关节(角色)动画

金小刚

Email: jin@cad.zju.edu.cn

浙江大学CAD&CG国家重点实验室

紫金港校区蒙民伟楼512

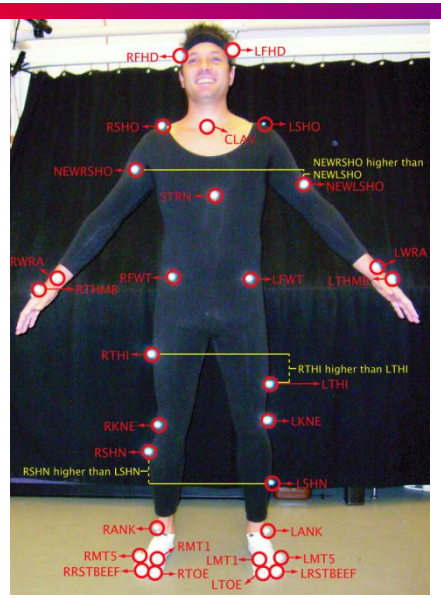
@Copyright Xiaogang Jin

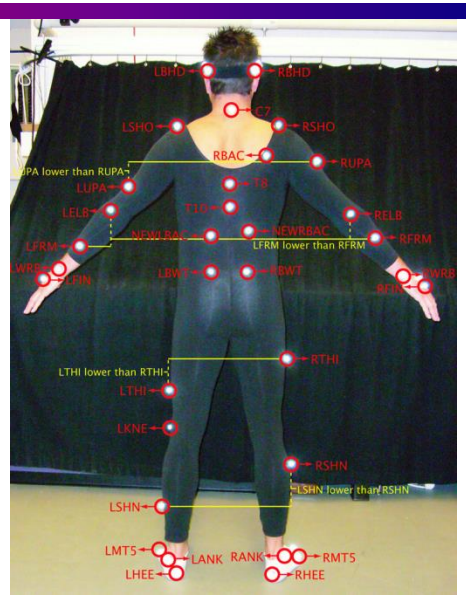
CMU Graphics Lab Motion Capture Database

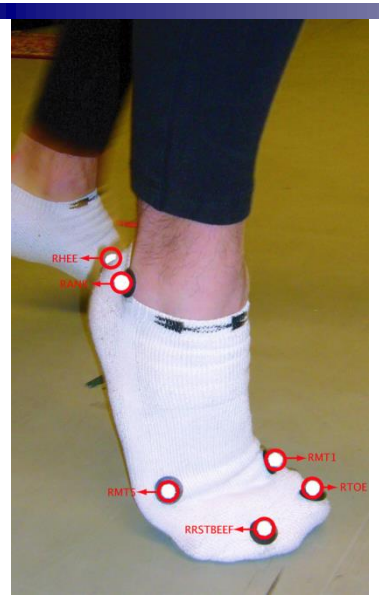
http://mocap.cs.cmu.edu/

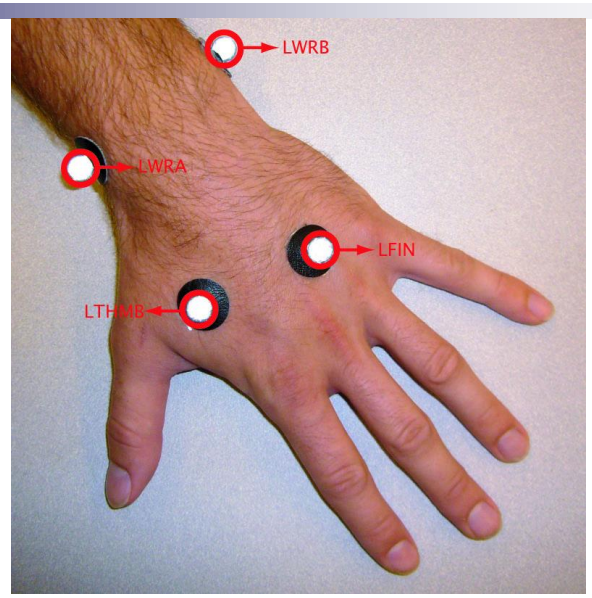
- Human Interaction
  two subjects
- Interaction with Environment
  playground, uneven terrain, ...
- Locomotion
  running, walking, ...
- Physical Activities & Sports
  basketball, dance, ...
- Situations & Scenarios
  common behaviors and expressions, pantomime, ...

CMU Graphics Lab Motion Capture Database









- **How is this data captured?**
  We have a Vicon motion capture system consisting of 12 infrared MX-40 cameras, each of which is capable of recording at 120 Hz with images of 4 megapixel resolution. Motions are captured in a working volume of approximately 3m x 8m. The capture subject wears 41 markers and a stylish black garment.
- 运动捕捉数据有多种格式：ASF/AMC, BVH

一些新的运动捕获数据库

- **https://babel.is.tue.mpg.de/**

  CVPR 2021, BABEL: Bodies, Action and Behavior with **English Labels (英文标签)**

- **https://amass.is.tue.mpg.de/** (smpl格式，带有mesh)

  ICCV 2019, AMASS: Archive of **M**otion Capture **A**s **S**urface **S**hapes

- **https://www.ianxmason.com/100style/** (带100种风格标签)

  100STYLE: This is the 100STYLE dataset, it contains over **4 million frames** of motion capture data for **100 different styles of locomotion**.

- **https://github.com/EricGuo5513/HumanML3D**

  CVPR 2022, 带文字标签的大型数据集

如何用这些Motion Capture数据?

- 离线(Off-line)
  - 通过滤波、逆向运动学进行处理
  - 建立运动数据库
    - 从中选择
    - 对运动数据进行混合(Blend)
    - 在线(on the fly)修改
- 在线(On-line) (表演动画)
  - 基于演员的实时表演直接驱动虚拟角色

运动捕获数据的处理

- 运动信号处理和运动变形(Motion warping)

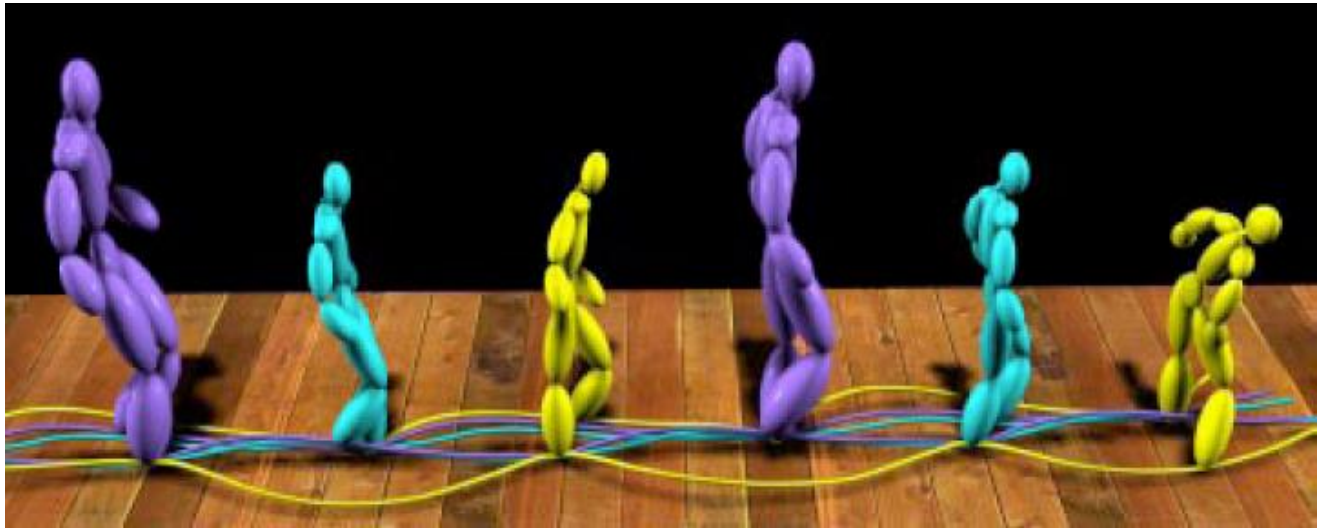
- 运动数据的重定向
  - 把运动映射到不匹配的虚拟角色上，并满足一些重要的约束条件
  - 通常具有相同的结构（with the same structure）

- 运动数据的合并
  - 把短的运动片段合并成长的运动序列

运动数据的信号处理

- 运动信号处理
  - 考虑信号频率如何刻画各种运动
    - 低频部分：表示基本运动(如走路)
    - 高频部分：特质运动(跛行)
- 运动的操作/编辑/变形
  - 对运动信号进行变形，使得满足用户指定的一些约束条件
- 把运动看成多维信号
- 低通滤波
  - 去噪
- 高通滤波
  - 风格的变换
- 高频信息可能是运动细节，而非噪声
- 通过滤波来进行运动的修改并非易事
  - 物理约束 (关节约束, 地面接触)?
  - 运动的自然性?

运动重定向



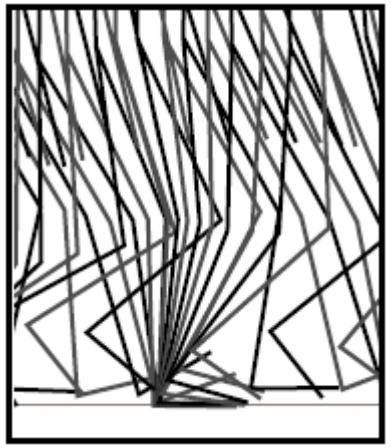
Michael Gleicher, “Retargetting motion to new characters ”, SIGGRAPH '98

- 把捕获的运动数据映射到不匹配的虚拟角色上，并满足一些重要的约束
- 约束（运动所应满足的性质）：
  - 避免脚穿透地板,避免自身穿透, 避免走路时的打滑；
- 生成一个与原始运动尽可能接近的新的运动，并强制满足约束条件；
- 新的运动可看成是一个时空(space-time)、非线性约束的优化问题。

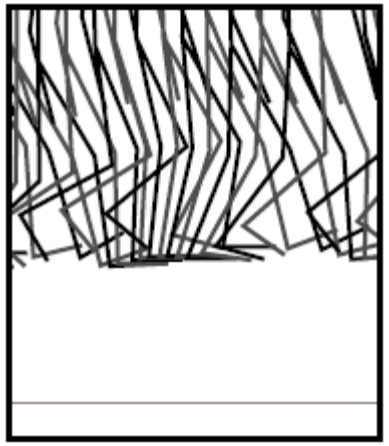
运动重定向

- $m(t) = m_o(t) + d(t)$
  - $m_o(t)$: 原始运动； $m(t)$: 重定向的运动
  - $d(t)$: 两个运动之间的距离
- 如果只是简单重用原始的运动，则目标角色**与其它物体之间的相互作用会失败**。
- 把运动重定向问题看成一个带约束的优化问题：
  - 找到运动$m(t)$, 满足函数$f(m(t)) \lozenge c$
    - $f(\cdot)$: 约束方程, c : 常数
    - $\lozenge \in \{\leq, \geq, =\}$
  - 最小化目标函数：$g(m) = \int_t (m(t)-m_o(t))^2 = \int_t d(t)^2$

原始运动



Naïve重定向运动

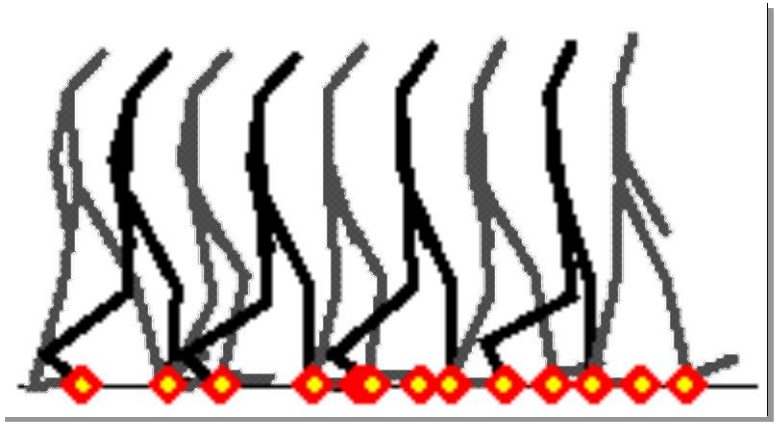


约束例子：Footskate

- **Footskate（打滑）**：是一种经常会发生的不真实现象；
- **Penetrations（贯穿）**：是一种经常会发生的不真实现象；

运动重定向

- 定义约束



- 加入平移偏移 (近似解)

