



清华大学
Tsinghua University

三维视觉：过去，现在和未来

刘焯斌

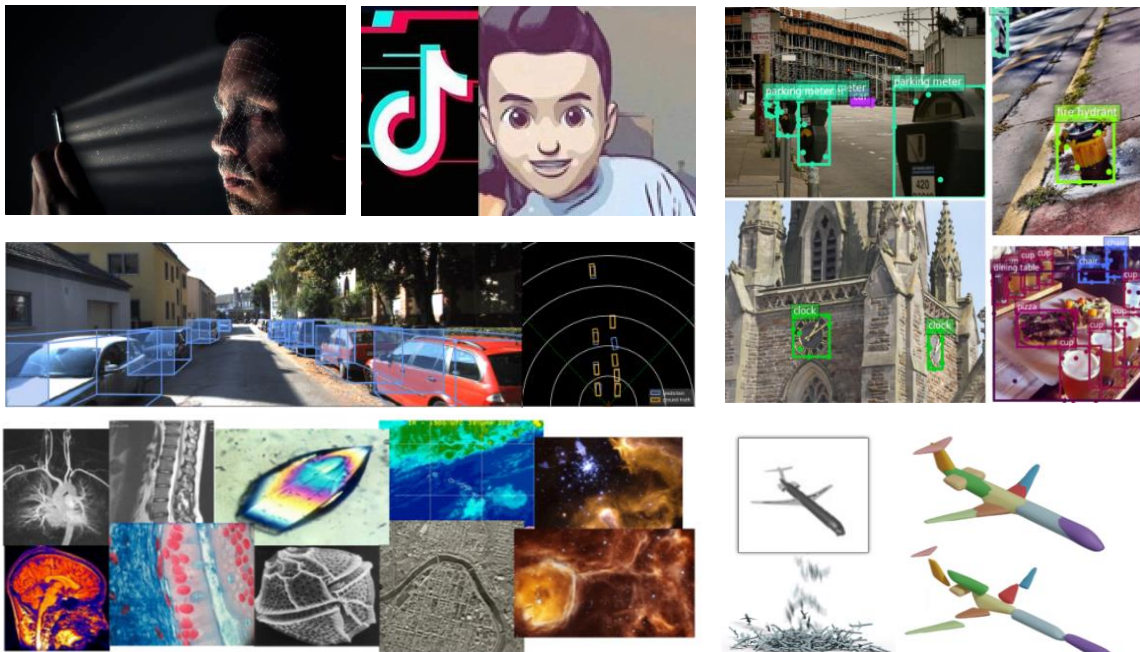
6/3/2021

主要内容

- **计算机视觉与三维视觉**
- **三维视觉的应用需求**
- **三维视觉的发展历史**
- **未来的三维视觉与计算机视觉**

计算机视觉

计算机视觉的产业价值：产业应用的技术基础



计算机视觉的需求无处不在

视觉市场规模飞速增长

计算机视觉是人工智能最重要的组成部分

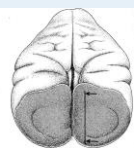
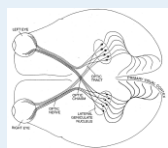
机构		2017年	增长率	2020年	增长率
国内机构	CAICT	80.2亿元	166%	600亿元	96%
	腾讯研究院	>70亿元	/	660亿元	~110%
	艾瑞咨询	40亿元	264%	725亿元	162.7%
	艾媒iiMedia	68亿元	124%	780亿元	125.5%
国外机构	Ganter	10.7亿美元	110%	110亿美元	117%
	CB Insights	13.5亿美元	超过150%	160亿美元	128%

计算机视觉

计算机视觉的起源与目标

理论基础

生物

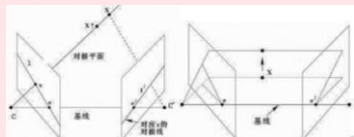
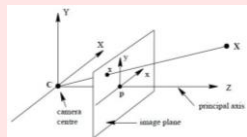


Hubel, D. H., & Wiesel, T. N. (1979). Brain mechanisms of vision. *Scientific American*

视觉分层理论的出现为计算机视觉奠定理论基础

视觉原理 → 语义理解

数学



Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision*. Cambridge university press.

多视图几何的出现为计算机视觉提供数学工具

测量几何 → 三维感知

计算机模拟人脑分析



用计算机模拟人对物理世界的视觉感知，启发对人脑与智能的理解

计算机视觉

计算机视觉的学术影响力

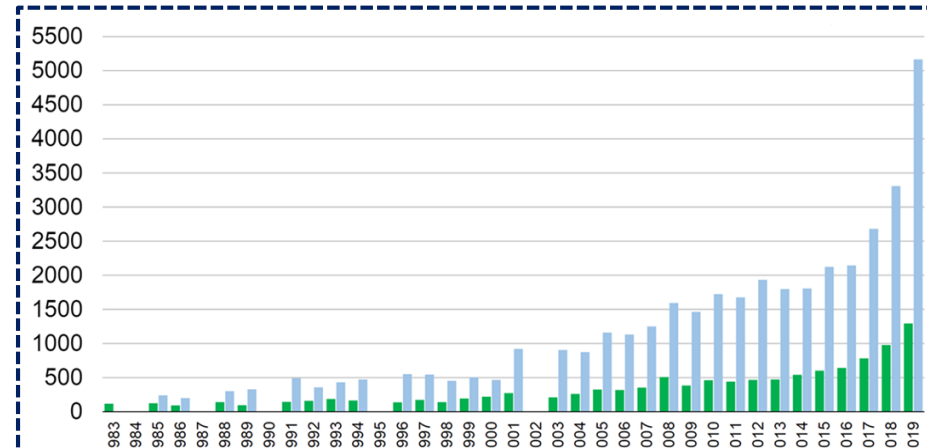
排名	期刊或会议	h5-指数	h5-中位数
1	Nature	376	552
2	The New England Journal of Medicine	365	639
3	Science	356	526
4	The Lancet	301	493
5	IEEE CVPR	299	509
6	Advanced Materials	273	369
7	Nature Communications	273	366
8	Cell	269	417

* h5 指数是指在过去整整 5 年中所发表文章的 h 指数。h 指在 2015-2019 年间发表的 h 篇文章每篇至少都被引用过 h 次的最大值。

谷歌学术发表2020年最新的学术期刊和会议影响力排名



CVPR会议各项指标逐年上升



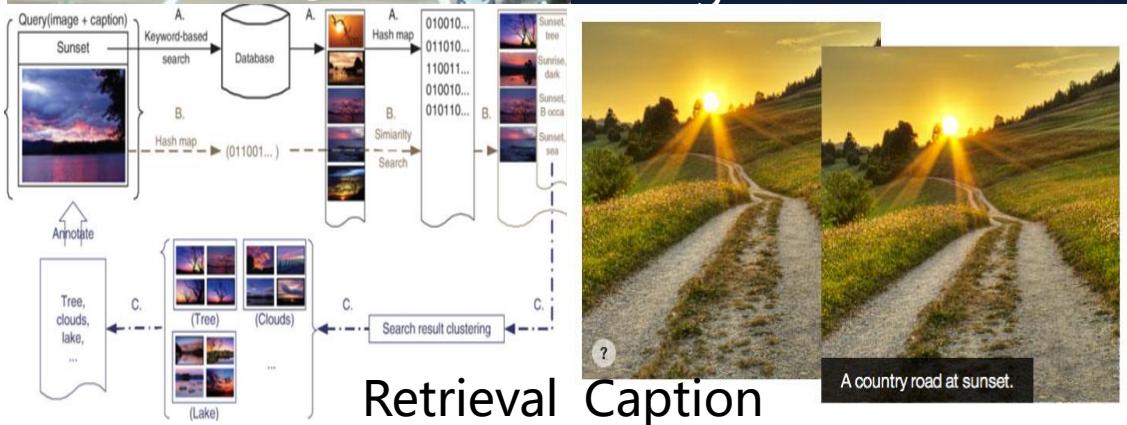
CVPR会议投稿量逐年上升

计算机视觉是国际学术热点，学术影响力名列前茅

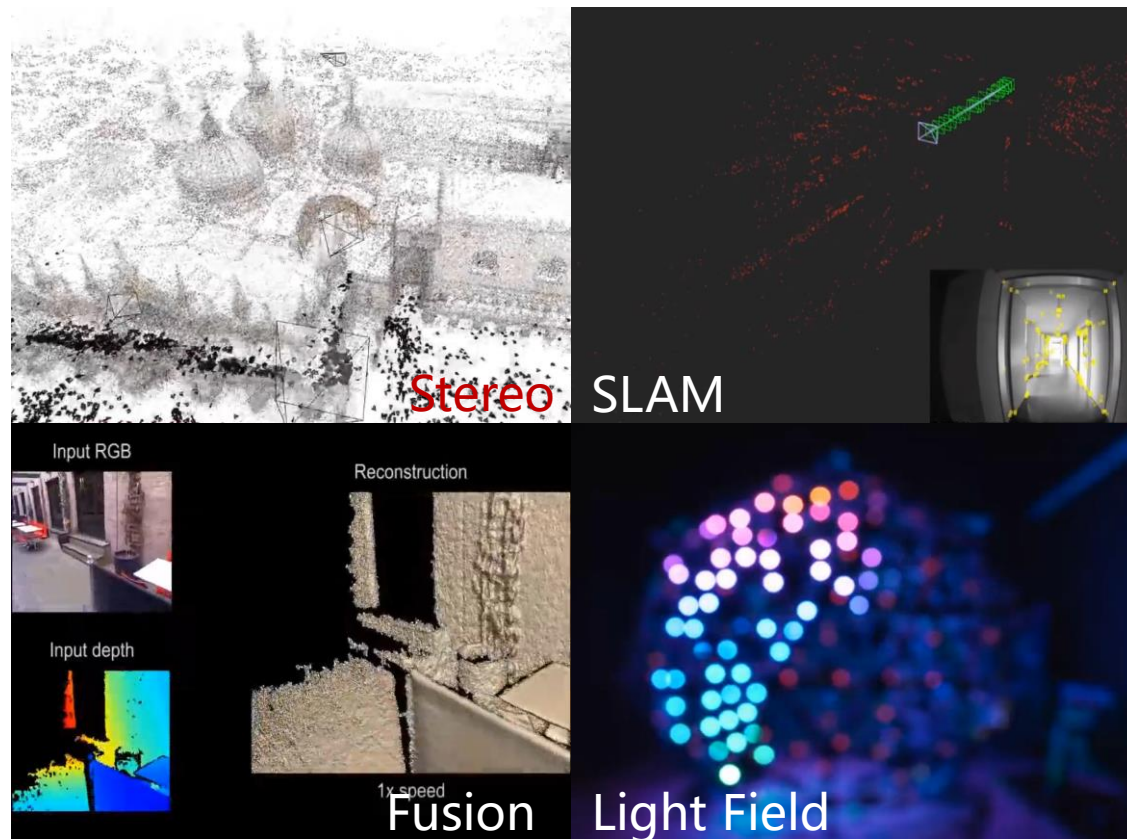
计算机视觉

计算机视觉的研究内容

语义理解



三维感知

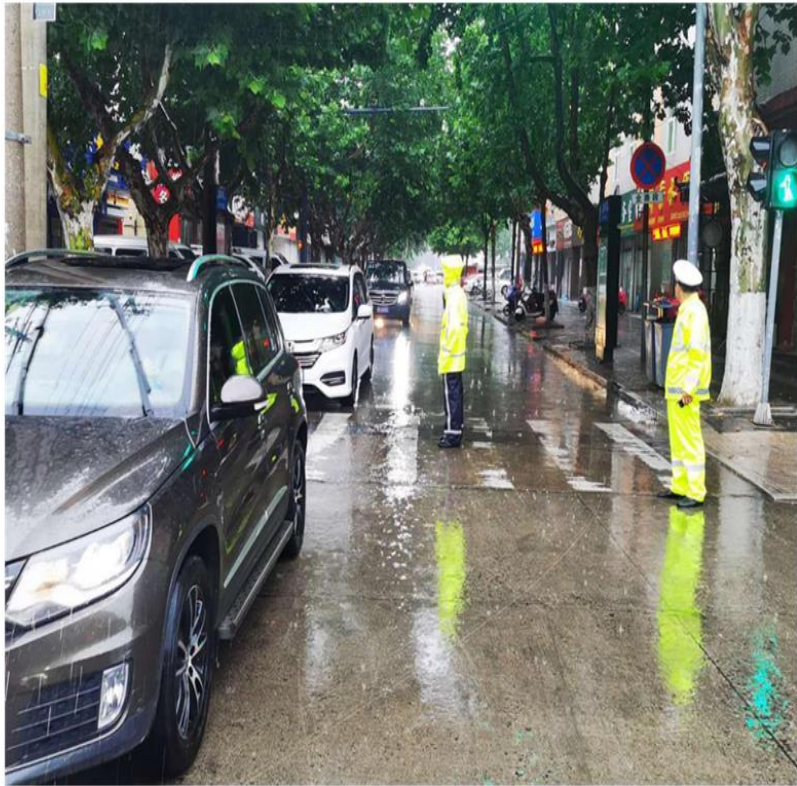


语义理解和三维感知是计算机视觉的两大板块

三维视觉

□ 三维视觉：从三维感知到三维理解

- 基于图片、视频以及各类深度传感器信息，采用几何、统计以及优化等数学工具对现实世界进行**三维测量、定位、建模及理解**。



- 前面是否有车？有多少车？
- 前面是否有交警或行人？
- 前面是否有障碍物？
- 交警离我有多远？
- 两车之间的距离多少？
- 交警看到什么？想什么？
- 下一秒会发生什么？
-

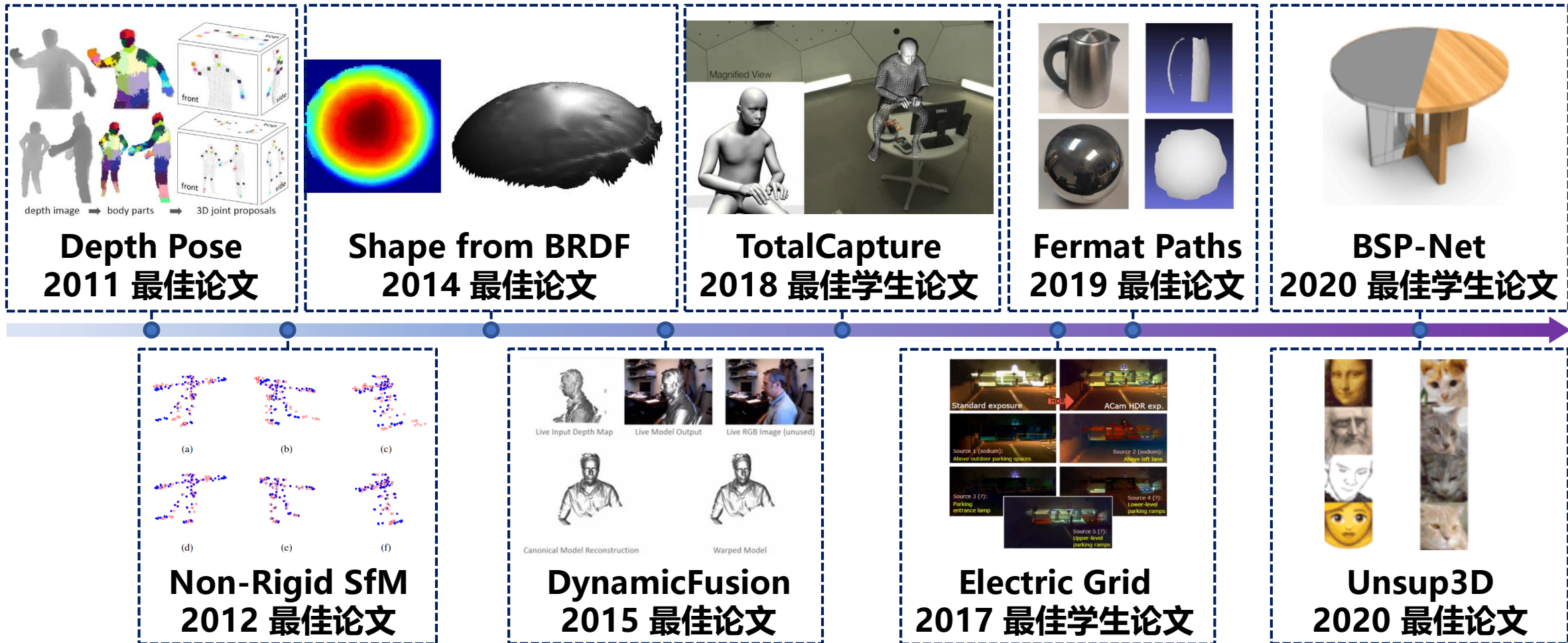
二维语义理解
可以完成

需要
三维视觉!

