

计算机图形学年鉴：研究现状、应用和未来

MSRA网络图形组 计算机科学讲义 昨天

【导读】计算机图形学是支持各种影视特效、三维动画影片、计算机游戏、虚拟现实以及大家手机上各种照片视频美化特效背后的技术基础，在本文中，微软亚洲研究院网络图形组深入解释了图形学的现状、发展和未来。

谈及“计算机图形学”，可能很多人会觉得很有距离感，或者和计算机视觉、图像处理等学科混淆。但是，如果告诉大家图形学技术是支持各种**影视特效、三维动画影片、计算机游戏、虚拟现实**以及大家手机上各种**照片视频美化特效**背后的技术基础，相信大家都不会再觉得陌生。

在计算机诞生后，如何**在计算机中有效地表达、处理以及显示三维信息**，很快变成了计算机应用研究中的一个重要问题。针对这一需求，计算机图形学在二十世纪六十年代应运而生。在过去的几十年中，计算机图形学得到了长足的发展，并深深地影响了很多产业的发展 and 人们的生活、工作和娱乐方式。

在**硬件**上，图形学的发展催生了专用图形处理器GPU(graphics processing unit)的产生与普及。在**软件**上，图形学的基本绘制流水线已成为操作系统的一部分，为各种计算机平台提供显示和图形处理。**应用**上，图形学催生了影视特效、三维动画影片、数据可视化、

计算机游戏、虚拟现实、计算机辅助设计和制造等一系列产业，并为这些产业的发展提供了核心技术和算法支持。

作为一个计算机应用学科，计算机图形学的内涵和外延在过去几十年里也在不断地演进和扩展。如果我们回顾计算机图形学年会 **ACM SIGGRAPH** 上过去十几年发表的论文，一方面会惊叹其中纷杂精彩的研究题目和每篇文章作者的奇思妙想，另一方面也难免感到有些迷失，似乎图形学仅仅是在不断追求新奇和炫目的视觉效果。这是在一个快速发展的应用学科中很多刚入门的学者都会有的困惑。

图形学研究的核心是什么？推动图形学发展的动力是什么？未来，随着计算机图形学的进一步发展，哪些应用场景将呼之欲出？伴随着这些新的应用场景、需求的出现，我们面临的技术挑战又是怎样的？在这篇文章中，我们试着对图形学的现状、发展和未来做一些思考，并尝试一一回答这些问题。

计算机图形学研究与应用现状

图形学的核心科学问题是在计算机中有效的表达和处理三维世界的各种属性。图形学所处理的三维信息既包括物理真实世界中的三维信息，也包含我们人类大脑通过想象产生的虚拟的三维信息。计算机图形作为一个中介，提供了这两个世界在计算机中的一个共同的表达和信息交流渠道。

在计算机图形学诞生之前，物理学家和数学家已经对真实三维世界进行了长期的研究，把我们观察到的世界有效的解构为核心的一些物理量和他们之间相互作用的规律。如图1所示，传统的图形学受物理学和数学启发，将三维对象分解为**几何**、**表观**、**行为**或者**动态**三种属性。其中几何描述三维对象的几何形状；表观描述三维对象的材料光照属性以及材料如何和光相互作用；行为则表达了一个三维对象的动态特性从而决定了对象的运动和其他物体的交互行为。