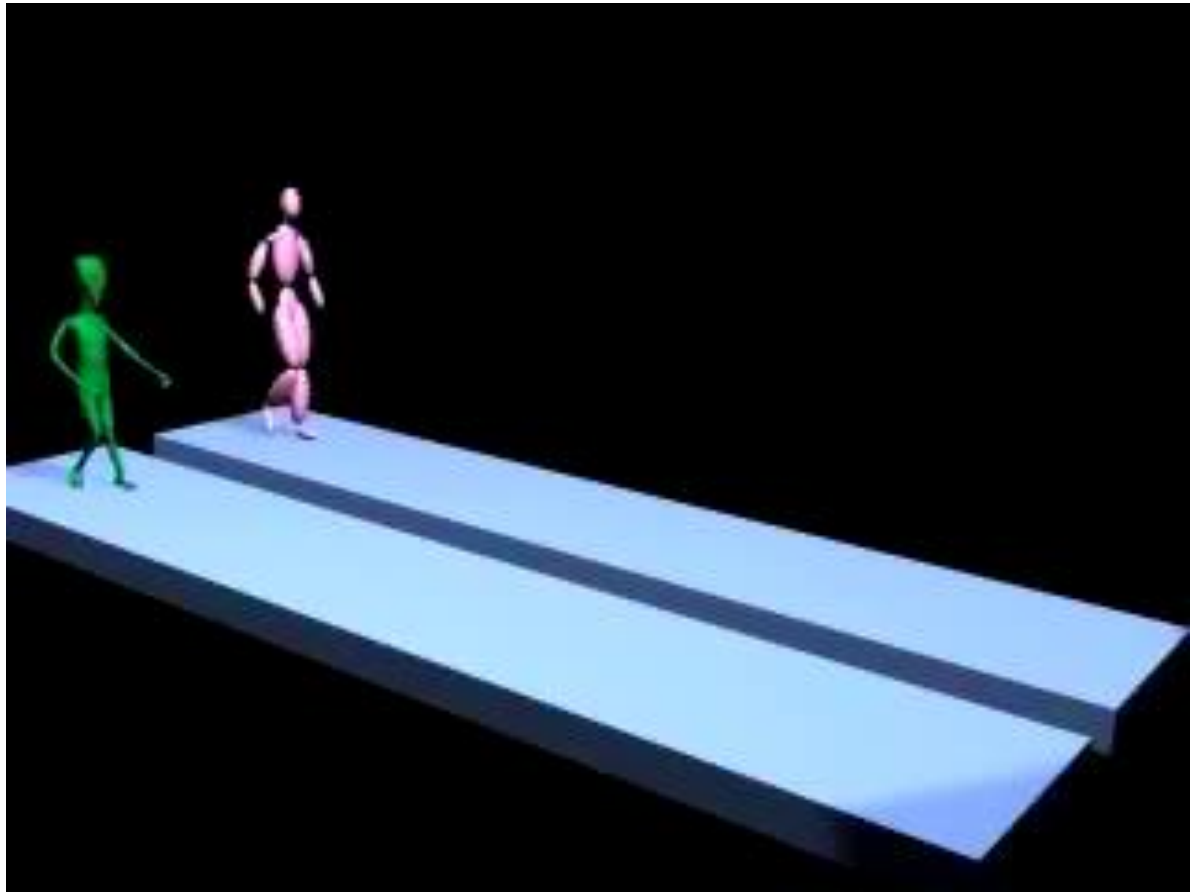
- 应用到新的角色

- 解非线性约束优化问题

把运动曲线应用于一个新的角色

- 定义运动的约束条件
- 对于整个运动，求解约束优化问题

运动重定向动画演示

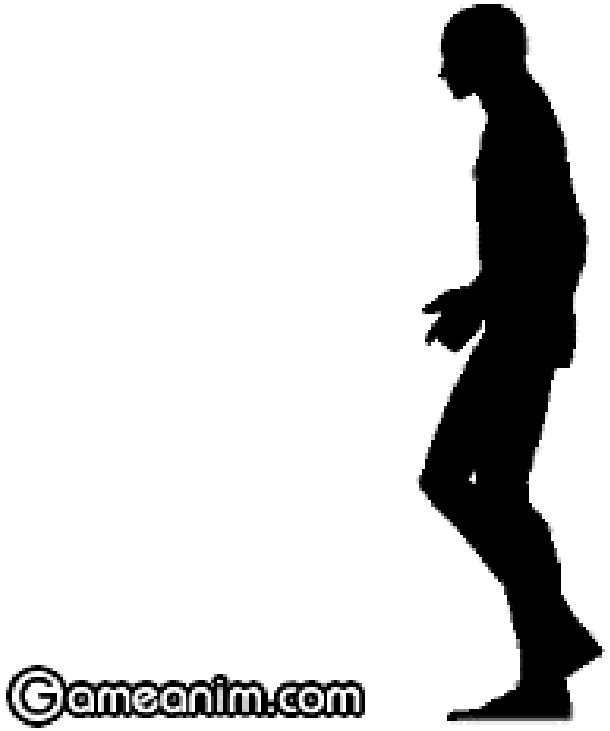


运动混合(Motion Blending)

- 把两段或多段运动混合在一起
  - 实际上是插值

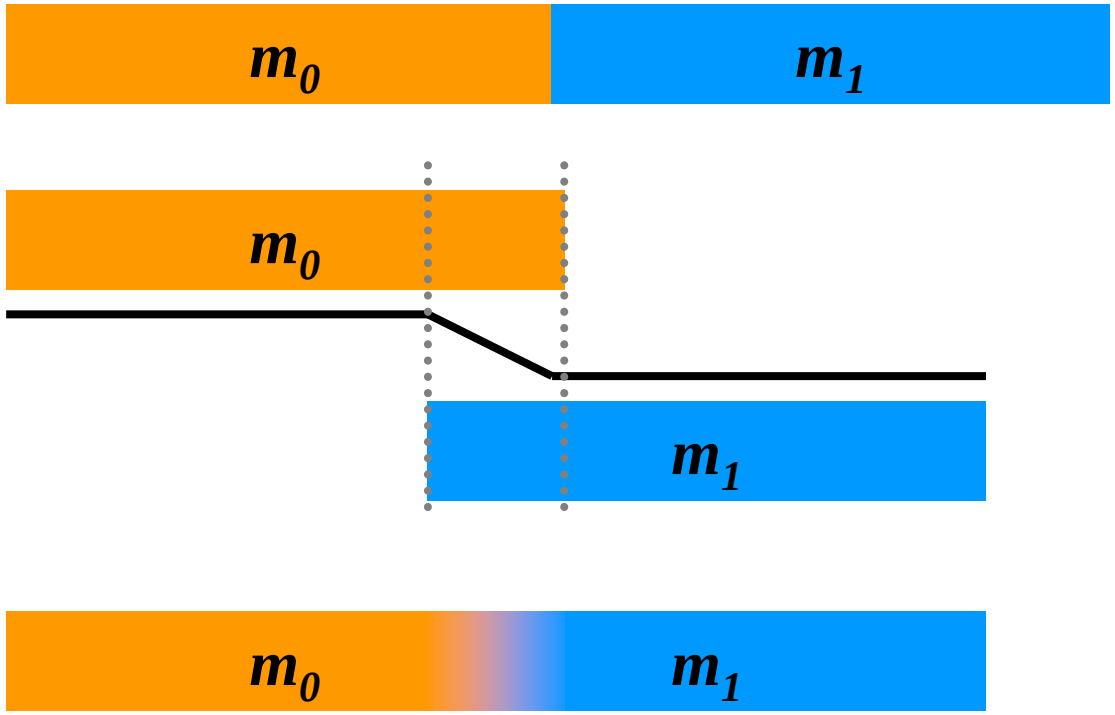
$$m(t) = a\, m_0(t) + (1\text{-}a)\, m_1(t)$$

这是一个逐帧的操作

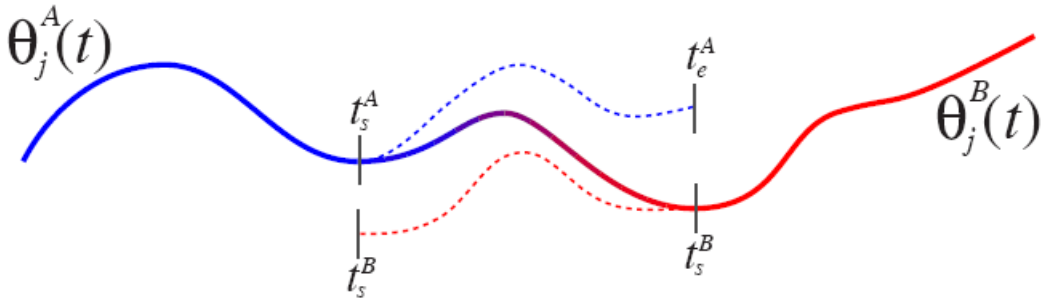


Blend Speed

运动过渡(Motion Transition)



- 运动捕获得到的通常是小的运动片段
- 需要把他们连接起来
- 如果运动是相近的，则很容易

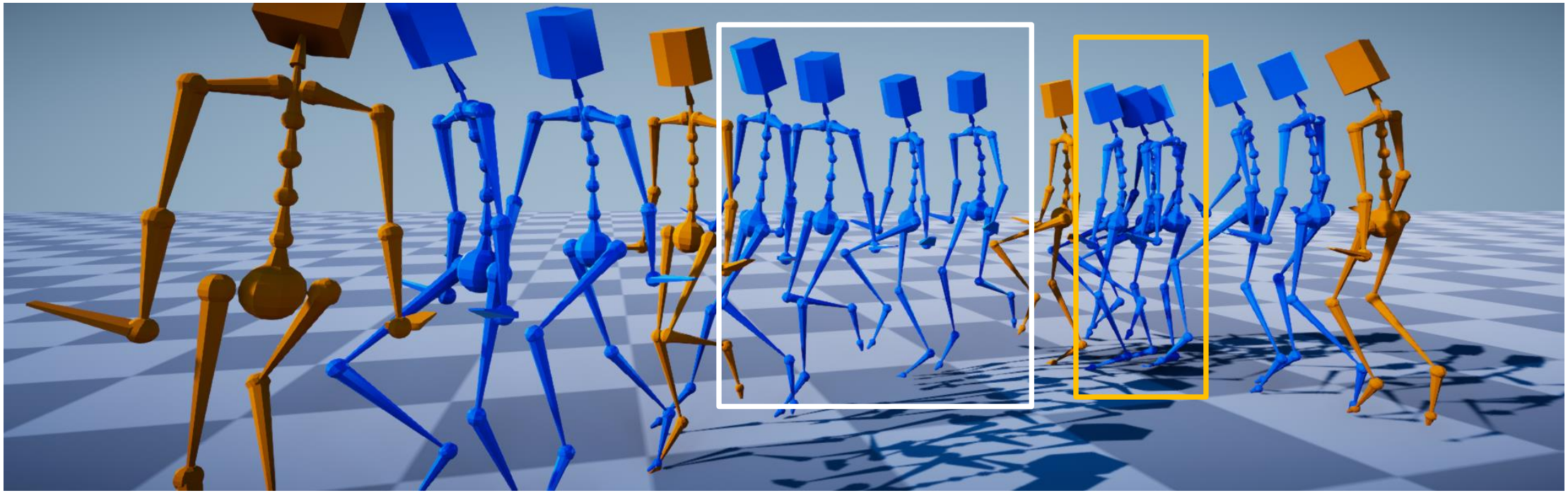


- 如果运动相差较远，则比较困难

运动过渡(MIB，Motion In Between)

Xiangjun Tang, He Wang, Bo Hu, Xu Gong, Ruifan Yi, Qilong Kou, Xiaogang Jin, “Real-time Controllable Motion Transition for Characters”, ACM Transactions on Graphics (Proc. Siggraph 2022), 2022, 41(4): Article 137.

运动过渡(MIB，Motion In Between)



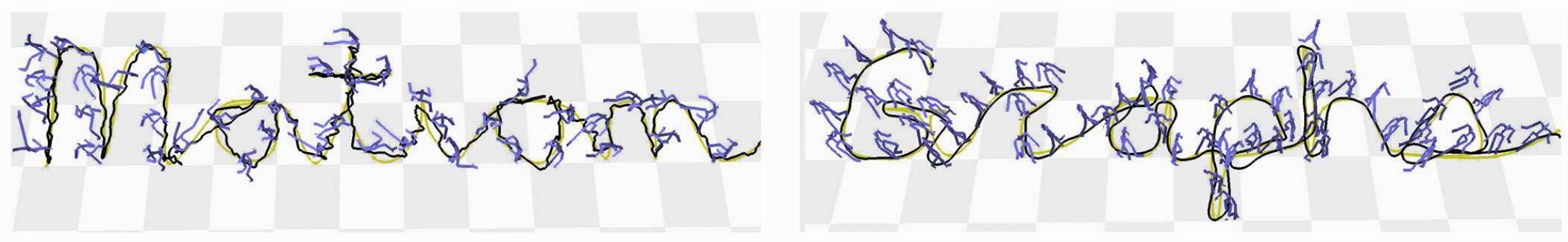
Xiangjun Tang, Linjun Wu, He Wang, Bo Hu, Xu Gong, Yuchen Liao, Songnan Li, Qilong Kou, Xiaogang Jin: RSMT: Real-time Stylized Motion Transition for Characters. *Proc. SIGGRAPH 2023*: 38:1-38:10

Real-time Stylized Motion Transition for Characters

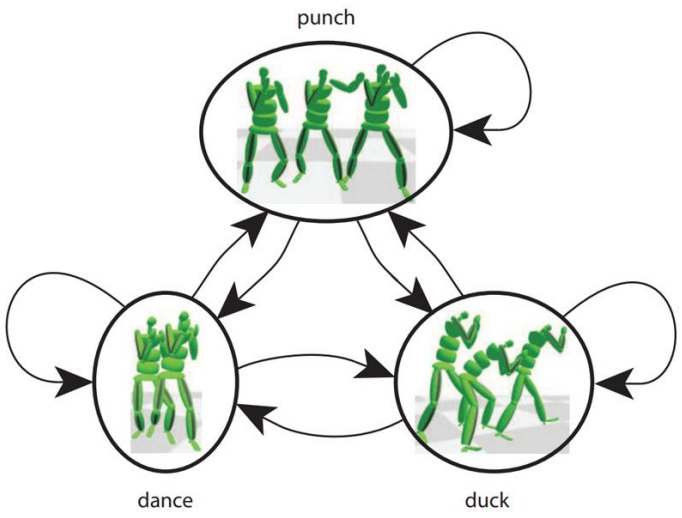
Anonymous author(s)