三维视觉是实现视觉智能的不可或缺的手段

三维视觉

三维视觉受到极大关注



近10年来CVPR最佳论文中，**三维视觉**占据半壁江山

主要内容

- 计算机视觉与三维视觉
- 三维视觉的应用需求
- 三维视觉的发展历史
- 未来的三维视觉与计算机视觉

三维视觉的应用需求

地球级数字新世界

*华为、微软、谷歌、苹果等企业纷纷投入地球级数字世界建设

人物
三维
重建



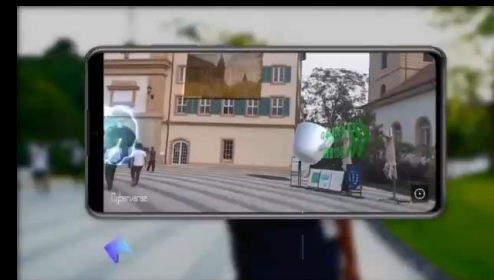
全息
广告
牌及
点评



AR
多人
互动
游戏



虚实
结合
城市
街景



三维
智能
助手



AR
步行
导航



厘米级3D精准空间定位，真实世界与物理世界的无缝融合，万物互联的智能世界

三维视觉的应用需求

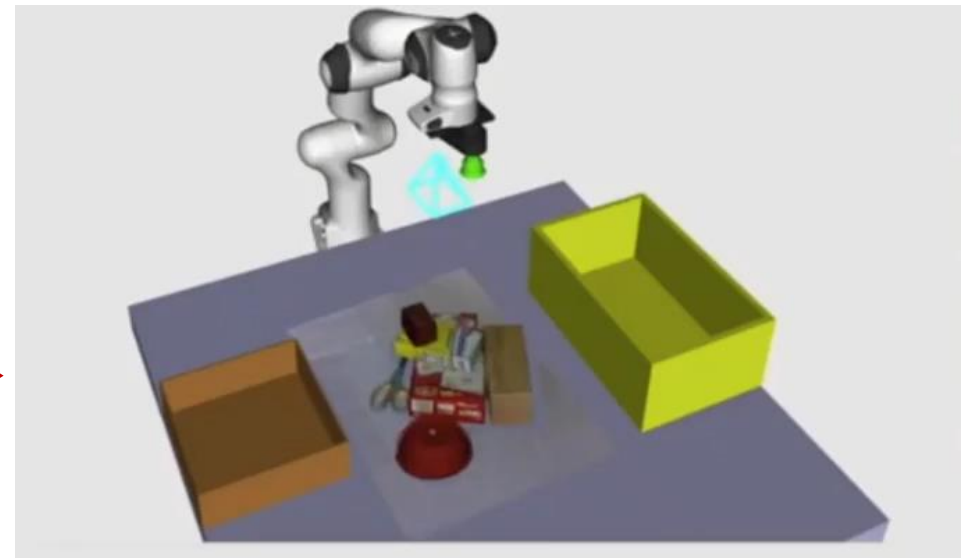
□ 机器人抓取系统

构建与现实同步的三维数字环境是机器人理解环境、自主作出判断的必要前提



3. 抓取

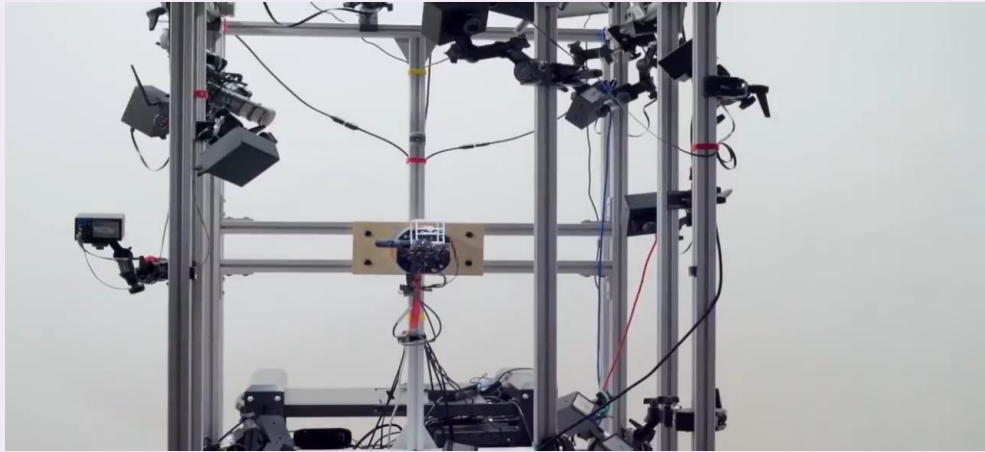
1.
三维
重建



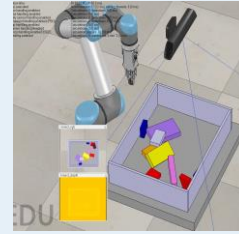
2. 预演、仿真、学习

三维重建是机器人观察世界、理解世界、与世界互动的必要手段

三维视觉的应用需求



基于三维视觉的仿生人手



Huaping Liu et al. II2020

基于三维重建的复杂环境交互



Siyan Dong et al. SIGGRAPH 2019

多机器人协同室内三维重建



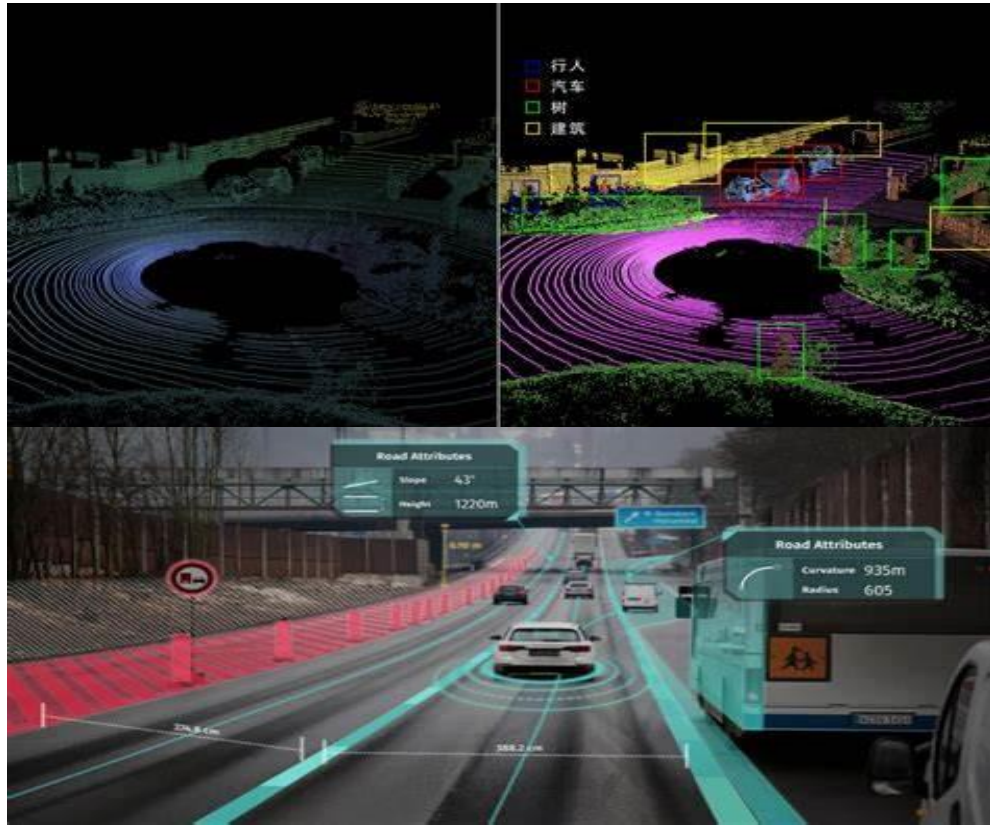
Dharmadhikari et al. ICRA20

基于三维重建的自主洞穴探索无人机

三维重建在各种机器人应用中扮演着不可或缺角色

三维视觉的应用需求

自动驾驶感知系统



高精地图构建与实时三维感知



行驶轨迹可视化与预测

精准三维地图构建与实时三维环境感知是自动驾驶智能决策的关键

三维视觉的应用需求

□ 全息交互与通信



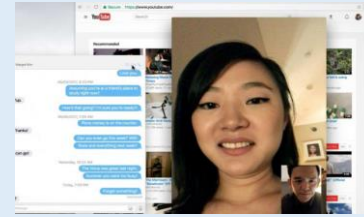
邮件
公元前1500



电报
1830年



语音电话
1860年



视频会议
2000年



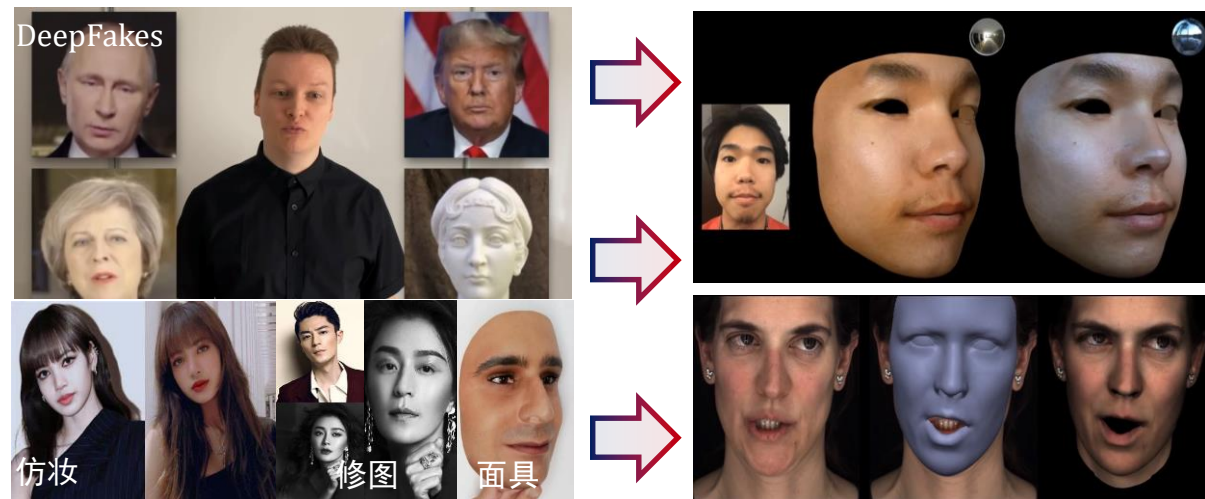
全息通讯
当前到未来



动态三维重建用于未来全息交互，是媒体通信技术的革新

三维视觉的应用需求

三维视觉驱动支付级人脸识别



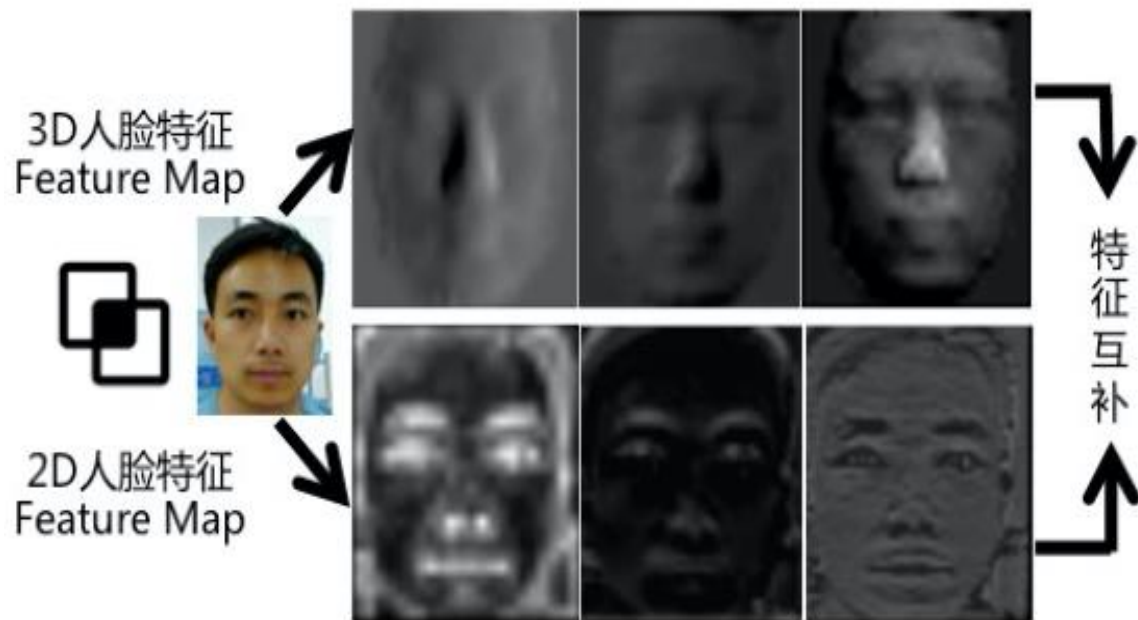
二维人脸识别困境

三维人脸识别

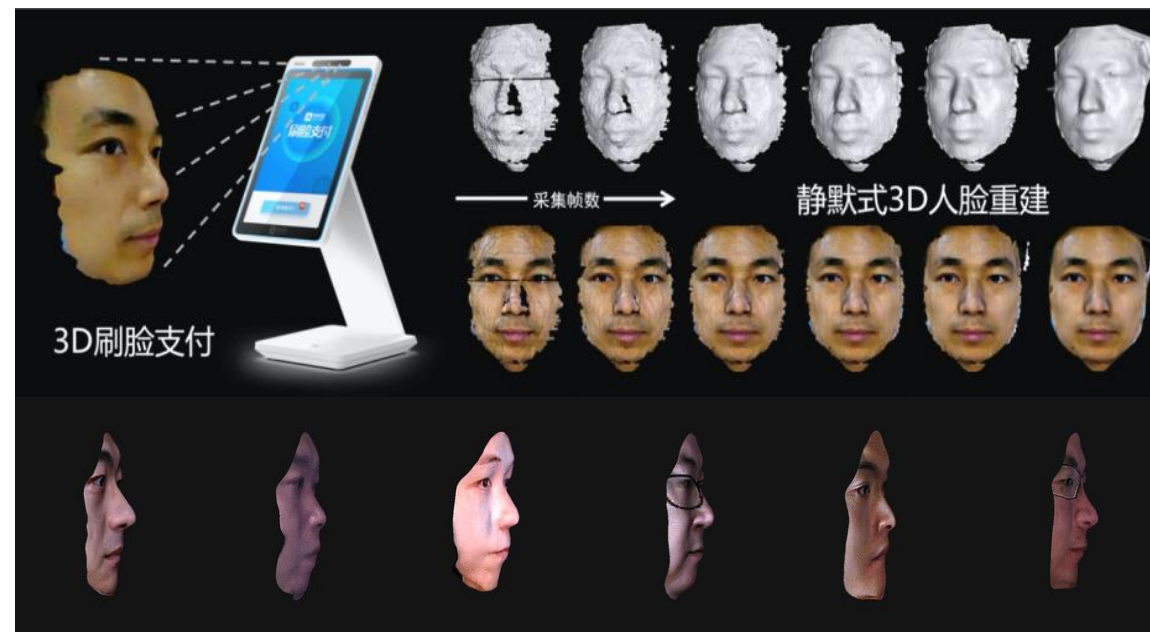
- 三维采集硬件在移动端快速发展 (3D collection hardware is rapidly developing on mobile devices)
- 三维识别克服二维困境，将刷脸这一高隐私、高风险、高价值的技术正式带向产业应用 (3D recognition overcomes 2D dilemmas, bringing facial payment technology with high privacy, high risk, and high value to formal industrial application)

三维视觉的应用需求

三维视觉应用于人脸识别



二维与三维结合的多模态人脸识别



三维人脸重建应用于人脸识别

与蚂蚁金服合作，实现了**三维**人脸识别，相比二维算法，精度显著提升：

- 1对1的人脸识别**错误率下降两个数量级**；1比N识别可支持的**数据规模扩大约30倍**
- 三维活体识别和高精度三维重建将进一步提升识别精度，保障**支付级别的安全性**

三维视觉的应用需求

动物模型行为计算

动物模型是**生命科学**的媒介

三维行为是动物固有属性

MONKEYS AND MICE ENLISTED TO FIGHT CORONAVIRUS
Animal models can reveal how infections develop, and aid drug and vaccine efforts.

By Ewen Callaway

With no sign that the coronavirus is going away, researchers are looking to animals to understand COVID-19. They are testing monkeys, mice and even ferrets to answer key questions about the disease and to fast-track potential drugs and vaccines for clinical trials.

Teams in China have reported initial findings from studies in which they infected monkeys and mice engineered to be susceptible to infection by the coronavirus, called SARS-CoV-2. And a team at the Australian Animal Health Laboratory in Geelong is studying the

Why does the coronavirus spread so easily?
A microscopic feature could make the virus more infectious than the SARS virus.

As the number of coronavirus infections passes 100,000 worldwide, researchers are racing to understand what makes it spread so easily.

A handful of genetic and structural analyses have identified a key feature of the virus — a protein on its surface — that might explain why it infects human cells much more readily than does the coronavirus that causes severe acute respiratory syndrome, or SARS.

Other groups are investigating the doorway through which the new coronavirus enters human tissues — a receptor on cell membranes. The cell



国际顶级学术期刊 **nature** 在3月份发文指出，建立猴子和小鼠等**动物模型**来对抗**新冠肺炎**



SMALL WONDER
Complete skull of diminutive dinosaur preserved in amber

Nature, 2020.03



建立模型



行为观测

宏观行为
社交距离
交互行为
传染特征

微观行为
病理表现



疫苗生产、病理研究




三维个体运动



三维群体运动



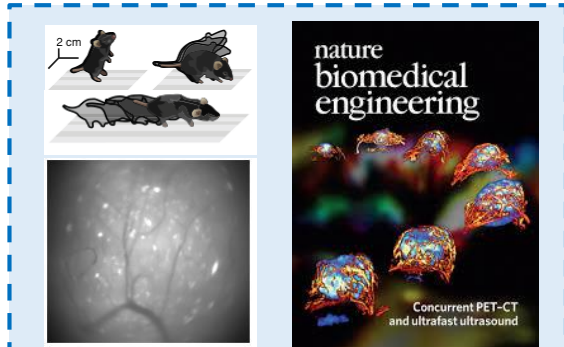
三维表情运动



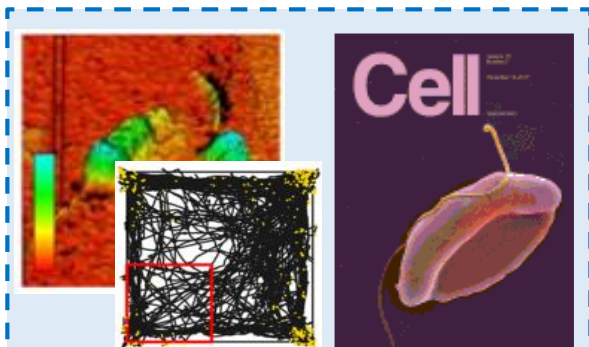
动物的**三维行为计算**是生命科学研究新工具

三维视觉的应用需求

近年来，动物三维行为计算受到广泛关注



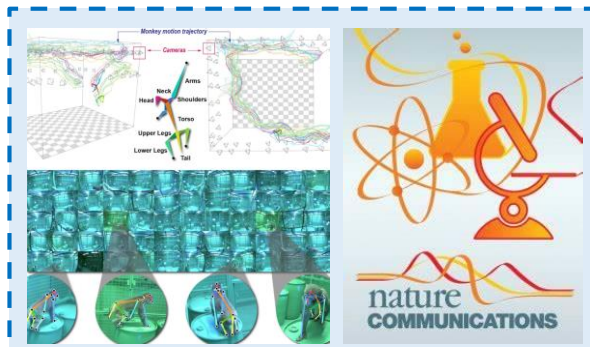
2018年：小鼠三维行为与大脑纹状体映射



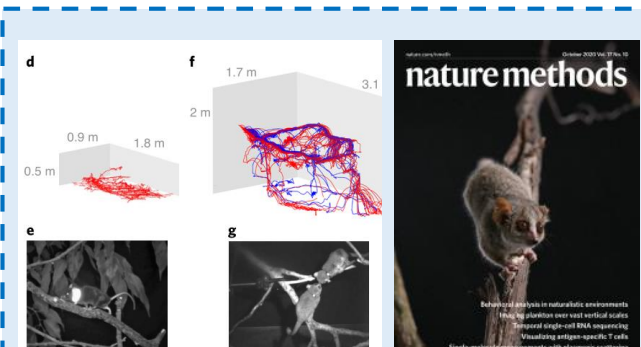
2019年：小鼠三维姿态与自闭症行为表现



2019年（封面文章）：猎豹三维姿态与物种保护



2020年：恒河猴三维姿态与三维聚类



2020年（封面文章）：尖嘴猴三维跟踪与大脑地图



2020年：动物表情行为与神经理解

未来

更高维 形态建模

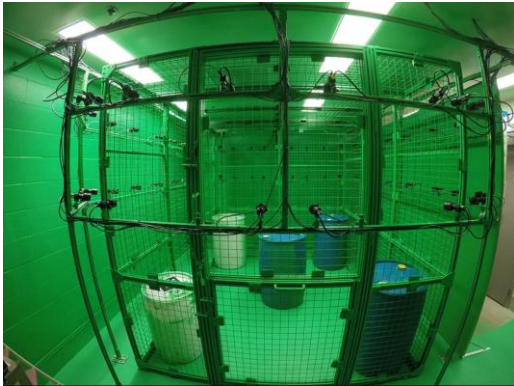
更丰富 交互运动

更深层 科学发现

三维视觉的应用需求

动物三维行为计算的价值

三维实验场景



三维行为计算

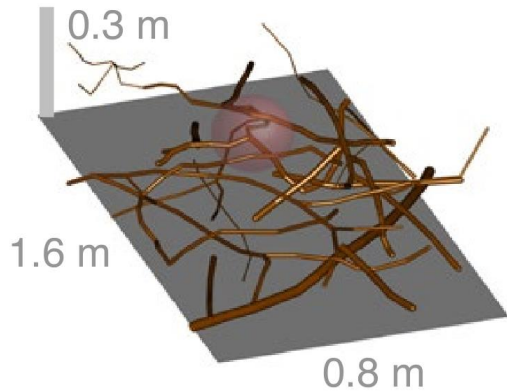
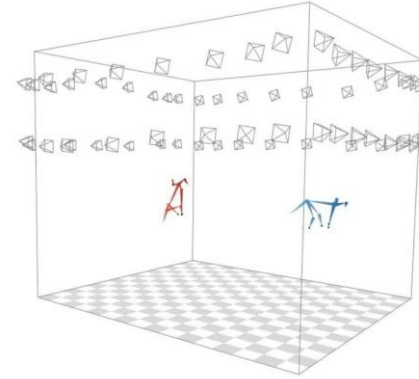


生命科学发现

明尼苏达大学

动物：恒河猴

结论：揭示灵长类动物的三维刻板行为的存在，以及三维交互行为的量化



瑞士日内瓦大学

动物：尖嘴猴/小鼠

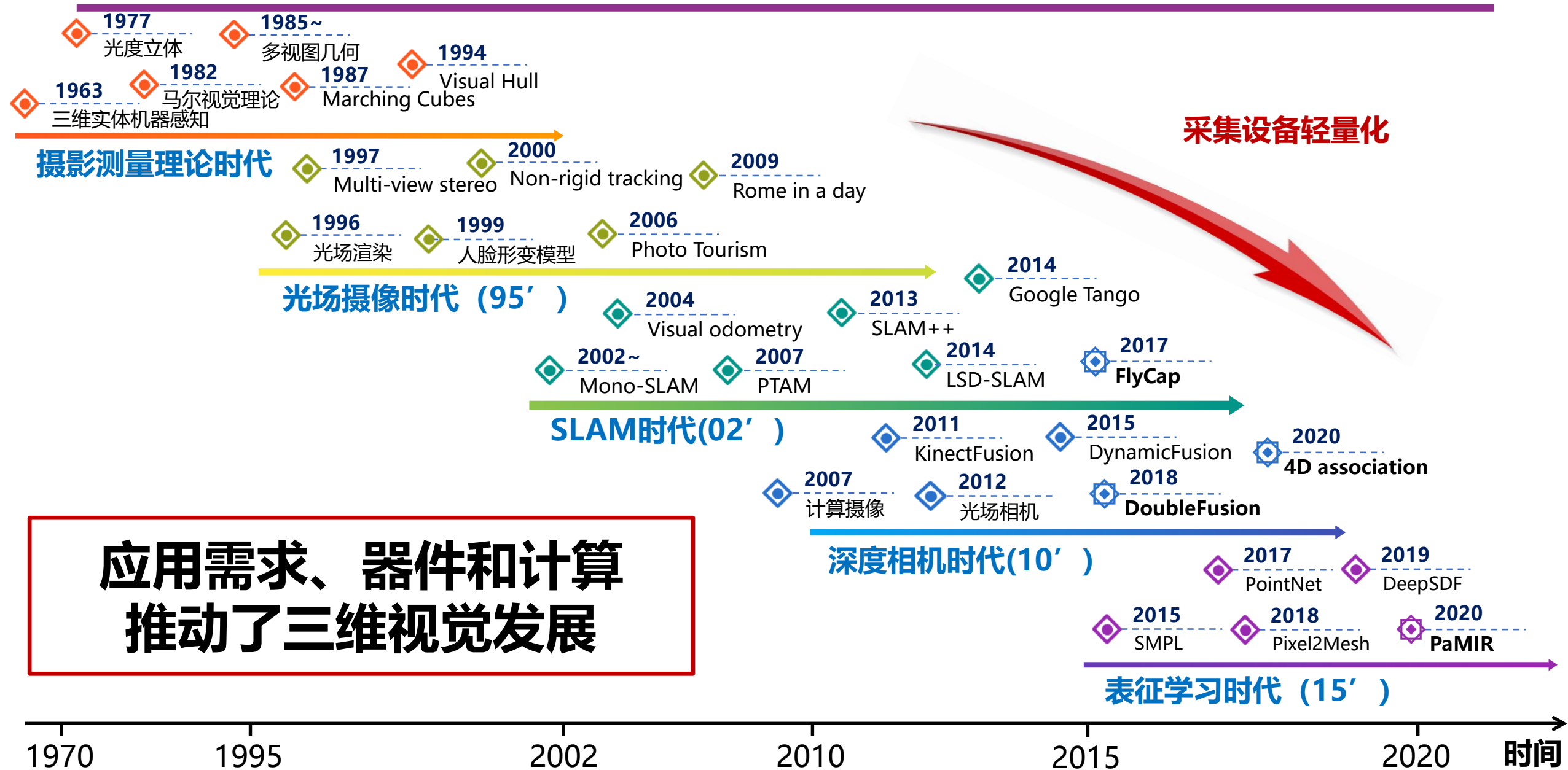
结论：在类自然环境下揭示了灵长类动物海马区有类似位置细胞的活动

三维行为计算新方法为生命科学研究带来新发现

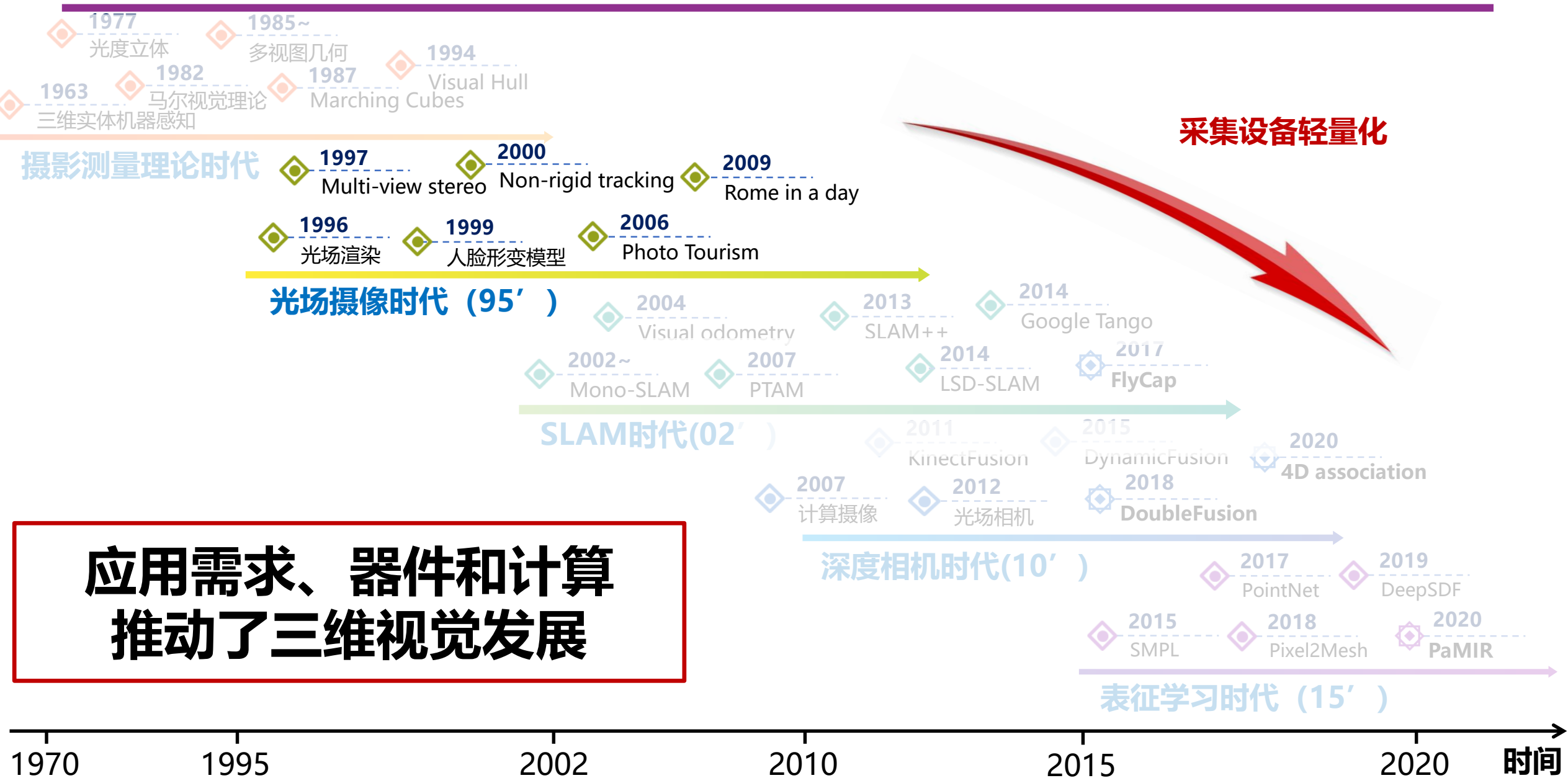
主要内容

- 计算机视觉与三维视觉
- 三维视觉的应用需求
- 三维视觉的发展历史
- 未来的三维视觉与计算机视觉

三维视觉的发展历史



三维视觉的发展历史

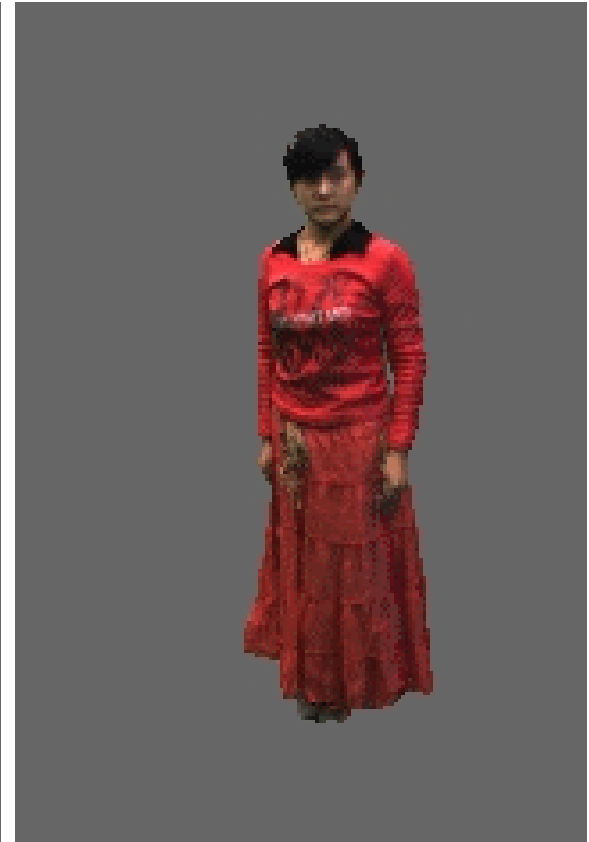
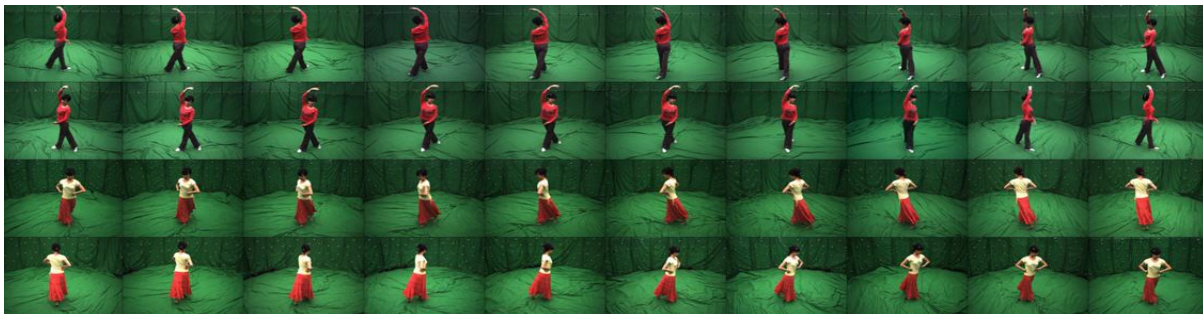


三维视觉：光场摄像时代

□ MVML Dome：密集光场三维数据获取的时代

2005年光场相机 2008年视角-光照协同采集装置

动态三维重建结果



构建40相机、680个光源组成的国际首个密集视点密集光照的采集装置

三维视觉：光场摄像时代

三维重建精度指标

Multi-View Stereo Evaluation • Datasets • Submit • Code

Acc. Threshold: 90%
Comp. Threshold: 1.25 mm

Data in new window Open Data Window
Data: View 1 and Ground Truth Image Size: Small

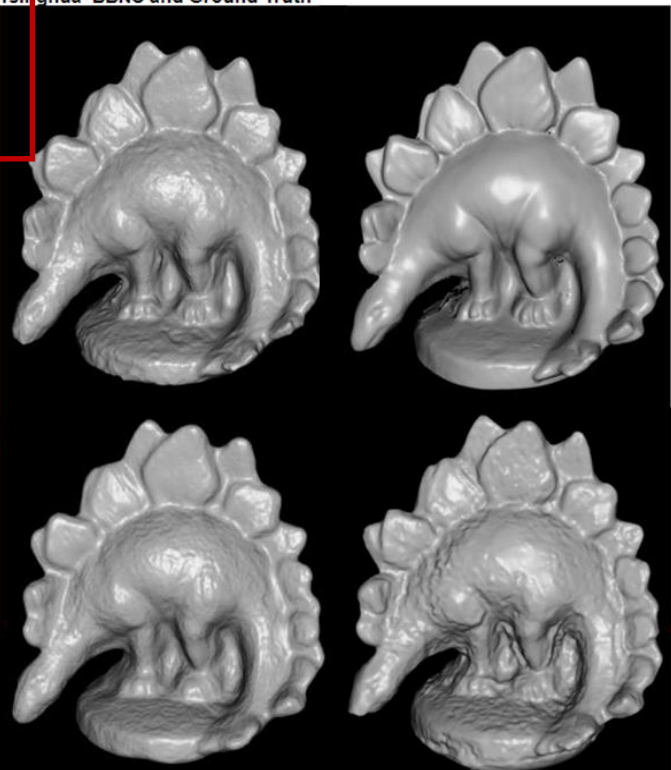
Tip: Mousing over any portion of a method's row will show its reference

Reference: Z. Xue, X. Cao, Y. Liu, Q. Dai, and N. Zhang. Continuous depth-map estimation with explicit occlusions for multi-view stereo. Submitted to CVIU 2011.

