;(2) 图象变形比图象绘制快得多,而且变形时间独立于场景复杂度;(3) 只需知道邻接图之间的相对轮廓信息,而不需要整体的一致性,例如图象可由手持相机捕获,并标上大致的全局坐标。但同时,这种方法也存在着立体视觉中固有的缺陷:(1) 由于场景有可能部分和全部地被遮挡,只能掌握场景有限的信息,导致在参考图中不可见而在新图中应该可见的区域出现空洞,如何填补这些空洞是一个难以解决的问题;(2) 由于只产生有限的深度分辨率(深度不连续),使得匹配处理出现误差。

### 2.2 基于视图插值的方法

上述视图合成方法能由参考图生成任意视点的新视图。与其不同,基于视图插值 (view interpolation, 也叫图象插值) 的方法要求新视点位于两参考图视点所决定的直线 (称基线) 上,于是新视图可由参考图线性插值产生。当然,如果有很多幅参考图,也可通过一系列的插值获得任意视点的图象。与视图合成方法相同,视图插值的方法同样要建立来自立体的对应关系,因而仍然面临由于遮挡和深度不连续所带来的问题。另外,许多视图插值方法在一般情况下不能精确地重构,即不能产生正确的透视投影,而只产生近似的中间视图。

Chen 和 Williams 所介绍的图象插值方法着重于对图形学中绘制速度的改进<sup>[11]</sup>,因此假设点的深度能够从 3D 场景模型中得到。另外,他们还讨论了一种简单的空洞填补方法,但未给出对深度不连续的处理办法。

Seitz 和 Dyer 基于仿射投影模型指出使用 monotonicity 规则可保证图象插值产生正确的合成视图<sup>[12]</sup>。他们的方法包含三个步骤:图象修正、线性

偏差插值和除修正。这种插值算法将亮度均匀的色块作为一个整体来进行匹配和移动。另外,他们还将该方法扩展到透视投影及对两幅以上的参考图象进行系列插值操作的情况<sup>[13]</sup>,并又提出了一种组合图象插值和图象变形的视图变形方法<sup>[14]</sup>。除此以外,还有一些其他基于视图插值的方法<sup>[15,16]</sup>。

### 2.3 基于图象拼合的方法

将同一场景的多张有重叠的图象组合成一幅较大图象的处理叫做拼合(mosaic)。图象拼合技术典型地被用于全景图的生成、图象稳定、增大图象分辨率、图象压缩及视频扩展等方面。在组合源图象之前,一项重要的任务是图象整合(image registration),它是把源图象中相互重叠的部分对齐所做的变换。关于图象整合技术的详细讨论可参阅<sup>[17]</sup>。

一般地,整合两幅图象的难度取决于需要估计的参数数目。两幅图象有相同视点的情况较容易处理,因为依赖于同一透视变换,因此需估计的参数最多只有8个。很多文章中都成功地实现了基于一个视点的全景图象拼合算法<sup>[8,18,19]</sup>,这也是PlenOptic Modeling和QuickTime VR<sup>[20]</sup>的基本思想。然而对于两幅图象有不同视点的情形就要难得多,因为除了估计决定相对相机轮廓的8个参数外,还必须估

计每个象素的深度。为解决这个问题,要么施加额外的约束,要么利用更多源图象来增加输入变量,以求获得独立于视点的场景深度表示<sup>[21,22]</sup>。一旦参数被估计出来,新的视图就能被合成,而且还可利用多个视图构造三维校正的拼合图象。