Submission ID: papers_434

运动图(Motion Graph)

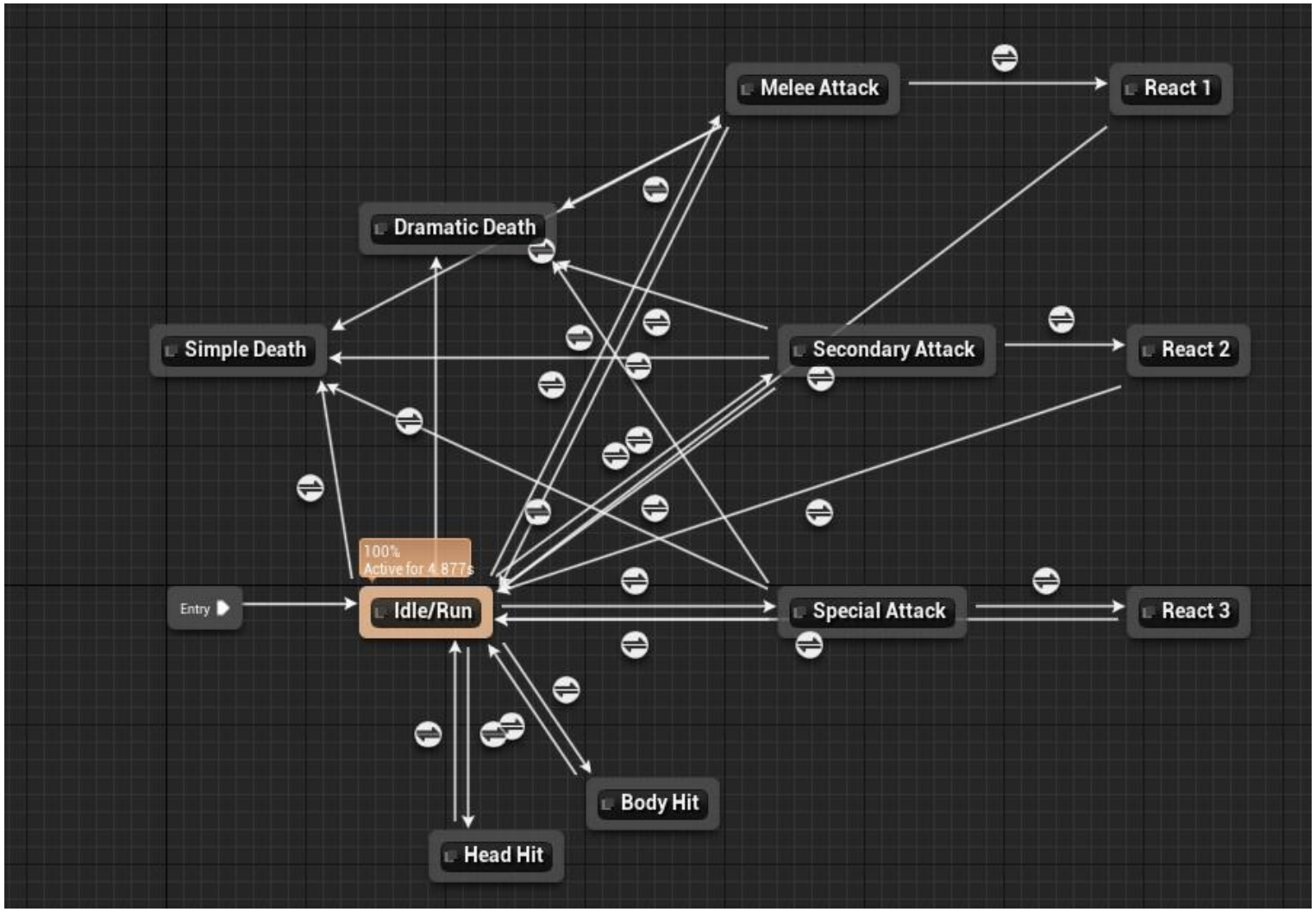


思想：将一组预先获取的运动片段通过特定的相似点连接成一个图结构，来生成新的连续动作。每个运动片段作为图的节点，节点之间的连接通过运动的相似性决定，连接处的平滑过渡依赖于帧的姿态和速度匹配。一旦构建了运动图，系统可以从中选取不同的路径，从而生成各种连续、自然的动画，而不需要重新录制新的动作。



Lucas Kovar, Michael Gleicher, Frédéric H. Pighin: Motion graphs. ACM Trans. Graph. 21 (3): 473-482 (2002)

运动图(Motion Graph)——例子



应用实例

《加勒比海盗3》

《指环王》系列电影中的古鲁姆的运动

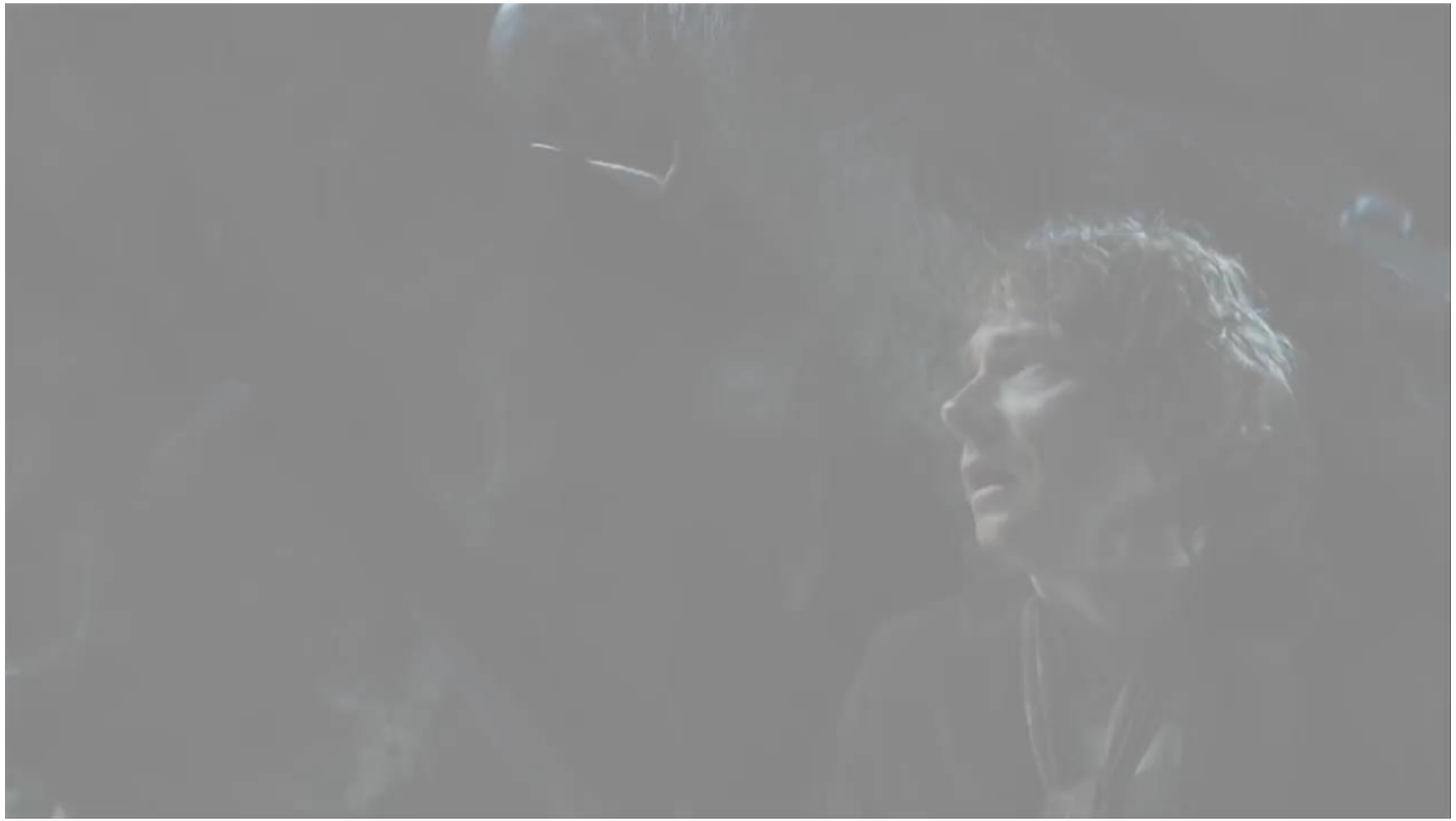
应用实例

表演者

树怪

02:36:00

Creating Gollum



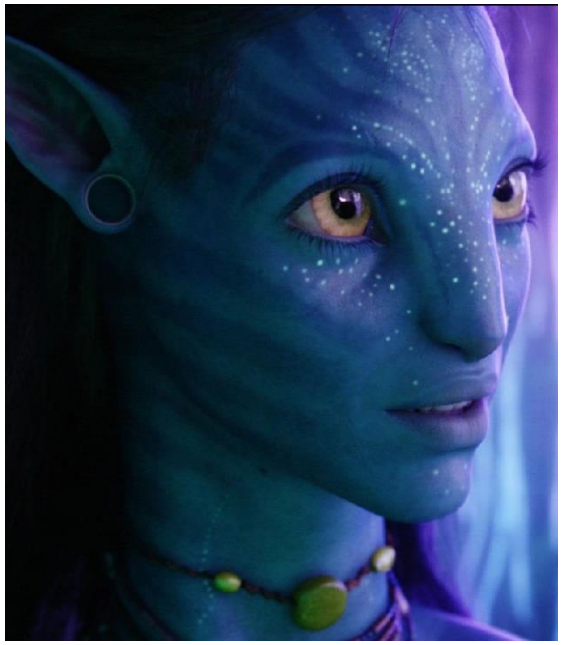
计算机木偶(Computer Puppetry)

Hyun Joon Shin, Jehee Lee, Sung Yong Shin, and Michael Gleicher. 2001. Computer puppetry: An importance-based approach. ACM Trans. Graph. 20, 2 (April 2001), 67-94.

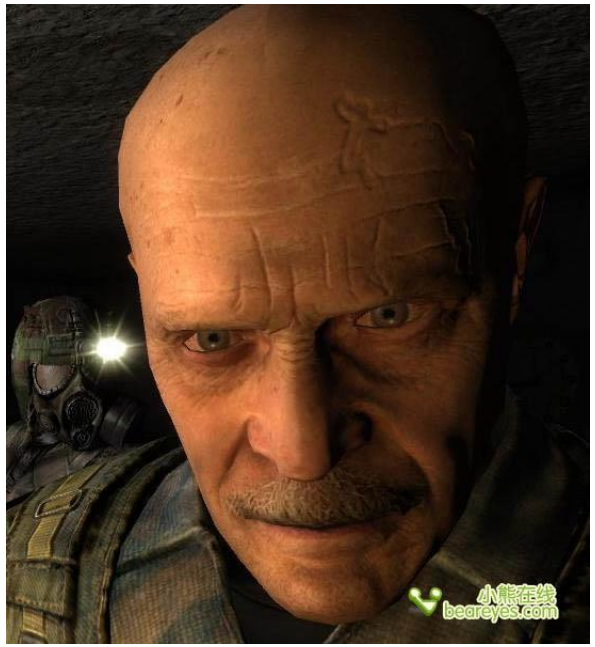
计算机木偶

人脸表情动画

- 人脸动画的应用领域



影视特效



游戏

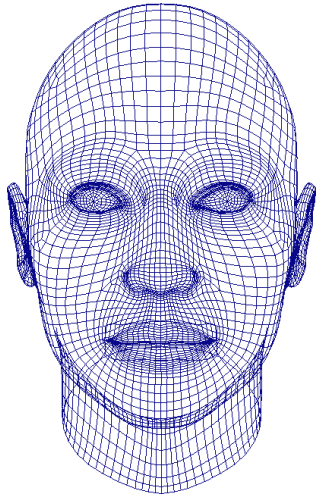
娱乐（Iphone X自制表情包）

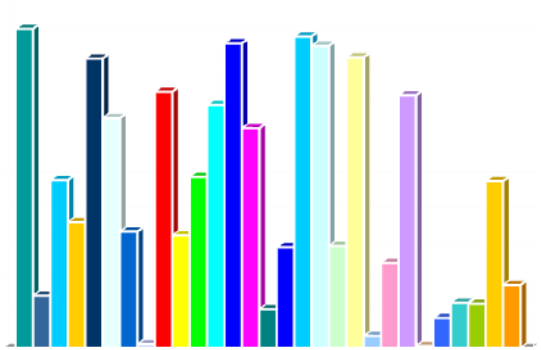
人脸表情动画

- 如何逼真地合成人脸动画?