Normalized Time (min.s): 0:03:32

Sort By	Acc [mm]	Comp [%]
Tsinghua_BBNC	0.3	99.1
Furukawa 3	0.37	99.2
Bradley	0.38	94.7
Furukawa 2	0.42	99.2
Depth Fusion	0.42	97.8
Zaharescu	0.45	99.2
Kan Li	0.47	97.1
Kolev3	0.48	98.6
Deng	0.49	97.8
Liu2	0.51	98.7
Hongxing	0.52	98.4
Kolev2	0.53	98.3
Song	0.54	95.5
Goesele	0.56	26.0
Furukawa	0.58	96.9
Liu	0.59	98.3
Hernandez	0.6	98.5
Jancosek-3DIM09	0.66	74.9
SurfEvolution	0.66	97.6
Chang	0.66	89.9
Guillemaut	0.68	98.0
Pons	0.71	97.7
Auclair	0.74	96.8
Gargallo	0.76	90.7
Delaunoy	0.89	93.9
Ladikos	0.89	95.0
Starck	1.01	90.7
Mouaddib	1.01	91.8
NIPS_829	1.07	91.0
Vogiatzis	1.18	90.8
Tran	1.26	89.3
Strecha	1.41	91.5

Tsinghua_BBNC and Ground Truth



Sort By	Acc [mm]	Comp [%]
Tsinghua_BBNC	0.3	99.1
Furukawa 3	0.37	99.2
Bradley	0.38	94.7
Furukawa 2	0.42	99.2

本项目所提算法

美国UIUC算法

加拿大UBC算法

美国UIUC算法

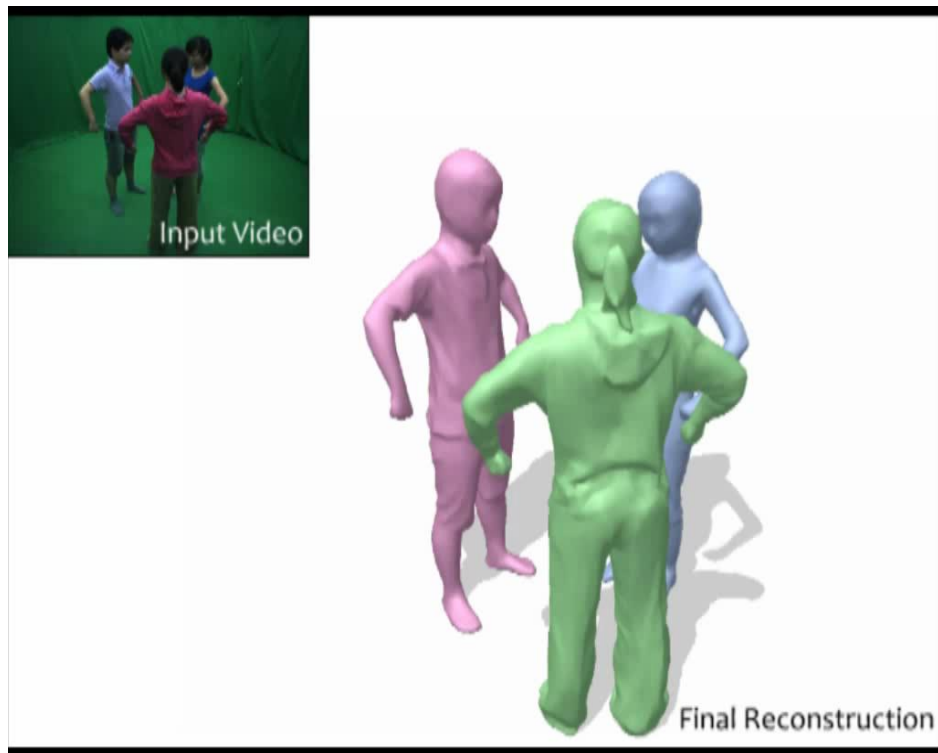
多视点相机三维重建在由美国自然科学基金会等机构组织的立体重建公开评测

精度连续7年排名第一

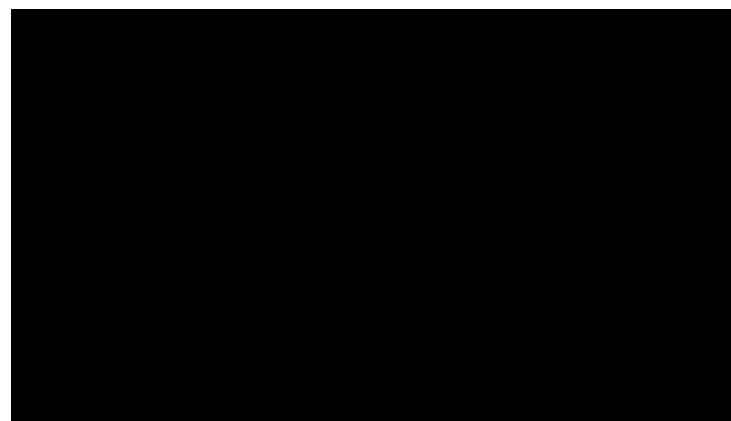
重建精度误差比第二名降低23%

三维视觉：光场摄像时代

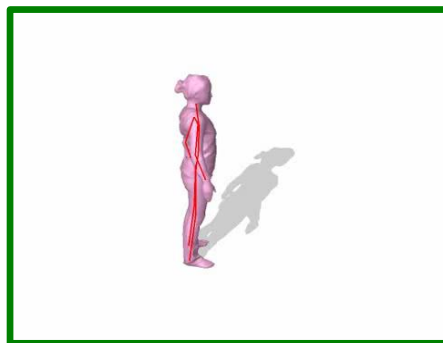
□ 多人体、变光照及可编辑动态三维重建



多动态对象重建(PAMI 2013)



动态三维光照
映射
(SIGGRAPH
2015)



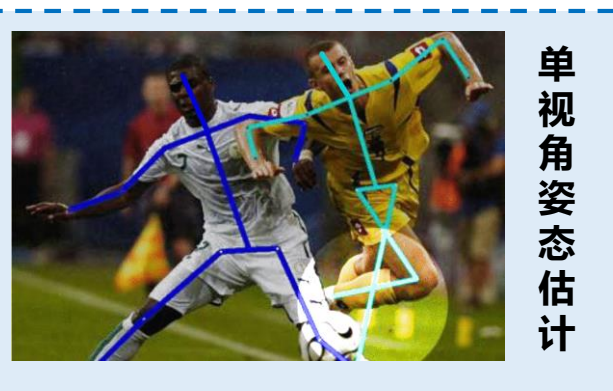
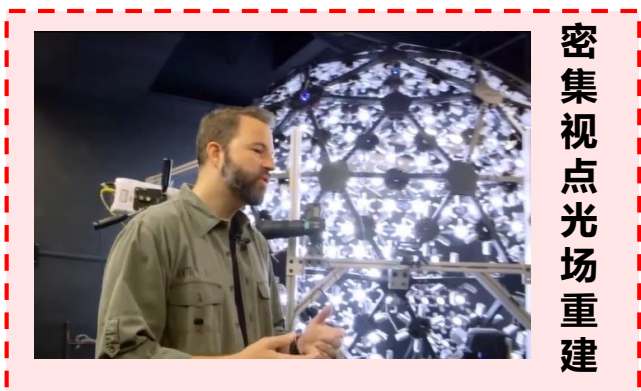
动态三维重建(CVPR 2011 Oral) 动态三维运动映射(SIGGRAPH2011)

率先实现多人动态三维重建、全自动运动映射与光照映射，发表SIGGRAPH、TPAMI、CVPR Oral等系列文章20余篇

三维视觉：光场摄像时代

应用背景

➤ 无标记动捕不仅赋能了更多应用场景，更是未来交互方式的全新变革



矛盾

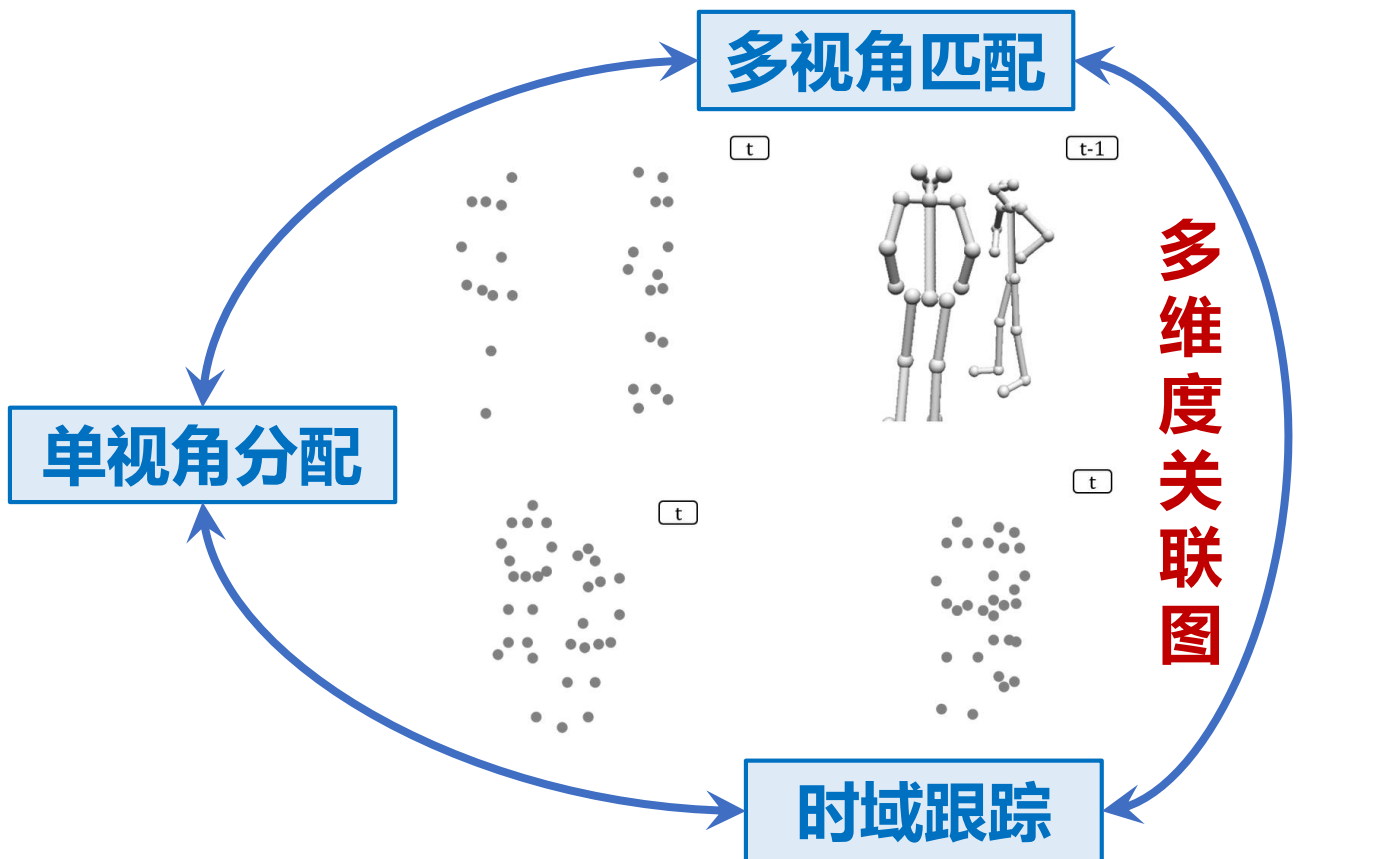
稠密视点系统运算繁重，场景受限

单视点方法无法分辨深度歧义性

核心问题：如何充分挖掘各视点信息，优化匹配策略，减少视角冗余

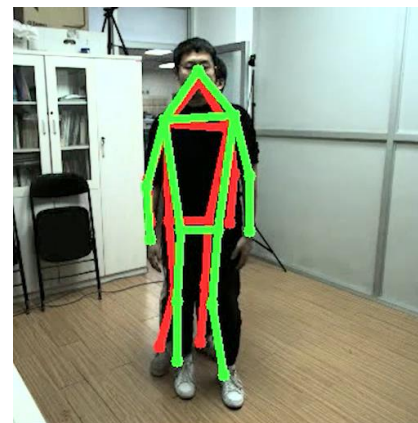
三维视觉：光场摄像时代

高维数据的视角-空间-时间解耦

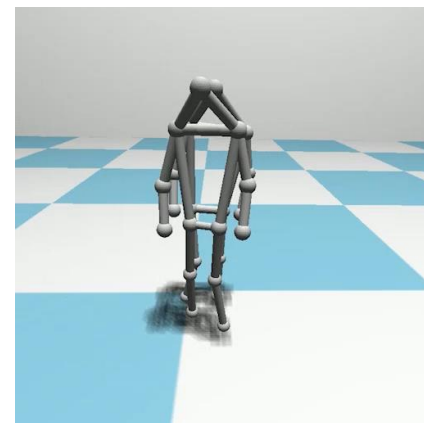


$$E(\mathcal{Z}) = w_p \sum q(z_{ij}^{mn}(c)) + w_m \sum q(z_{ii}^{mn}(c_p, c_q)) + w_t \sum q(z_i^{mk}(c))$$

4D多视时空聚类解耦模型及其实时优化方法

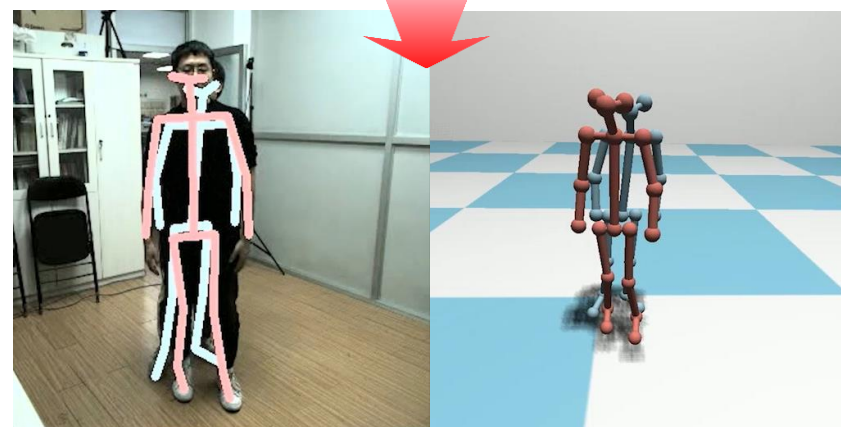


单视角划分错误



导致重建质量差

序贯式：孤立逐维进行匹配搜索

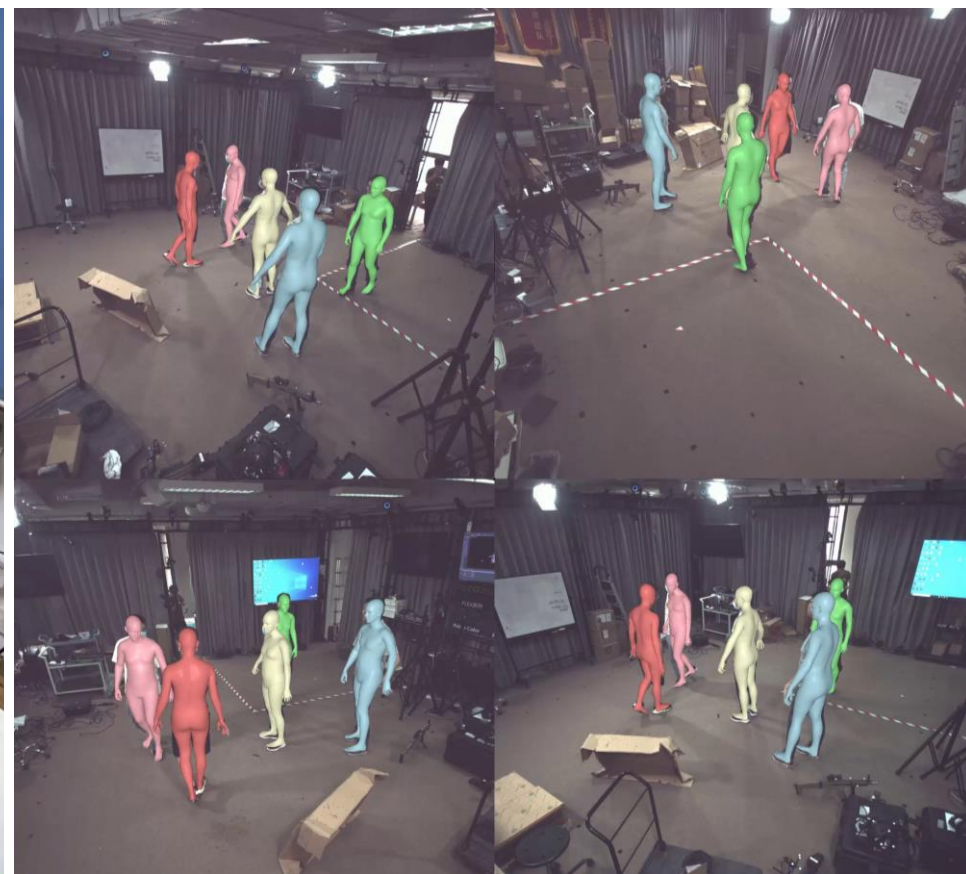


协同优化提升性能

协同式：闭环联合优化各维信息

三维视觉：光场摄像时代

□ 实时多人运动捕捉：4D多视时空聚类解耦



Zhang et al. 4D Association Graph for Realtime Multi-person Motion Capture Using Multiple Video Cameras, CVPR Oral 2020, 可应用于**数字人虚拟主播**、**细粒度行为监测**、**智慧体育**等

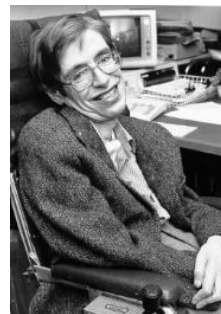
三维视觉：光场摄影时代

小型猪三维行为计算与渐冻人症

ALS (Amyotrophic lateral sclerosis, 肌萎缩性脊髓侧索硬化症), 俗称“渐冻人症”, 是典型的神经退行性疾病, 五种最常见的运动神经元疾病之一。目前, 中医西医对此病都无能为力。ALS的发病与多种基因突变有关, 例如SOD1, FUS, TDP-43等。

目前, 早期行为学症状仍是空白

霍金, 著名物理学家



纽约洋基队传奇一垒 Lou Gehrig



渐冻人症患者

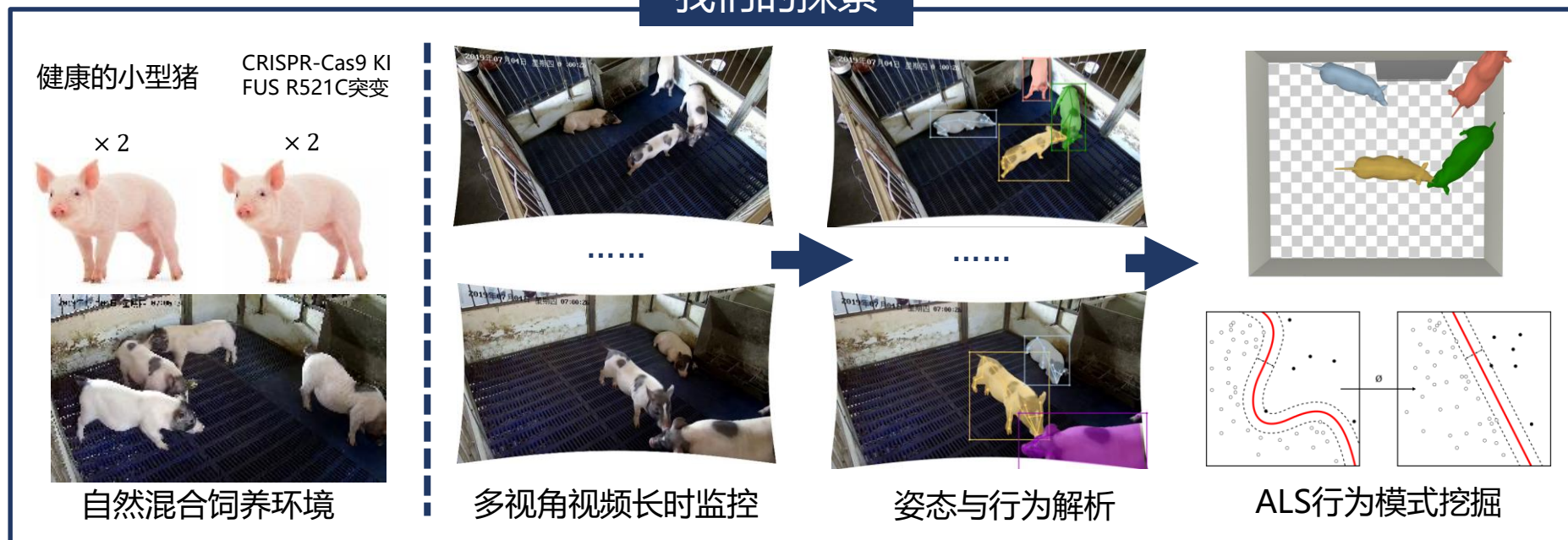
当前的问题

(1) 小型动物内源性突变难以还原疾病

(2) 缺少对疾病早期表现的长时准确观测

(3) 人工设置环境带来动物行为的“偏见”