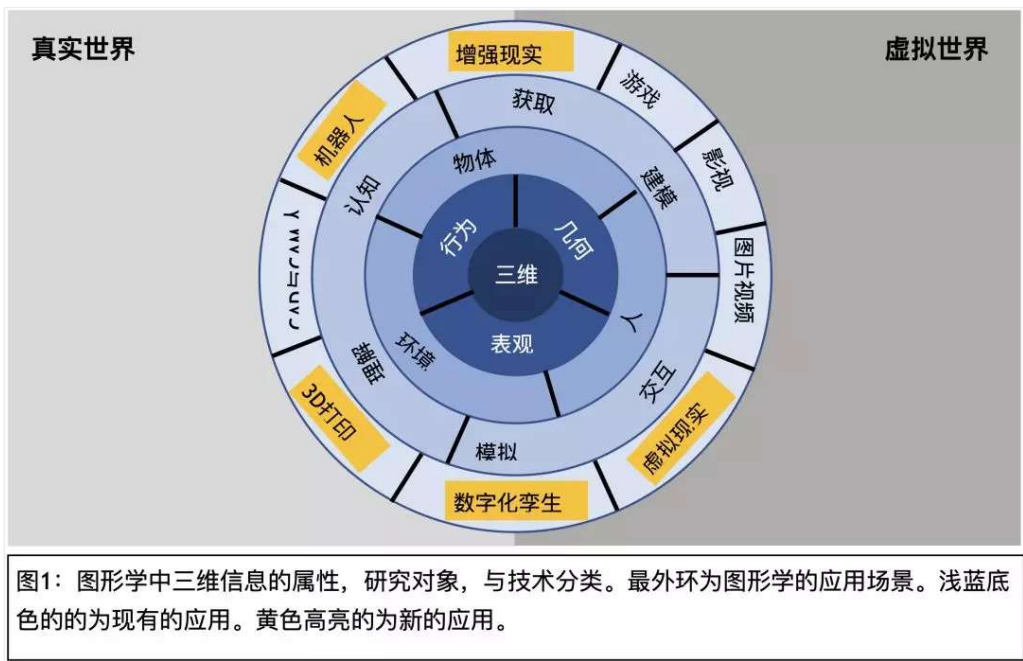
在这个基础上，针对不同对象特性和应用要求的不同，图形学研究中具体的三维对象又可以大致分为**物体**、**人**（包括类人的角色 character）以及**环境**三部分。

针对这些三维对象的不同三维信息（几何、表观、行为），我们把图形学的研究方向和技术也可以大致分为三个大类：

一是获取和建模。主要研究如何有效地构建、编辑、处理不同的三维信息在计算机中的表达，以及如何从真实世界中有效地获取相应的三维信息。这既包括三维几何建模和几何处理这一研究方向，也包含材质和光照建模、人体建模、动作捕捉这些研究课题。

二是理解和认知。主要研究如何识别、分析并抽取三维信息中对应的语义和结构信息。这个方向有很多图形学和计算机视觉共同感兴趣的研究课题，如三维物体识别、检索、场景识别、分割以及人体姿态识别跟踪、人脸表情识别跟踪等。

三是模拟和交互。主要研究如何处理和模拟不同三维对象之间的相互作用和交互过程。这既包含流体模拟和物理仿真，也包含绘制、人体动画、人脸动画等方面的研究。



在应用层面，图1中最外环黑色字展示了计算机图形学的经典应用场景，图形学的早期发展来源于使用计算机设计真实世界产品的需求，如汽车外形。因此，计算机辅助设计和制造成为了计算机图形学在真实世界的核心应用场景。随着图形学的发展，创建虚拟场景实现人类的想象，成为了图形学在虚拟世界的核心应用场景，产生了游戏、影视特效等应用场景。随着相机的普及，图片和视频的编辑也成为图形学中一个重要的横跨虚拟世界和真实世界的重要应用。

有了上面的“洋葱”结构，我们就可以对每一个图形学论文或者研究热点，通过其研究对象、三维属性和所属技术对其进行归类，比如**绘制 (Rendering)** 算法的研究是对场景的**表观属性**进行**模拟和交互**的研究：算法通过研究光与环境的交互作用，生成真实感的图像。

对计算机图形学发展模式的回顾和思考

回顾和思考过去几十年来图形学的发展，我们发现图形学研究的核心对象和科学问题并没有发生根本性的变化。但是技术和三维信息的表达却在不断的发展更新。而这些技术的发展往往发端于新的硬件设备的出现和普及。