人脸表情动画

- 逼真人脸表情合成
  - 人脸参数化(Rigging): $E=E(w)$。构造一个具有较低维度的并且具有直观视觉意义的参数空间，并给定参数空间到表情空间的映射关系。





N_{points}×3维
人脸表情空间(E)

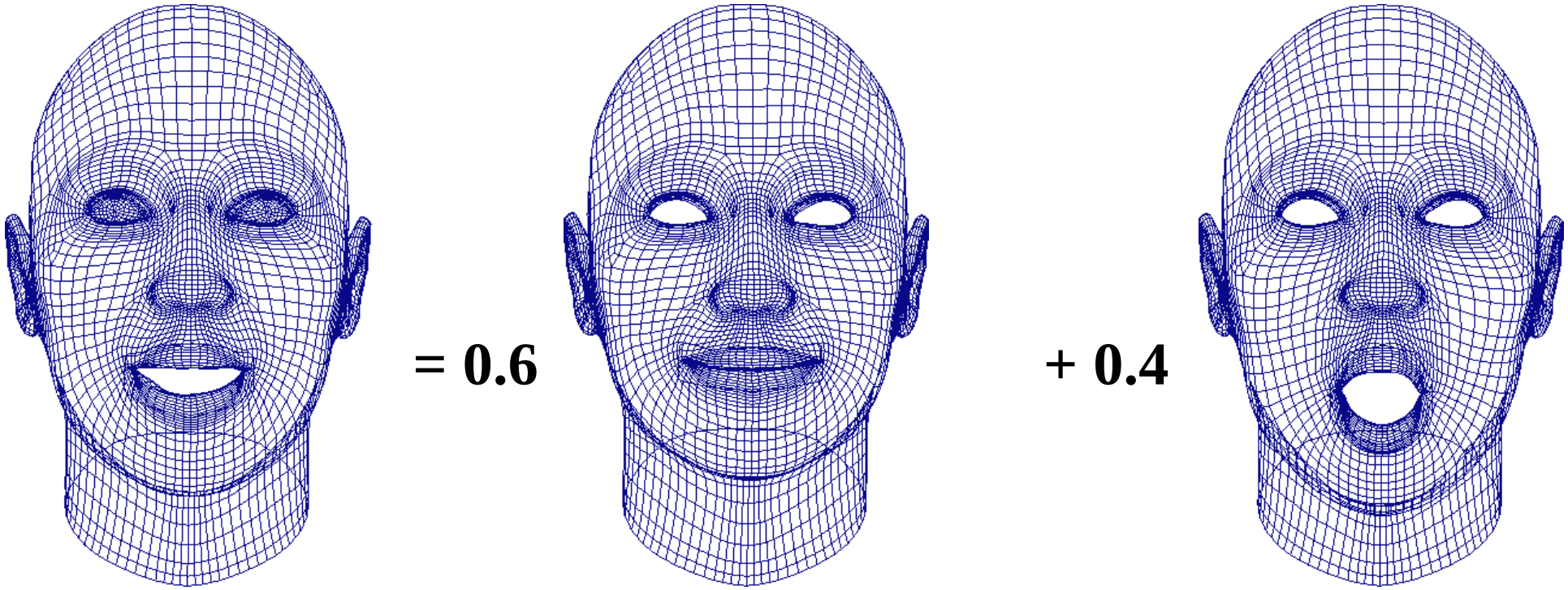
N_w×3维
人脸表情空间(w)

$N_{points} >> N_w$

  - 调整人脸表情参数，合成逼真人脸动画

人脸表情动画

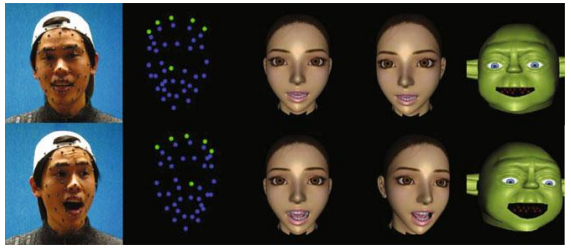
- 逼真人脸表情合成
  - 人脸参数化(Rigging)：
    - 基于Blendshape的人脸表情参数化（直观）



人脸表情动画

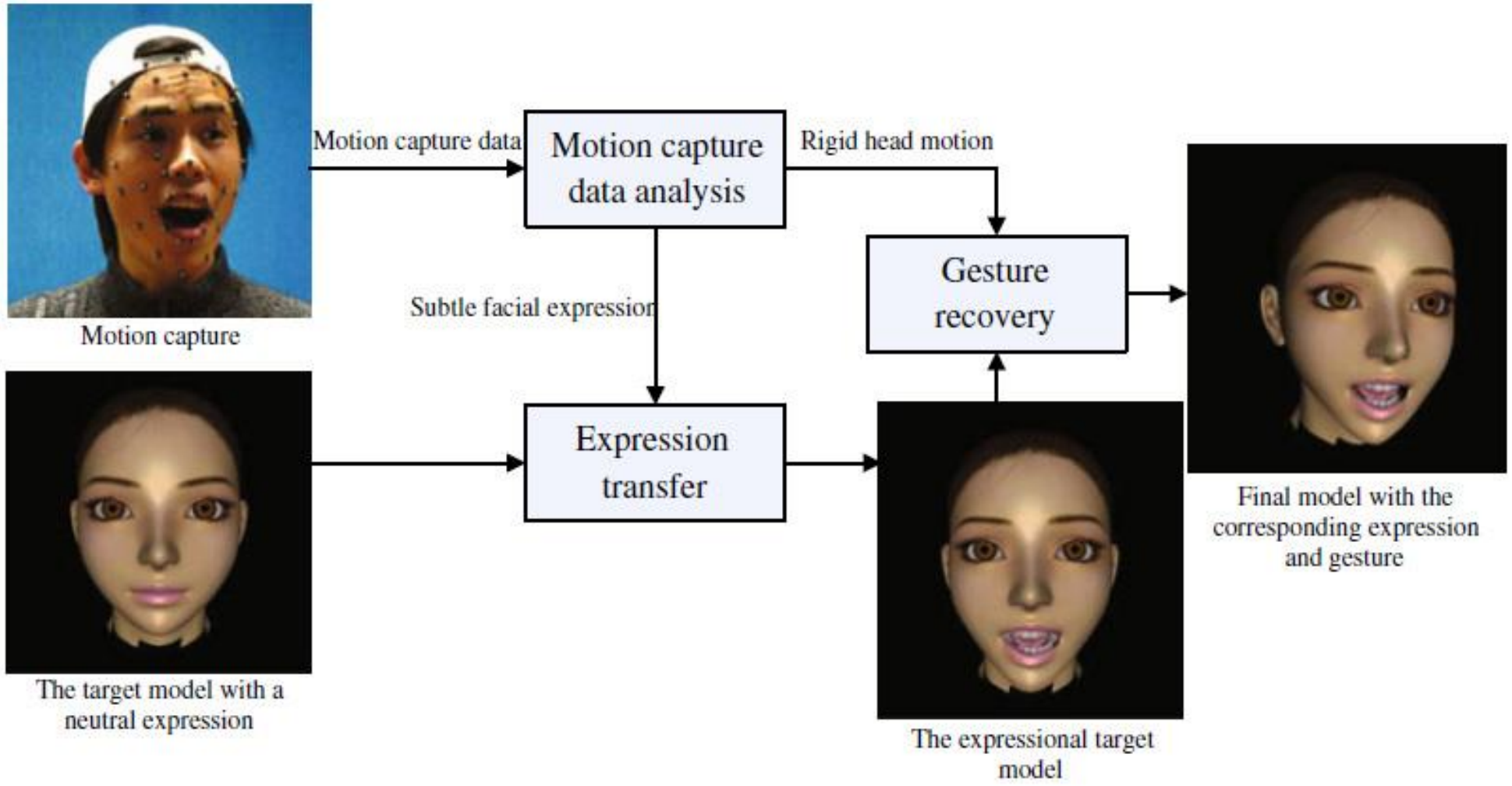
- 人脸表情参数化面临的挑战
  - “对于世界顶级动画工作室Weta而言，人脸表情的参数化是最具有挑战性的工作，需要反复的繁琐的手工操作”——James Cameron, [Cinefex]
  - “在影片《阿凡达》的制作过程中，Weta额外花费了一年的时间对原来设计的人脸表情参数化方法进行了反复的修正，以达到影片要求的效果”——Joe Letteri, [Cinefex]

基于拉普拉斯变形的人脸表情合成

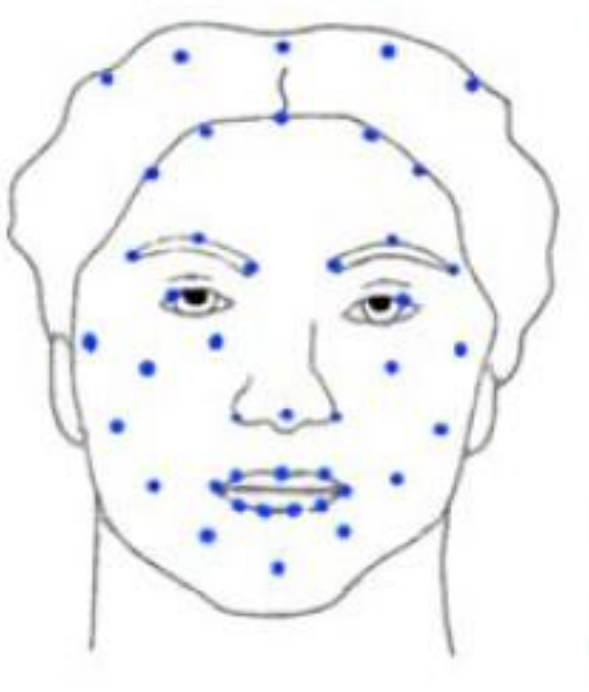


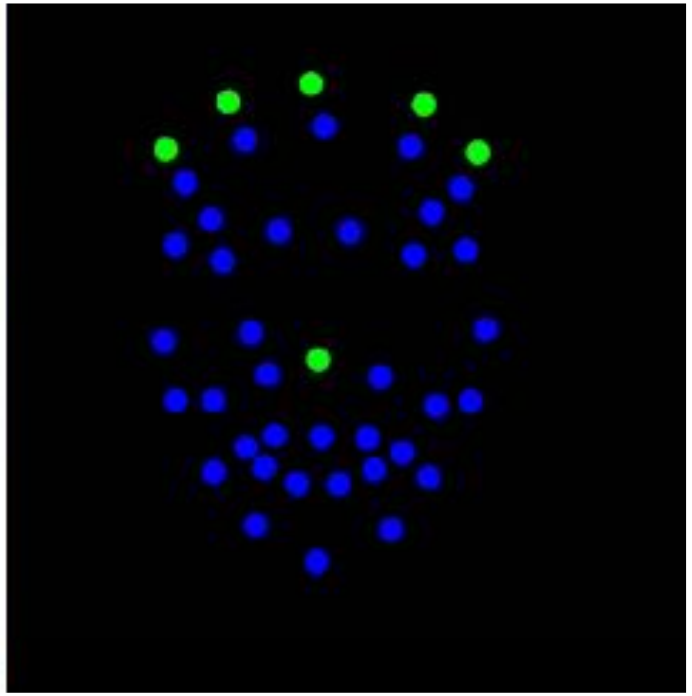
- 将采集的表演者真实人脸表情变化迁移到一个虚拟的中性目标人脸模型上。主要包括以下六个步骤：
  1. 使用运动捕获设备精确地采集表演者的脸部特征部位的运动；
  2. 将捕获的脸部特征部位的运动数据分解为脸部表情运动和头部刚体运动；
  3. 载入中性目标人脸模型；
  4. 根据中性目标人脸模型的几何结构与拓扑结构，计算中性目标人脸模型中各顶点的拉普拉斯坐标；
  5. 将分解所得的脸部表情运动迁移到中性目标人脸模型上，使得中性目标人脸模型具有与表演者一致的表情；
  6. 将分解所得的头部刚体运动迁移到具有与表演者一致表情的目标人脸模型上，使得最终的目标人脸模型具有与表演者一致的面部表情与头部姿态。

基于拉普拉斯变形的人脸表情合成框架



人脸表情捕获





我们放置5个标记点在表演者的前额 、1个标记点在鼻尖来记录头部运动

采用VICON光学运动捕获系统，基于MPEG-4标准定义，来获得高逼真的脸部运动数据

脸部运动数据预处理——脸部运动分解

- 头部刚体运动 $\mathbf{A}_t$ (包括旋转 $\mathbf{R}_t$ 和平移矩阵 $\mathbf{T}_t$) 和脸部表情 $\mathbf{F}_t'$ 之间的关系可用下面的公式来表示:

$$\mathbf{F}_t = \mathbf{R}_t \mathbf{F}_t' + \mathbf{T}_t$$

$\mathbf{F}_t = \{\mathbf{P}_t^1, \cdots, \mathbf{P}_t^{42}\}$ ：刚体标记点的位置

脸部运动数据预处理——脸部运动分解

- 但是，在运动捕获过程中，刚体标记点会有微小的非刚体运动

- 因此，我们在最小二乘意义上计算近似**最优的刚体变换**$\mathbf{A}_t$。对应于求解如下问题的极小化问题：

$$\arg\min_{\mathbf{A}_t} \sum_{i=1}^{6} \left\| \mathbf{A}_t\left(\mathbf{P}_{t-1}^{i}\right) - \mathbf{P}_t^{i} \right\|^2$$

Demo

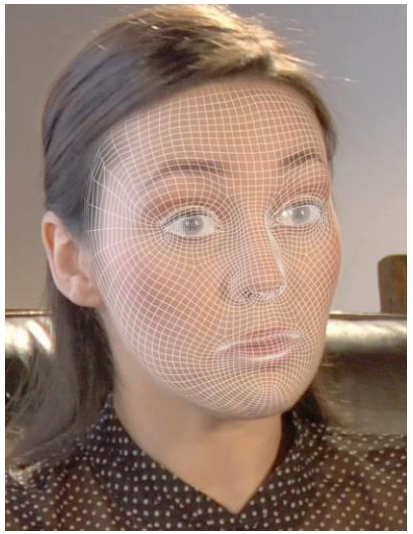
Facial Expression Synthesis
using Laplacian Deformation

The Digital Emily Project:

——真实感人脸建模和动画

- Paul Debevec, University of Southern California
  http://ict.debevec.org/~debevec/