我们的探索

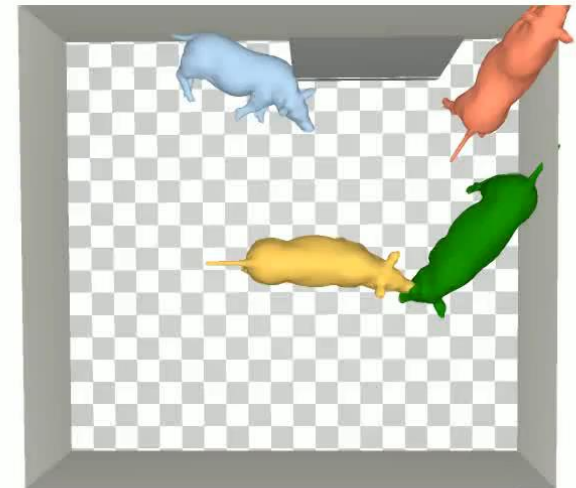
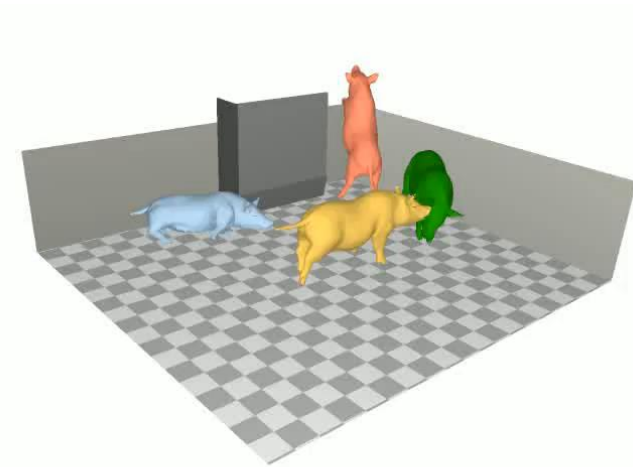
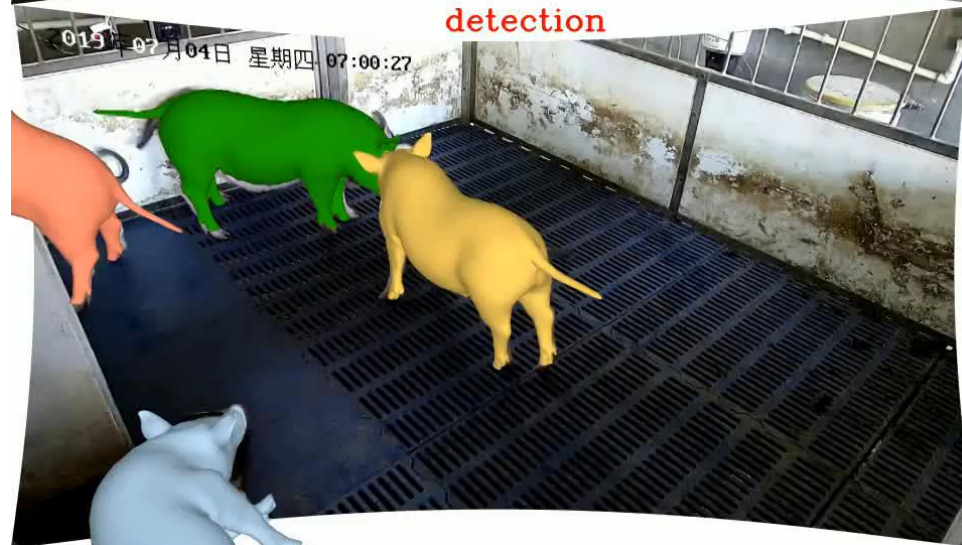
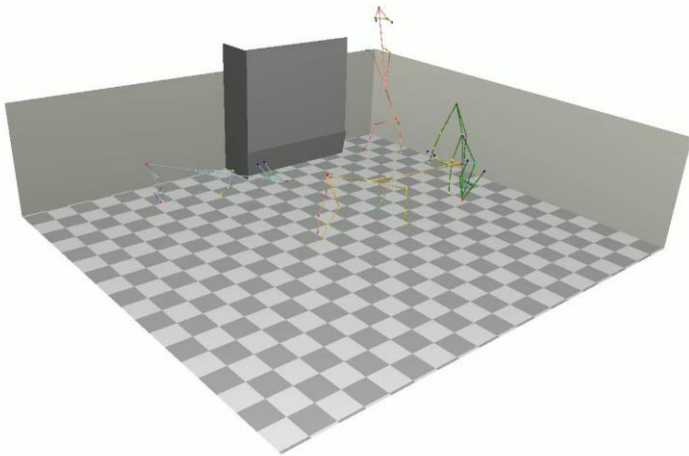


三维视觉：光场摄影时代

三维行为计算探索渐冻症早期症状

三维表面运动

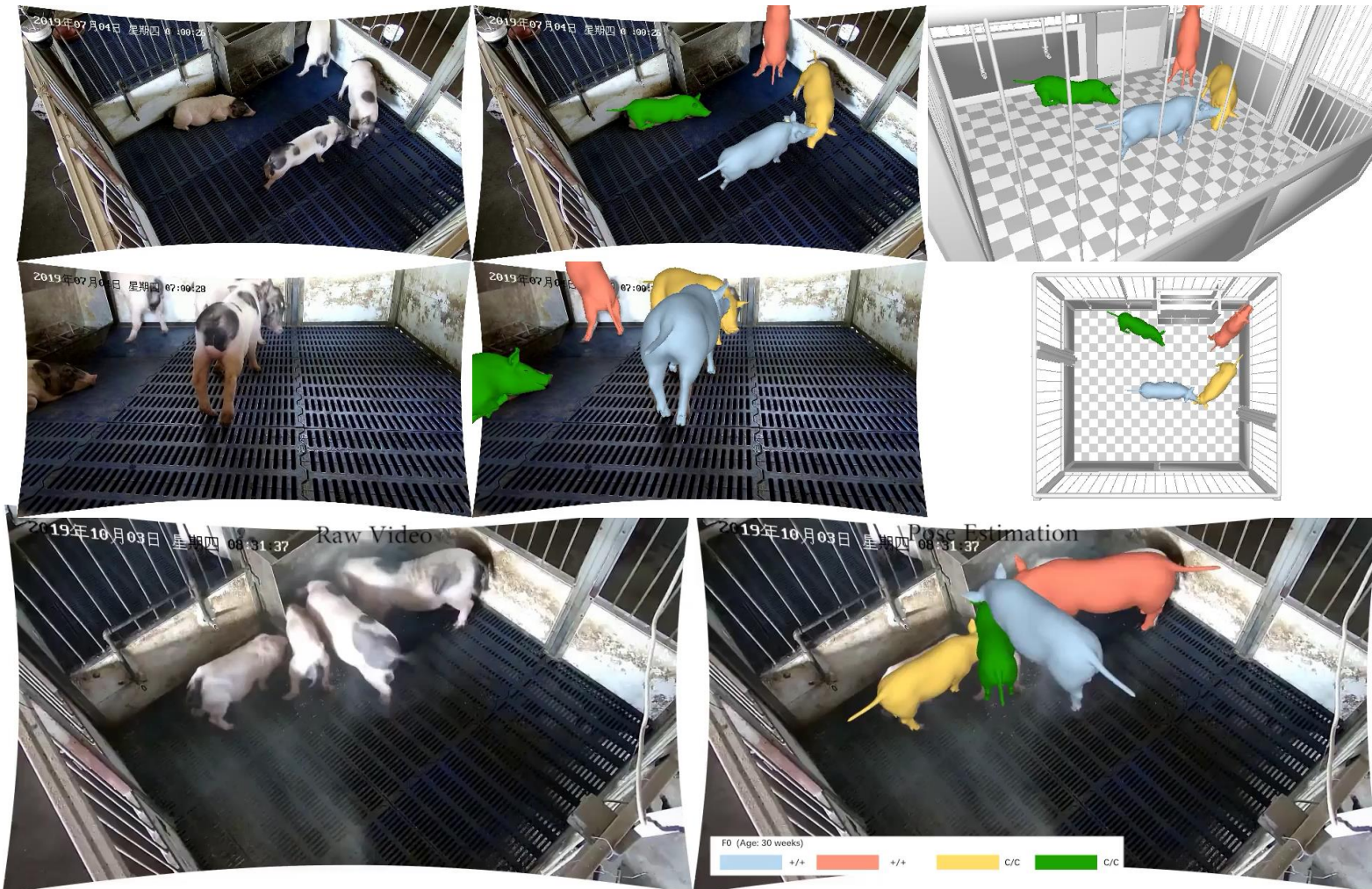
三维场景与轨迹



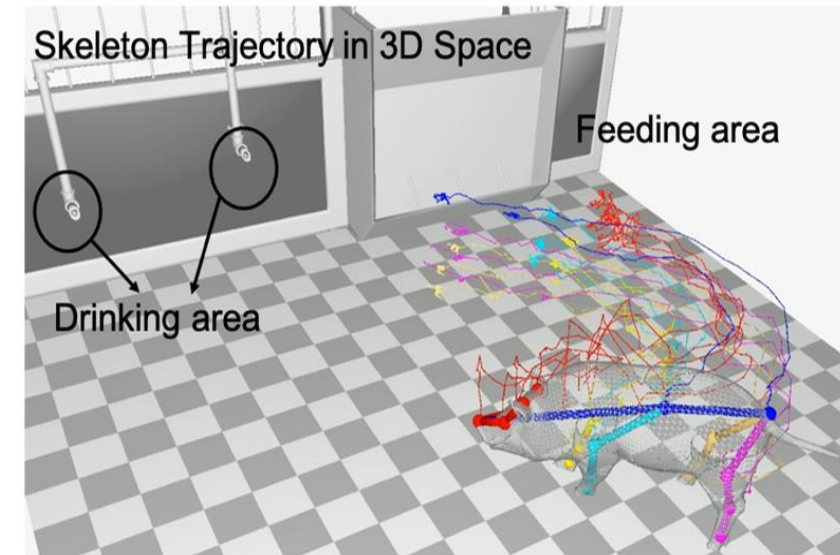
三维复杂行为计算助力神经退行性疾病的诊断

三维视觉：光场摄影时代

三维行为计算探索渐冻症早期症状



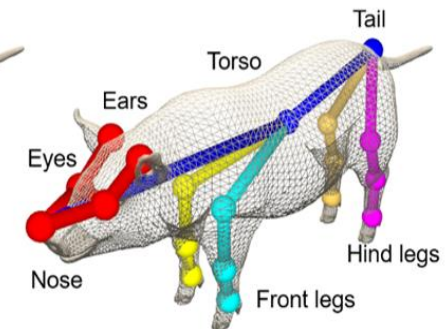
三维表面运动



Surface Model



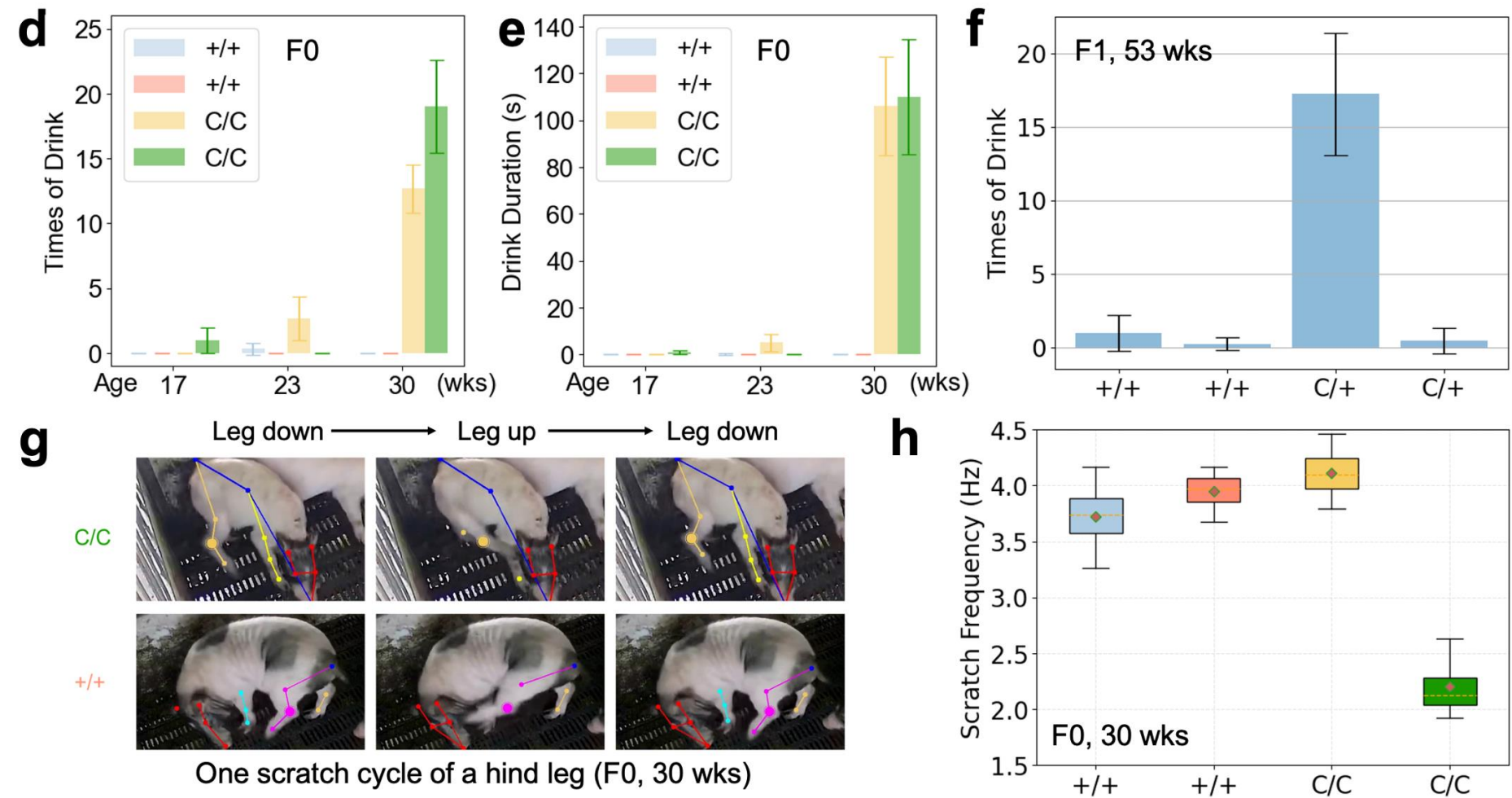
Skeleton



4只广西巴马小猪，2只健康，2只经基因突变，25个监控相机，时长1000小时

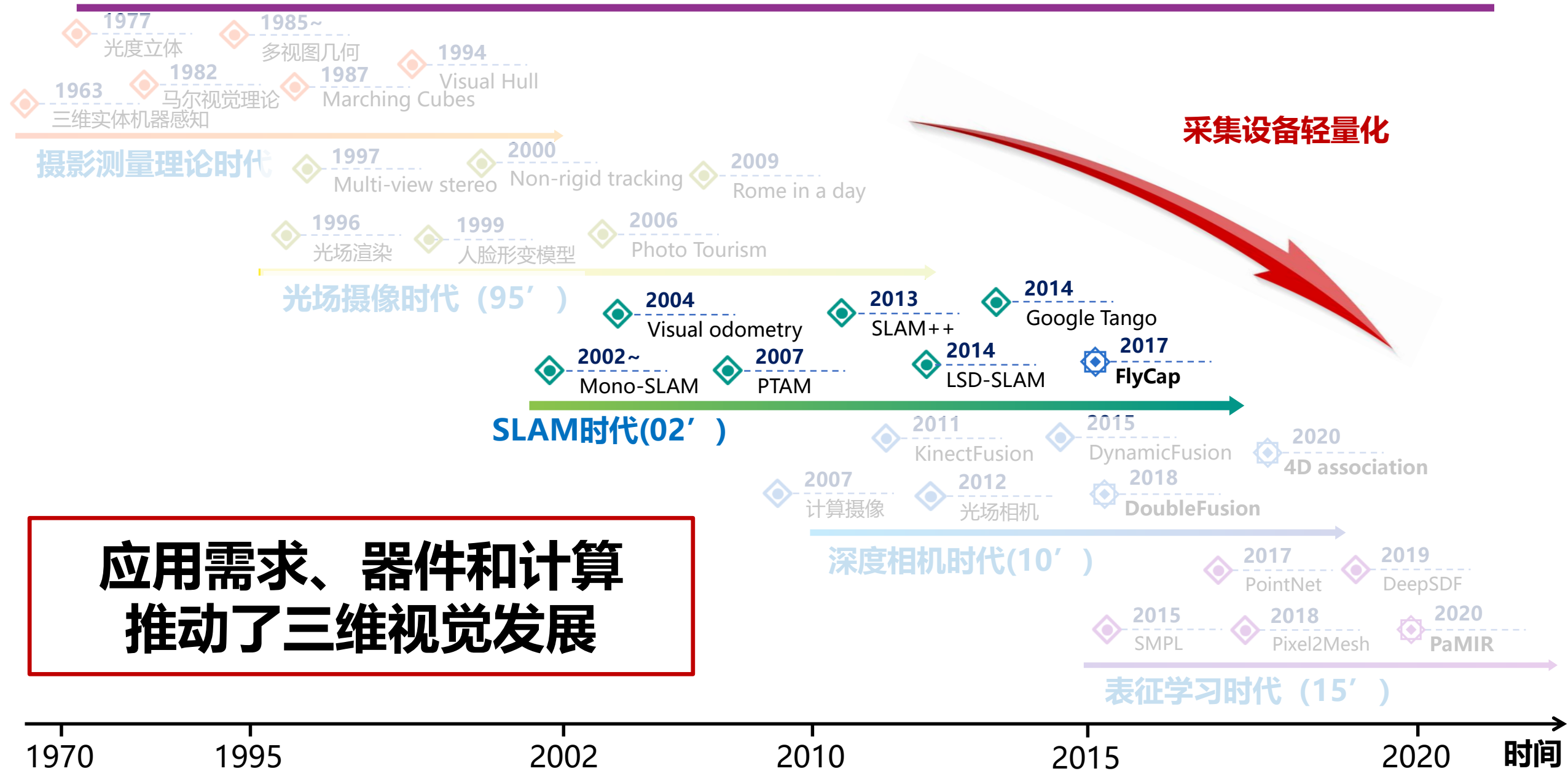
三维视觉：光场摄影时代

三维行为模式挖掘与分析



已分析的行为：喝水频率、喝水时长、侧卧挠腿频率

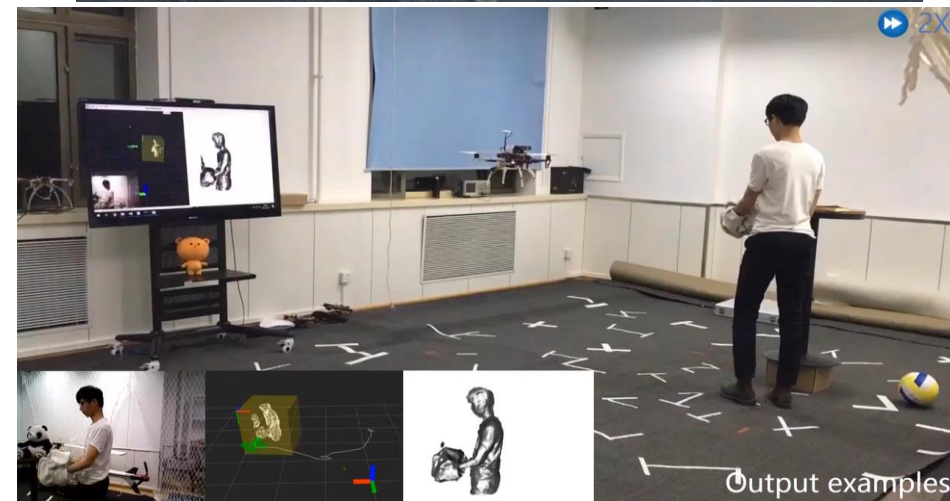
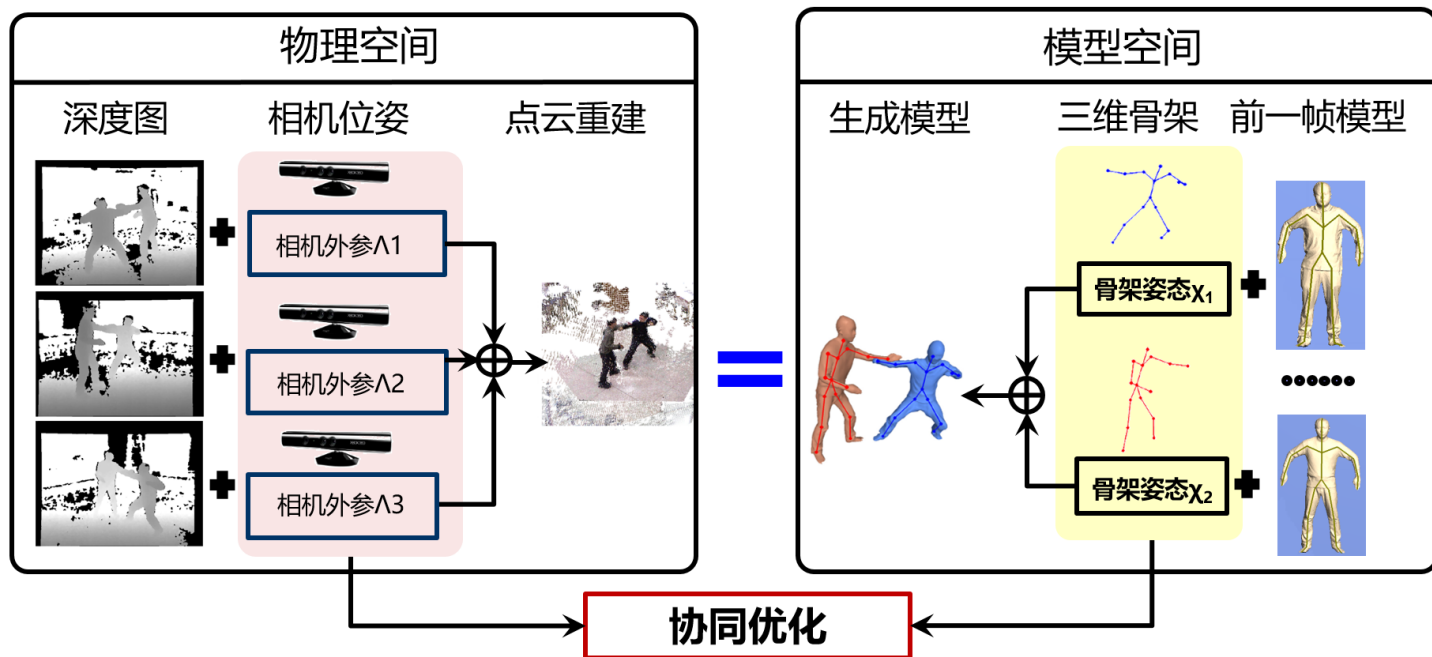
三维视觉的发展历史



三维视觉：相机定位时代

多手持或飞行相机动态重建方法

$$\arg \min_{X, \Lambda} \left\{ \sum_{(p, v) \in Z_{pv}} \frac{\|p(\Lambda) - v(X)\|^2}{\|Z_{pv}\|} + \sum_{(p, v) \in Z_{pg}} \frac{\|p(\Lambda) - v\|^2}{\|Z_{pg}\|} + \sum_{(p, p') \in Z_s} \frac{\|p(\Lambda) - p'\|^2}{\|Z_s\|} \right\}$$



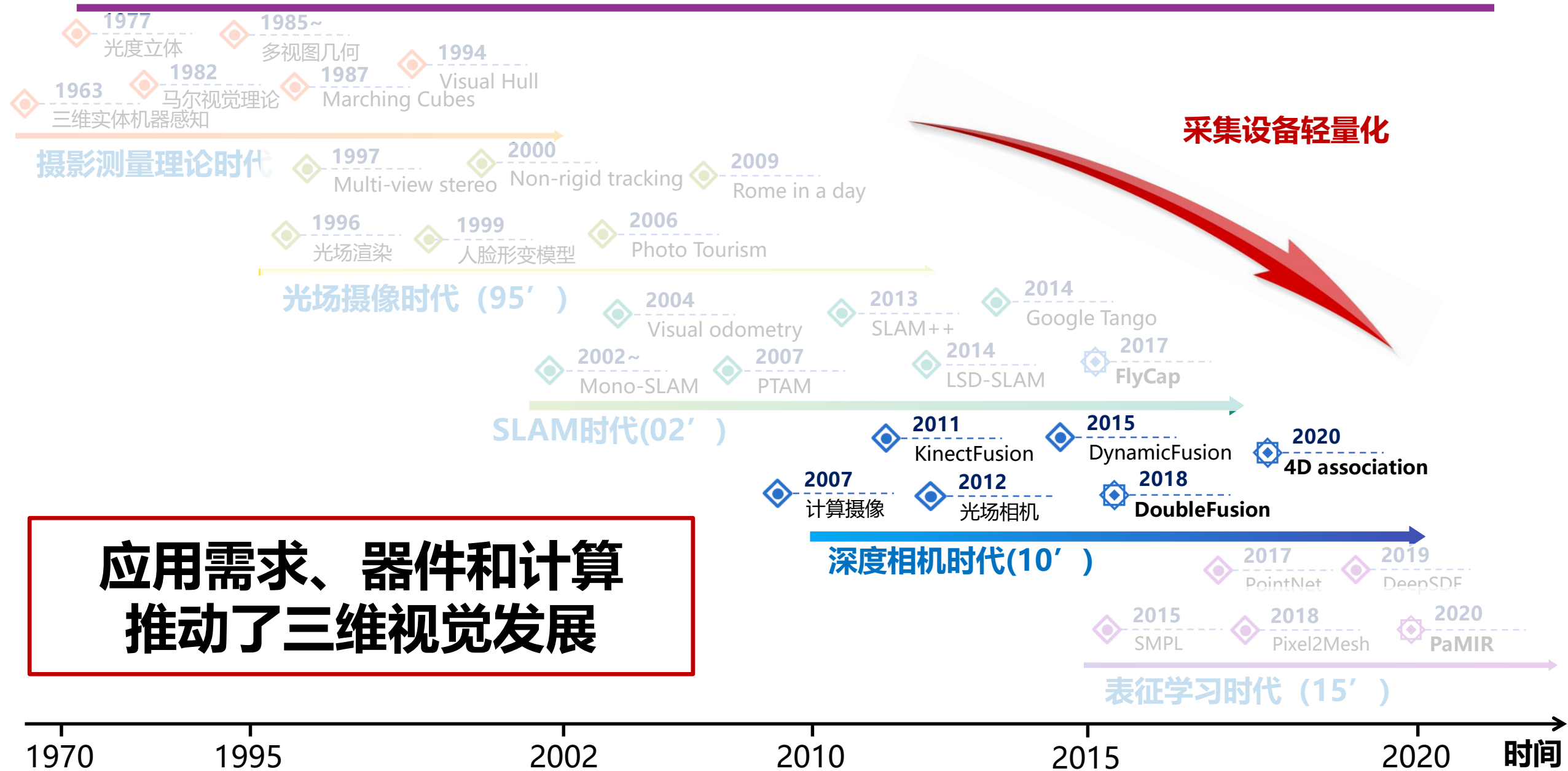
Xu et al. **UnstructuredFusion**: Realtime 4D Geometry and Texture Reconstruction using Commercial RGBD Cameras, IEEE **TPAMI** 2020