如图2所示，新的硬件设备的出现一方面往往引发了新的应用需求，或使得某个应用的技术成本急剧的下降。另一方面带来了新的数据和技术问题，从而引发了新的研究方向和技术，推动了对图形对象表达的更新和研究方法的更新。而这些技术的发展又反过来进一步推动了硬件的发展和应用的普及，从而带动整个领域的快速迭代发展。光栅化图形显示技术的出现，GPU图形学流水线的提出，可编程GPU的出现，三维扫描仪的出现，图像采集设备的出现和普及，是过去几十年图形学发展几次浪潮的背后缘起。

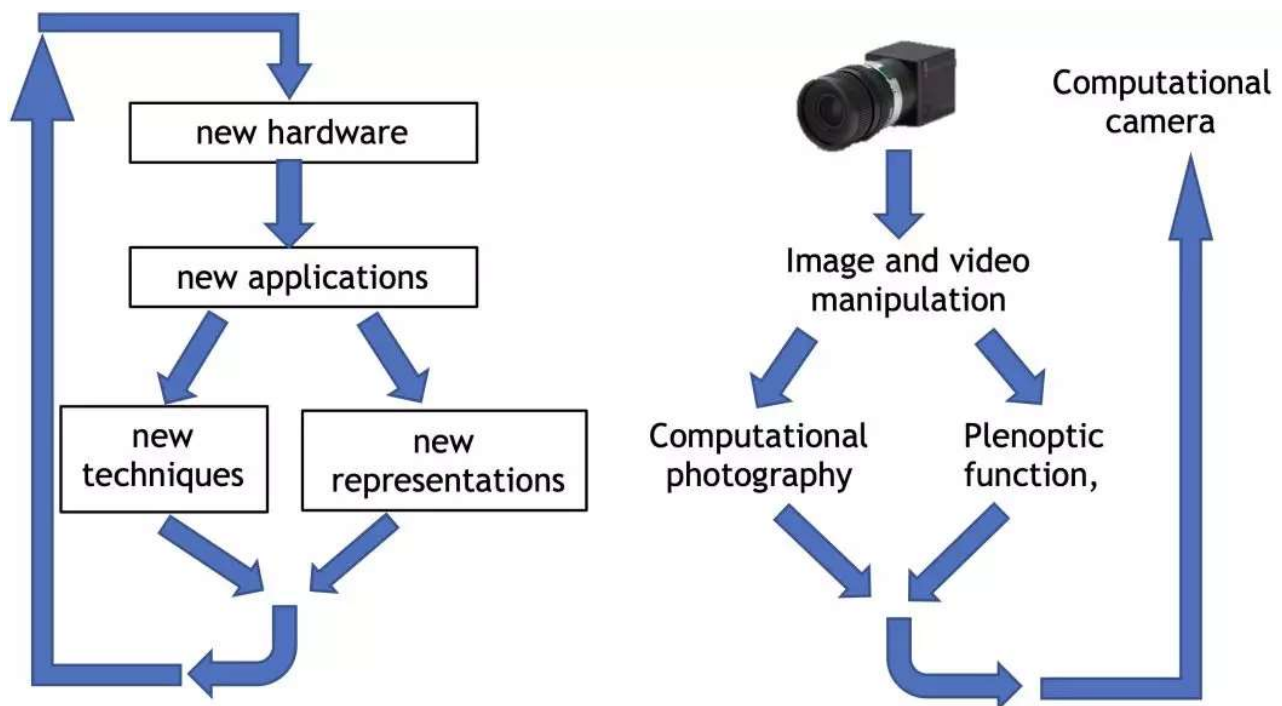
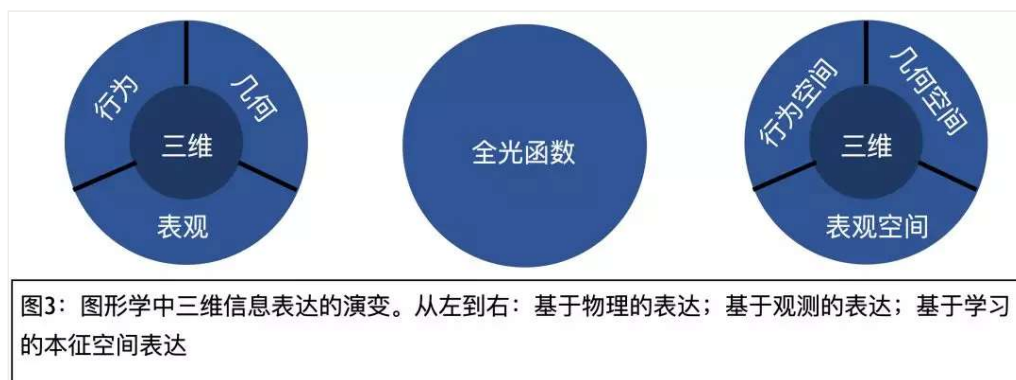