- 照片真实(photo-real?)
  - 人脸在表现情绪上非常重要，不真实的人脸会使得观众分心
  - 无缝结合真实和虚拟的角色
  - 数字替身(digital doubles)：演员无法直接拍摄
  - 操纵人的年龄：甚至把死去的演员复活
- 真实感人脸动画



动画流程

从视频数据中抽取
演员的表情和运动

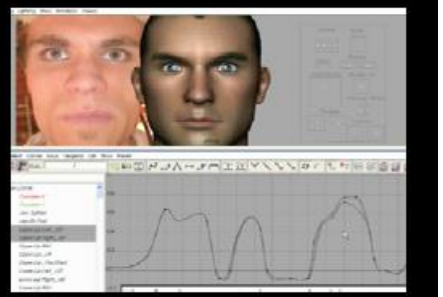
前述过程整合到动
画师的工作流程，
需要时可用关键帧
动画进行编辑

表演捕获

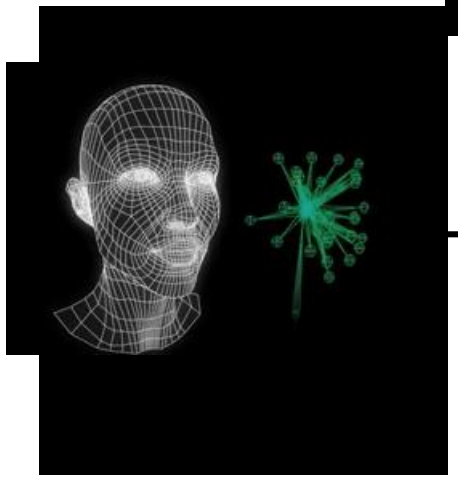
表演分析

Retargeting

关键帧动画



Rigging

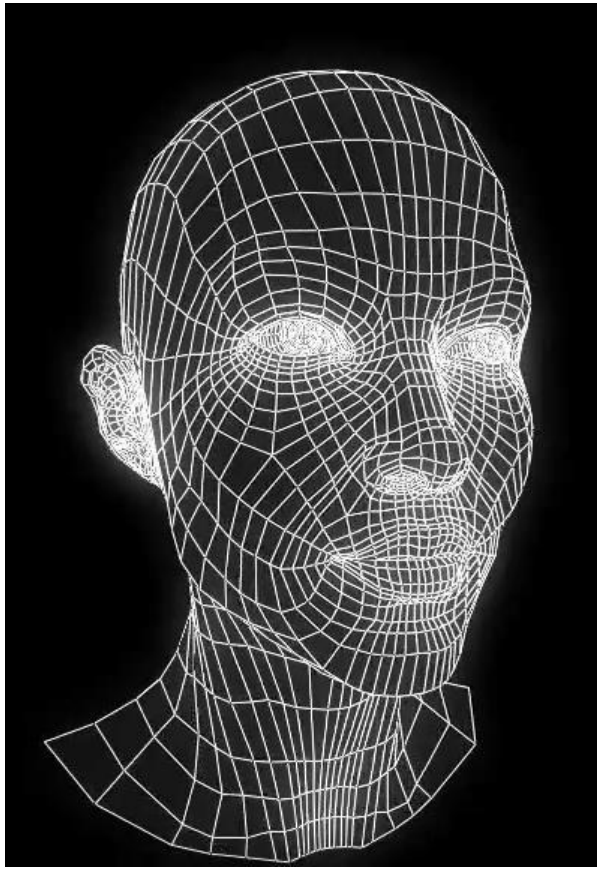


建立一个可变
形的三维角色

把表情迁移到CG角色

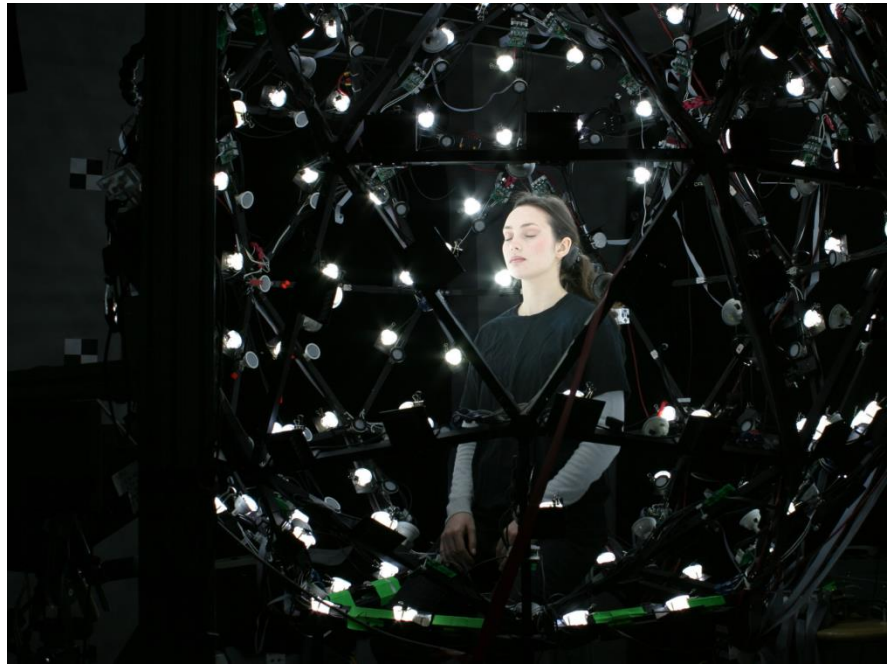
Facial Rigging

- **挑战性的问题：**
  - 专业级的艺术
  - 真实的视觉表现
    - 绘制后，结果应真假难辨
  - 真实的运动
    - 运动必须真实，能涵盖绝大部分人脸表情



三维模型的获取

- 为了生成照片真实的三维模型，需要角色的高精度三维扫描模型
  - ✓ 三维几何
  - ✓ 纹理
  - ✓ 人脸皮肤的反射属性
  - ✓ 各种各样的表情

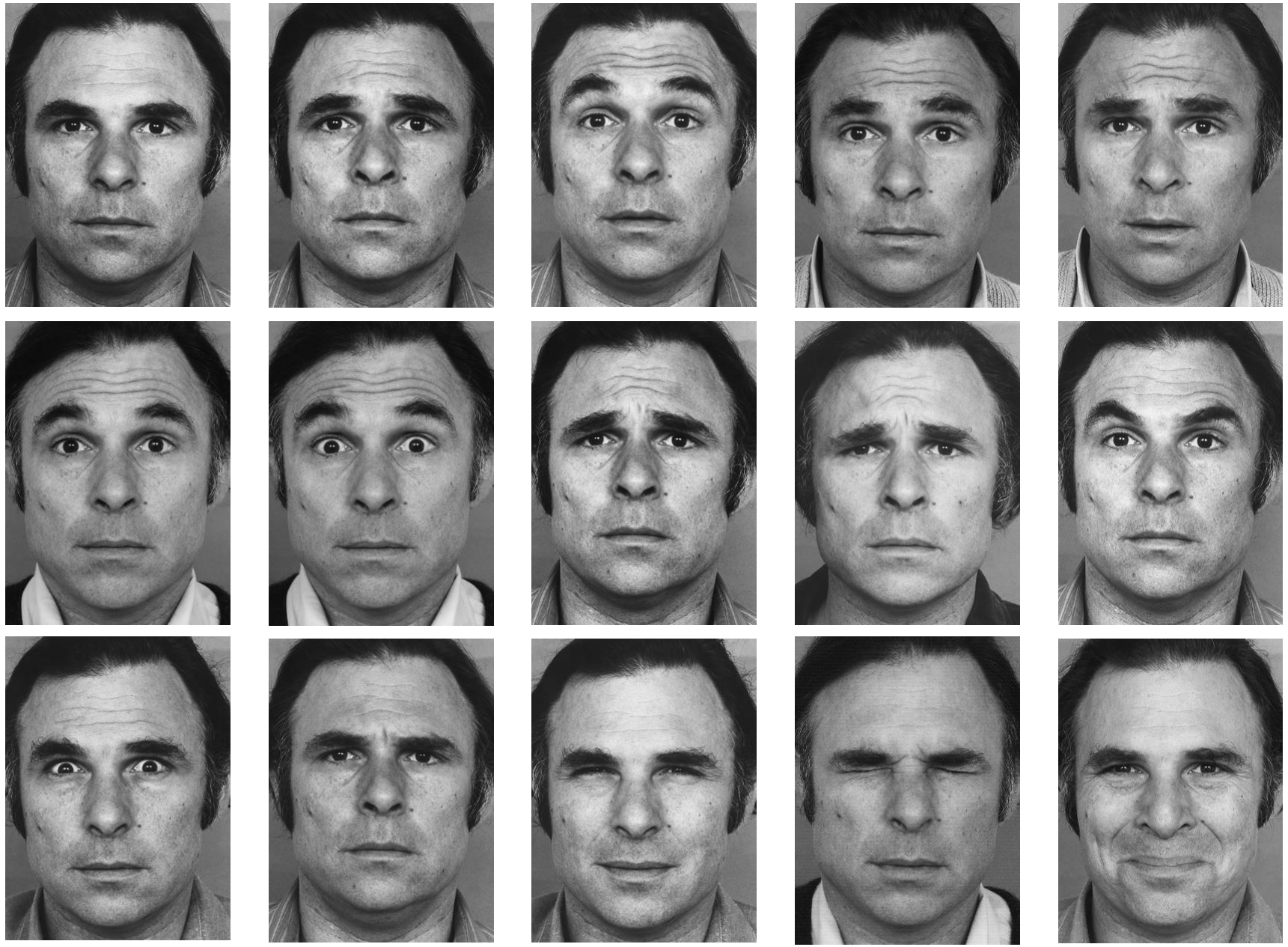


面部动作编码系统FACS

——Facial Action Coding System

- 采用由Paul Ekman提出的面部动作编码系统FACS
- 该系统最初的设计用途是：给心理学家提供一种表达人脸表情的通用语言
- 该系统通过一系列运动单元编码了所有可能的人脸动作，由这些单元可生成所有的人脸运动
- 因为它们很好地对应了人脸下层的肌肉运动，可用这些单元来设计人脸的绑定