Xu et al. **FlyFusion**: Realtime Dynamic Scene Reconstruction Using a Flying Depth Camera, IEEE **TVCG** 2021

Xu et al. **FlyCap**: Markerless Motion Capture Using Multiple Autonomous Flying Cameras, IEEE **TVCG** 2018

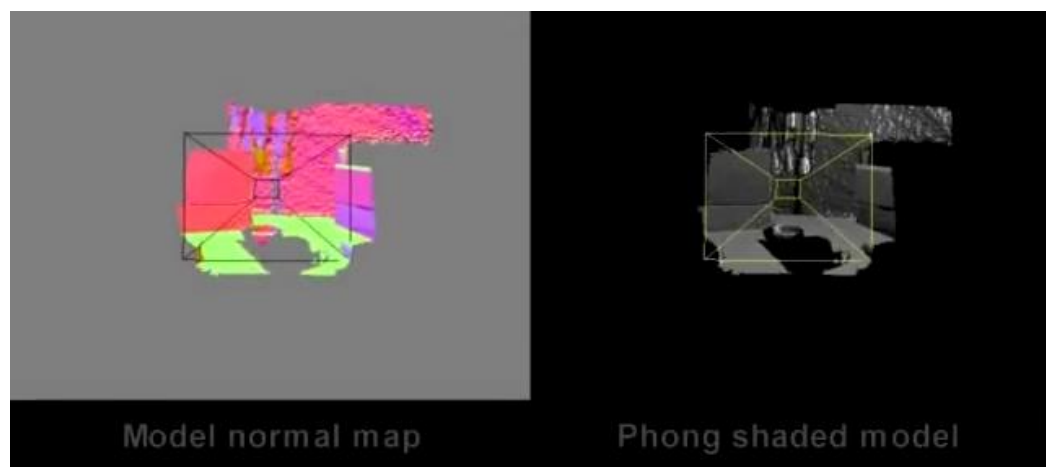
首次实现面向动态场景的三维重建、相机协同定位和自适应视点优化

三维视觉的发展历史



三维视觉：深度相机时代

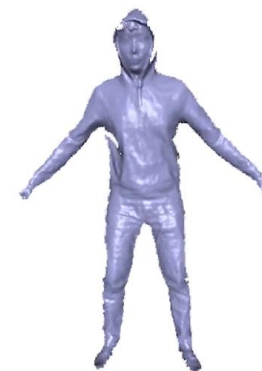
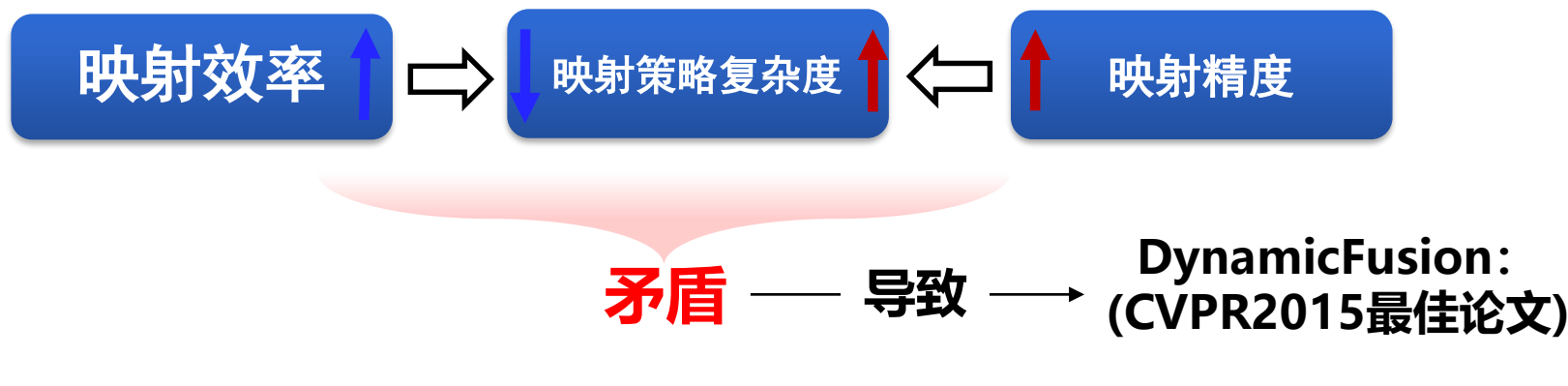
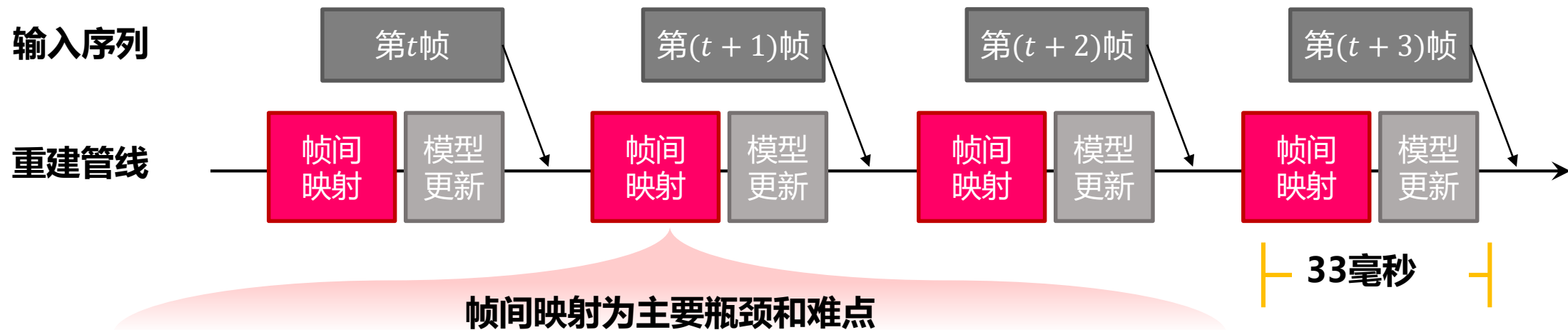
□ 开启轻量级采集-实时三维数据获取时代



已有重建方法仅实现时空局部映射，难以重建复杂运动

三维视觉：深度相机时代

□ 动态场景实时融合重建



核心问题：不完整融合表面之间的实时、精准非线性映射问题

三维视觉：深度相机时代

□ 双表面协同动态重建方法

已有国际前沿：**单表面重建**

KinectFusion
(2011, 引用量3873)

Fusion4D
(2016, 微软)

DynamicFusion
(2015, CVPR最佳论文)

Motion2Fusion
(2016, 谷歌)

提出方法：**双表面协同重建**

DoubleFusion
(CVPR 2018 Oral)

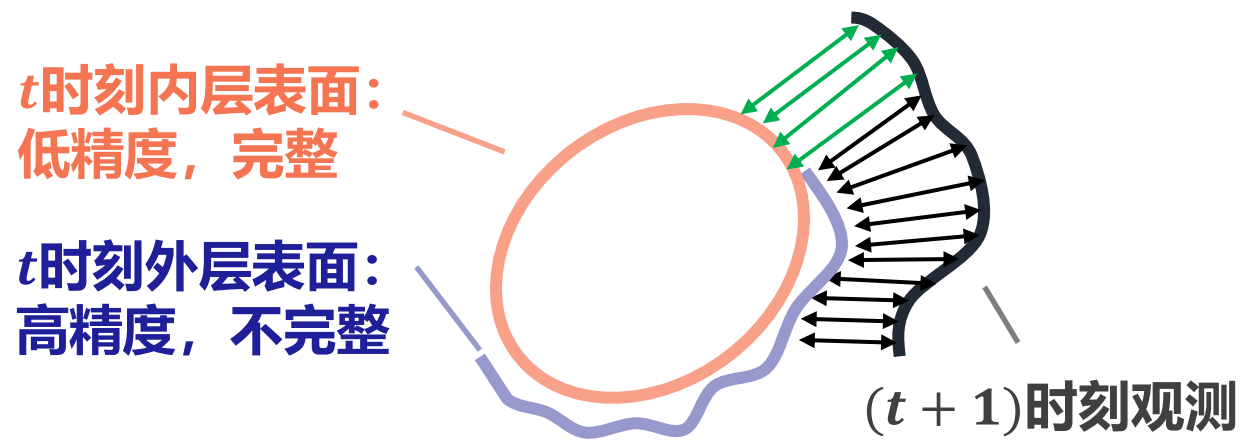
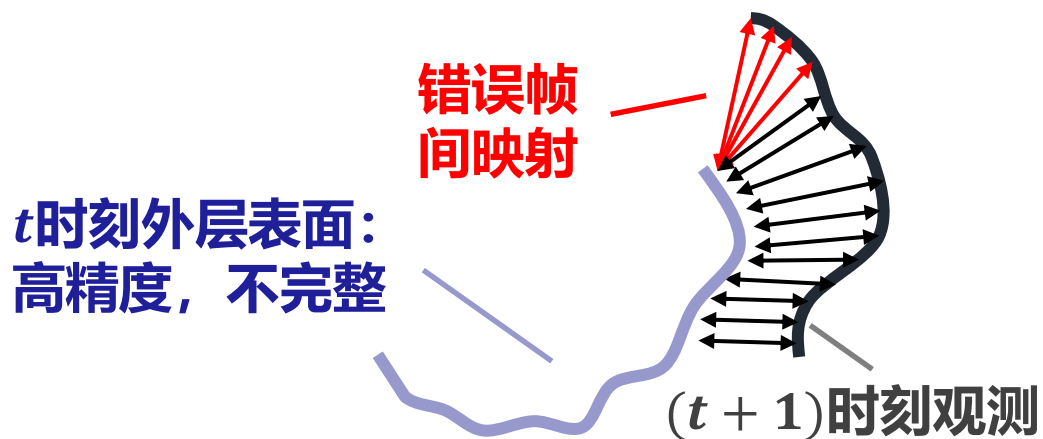
BodyFusion
(ICCV 2019)

3DPortrait
(CVPR 2020 Oral)

HybridFusion
(ECCV 2018)

UnstructFusion
(PAMI 2019)

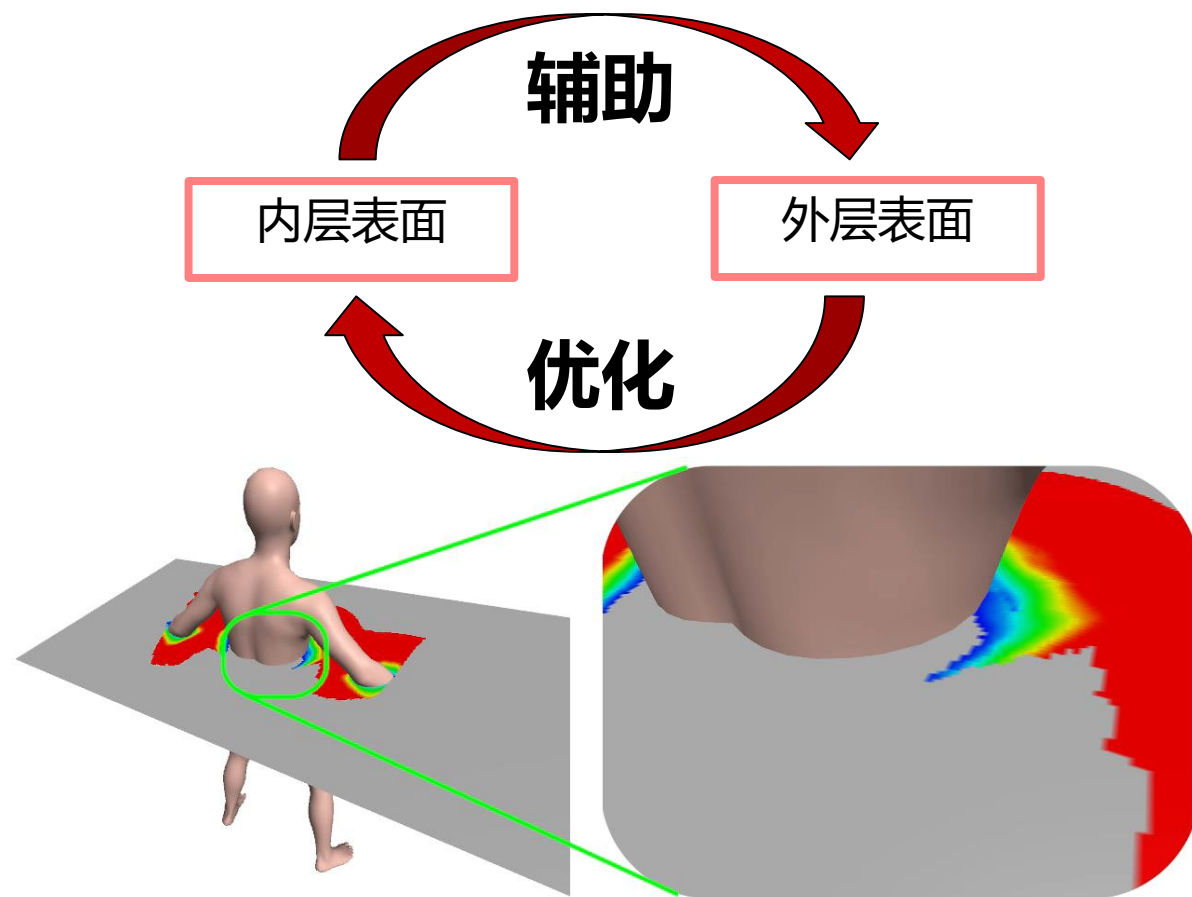
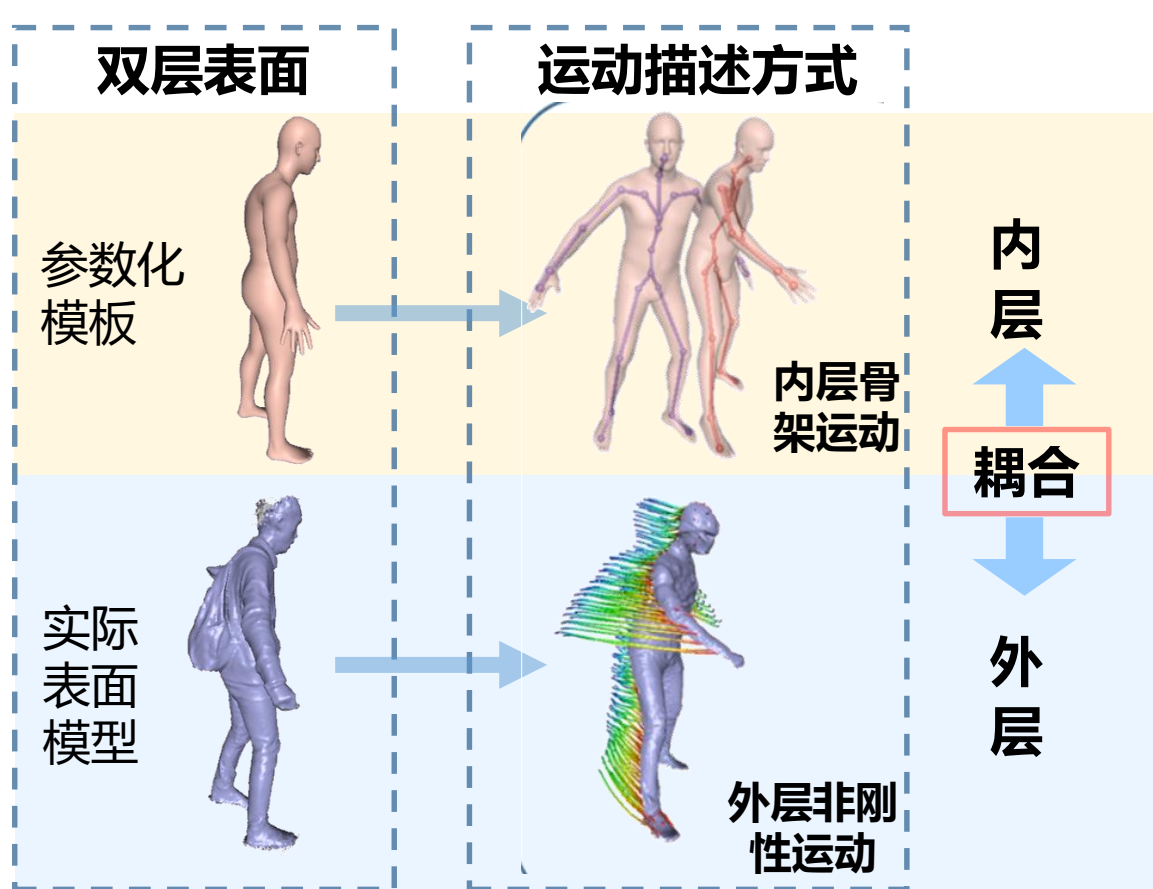
Function4D
(CVPR 2021 Oral)



建立了双层表面动态三维重建技术体系，支撑实时精准的帧间非线性映射及动态融合重建

三维视觉：深度相机时代

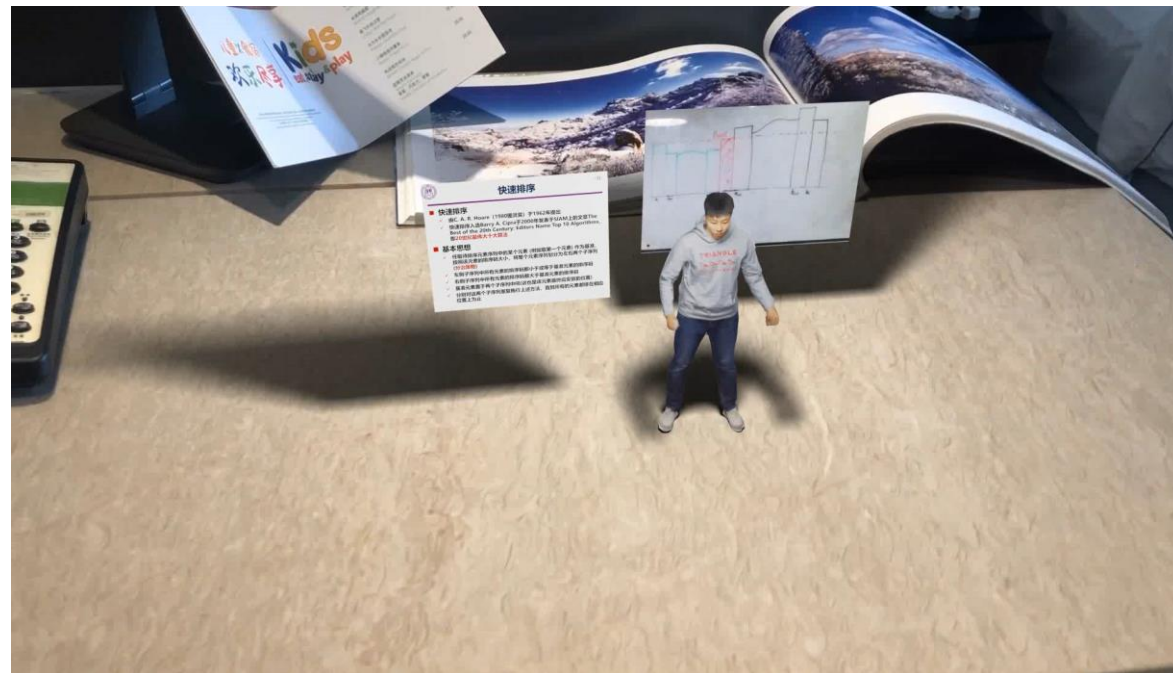
□ 双表面协同动态重建方法



利用人体动态三维信息的**多层语义化**特性，建立双层表面实现**高效时空全局映射**

三维视觉：深度相机时代

□ 技术突破1：提出了双表面协同动态重建方法



Yu et al. DoubleFusion: Real-time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor, **CVPR Oral**, 2018
Yu et al. DoubleFusion: Real-time Capture of Human Performances with Inner Body Shapes from a Single Depth Sensor, **IEEE TPAMI**, 2020

国际首个单深度相机的实时人体动态三维重建方法，可提供全息AR远程教学应用

三维视觉：深度相机时代

TC

Startups
Apps
Gadgets
Events
Videos
—
Crunchbase
More
Search:
Samsung
Tesla
Apple
Disrupt SF 2018

Login / Sign up

腾讯网 要闻 推荐 财经 娱乐 体育 时尚

What's under those clothes? This system tracks body shapes in real time

Devin Coldwey @techcrunch / Jun 19, 2018

Comment



双重融合 (DoubleFusion) 揭示一个人衣服底下的身体的形状和位置

2018-06-22 10:09:52 来源：网易科技报道

参与回复

微信

QQ空间

据国外科技博客TechCrunch报道，随着增强现实技术日益火红，深度跟踪摄像头即将出现在旗舰级手机上，现在是时候改进计算机跟踪它们所看到的人的运动轨迹的方式了——即使这意味着要“剥光”他们的衣服。有一种新的计算机视觉系统可以做到这一点，听起来可能有点让人毛骨悚然，但它肯定有它的用处。

凤凰网 科技 凤凰网科技 > 产品 > 正文

实时人体形状追踪系统：隔着衣服看穿你

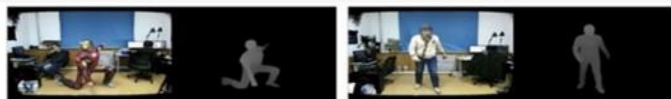
2018年06月22日 15:03:35

来源：201中关村在线

人参与 评论

原标题：实时人体形状追踪系统：隔着衣服看穿你

日前，有一套名为“DoubleFusion”的实时人体形状追踪系统出现在人们的视野中，这套系统最强的地方在于可以实时追踪到人的衣服下面都有什么，也就是说你在它的面前将毫无遮挡。



动点科技

硬件创业者团队招募中！2018年马上报名参赛，奖金丰厚！7月15-18日，北上广深杭，5天，0门槛！

首页 视频 初创公司报道 涂鉴 区块链 NodeBang 科技快讯 更多 专题

结合深度数据和骨架，清华大学研究单深度相机实时动捕技术 DoubleFusion

2018/06/20



结合深度数据和骨架，清华大学研究单深度相机实时动捕技术 DoubleFusion

编辑 颜映华 · 分类 快讯 / 论文 / 输入 · 2018年06月19日 15:32:59

文章相关引用及参考：techcrunch

你知道为什么在动捕设置中，演员必须穿上紧身衣吗？

（映维网 2018年06月19日）随着增强现实技术的日渐流行和深度跟踪摄像头开始登陆各款旗舰手机，提高计算机对人体运动追踪的时机已经成熟。现在映维网了解到，一项全新的计算机视觉系统甚至可以隔着衣服抽离人体运动。

一个基本的问题是，如果你要在电影或AR游戏中捕捉运动中的人体，衣服所造成的模糊感将会令人感到沮丧。你知道为什么在动捕设置中，演员必须穿上紧身衣吗？因为只有这样系统才能更准确地识别各肢体的位置。