图2：对图形学发展模式的一些思考。

这里我们以基于图像的绘制和光场表达的出现为例对上述的发展模式做一个分析。传统图形学中，所有的研究对象和属性基本是基于物理表达。在这一表达下，几何和物理过程成为了各个研究方向的基础。

从20世纪后期开始，随着图像捕捉设备的快速发展，人们有机会对真实世界进行大量的图像采集。这些大量的图像一方面需要研究者研发有效的图像编辑，分析和解构技术。另一方面，也使得研究者开始探索是否可以抛开背后的物理机制，直接基于三维世界的这些观察建立新的表达。由此催生了基于**全光函数**的表达和基于图像的绘制技术。



这里，全光函数是一个高维函数，记录了在一个三维场景的任意一点 (x,y,z) 沿任给方向 (θ,ϕ) ，在某一时间 t ，在每一波长 λ 上的光强。在真实世界中，虽然每种我们可以观察到的视觉现象都可以解构为以上的三维基本属性及其相互作用，但是我们的人眼和图像传感器可以观测到的却是光线，即 **全光函数（Plenoptic Function）** $(x,y,z,\theta,\phi,t,\lambda)$ 的一个采样。

图像的表达和绘制技术的进步，催生了计算摄像学的发展，反过来促进了新的摄像设备的诞生和发展，并进一步促进了图形学中对全光表达函数的采样与重构、分析与编辑、认知与理解三个方向的研究。这一迭代发展过程，从根本上将三维信息的表达由基于物理的表达推广到新的基于观测的表达，从而拓展了研究方法，并将图形学的研究领域从传统三维几何扩展到了图像和视频，并且和计算机视觉、图像和视频处理、光学成像等学科产生了新的交叉。

计算机图形学的未来：设备和硬件

展望未来，我们认为，上述图形学发展的模式还会继续。硬件的发展和革新，会不断促进了新的图形技术和应用产生和迭代发展。在这个

过程中，图形学也在不断地结合计算机视觉、光学、信号处理与机器学习等学科的最新研究成果，来解决图形学中的研究问题。下面，我们就从各个层面对计算机图形学的未来进行一些大胆的展望。

在硬件设备方面，我们认为下面的这些硬件会迎来新的发展并带来图形学技术和应用的革命性进展。

- **三维显示**。提供高分辨率，高动态范围的全三维显示。包括近眼的光场显示设备，或者多焦平面显示设备。或者远场的全沉浸式的光场显示设备。
- **深度相机**。提供和现有的彩色相机相匹配的高分辨率，高帧率，低功耗，低噪声的深度相机。
- **多自由度机械手和类人软体机器人装置**。提供低成本，高精度，编程可控的多自由度机械手以及具有类人外形的软体机器人。
- **三维打印机**。提供同时支持多种打印材料，高精度，低价格，快速的三维打印。
- **IOT与传感器**。提供小型、省电、低成本的能测量真实世界各种物理参数的传感器与实时的数据收集。
- **力学捕捉与反馈设备**。提供精确的，具有高空间分辨率和力分辨率的触觉输入输出。