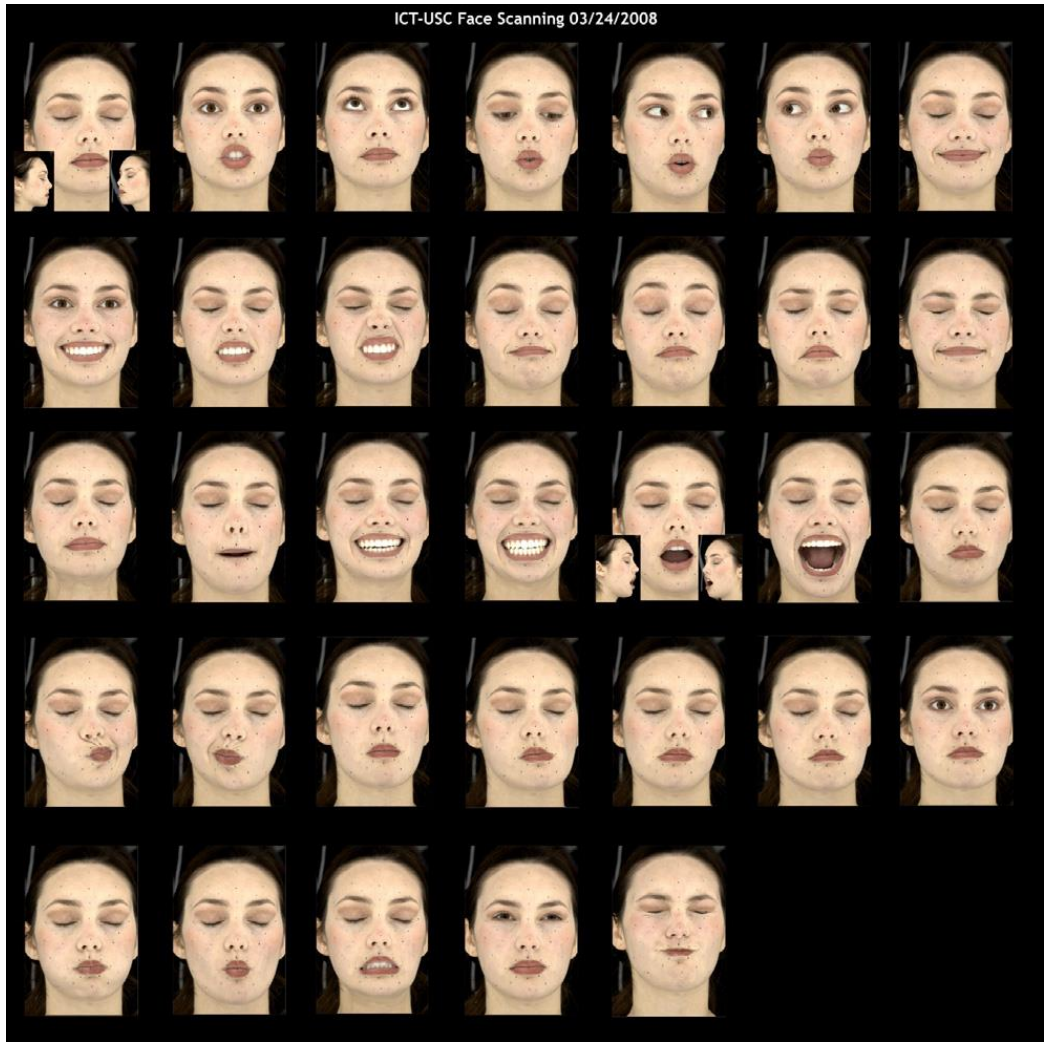
一些典型的FACS单元



面部动作编码系统FACS

——Facial Action Coding System

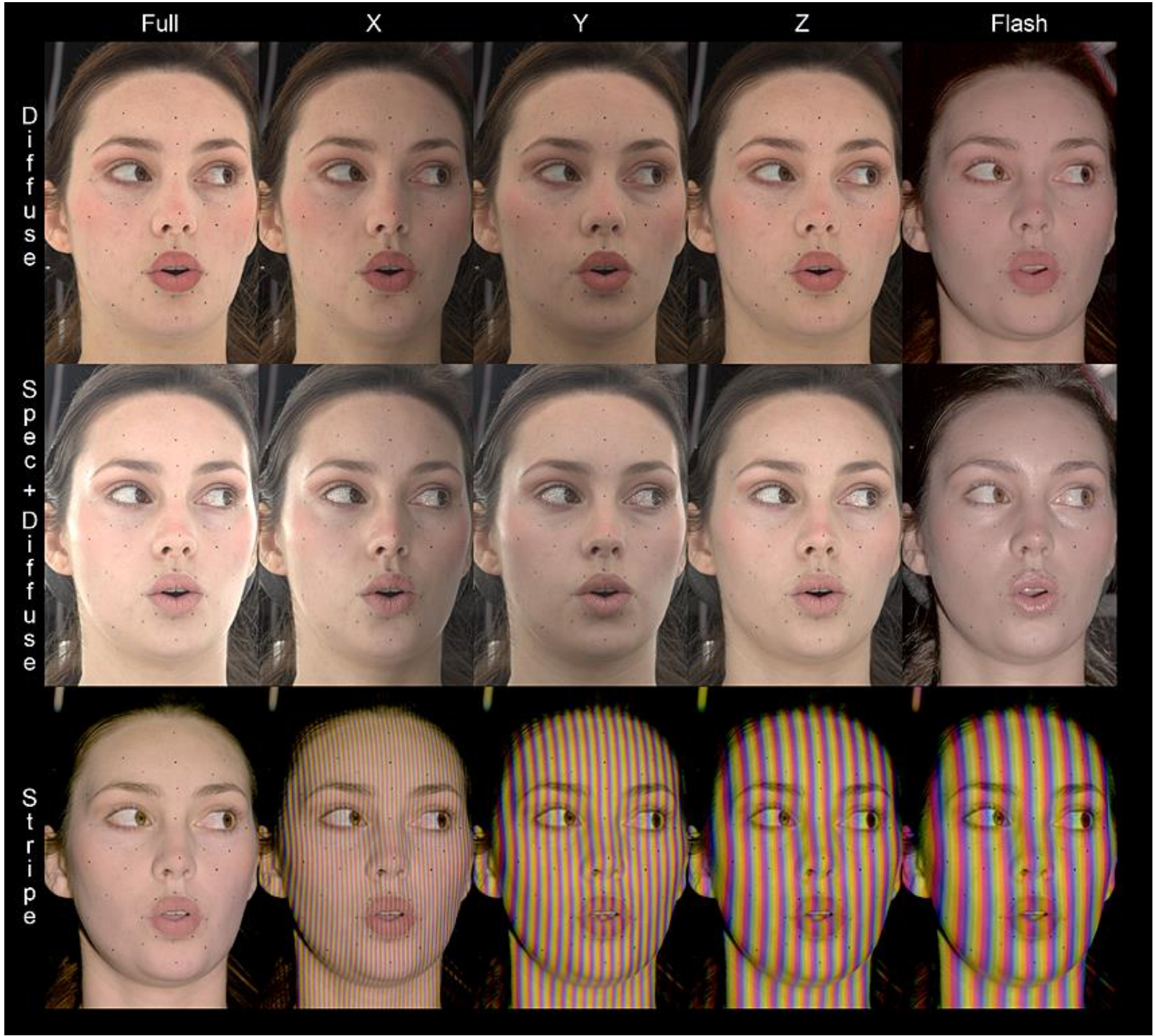
- 为了建立Emily模型，基于FACS系统，我们定义了一系列表情(33个)；
- 在捕获过程中，每个表情保持几秒钟；



女演员：Emily O’Brien
《不安分的青春(The Young and the Restless)》中Jana Hawkes的扮演者

Light Stage 5, USC ICT

156 white LED light sources



黑点用来对齐
(因为人会动！)