即便是相当家身的牛仔裤，这都会影响计算机对身体位置的捕捉。



腾讯云 资讯 快讯 区块链 政府双创 投融资 活动 她说 发现



2018-06-20 07:30:00
逆天的双融合技术：若非裸奔，不然无法逃脱它的数字眼

DoubleFusion：深度和骨架结合的解决方案

— 2018 —

06/23

23.19

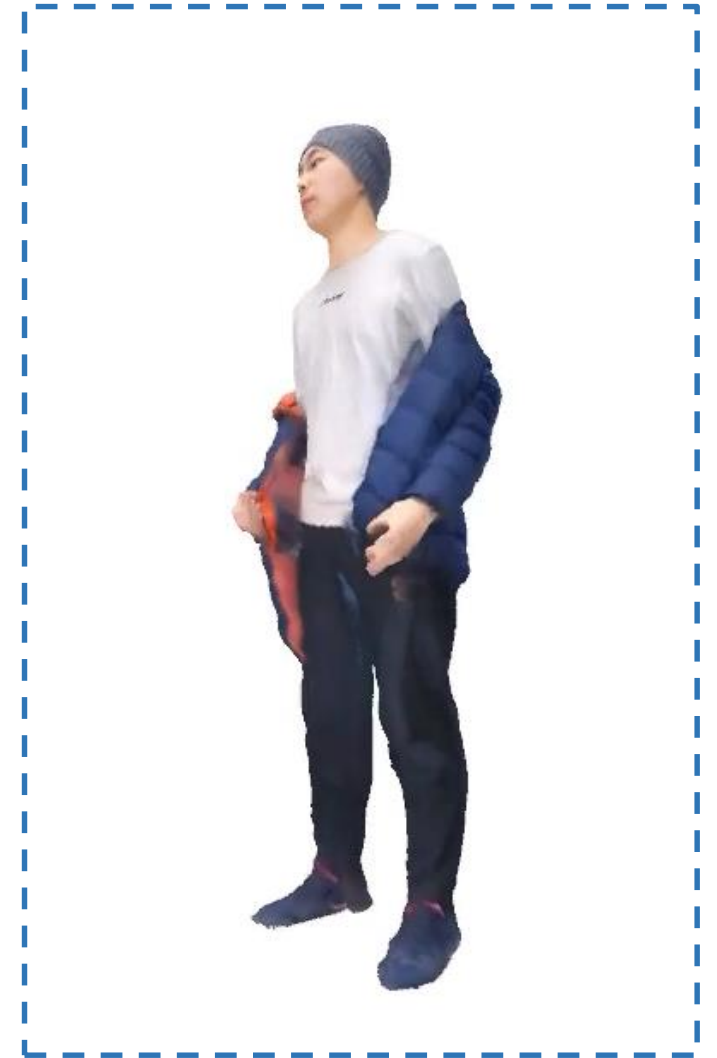
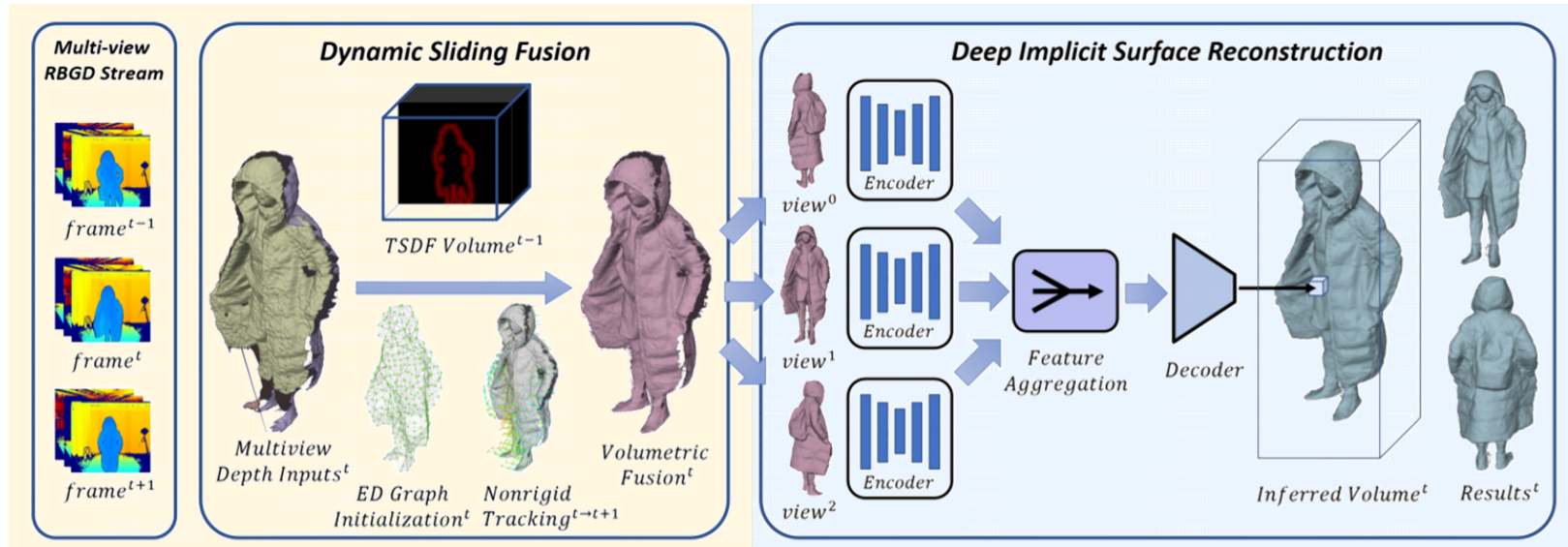
AR 版

我们已经介绍过很多单目的动作捕捉方案，最近的单目动捕方案可以说大同小异，基本没有什么区别，都是利用卷积神经网络识别对象，估算骨骼模型，再在此基础上进行这些解决方案的困难也都类似，例如老大难的遮挡问题，脚踝处的识别和骨骼模型往往等等。

全球十余家媒体报道本项目单深度相机实时三维重建DoubleFusion成果，包括美国互联网产业的风向标TechCrunch，国际知名的技术评论网站凤凰网等

三维视觉：深度相机时代

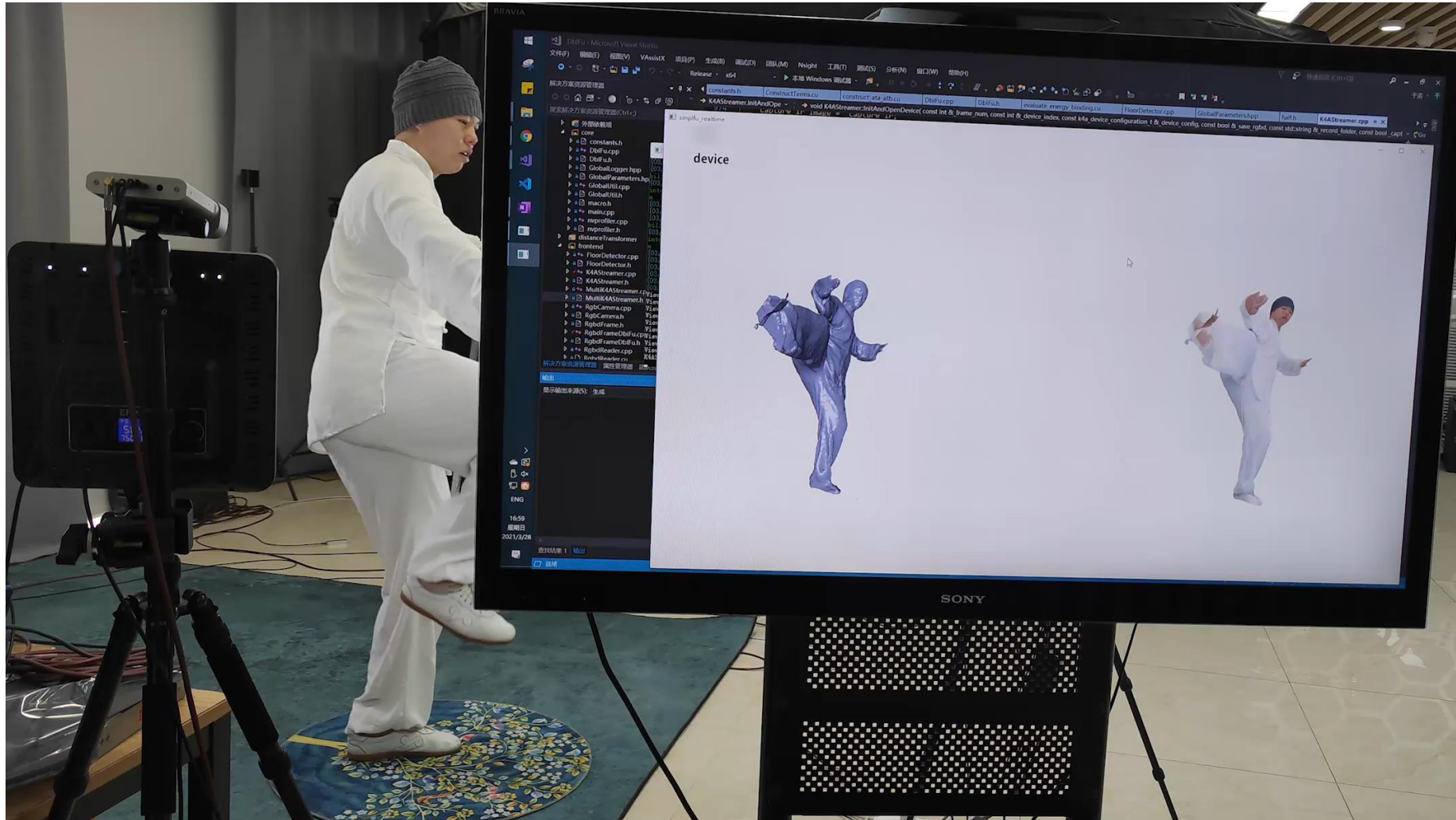
隐式函数表面融合方法



对任意服饰人体动态、人体物体交互、人体人体交互场景的扩展

Function4D: Real-time Human Volumetric Capture from Very Sparse Consumer RGBD Sensors, CVPR 2021 Oral

三维视觉：深度相机时代



Yu et al. Function4D: Real-time Human Volumetric Capture from Very Sparse Consumer RGBD Sensors, CVPR Oral, 2020
仅需4个商用深度相机，实现高质量实时三维重建，系统成本5万元以内

三维视觉：深度相机时代

□ 隐式函数表面融合方法



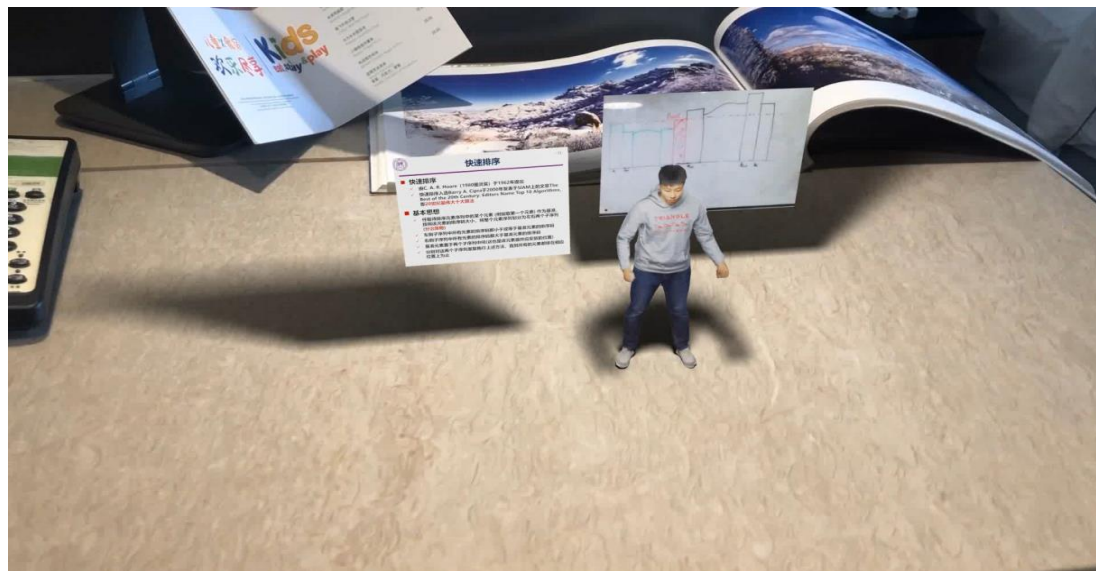
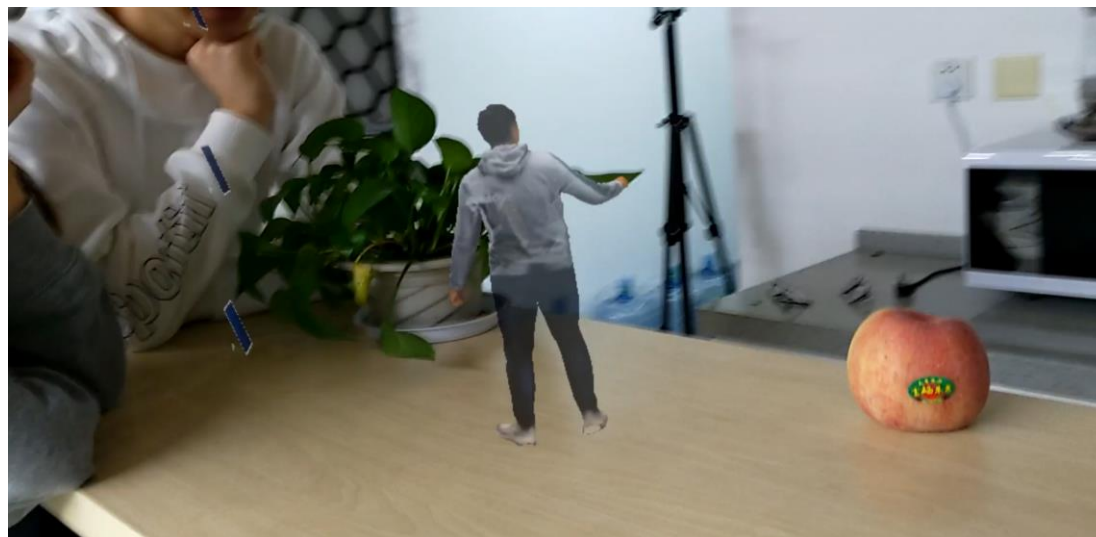
× 8

微软Holoportation 2016
8组共24个相机，成本逾百万



Yu et al. Function4D: Real-time Human Volumetric Capture from Very Sparse Consumer RGBD Sensors, CVPR Oral, 2020
仅需4个商用深度相机，实现高质量实时三维重建，系统成本5万元以内

三维视觉：深度相机时代



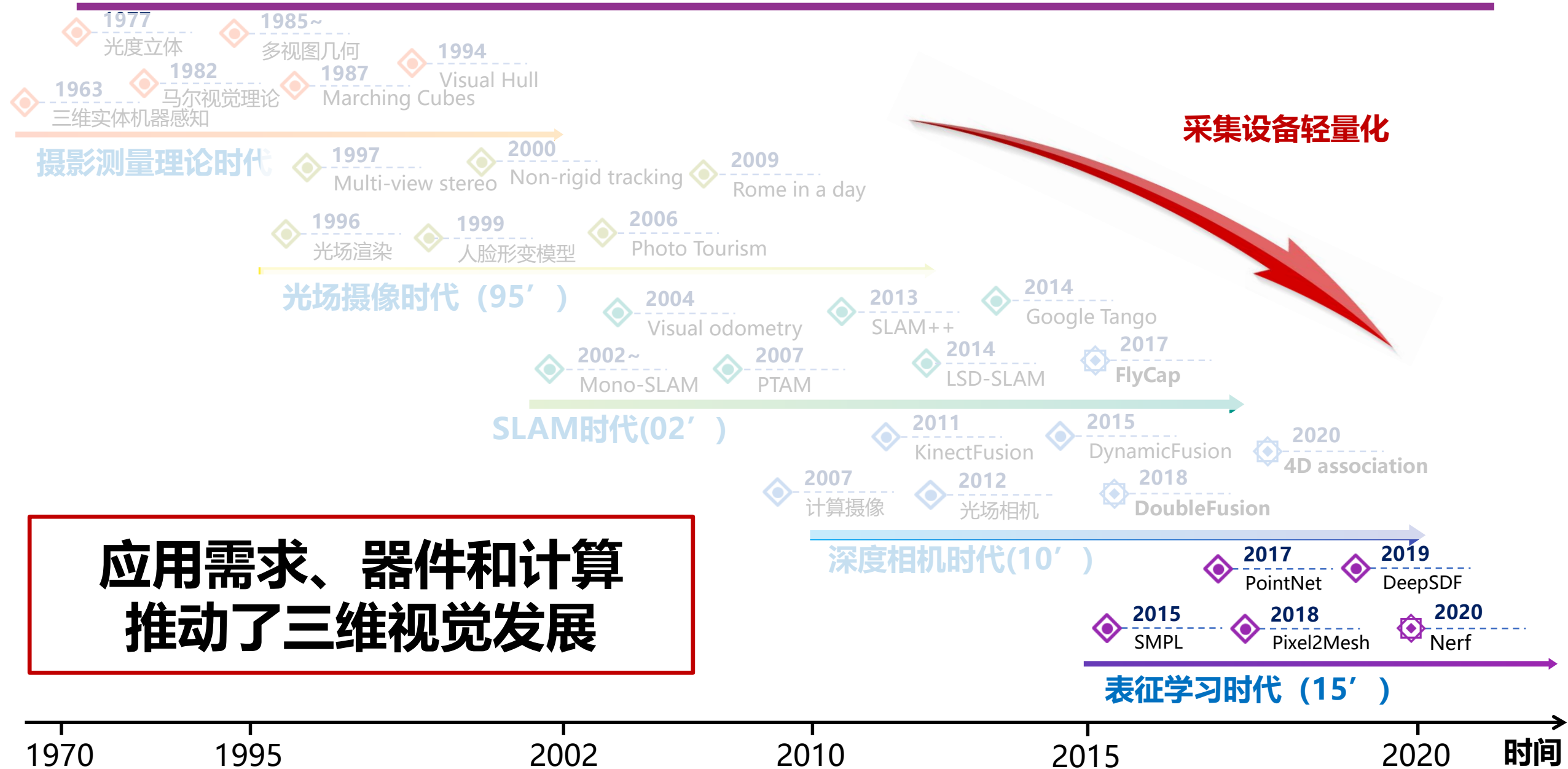
拟建立全息互动内容互联网社区平台，分享创意性AR短视频

三维视觉：深度相机时代



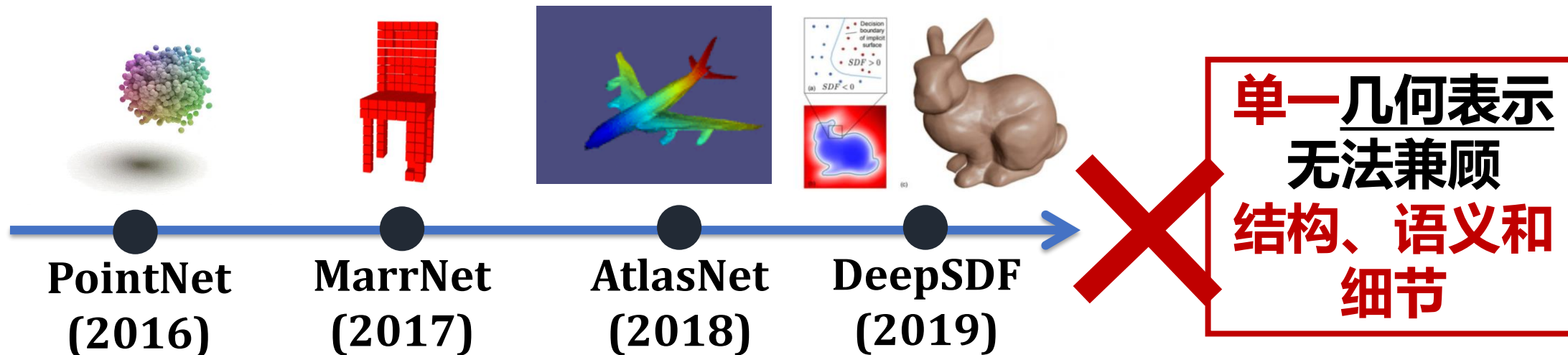
近3年来，受华为、苹果、海信、平安、字节跳动、商汤、快手、同方威视等19家公司技术许可购买或横向课题资助

三维视觉的发展历史



三维视觉：表征学习时代

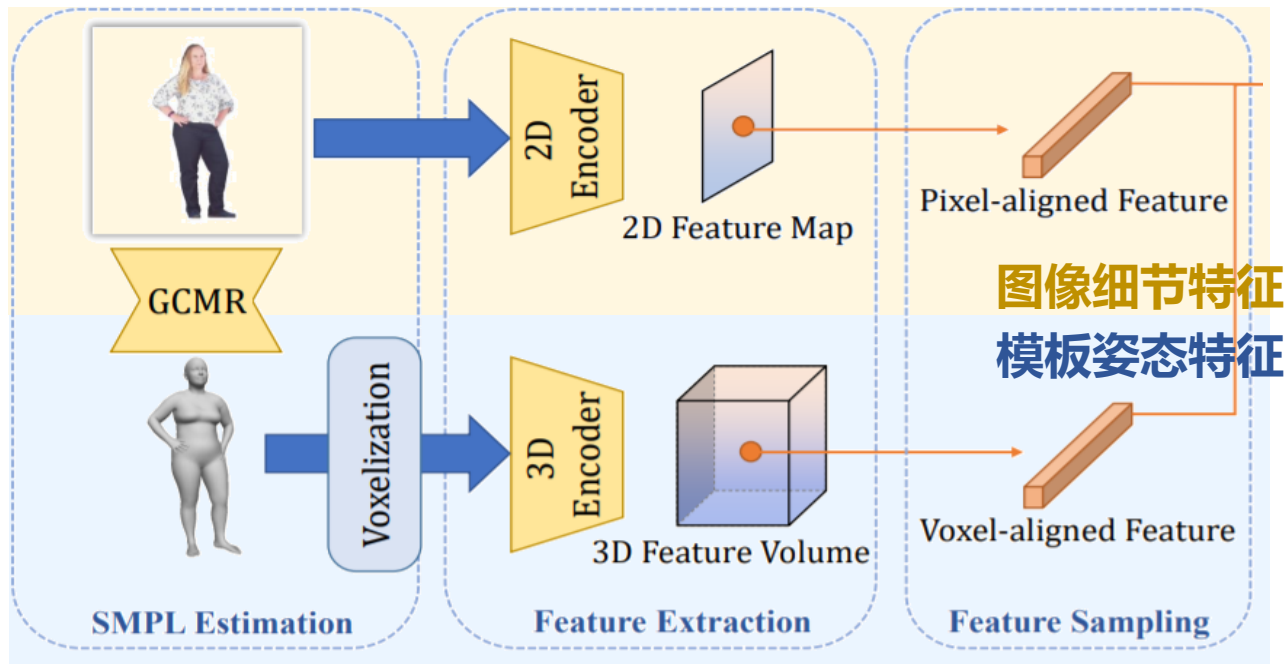
□ 二维表征学习被推广到三维，单图像三维重建时代正式到来



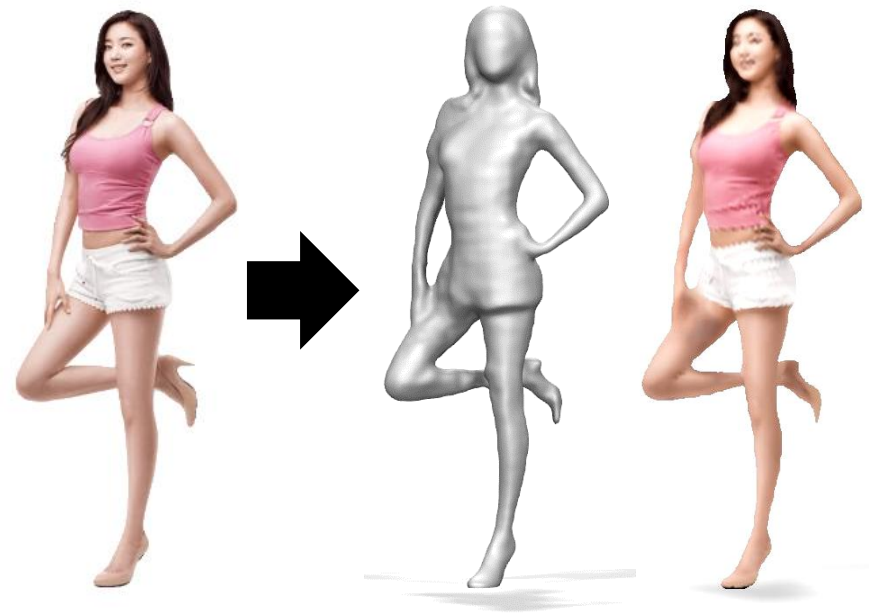
参数化模板	体素	网格	符号距离场	点云	光场
无法处理拓扑变化	数据量大 丢失表面细节	结构不定 难以统一	定义复杂, 丢失语义对应关系	丢失结构信息	数据量大 丢失几何信息

三维视觉：表征学习时代

多表征协同重建方法



单图像重建结果



**多表征重建（参数化模板与隐式函数），建立多表征的协同映射机制，
解决单一表征局限，提高单图像重建精度和鲁棒性**

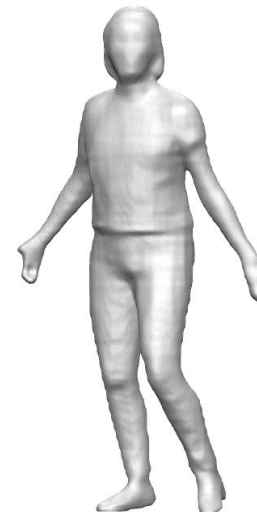
Zheng et al. 3D Human Reconstruction from a Single Image, IEEE ICCV (Oral) 2019.

Zheng et al. PaMIR: Parametric Model-Conditioned Implicit Representation for Image-based Human Reconstruction. T-PAMI 2020

三维视觉：表征学习时代

□ 多表征协同重建方法

单图像重建结果



**多表征重建（参数化模板与隐式函数），建立多表征的协同映射机制，
解决单一表征局限，提高单图像重建精度和鲁棒性**

Zheng et al. 3D Human Reconstruction from a Single Image, IEEE ICCV (Oral) 2019.

Zheng et al. PaMIR: Parametric Model-Conditioned Implicit Representation for Image-based Human Reconstruction.