事实上,图象拼合技术的真正意义并不在于把它作为一种独立的IBR方法,而是在于它往往是其他IBR方法中所要碰到的一个技术问题。

### 2.4 基于分层表示的方法

还有一种方法与图象拼合正好相反,叫做分层表示(layered representation)。它针对一视频系列将3D场景分成运动独立的、由仿射运动模型描述的不同层次<sup>[22,23]</sup>。每一层都产生一2D图象流和一2D变换流(统称脚本—sprite),并最终组合到显示屏幕上(见图2)。在SIGGRAPH'97上,Lengyel和Snyder介绍了一种将分层技术用于实时计算机图形学中的新方法——利用连贯层次的绘制(Rendering With Coherent layers)<sup>[24]</sup>。他们的分解充分利用了帧与帧之间的连贯性,不仅针对动态的几何对象,而且能分解光照模型中不同的光照效果(如阴影、高光、反射等等)。每一层都可单独控制其刷新频率、空间分辨率及绘制质量参数等,并根据它可用的资源权衡其逼真度。

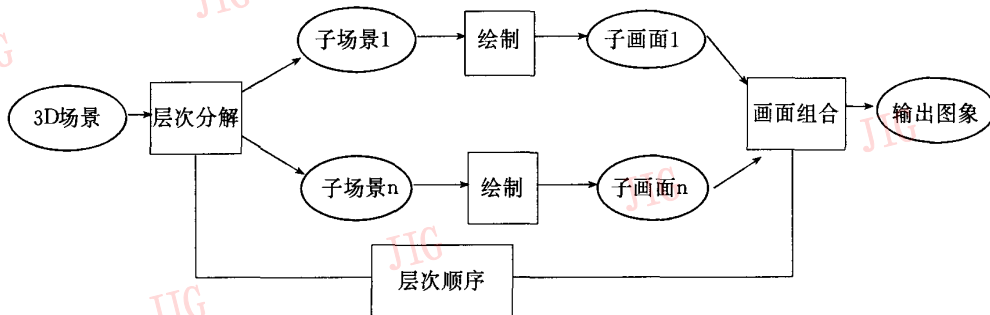


图2 基于分层表示的绘制流程

很明显分层表示对于实时计算机图形绘制非常重要,表现在:(1)分层绘制不仅将快速移动的前景对象与变化较慢的背景层次区分开来,而且还可将不同的光照效果区分开来,从而更好地利用帧间连贯性;(2)分层绘制使绘制资源的分配更优化,使资源集中分配给比较重要的层次;(3)分层绘制能自然地某些二维元素,如覆盖视频、离线绘制脚本或手工动画元素集成到三维场景中。

### 2.5 基于全视函数的方法

全视函数(PlenOptic Function)是由Adelson和Bergen命名的<sup>[25]</sup>。简单地讲,它描述了从空间任意点所看到的全部信息。用计算机图形学术语,它描述了给定场景中所有可能的环境映照(environment map)集合。基于全视函数的IBR方法试图捕获空间任意区域内的完全光流。这种方法从一些有向的离散样本中重构连续的全视函数,然后通过在新的视

点位置重新取样该函数来绘制新的视图。如果考虑视点处  $360^\circ$  的球形环境映照,该函数可描述成以下等式:

$$P(V) = \sum_{k=0}^{M-1} P_{V_k} B_k(V)$$

其中  $V$  是视点,  $P_{V_k}$  是函数在视点  $V_k$  的取样,  $B_k$  是相应于视点的基函数。但是球形环境映照不适合计算机表示。

PlenOptic Modeling 是第一个基于全视函数的系统<sup>[5]</sup>,作者将该函数描述成有三维空间位置 ( $V_x, V_y, V_z$ ) 和二维方向 ( $\theta, \varphi$ ) 决定的参数函数  $P(\theta, \varphi, V_x, V_y, V_z)$ 。这是一个柱面全景图的集合。这种表示在所有五维上都存在大量的冗余,并且如前所述方法中还要解决立体对应问题。

SIGGRAPH'96 上提出了两种非常近似的方法:光场绘制<sup>[26]</sup>和 Lumigraph<sup>[27]</sup>。基于自由空间中沿一条光线传递的光辐射不变的假设,它们把 5D 全视函数简化为描述离开或进入一封闭自由空间(如空立方体)的完全光流分布(如图 3 所示)。这是一个 4D 函数  $L(u, v, s, t)$ 。这两种方法的最大优点在于:(1)可以不需要立体对应关系,从而避免了立体视觉所面临的问题;(2)由于只考虑视流信息,因此不必对反射属性作假设;(3)表示模式简单有效,既便于控制和计算,又能够均匀取样。这两种方法的缺点是它们只解决了没有遮挡的、光照固定的、静态对象的表示及绘制问题;而且没有解决完全的虚拟环境漫游问题。