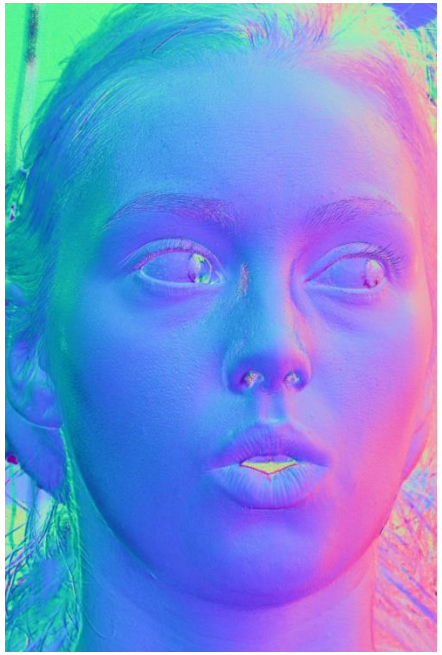
Capture time:
3 seconds

每个表情拍摄15张

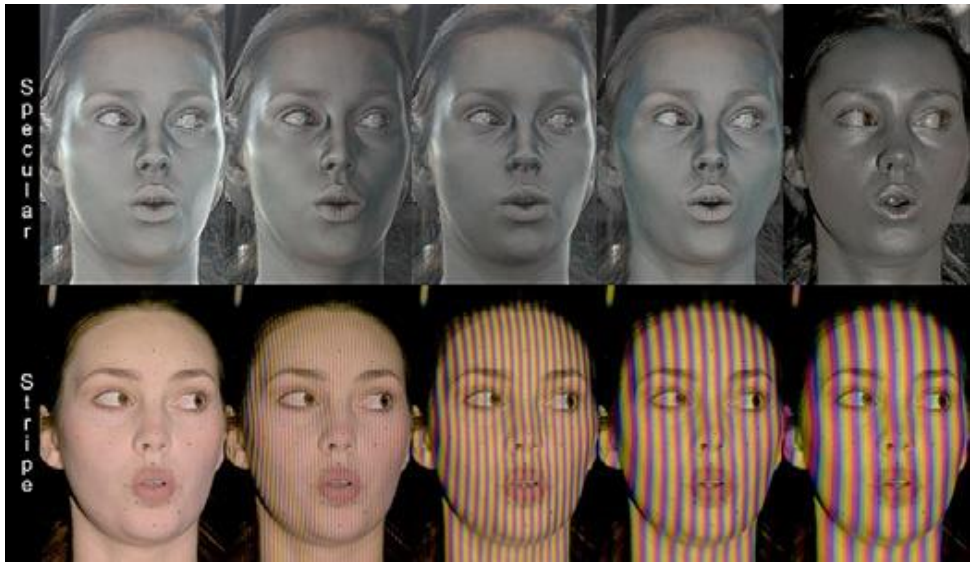
2K / 4K resolution

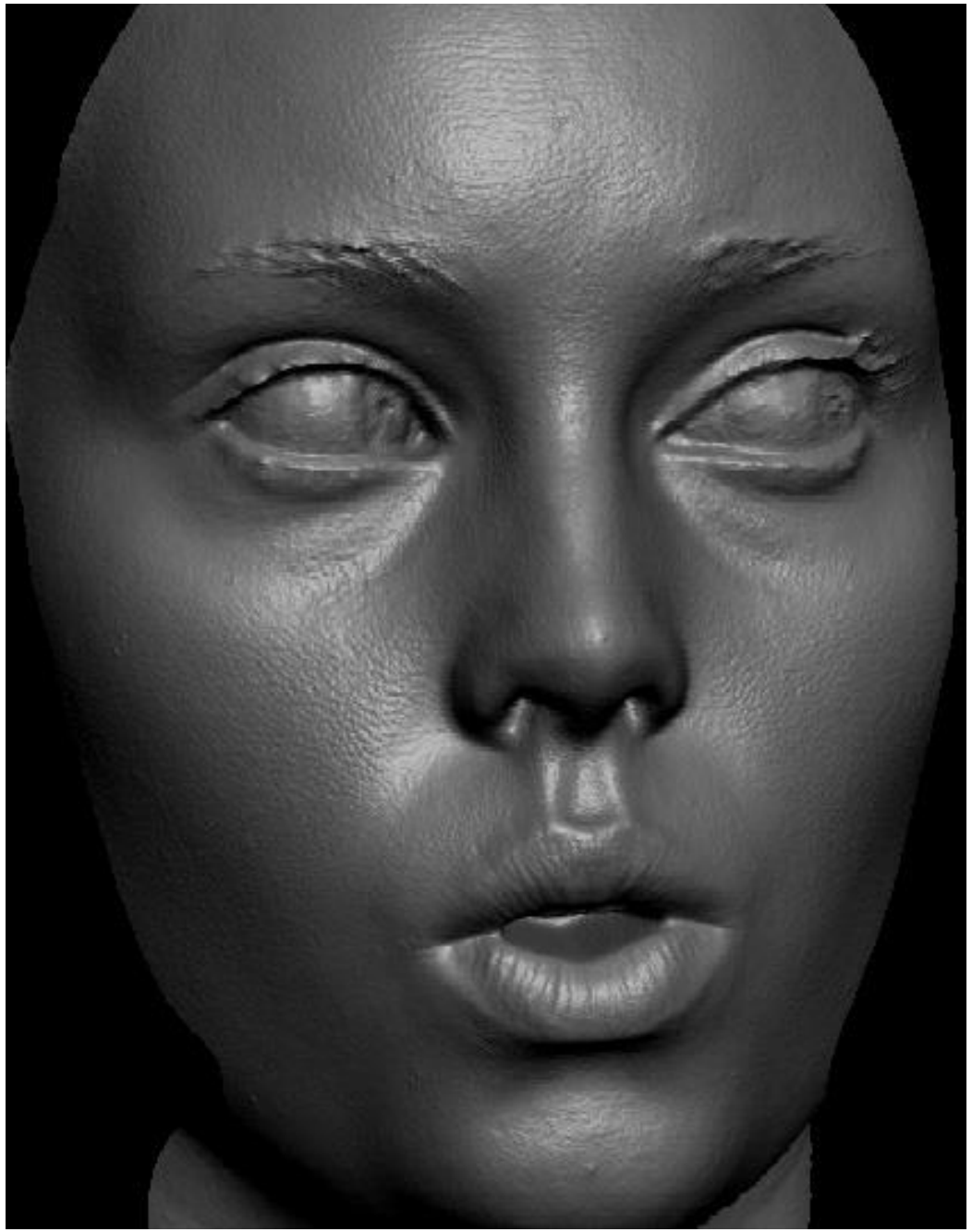
Diffuse (low-res) geometry

从扫描的这组图像中，可以鲁棒地恢复左右眼视域像素的对应关系，并根据相机标定，可重建Emily的人脸几何模型



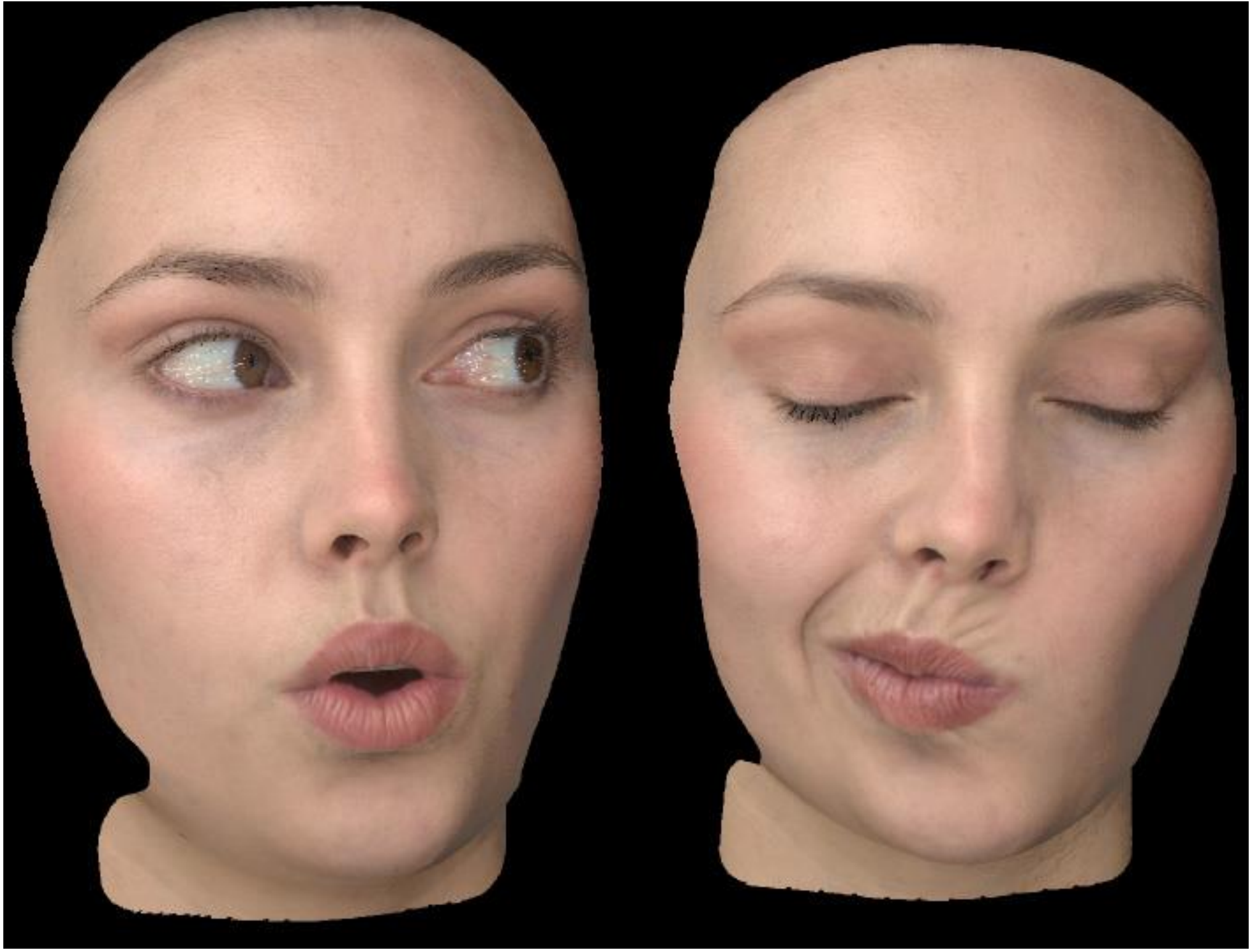
Specular (high-res) geometry



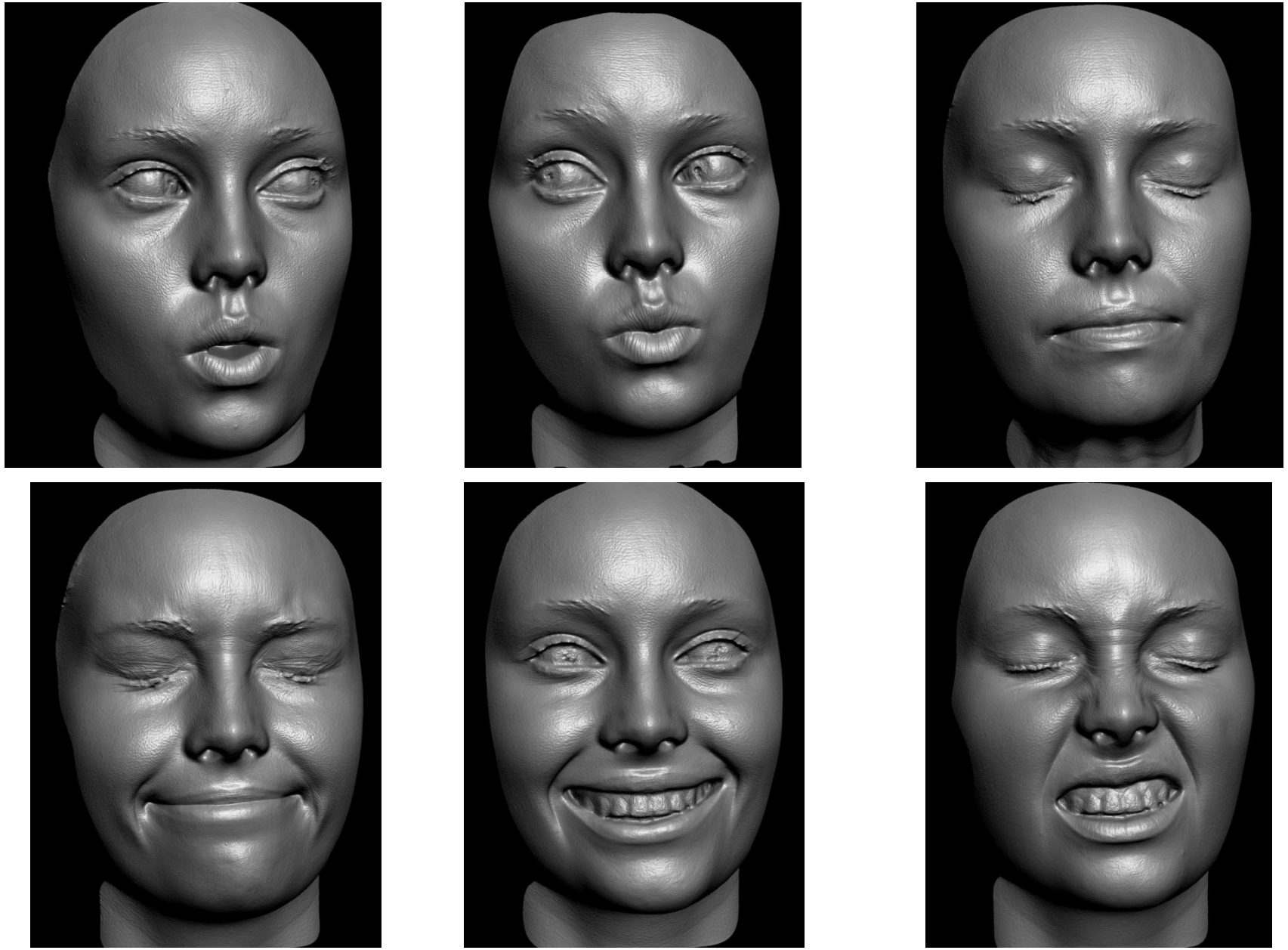


Diffuse texture

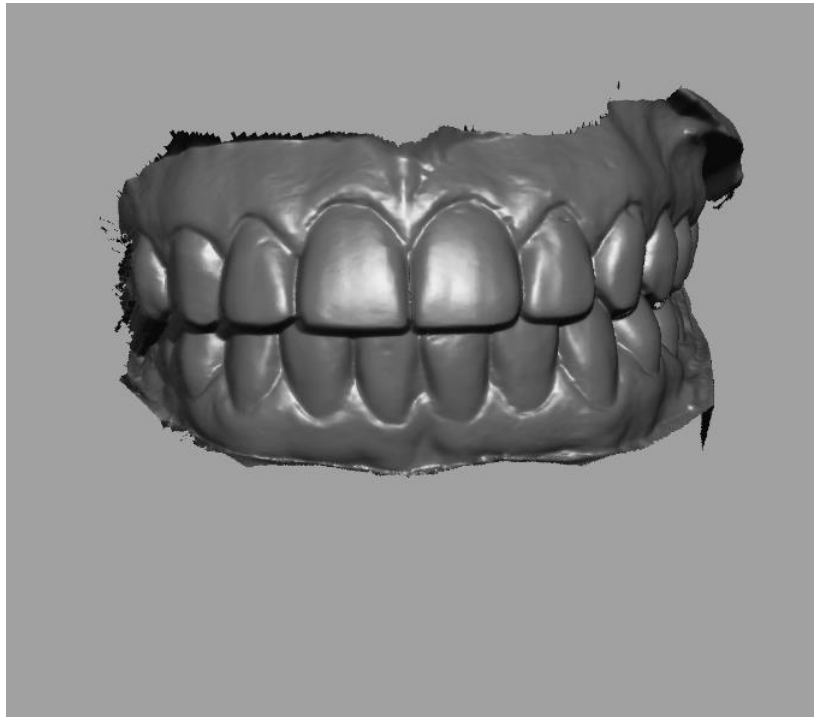
Hybrid normals

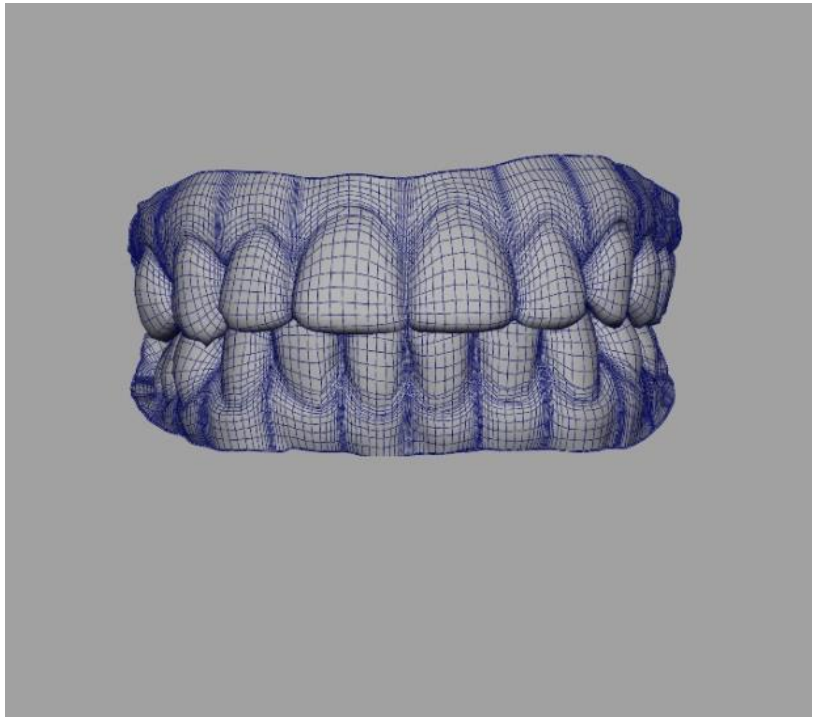


Emily的一些FACS单元



牙齿的处理(Teeth)





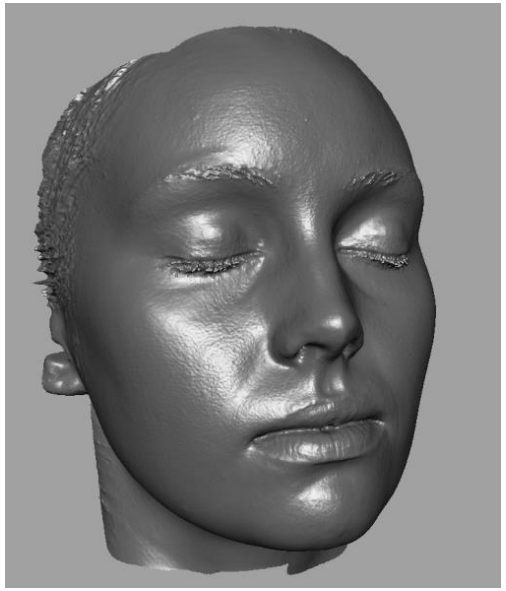
角色构建(Character Construction)

- 目标：用一系列的三维扫描表情模型构建一个绑定好的人脸

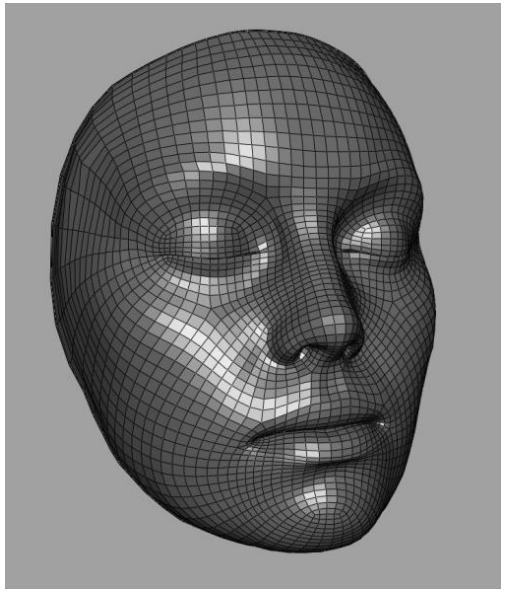
- 挑战性的问题：
  - 网格模型存在缺陷
  - 不一致的扫描覆盖
  - 缺少对应关系
  - 物理上未对应的区域

角色构建(Character Construction)

- 扫描得到模型的网格拓扑结构为任意的，而且每次扫描结果也是不一样的。
- 我们需要对中性的扫描表情进行重新网格化(Re-mesh)以适合于绑定，重新网格化由一个专业美工来完成。
- 然后需要定义该中性网格模型和每个不同表情的扫描模型之间稠密的对应关系，从而建立Blend Shapes。
- 中性模型的网格分辨率通常比扫描模型的分辨率低很多。其精细细节通过纹理和带动画的位移贴图来恢复。