计算机图形学的未来：应用场景

随着上述硬件设备的发展和普及，以及计算机视觉和机器学习技术的进步，图形学的应用场景将得到更大的扩展。如图1黄色高亮部分所

示，面向真实世界，机器人和三维打印将成为新的应用场景。面向虚拟世界、虚拟现实，混合可视媒体将成为新兴的应用场景，带给人们更好的娱乐体验，释放人类的想象力。在真实世界和虚拟世界之间，增强现实将虚拟信息融合进真实世界，并增强人类在真实世界的体验；数字化孪生则产生真实世界在虚拟世界的镜像，方便我们更好地管理规划真实世界。下面，我们将讨论每个应用场景，和它们对相关图形学技术的需求。

机器人

随着机械硬件，传感器设备以及人工智能技术的进步，多用途的机器人将逐渐被应用到不同的真实世界场景中，自动化或半自动化地帮助人类完成各种任务。自动驾驶可以被认为是这一场景中一个应用。机器人为了在不断变化的三维场景中完成给定任务，不仅需要实时重建不断变化的三维场景的几何，还需要识别真实场景中的物体的类别和物理特性，从而预测物体的运动并决定自己的运动。同时，机器人自身也需要实时的动态模拟技术来准确地规划和预测自己的运动，和环境中物体进行交互，从而最终完成任务。

三维打印

三维打印硬件的发展使得生产复杂几何形状和不同几何形状的成本显著下降。和传统的减材制造不同，三维打印可以精确地控制三维形体中每个体素的材质构成，从而可以产生更为丰富的设计和功能。为了支持三维打印，图形学技术需要将设计与物理模拟更好地结合在一

起，提供一体化的端到端解决方案。通过高效的计算模拟和逆向优化，帮助设计师和制造者快速地设计产品的三维形状和内部材质分布，从而达到所需要的功能。

虚拟现实

虚拟现实技术作为一类新的媒体，提供了全新的沉浸式体验，在教育、游戏等方面具有重要的应用。为了达到更好的虚拟现实体验，我们不仅需要图形学渲染技术的进步，也需要更好的物理模拟技术和交互技术，提供视觉外其他物理特性，如触觉和听觉的建模和实时渲染。更为根本的是，如何更加快速地生成高质量的三维虚拟内容，以及如何在虚拟环境中与不同的虚拟内容进行有效的交互，是虚拟现实应用得以成功和普及的关键。

增强现实

增强现实和混合现实系统通过将虚拟三维内容叠加在真实场景中，从而实现了虚拟信息和真实世界的融合，提高了人们在真实世界的工作效率，提供了个性化的环境和更好的生活体验。某种程度上，可以将增强现实理解为新一代的精确GPS定位系统。它可以提供在场景中的实时精确三维定位和实时的三维地图构建服务。为了实现这一目标，三维场景的实时捕捉建模（包括几何、表观、物理特性和行为），分析和理解将成为这一应用场景背后的核心技术。

数字化孪生

和增强现实将虚拟信息叠加在真实世界相反，数字化孪生尝试建立真实世界在计算机中的虚拟镜像，并实时地记录预测真实世界的所有变化。结合IoT和传感器技术的发展，数字化孪生技术将提供真实环境的完整数字化，从而实现对真实世界的高效信息分析和控制。同时，数字化孪生将为机器学习技术提供更多的数据和训练环境。为实现这一目标，我们需要研究更加有效的三维建模和捕捉技术，以及实时的物理模拟技术。

计算机图形学的未来：技术挑战

上文中，我们看到了未来计算机图形学的应用场景，并讨论了每个应用场景所需要的关键技术。这些需求也为图形学的发展提出了一系列的研究问题与挑战：

- **高效高质量的三维内容创作系统**

虽然已有的图形学算法和系统可以让艺术家创作出具有高度真实感的虚拟环境和栩栩如生的人物及其动态，这一过程仍然需要大量时间、专业技巧以及昂贵复杂的设备。发展高效高质量的三维内容生成算法和创作系统是图形学研究中一个永恒的任务，也是虚拟现实、数字化孪生以及新一代的混合媒体等应用场景得以实现的关键技术。