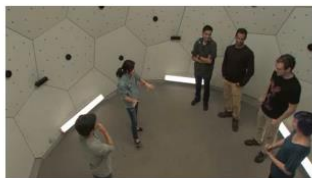
T-PAMI 2021



三维视觉：表征学习时代

□ 稀疏多视点下的多人体重建：多表征学习重建

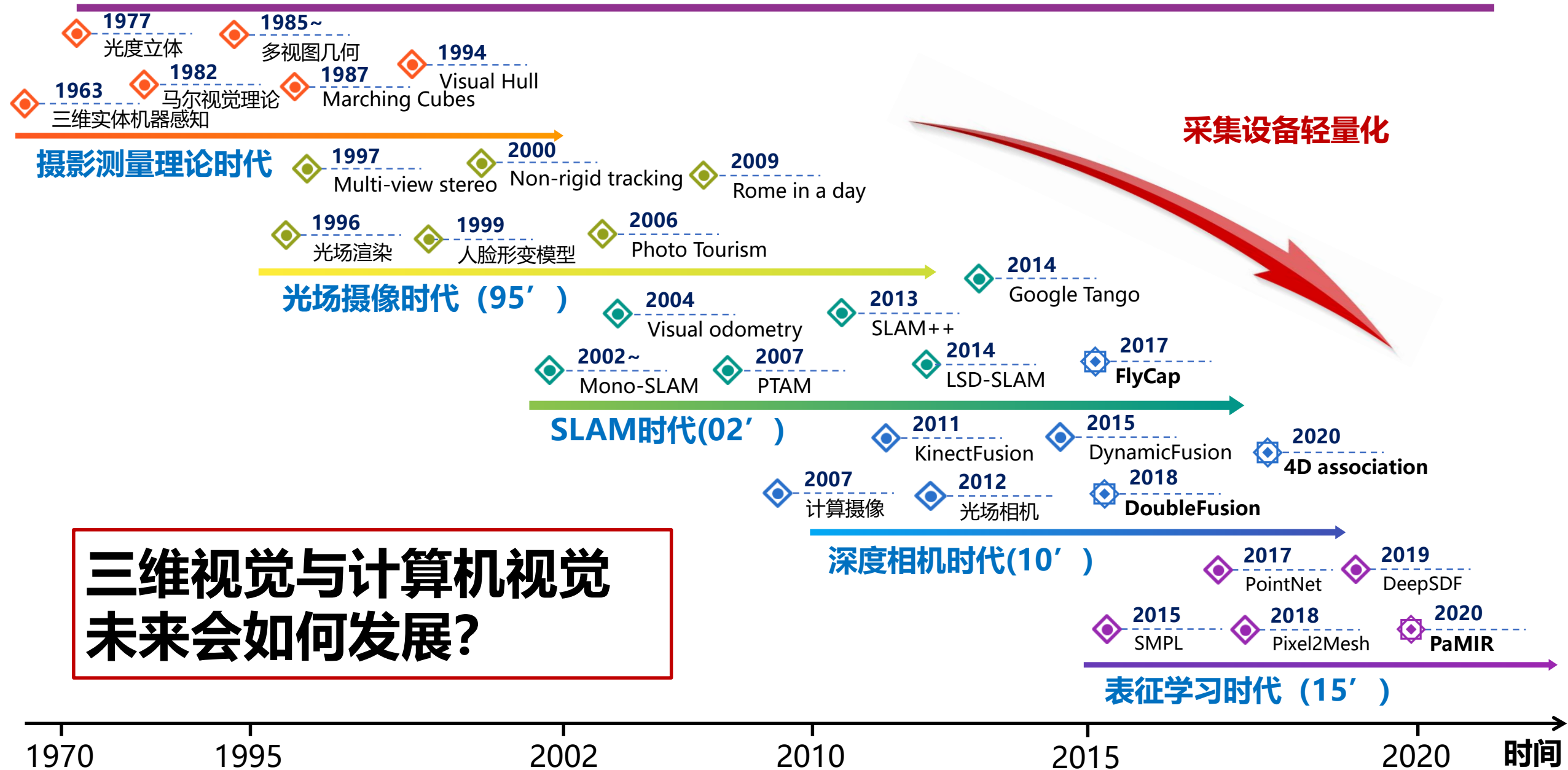
Multi-person Reconstruction



2 of 8 input views
(Total Capture dataset)



三维视觉的发展历史

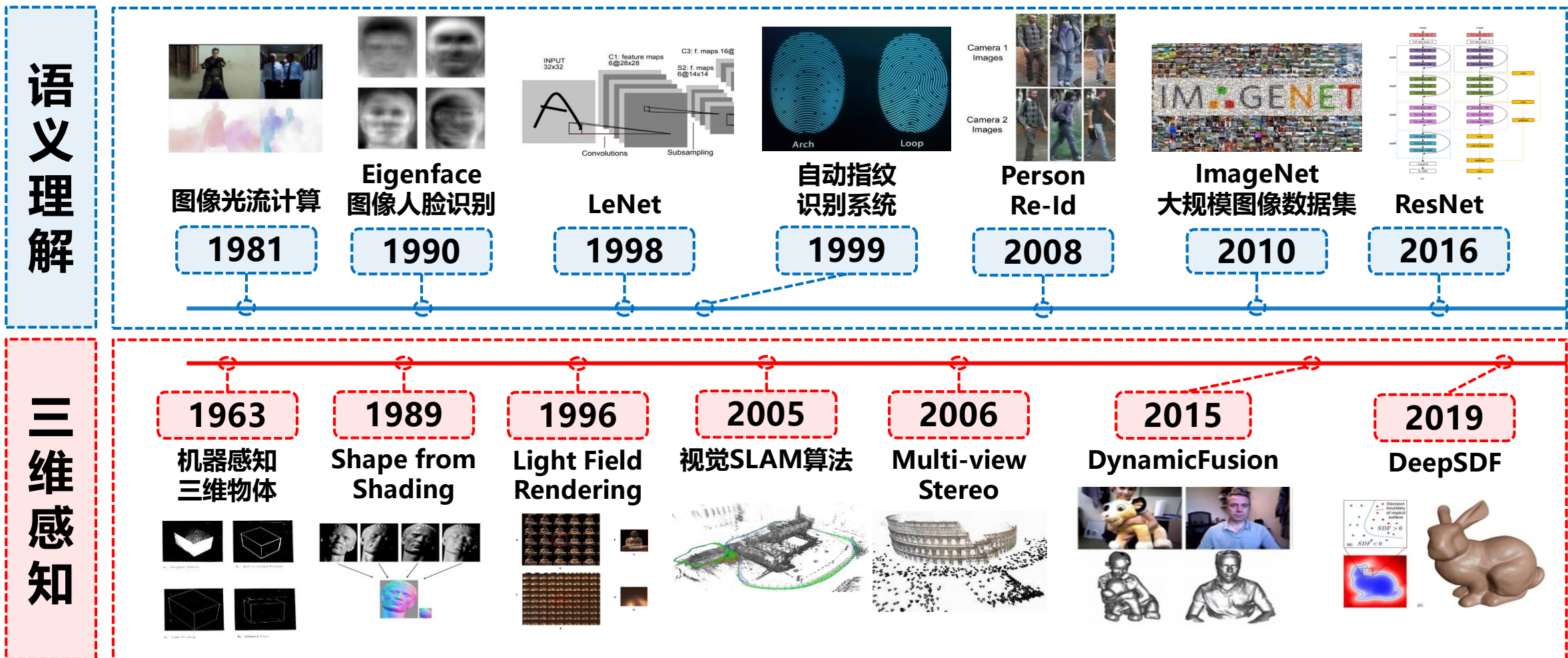


主要内容

- 计算机视觉与三维视觉
- 三维视觉的应用需求
- 三维视觉的发展历史
- 未来的三维视觉与计算机视觉

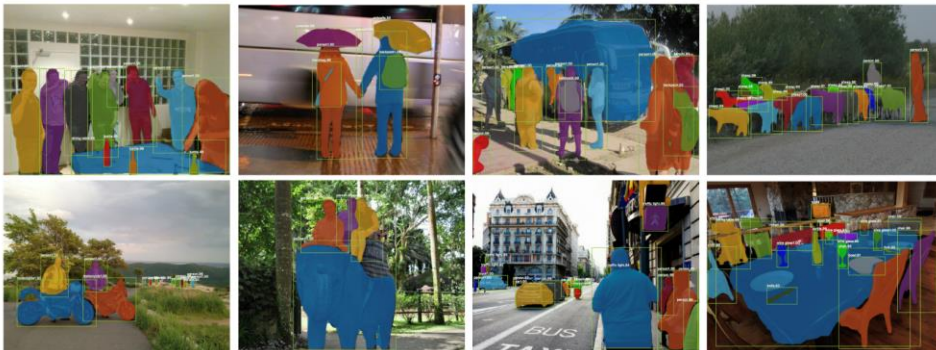
计算机视觉的未来

现状：语义理解和三维感知的并行独立发展



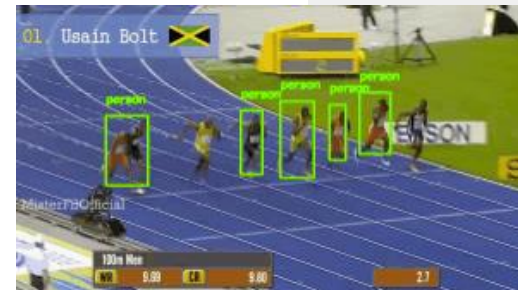
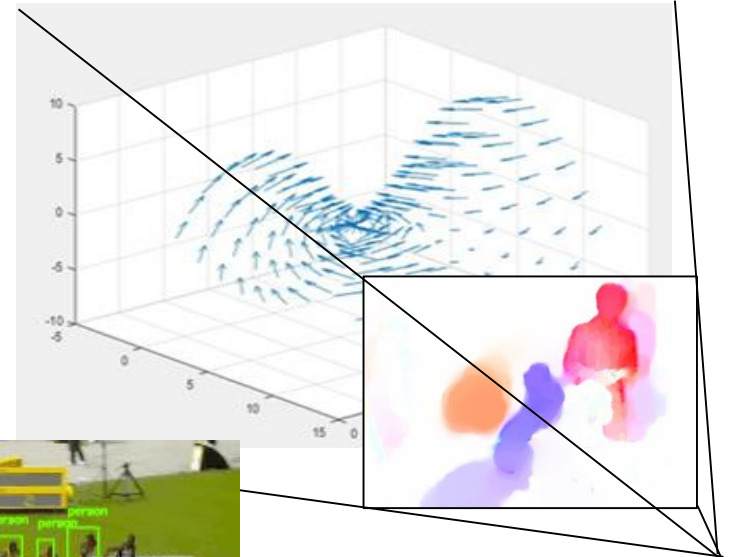
计算机视觉的未来

□ 主流的场景视觉理解



二维目标识别

只是离散的像素分类问题



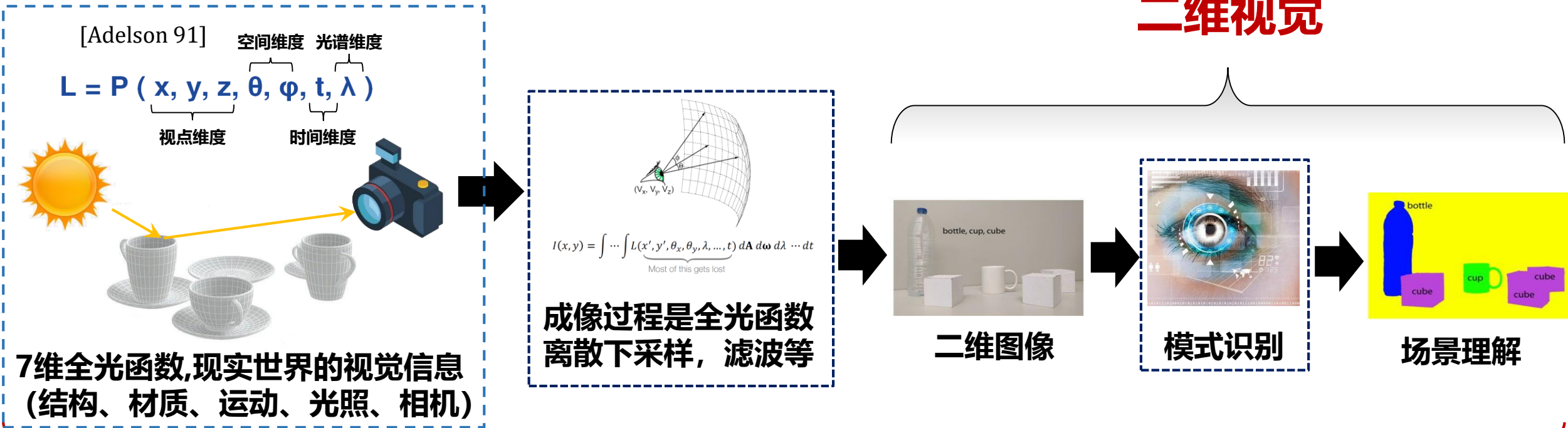
二维运动跟踪

投影导致信息降维

二维场景理解不足以获得连续的世界的本质

计算机视觉的未来

思考：三维视觉与二维视觉的关系



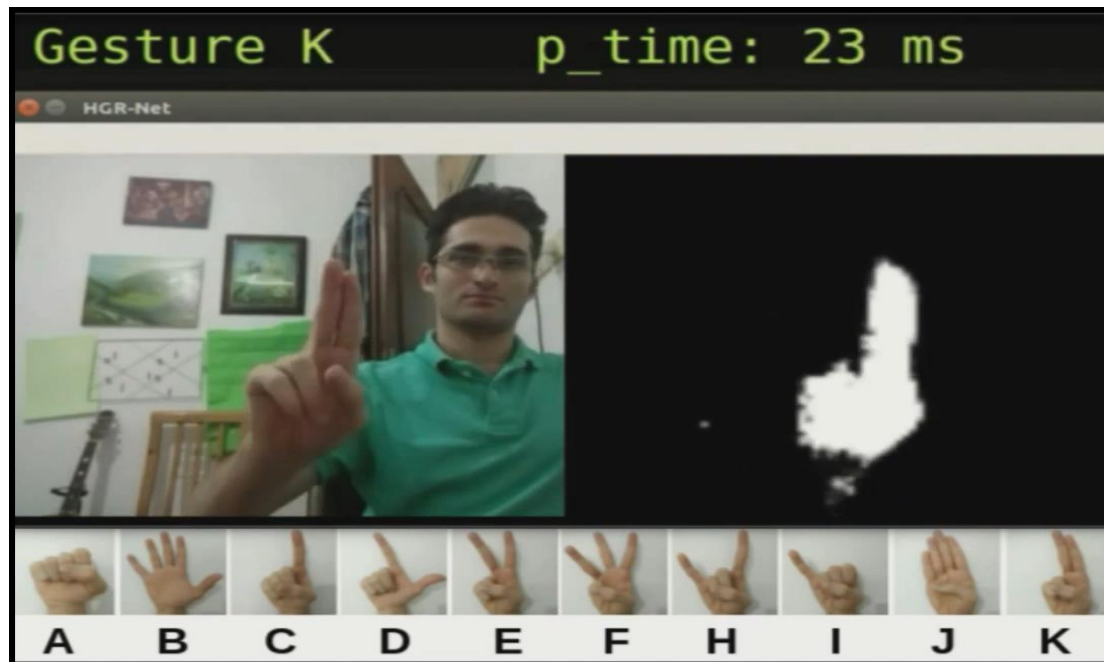
三维视觉

相比于二维视觉, 三维视觉建模了视觉成像机理, 更接近视觉世界本质

计算机视觉的未来

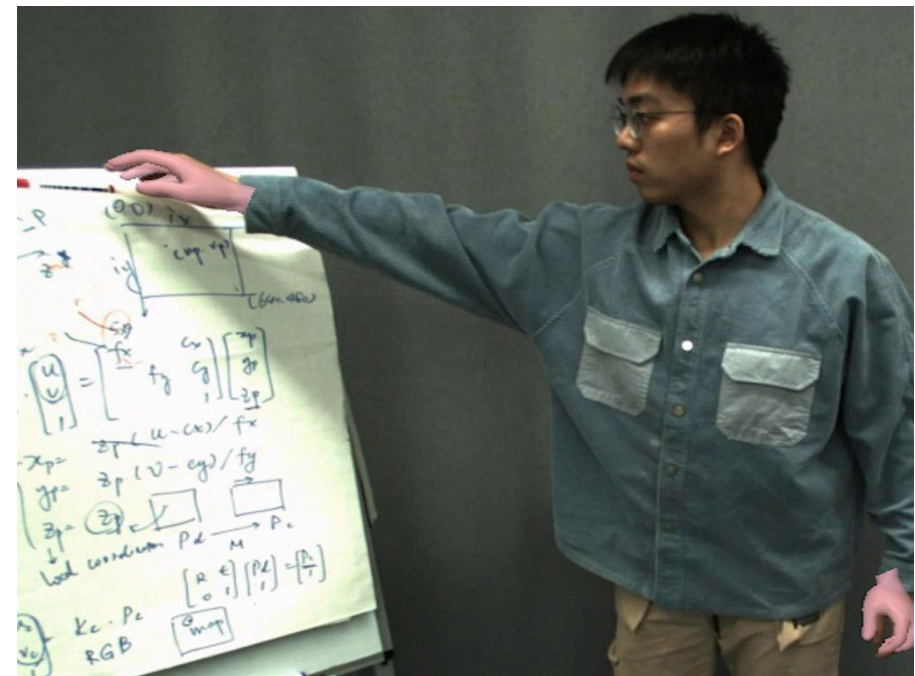
□ 二维到三维的发展：以人体/人手姿态检测为例

提取少量特征进行离散的分类



二维视觉（分类）

获取完整的信息并呈现连续的姿势

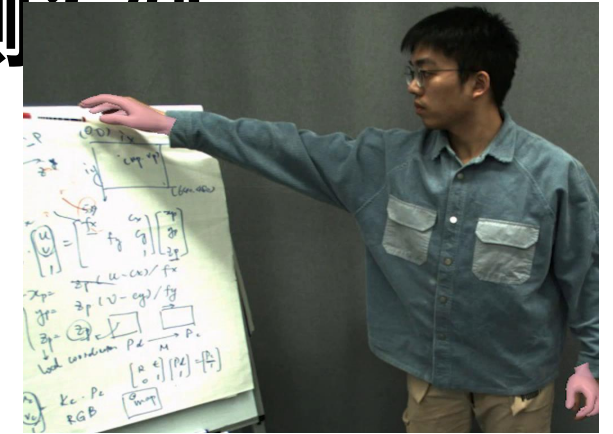
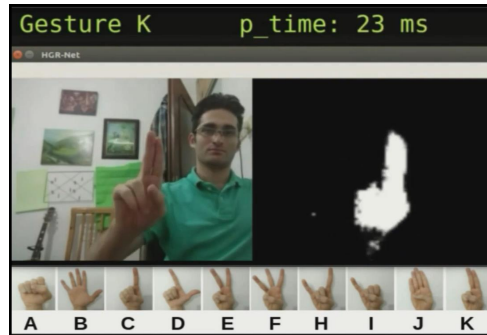