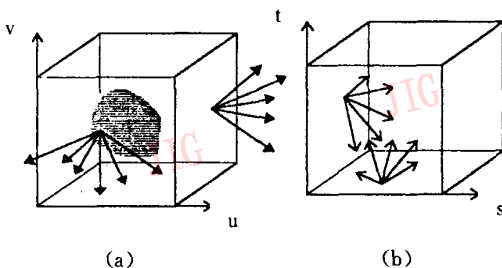


图3 离开和进入一封闭自由空间的光流

不可否认,以上所介绍的这几类方法对改善场景的表示、加速图形的绘制进行了有益的探索,为进一步研究 IBR 方法奠定了基础。

### 3 讨论和进一步的研究工作

本文总结并讨论了基于图象的多种绘制方法。根据场景表示模式的不同,我们把这些方法简单地

分成以上 5 类。这些方法各有特点,但都不是对虚拟环境实时漫游的一个完全的解决。因此,IBR 方法还有待进一步研究和解决。

在我们的基金项目研究中,正在探索更有效实用的 IBR 方法,主要的研究工作有以下几方面:

(1) 完全基于图象的光场绘制方法。我们在 Levoy<sup>[26]</sup>方法的基础上做了改进,采用多光场连接的表示方法,使光场表示能够支持虚拟场景内部的完全漫游。另外,在绘制时采用从压缩纹理直接绘制以及基于金字塔形式的纹理表示方法,有效地节约了存储开销,同时又不太影响绘制速度和质量。

(2) 基于立体视觉的混合光场表示方法。结合基于立体视觉的混合表示方法<sup>[10]</sup>和基于手持相机采样的光场绘制<sup>[27]</sup>两种方法来建立对建筑场景外观的光场表示,它支持由少数手持相机捕获图象的光场重建。从粗略的几何模型中可得到较精确的深度信息,它用于光场数据的校正;同时利用基于模型的图象变形和拼合技术可以大大丰富数据场,使光场重建更有效。

除以上讨论的工作外,我们还对以下几个方面感兴趣:

#### (3) 图象拼合算法

基于同一视点的图象拼合算法比较简单,可以用它来建立柱面全景,研究基于柱面全景的光场表示。对于不同视点的图象拼合算法已在上面第二种方法中涉及到,那是借助于模型实现的。由于稀疏的样本图象,因此面临选择参考图象的问题。这个方面有待进一步研究。

#### (4) 可变光照的光场算法研究

到目前为止,光场的研究多是考虑在固定的光照条件下。如果在时间参数  $t$  变化的情况下研究光场,则必须考虑光照的改变。另外,当强调场景中某个对象时,也需要改变场景的照明。为此,需要取样并存储同一视域方向上每个网格点在光照改变时的多个全视函数值。绘制时通过有效算法重构视图,再用于光场绘制中的纹理影射操作。这方面的工作已有一些作者在研究<sup>[28]</sup>。

### 参考文献

- 1 Lippman A. Movie Maps: An Application of the Optical Videodisc to Computer Graphics. In Computer Graphics (SIGGRAPH'80), 1980, 32~43.
- 2 Miller G E, Hoffert S E, Chen E, et al. The Virtual Museum: Interactive 3D Navigation of a Multimedia Database. The Journal of Visualization and Computer Animation, 1992(3): 183~197.