为了实现这个目标，我们需要在以下三个方面的研究取得突破：一是研发新一代的捕捉硬件系统和算法，使得普通用户越来越容易从真实

世界中快速地捕捉所需要的三维内容；二是利用三维内容属性的本征属性，从用户的少量输入如草图、照片、视频中构建符合用户需求的三维内容；三是利用机器学习技术，如对抗神经网络（GAN），直接从已有的大量数据中生成新的三维内容。

- **三维世界的实时理解与分析**

实时地对我们所处的三维世界进行理解，识别出场景中物体和人，推断物体和人之间的空间关系与约束，以及人的动作，是增强现实和机器人应用场景中的核心技术。

在计算机视觉领域，由于大量标注数据的出现和深度学习技术的发展，图像和视频的理解与分析工作取得了飞速的发展。但是三维世界的理解和分析工作仍旧处于起始阶段。

一方面，三维内容由于获取困难，可使用的标注数据少，数据噪声大。另一方面，三维数据表达多样、维度更高。这些特点也对三维世界的分析理解算法的实时性、鲁棒性提出了更大的挑战。如何研发适合三维内容的通用表达和机器学习算法，结合已有的图片和视频信息进行三维世界的理解和分析也是这一领域未来研究的重点。

- **大规模可扩展的实时模拟技术**

在真实世界中，不同物体的运动和相互作用构成了世界复杂的动态。而在人类社会中，人的行为和交互则更为复杂。模拟这些复杂的动态

和交互是图形学中一个重要任务，也是三维打印、机器人、数字化孪生应用背后的重要技术支撑。现有的图形学技术发展了一系列快速的技术来模拟环境、物体和人的运动和复杂交互。然而，这些算法仍然存在复杂性高、计算不稳定、收敛慢的问题。寻找适用于不同场景的更为通用的模拟算法，发展快速数值解法，将深度学习技术用来加速优化求解，以及利用增强学习技术进行运动的规划都是这一领域下一阶段的研究重点。

■ 人机交互与图形学的深度结合

随着新型传感器、穿戴设备、VR/AR/MR设备的迅猛发展，人们有着更多的方式与机器打交道。这些新的输入输出方式也为图形学研究带来新的挑战。如何使用多元异构的数字输入信息来指导生成符合用户期望的三维影像与世界，如何针对不同设备设计便捷的输入方式与交互手段，如何协同多用户的操作并实时提供数字上和物理上的真实反馈、如何动态调整已有算法以适配用户的不断更新的个性化要求等问题，都值得图形学研究者与从业者积极探讨与深入研究。

计算机图形学的未来：研究趋势

通过上述内容，我们可以很容易地看到，每个新的图形学应用场景都不可能由单一的一个图形学技术来解决。为此，我们不仅需要在每个研究方向上进一步努力，更需要借鉴最新的机器学习技术和计算机视觉技术，以及本领域的其他研究方向的技术和算法，才能最终解决问题。

从基于物理和观察的表达到基于学习的本征表达

现有的图形学可以对单个三维对象的三维属性进行有效的表达和处理。但是对于所有的三维对象构成的三维属性空间，例如所有特定人造物体（椅子）的三维形状空间，所有真实世界表面材质的空间，或一个场景中所有光照传输路径的空间，我们仍然缺乏有效的研究和表达。随着数据的增多和机器学习算法的应用，这方面的研究慢慢成为可能。这些研究会导致三维信息新的表达形式的出现，即基于机器学习的三维形状、材质、行为等属性空间的本征表达。这一表达会与传统基于物理的表达与基于观测的表达共存。这一研究将成为图形学的一个基础理论问题，并对我们研发高效的三维内容建模、模拟和识别理解算法都具有极为重要的意义。基于这一全新的表达，在图形学研究中，如何有效的结合这一全新的表达和已有的表达，如何形成不同表达间的映射与转换，也会成为一个重要的研究问题。