三维视觉（单图像重建）

从二维到三维，提高了手势识别细粒度

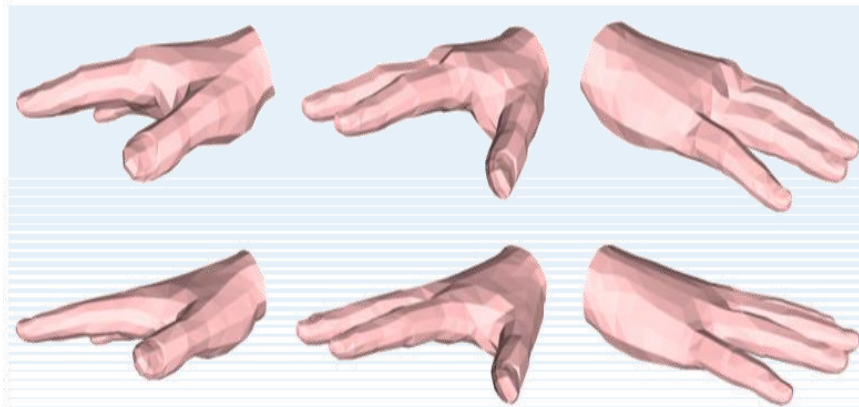
计算机视觉的未来

□ 二维到三维的发展：以人体/人手姿态检测为例



三维
表征
技术

人手几何模板



人手多视角数据集



三维
采集
技术

三维技术提供**知识和数据**，实现**完整、紧致、语义化、连续**的智能视觉感知

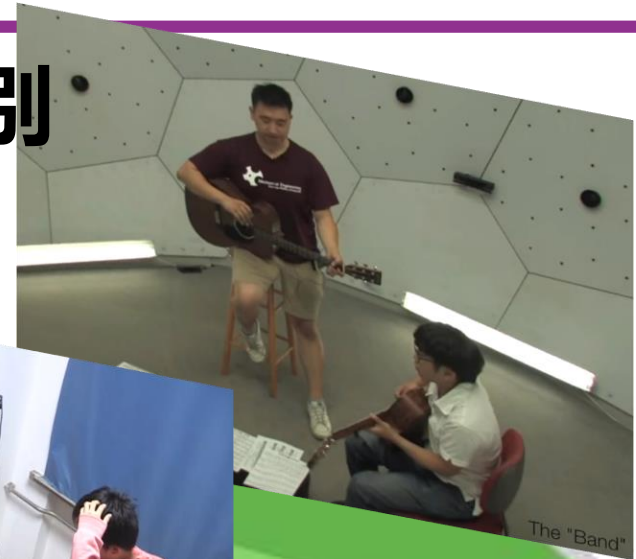
计算机视觉的未来

二维到三维的发展：人体姿态动作识别 三维骨架检测

三维
型体
运动
捕捉

二维骨架检测

二维姿态识别



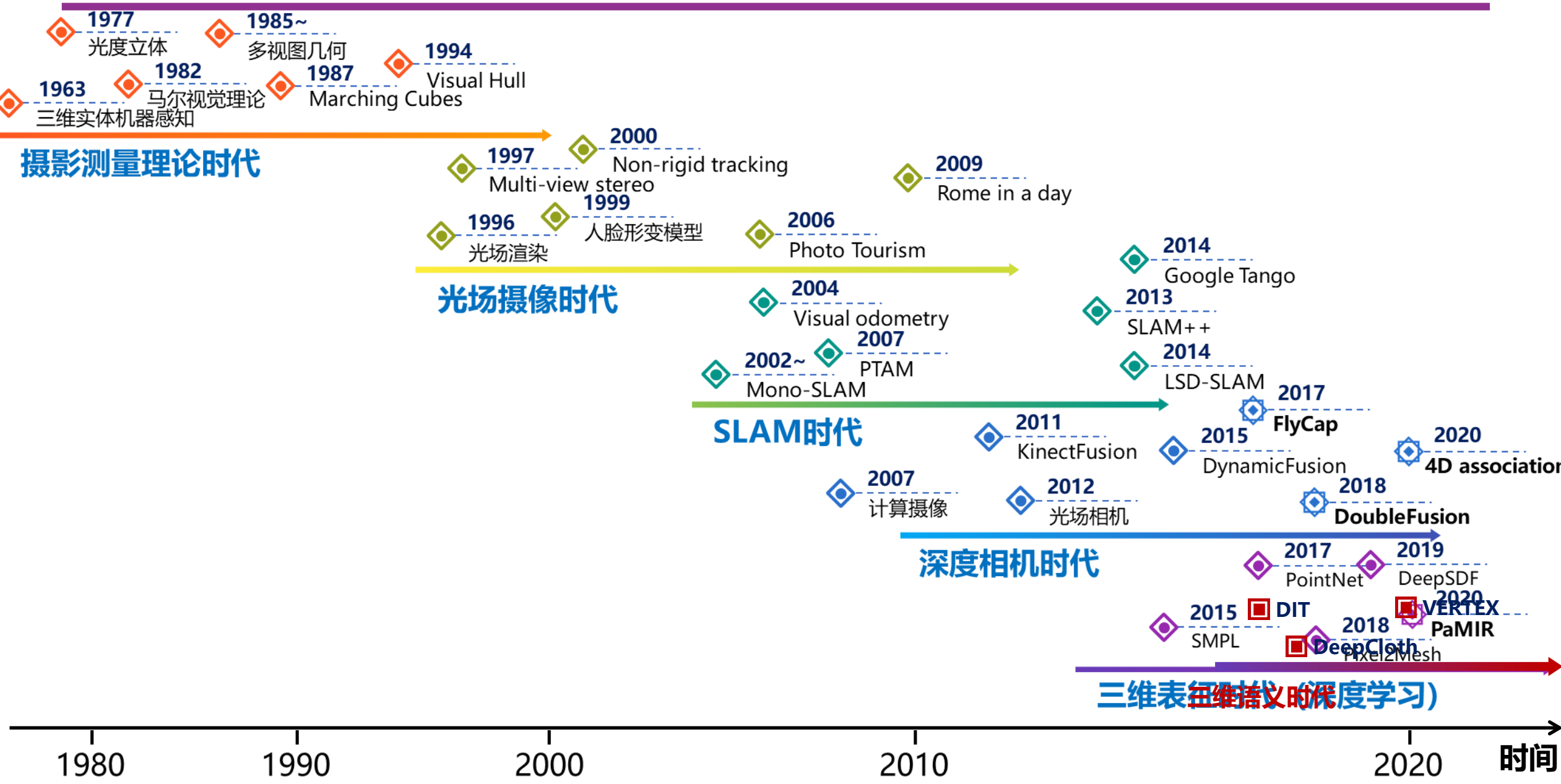
三维



二维

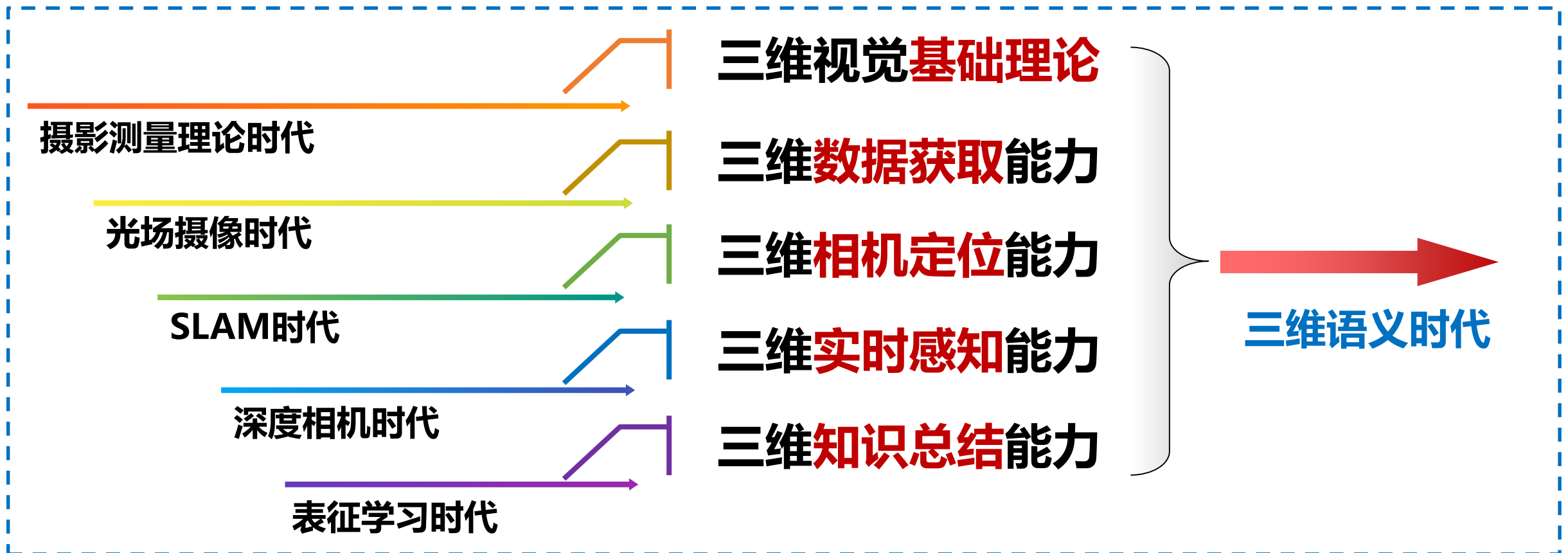
三维重建是深度聚合多个
二维视觉任务的关键载体

计算机视觉的未来



三维语义时代

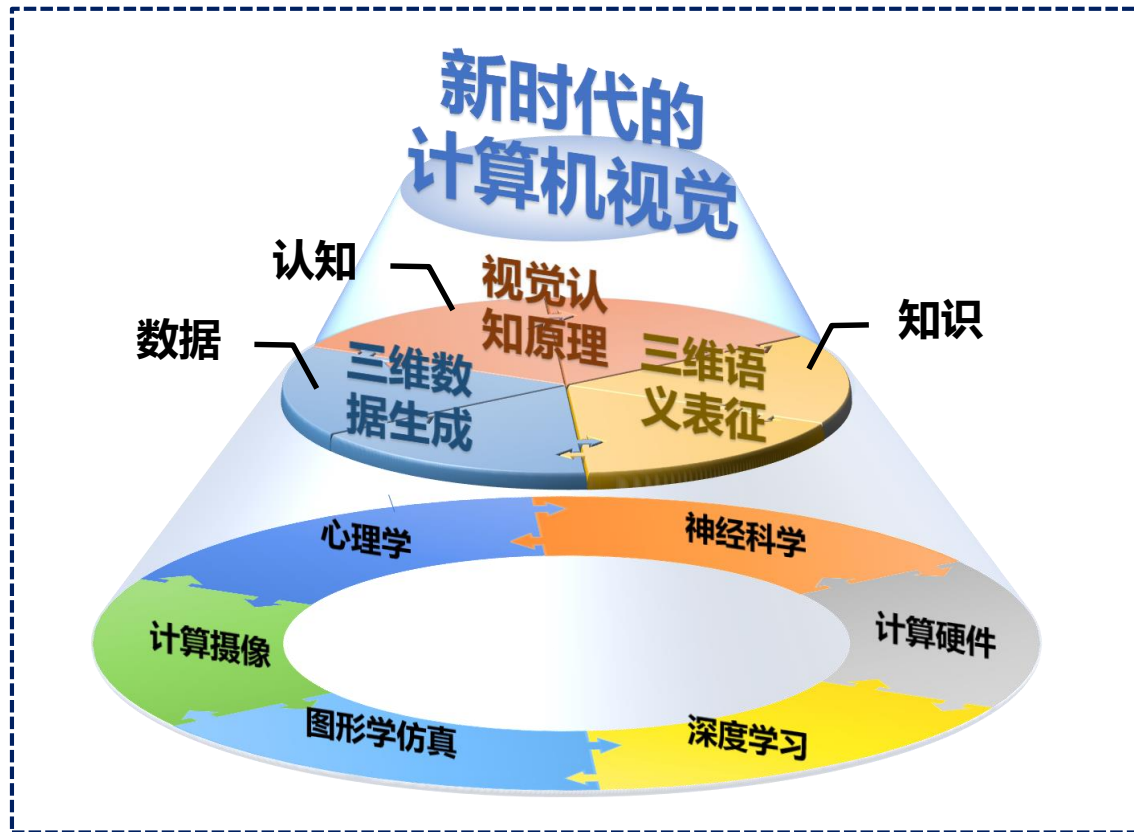
□ 已有时代与三维语义时代的关系



已有三维视觉的理论和技術赋能三维语义时代，实现新一代智能视觉感知

三维语义时代

□ 三维感知技术能够为场景理解提供更多先验



三维语义时代

定义：

带有语义的三维，带有三维的语义

特点：

- 蕴含连续高维知识的三维表征
- 可编辑可交互
- 未来目标：单一图像全景三维语义重建

数据 + 知识 + 认知 = 未来视觉智能（三维语义智能）

必须将目标识别与三维感知结合，才能触及计算机视觉最终目标

三维语义时代

二维语义：离散化、单一属性

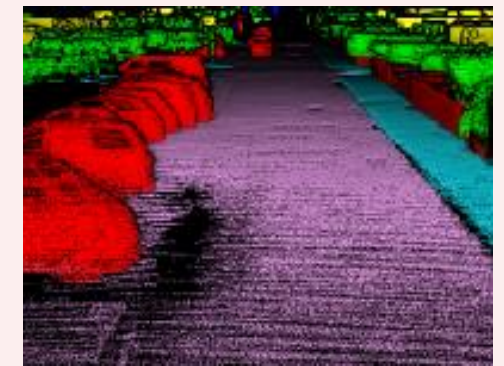
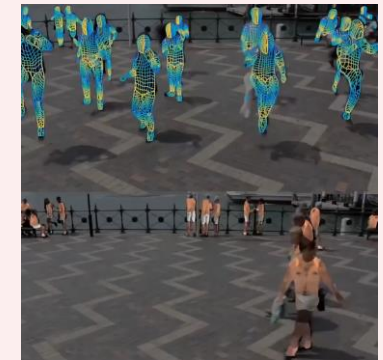


三维语义：连续化、多维度属性

二维



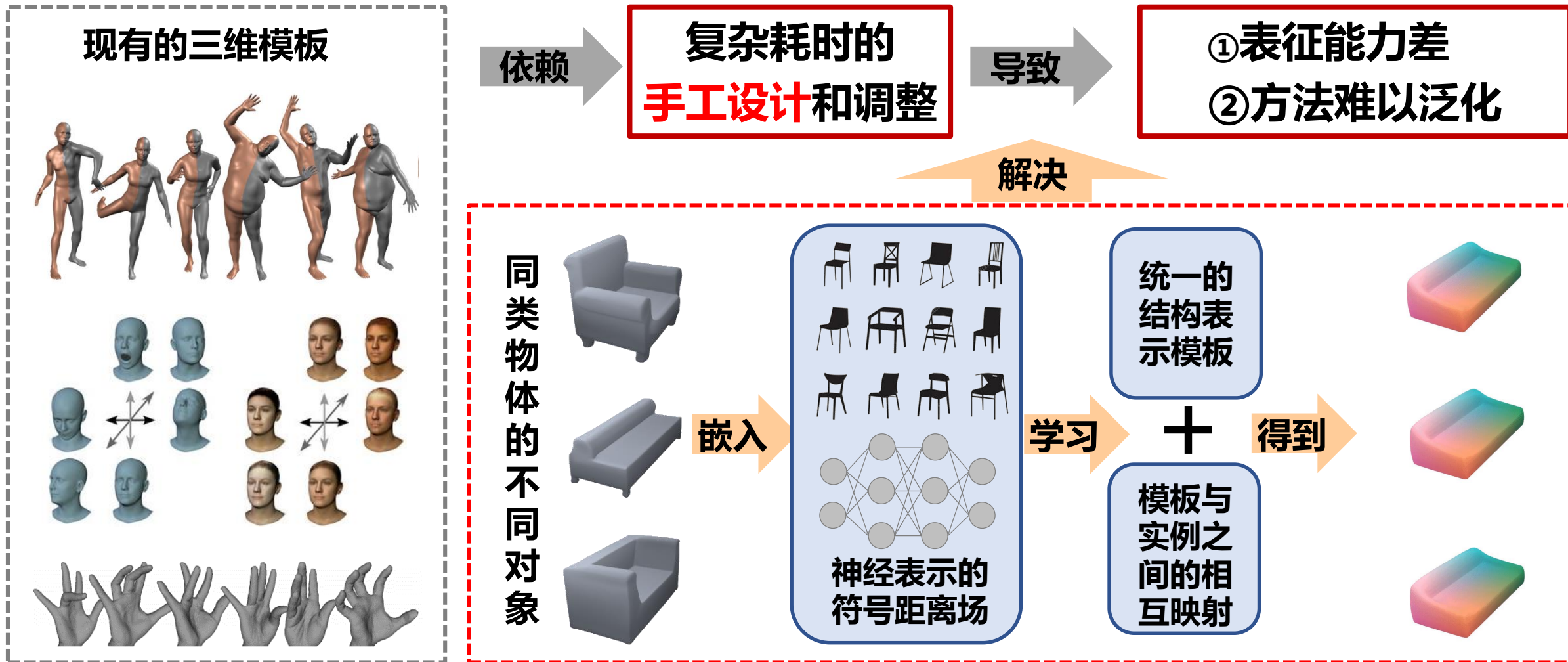
三维



三维语义具有天然的连续性特征，提供了更准确、更高密度的语义信息

三维语义时代

□ 隐式模板——几何模板的自监督学习方法



三维语义时代

隐式模板—几何模板的自监督学习方法

语义理解

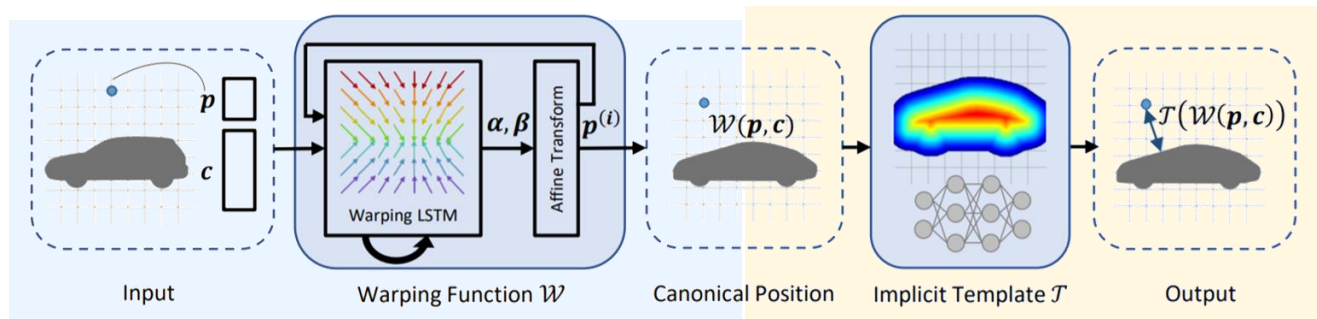
(图像信息冗余, 还原三维世界复杂耗时)

三维感知

(仅考虑建立可视化模型, 难以场景理解)

单一语义几何表示

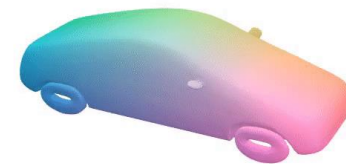
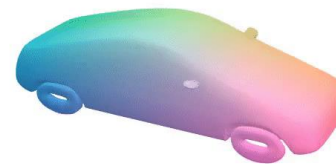
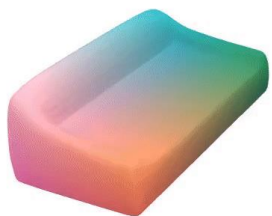
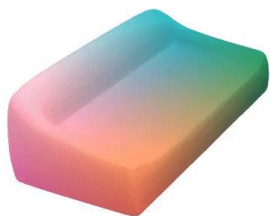
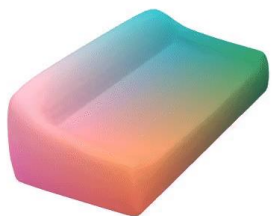
语义化三维建模, 支撑完整的场景理解



Geometry



Correspondences

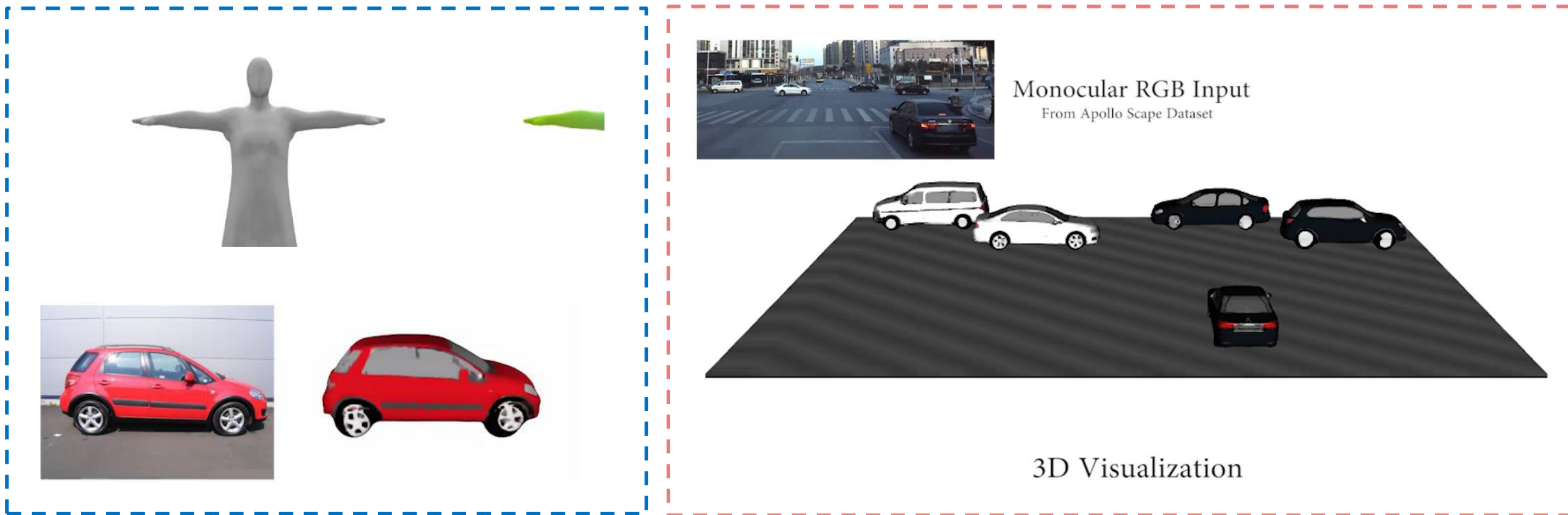


实现对各类不同三维对象的几何结构和相互映射机制的自动提取

Zheng et al. Deep Implicit Templates for Shape Representation. CVPR 2021 Oral

三维语义时代

□ 隐式模板—几何模板的自监督学习方法



实现车辆的语义化纹理和结构重建，可应用于自动驾驶领域道路场景动态可视化

Zhao et al. VERTEX: VEHICLE Reconstruction and TEXtural Estimation Using Deep Implicit Semantic Template Mapping, Arxiv

三维语义时代

□ 语义可编辑的衣物表征与重建方法

三维语义智能编辑

单图像重建与语义智能编辑

三维语义衣服表征、重建与编辑潜在应用于智能服装与纺织产业

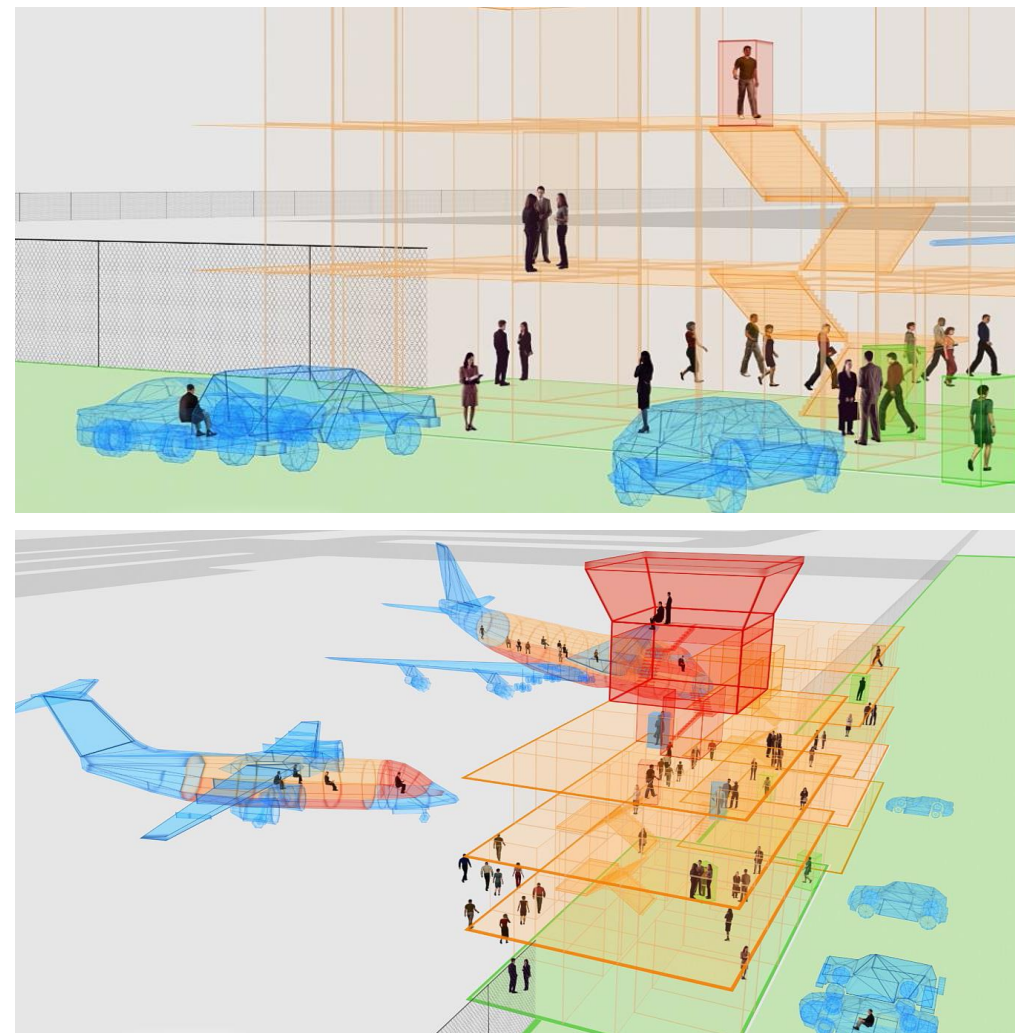
Su et al. DeepCloth: Neural Garment Representation for Shape and Style Editing. Arxiv.

三维语义时代

□ 单图像全景三维语义重建

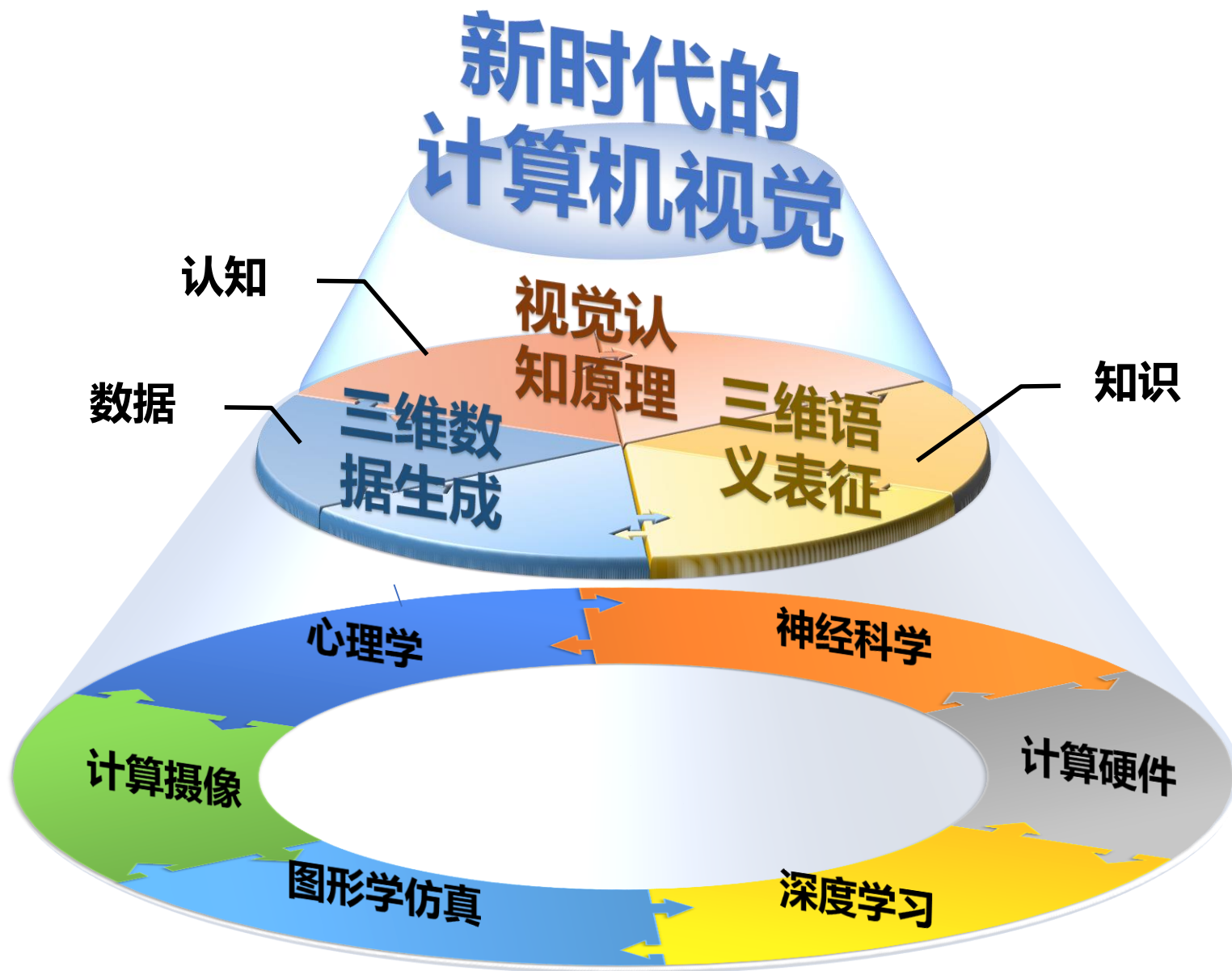


最终目标：将场景所有对象三维语义重建



单图像全景重建，无须二维检测识别，直接对图像进行三维语义重建

总结



- 三维视觉在人工智能各产业具有重要地位
- 三维视觉是计算机视觉的必然趋势
- 知识+数据+认知 驱动计算机视觉发展
- 三维语义重建是三维视觉未来十年的热点

《中国图象图形学报》 “三维视觉与智能图形” 专刊



➤ 专刊主编

- 陈宝权 教授 北京大学
- 虞晶怡 教授 上海科技大学
- 周 昆 教授 浙江大学

➤ 专刊编委

- 郭裕兰 副教授 国防科技大学
- 黄 惠 教 授 深圳大学
- 刘利刚 教 授 中国科学技术大学
- 刘焯斌 副教授 清华大学



扫描了解 **专刊详情**

- 徐 凯 教 授 国防科技大学
- 章国锋 教 授 浙江大学
- 周晓巍 研究员 浙江大学

➤ 截稿时间：2021年6月30日

➤ 同行评议返回时间：2021年8月中旬

➤ 录用通知时间：2021年9月中旬

➤ 拟出版时间：2022年2月