- 3 Ripley D G DVI-a Digital Multimedia Technology. Communications of the ACM. 1989,32(7): 811~822.
- 4 Laveau S,Faugeras O. 3-D scene representation as a collection of images. In 12th International Conference on Pattern Recognition (ICPR'94), volume 1, pages 689-691, Jerusalem, Israel, October 1994. INRIA Technical Report No. 2205.
- 5 McMillan L,Bishop G. Plenoptic modeling: An image-based rendering system. In Computer Graphics (SIGGRAPH'95), 1995, 39~46.
- 6 Fuchs H, bishop G, Arthur K, et al. Virtual space teleconferencing using a sea of cameras. In First International Symposium on Medical Robotics and Computer Assisted Surgery, Pittsburgh, PA, September 1994.
- 7 Kanada T, Narayanan P J, Rander P W. Virtualized reality: Concepts and early results. In IEEE Workshop on Representations of Visual Scenes, pages Cambridge, MA, 1995,69~76.
- 8 Kang S B,Szeliski R. 3-D scene data recovery using omnidirectional multibaseline stereo. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'96). San Francisco, CA, IEEE Computer Society Press. 1996,364~370.
- 9 Koch R. 3-D surface reconstruction from stereoscopic image sequencerink. In Figth International Conferende on Computer Vision (ICCV'95), MIT, Cambridge, MA, IEEE Computer Society Press. 1996,109~114.
- 10 Debevec P E, Taylor C J, Malik J. Modeling and rendering architecture from photographs; A hybrid geometry-and image-based approach. In Computer Graphics (SIGGRAPH'96), 1996,11~20.
- 11 Chen S E, Williams L. View interpolation for image synthesis. In Computer Graphics (SIGGRAPH'93), 1993,279~288.
- 12 Seitz S,Dyer C. Physically-valid view synthesis by image interpolation. In IEEE Workshop on Representations of Visual Scenes., Cambridge, MA, 1995,18~25.
- 13 Seitz S, Dyer C. Toward image-based scene representation using view morphing. In 13th International Conference on Pattern Recognition (ICPR '96). Vienna, Austria, 1996,1:84~89.
- 14 Seitz S, Dyer C. View morphing: Synthesizing 3D metamorphoses using image transforms. In Computer Graphics (SIGGRAPH'96), 1996,21~30.
- 15 Katayama A, et al. A viewpoint dependent stereoscopic display using interpolation of multi-viewpoint images. In SPIE Stereoscopic Displays and Virtual Reality Systems II, Febuiary 1995, 2409:11~20.
- 16 Werner T, et al. Rendering real-world objects using view interpolation. In Fifth International Conference on Computer Vision (ICCV'95), MIT, Cambridge, MA, IEEE Computer Society Press. 1995,957~962.
- 17 Brown L G. A survey of image regis tration techniques. Computing Surveys, 1992,24(4):325~376.
- 18 Szeliski R. Image mosaicing for tele-reality applications. In IEEE Workshop on Applications of Computer Vision (WACV'94). Sarasoda, IEEE Computer Society Press,1994,44~53.
- 19 Szeliski R. Video mosaics for virtual environments. IEEE Computer Graphics and Applications, 1996,16(2): 22~30.
- 20 Chen S E. QuickTime VR-an image-based approach to virtual environment navigation. In Computer Graphics (SIGGRAPH'95), 1995,29~38.
- 21 Kumar R, et al. Representation of scenes from collections of images. In IEEE Workshop on Representations of Visual Scenes, Cambridge, MA, 1995,10~17.
- 22 Sawhney H S, et al. Model-based 2D&3D dominant motion estimation for mosaicing and video representation. In Fifth International Conference on Computer Vision (ICCV'95), MIT, Cambridge, MA, IEEE Computer Society Press. 1995,583~590.
- 23 Adelson E H. Layered representations for vision and video. In IEEE Workshop on Representatons of Visual Scenes, Cambridge, MA, 1995,3~9.
- 24 Lengyel J, Snyder J. Rendering With Conherent Layers. In Computer Graphics (SIGGRAPH'97), August 1997.
- 25 Adelson E H, Bergen J R. The plenoptic function and the elements of early vision. In M. Landy and J. Movshon, editors, Computational Models of visual Processing, chapter 1. MIT Press, Cambridge, MA, 1991.
- 26 Levoy M, Hanrahan P. Light field rendering. In Computer Graphics (SIGGRAPH '96), 1996,31~42.
- 27 Gortler S J, Grzeszczuk R, Szekiski R Cohen M F. The Lumigraph. In Computer Graphics ( SIGGRAPH'96 ), 1996,43~54.
- 28 Wong T T. Image-based Rendering with Controllable Illumination. the 8th Eurographics Workshop on Rendering, June 1997.