在技术层，基于这一新的表达，机器学习技术将利用三维数据的本征属性而不是物理属性来解决三维内容捕捉、生成、处理和模拟问题。如何将机器学习技术有效地用于高维的三维图形数据，结合用户的交互输入，是目前研究的热点。进一步将原有基于物理的方法与机器学习的方法有效结合，充分利用两者的优势，也是图形学研究中一个广受关注和需要解决的重要问题。

从属性的单一表达到属性的统一表达与融合

现有的图形学对不同的三维属性（形状、表观、动态）具有各自单独的表达和不同的处理方法。然而，一个三维对象（如物体）的表观、形状和行为并不是任意组合的，属性之间也具有一定的约束和相关性。比如，一个木制的椅子由于材质的限制，其椅腿的粗细和细节不可能是任意的。而它的表观、重量和可能的运动特性也与木头材质属性紧密相关。如何得到所有三维属性更为简洁一致的表达是图形学研究中的一个基础问题。

在技术和应用层面，针对每个特定的应用和问题我们研发了可用的算法和解决方案。然而，这些算法或工具集是彼此孤立的。在目前的实际应用中，我们需要具有领域知识的人将这些算法放在一起，辅以大量的人工和反复使用，调整修改每个工具的结果，反复迭代来达到最终的目标。举例来说，为了设计一个像章鱼爪子的软体变形机械手，设计者可能需要先用造型软件设计机器手外形，然后运行仿真模拟软件计算力学特性。设计好后，再运行三维打印软件进行制造。制造好以后，进行实际测试。由于每个模块不知道最终的目标，由此带来的误差需要人工反复修改并重复这一过程。

为此，我们需要将不同的技术方案，如几何设计、仿真模拟、三维打印等有机地集成到一起，将每一步的约束引入到其他算法中，并允许所有的算法在统一的逆向优化反馈框架下进行自动迭代，快速地生成满足设计要求的结果。在最近几年，越来越多的研究尝试将不同的技术融合在一起，形成一个端到端的解决方案，这也成为了图形学研究的一个趋势。

从基础工具集到智能系统的演进

图形学技术的一个重要目标是将用户的抽象设计意图变成具体的三维对象。设计的最终目的是满足一定的物理功能或故事情节的视觉展现。在功能和情节的约束下，最终得到美观、成本合理的物理设计和视觉作品。目前，针对每个环节，已有的图形学技术实现了基本的工具，可以帮助用户完成形状、表观、动态等底层三维属性特性的生成、编辑，以及物理特性的模拟。然而，由于图形学的属性和对象表达缺乏对物理功能和语义的有效描述，使得现有的图形学技术工具集无法帮助用户有效的将高层抽象的功能和情节描述转化为具体的三维属性和表达。随着机器学习的引入，图形学技术的集成和表达的融合，图形学研究将慢慢从三维信息的基础设计和表达工具向高层语义的目标进发，最终实现从用户的高层语义描述自动生成三维内容的最终目标。

放眼未来，随着计算机图形学的进一步发展，计算机辅助设计和制造技术的进步，带有传感器的三维打印的个性化产品和机器人将被广泛应用于人类的实际生活和现实世界中。而真实环境的数据化孪生也将在计算机中实时地监控着真实环境的动态变化，规划协调机器人高效地完成不同任务。而在虚拟世界中，随着内容创作工具的进步，每个人的艺术天分都可以得到充分发挥，从而自由地创建自己的虚拟世界、游戏和虚拟化身。随着下一代的虚拟现实设备和增强现实设备的出现，真实和虚拟的世界会得到更好的融合，新一代的人类将不需要再区分真实世界和虚拟世界。人、计算机（机器人和虚拟世界）和真

实的物理世界将和谐高效地融合在一起，带给人类一个超现实的世界。

本文作者：童欣、刘洋、董悦。

作者感谢与网络图形组各位同事的日常讨论，以及与美国德克萨斯A&M大学柴金祥教授、浙江大学周昆教授的讨论所带来的启发。

* 本文经授权转载自新智元，原文链接



计算机科学讲义
助力学术写作出版！

扫描左侧二维码
关注我们