**徐 丹** 云南大学计算机系讲师, 现在浙江大学 CAD&CG 国家重点实验室攻读博士学位。1990 年获云南大学数学系学士学位, 1993 年获浙江大学计算机系硕士学位。主要从事计算机图形学、多媒体及虚拟现实方面的研究。

**潘志庚** 1990 年毕业于南京大学计算机系, 获硕士学位。1993 年毕业于浙江大学计算机系, 获博士学位。现为浙江大学 CAD&CG 国家重点实验室研究员, 研究方向为分布式图形、虚拟现实和多媒体计算技术。

**石教英** 教授, 博士生导师, 现任中国图象图形学会副理事长。研究方向为虚拟现实、多媒体和科学计算可视化。

## Image-based Rendering Techniques in Virtual Reality

Xu Dan, Pan Zhigeng, Shi Jiaoying

(State Key Lab. of CAD&CG Zhejiang University, Hangzhou 310027)

**Abstract** Virtual Reality is a high fidelity interface that simulates what you see, hear and act in the nature. Image-based rendering is a much attractive topic in last years or two. It has become an important technique which supports modeling and real-time rendering in the VR applications. The paper discusses thoroughly many present image-based rendering methods and gives a catalogue of them. Moreover, we introduce simply the researches we are undergoing and the future works interested.

**Keywords** Image-based rendering (IBR), Real-time graphics generation, Virtual reality (VR).

## AVS 高级体视化软件系统进入中国市场

1998 年 10 月 28 日, 美国 AVS (Advanced Visual System) 公司和美国落矶山地球物理公司在北京友谊宾馆举办了高级体视化软件产品发布及演示会。AVS 公司总裁 Harry Cochran 和副总裁 John Sheehan 在会上全面介绍了 AVS 公司为适应硬件平台、操作系统、网络和通信方面的飞速发展而开发的一系列体视化产品, 包括多维多平台的开发环境 AVS/Express 开发版、多维数据可视化和分析的 AVS/Express 可视化版和对于最终用户的原始数据可视化软件 AVS5 等。同时演示了 AVS 体视化软件系统在医学图象、油气开发、航空航天、遥感、地理信息系统、通讯和金融等方面的应用。AVS 公司总裁宣布在华语地区, AVS 产品由美国落矶山地球物理公司唯一代理, 落矶山公司总裁夏恒仁博士在会上接受了委托书。我国的医学、石油、地质、电子、计算机、航空航天界科研单位和公司的代表共 300 人参加了会议。

## 计算机工程与应用

● **简介:** 本刊由华北计算技术研究所主办, 创刊于 1964 年, 属中国计算机学会会刊、中国电子学会一级会刊、计算机工程与应用学会学报、计算机中文核心期刊、中国科学论文统计用刊。是一本集学术与应用为一体的有较高权威性的计算机科技刊物。

● **主要栏目:** 博士论坛、数据库、网络与通信、研究探讨、开发设计、工程应用等。

● **读者范围:** 科研院所; 大专院校; 计算机工程项目规划、设计、实施人员; 计算机应用人员。

● **订阅办法:** 本刊 1999 年为月刊, 大 16 开, 每期正文 128 页, 定价 12.5 元, 全年 150 元, 邮发代号为 82—605, 每月 9 日版, 全国各地邮局均可订阅。

联系地址: 北京市 619 信箱 26 分箱

邮 编: 100083

电 话: (010)62323668