中性的扫描表情

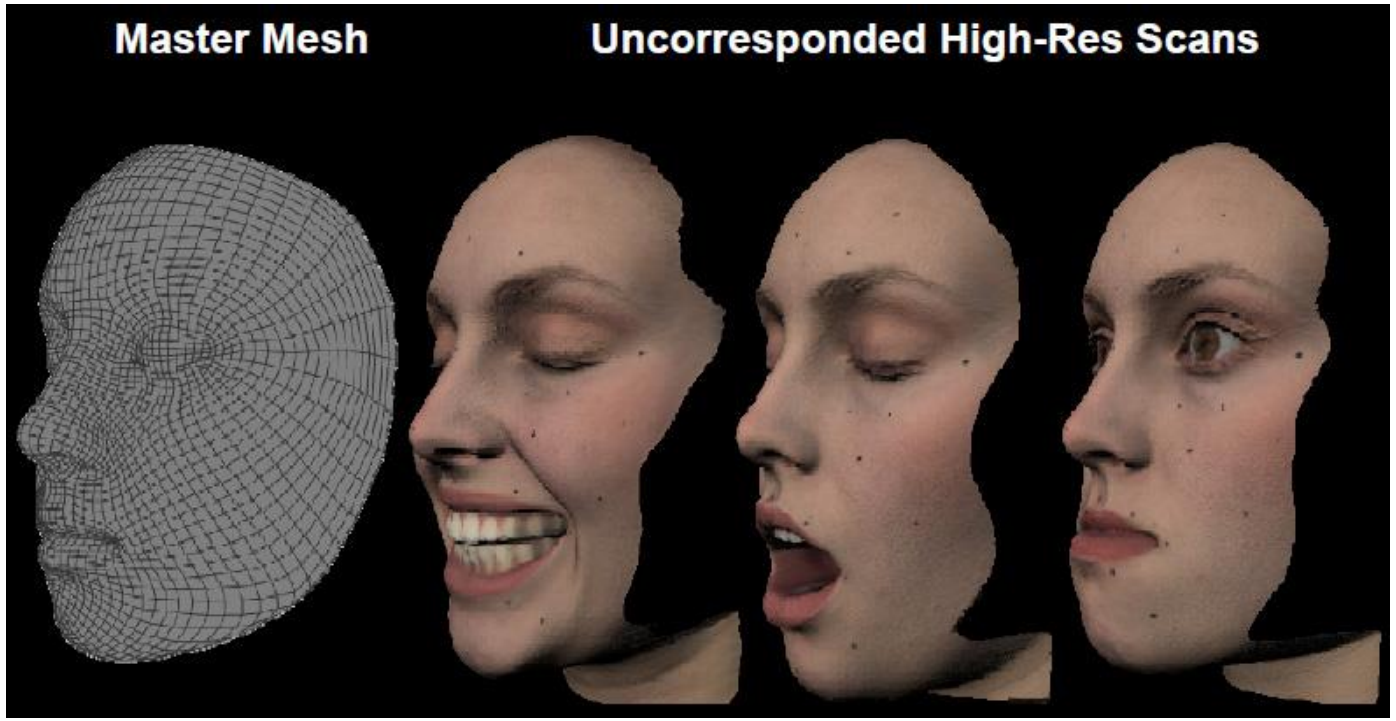


4000 polys done in zBrush/Maya

Re-mesh后的中性网格模型——主网格

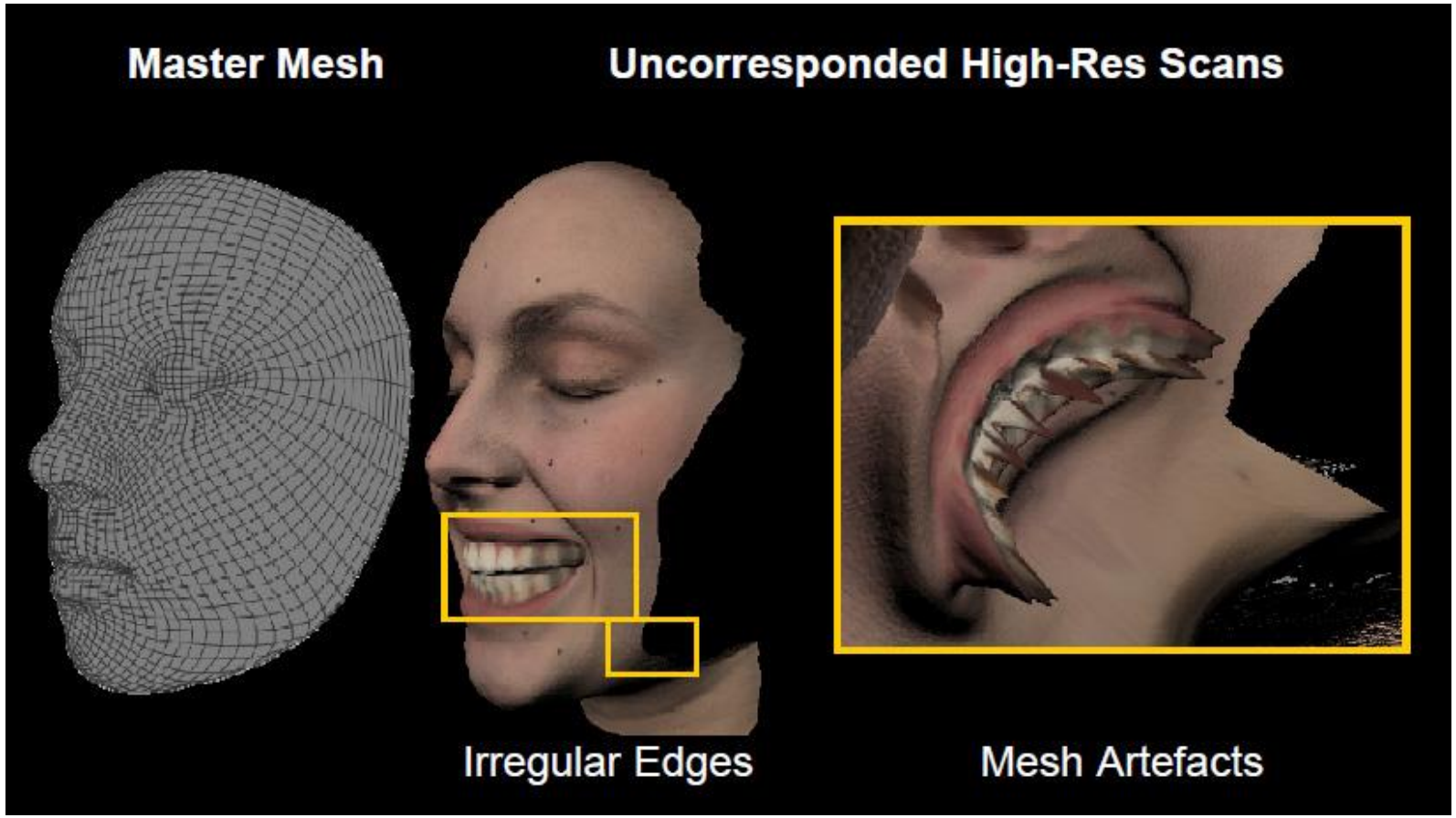
角色构建——Blend Shapes模型的构建

- 如何从扫描得到的原始模型建立Blend Shapes?
  - Mater Mesh与扫描模型会有较大的偏差
  - 我们需要把每个扫描模型映射到手工建立的主网格上
  - 最重要的是找到主网格每个顶点和扫描表情某个点的对应关系



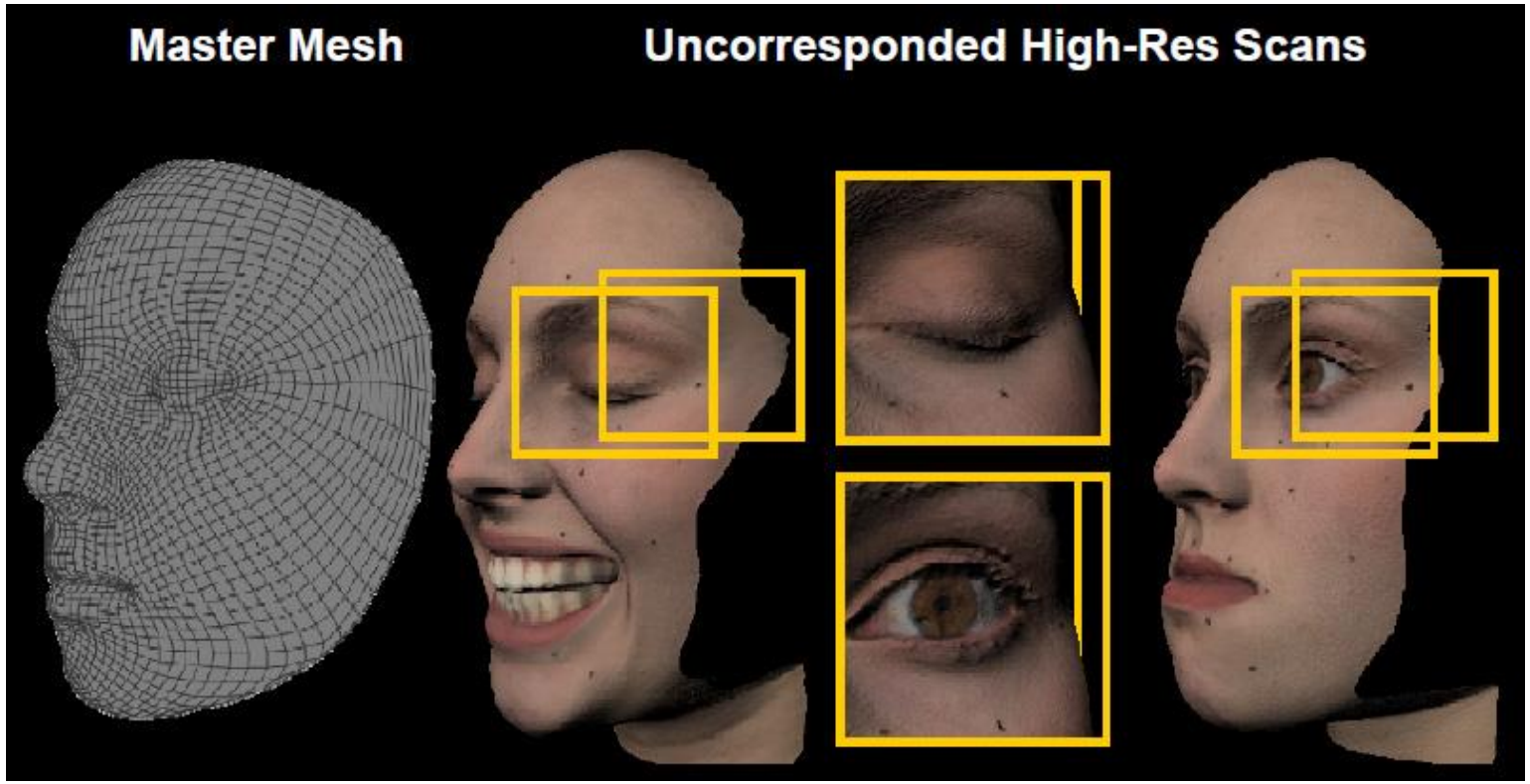
角色构建——Blend Shapes模型的构建

- 必需处理模型的缺陷
  - 模型的边界可能不规则、较差的三角化;
  - 牙齿和眼睛区域的缺陷;
  - 缺陷使得建立模型之间的对应关系相当困难;



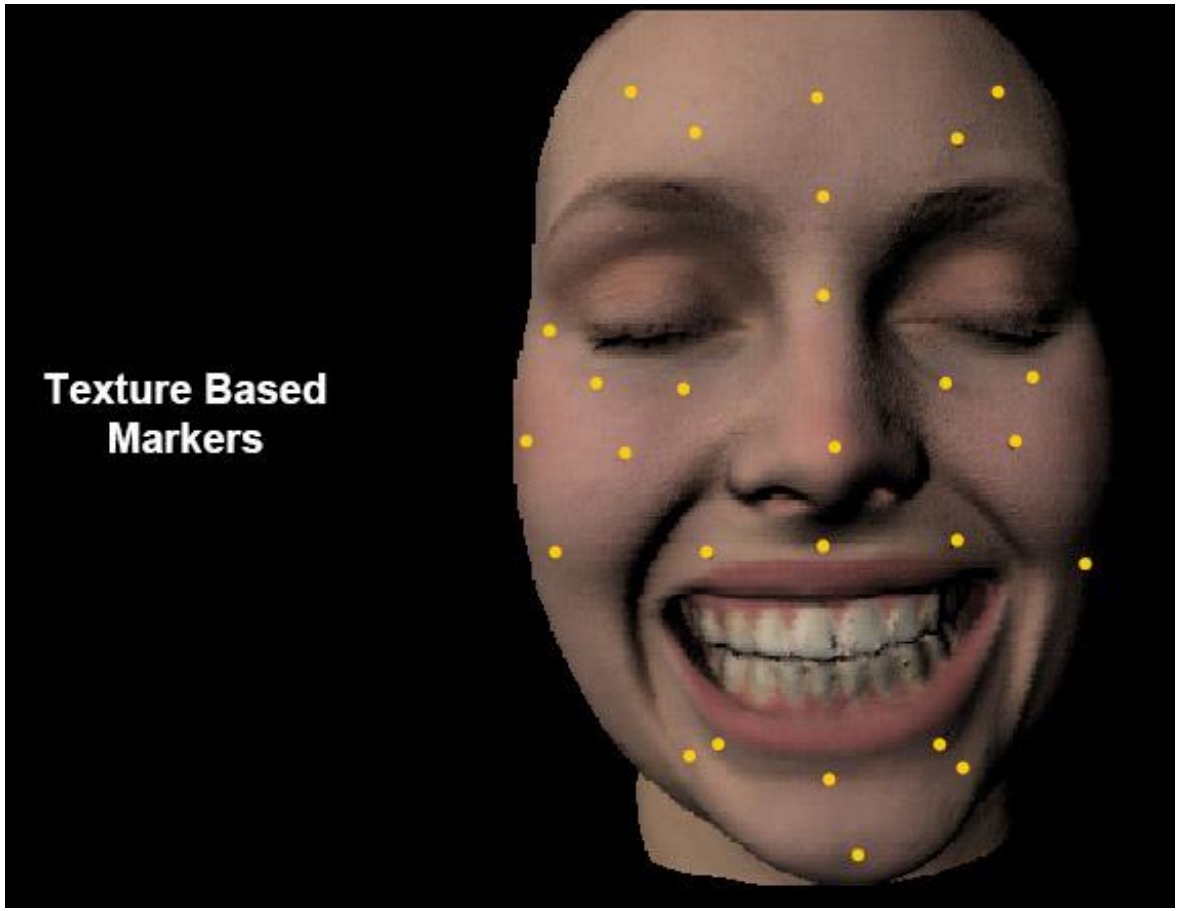
角色构建——Blend Shapes模型的构建

- 扫描模型会有不一致的覆盖
  - 注意对应的黄框区域;
  - 主模型没有脖子信息，而扫描模型有;
  - 眼睛、眼帘区域，牙齿区域;



角色构建——Blend Shapes模型的构建

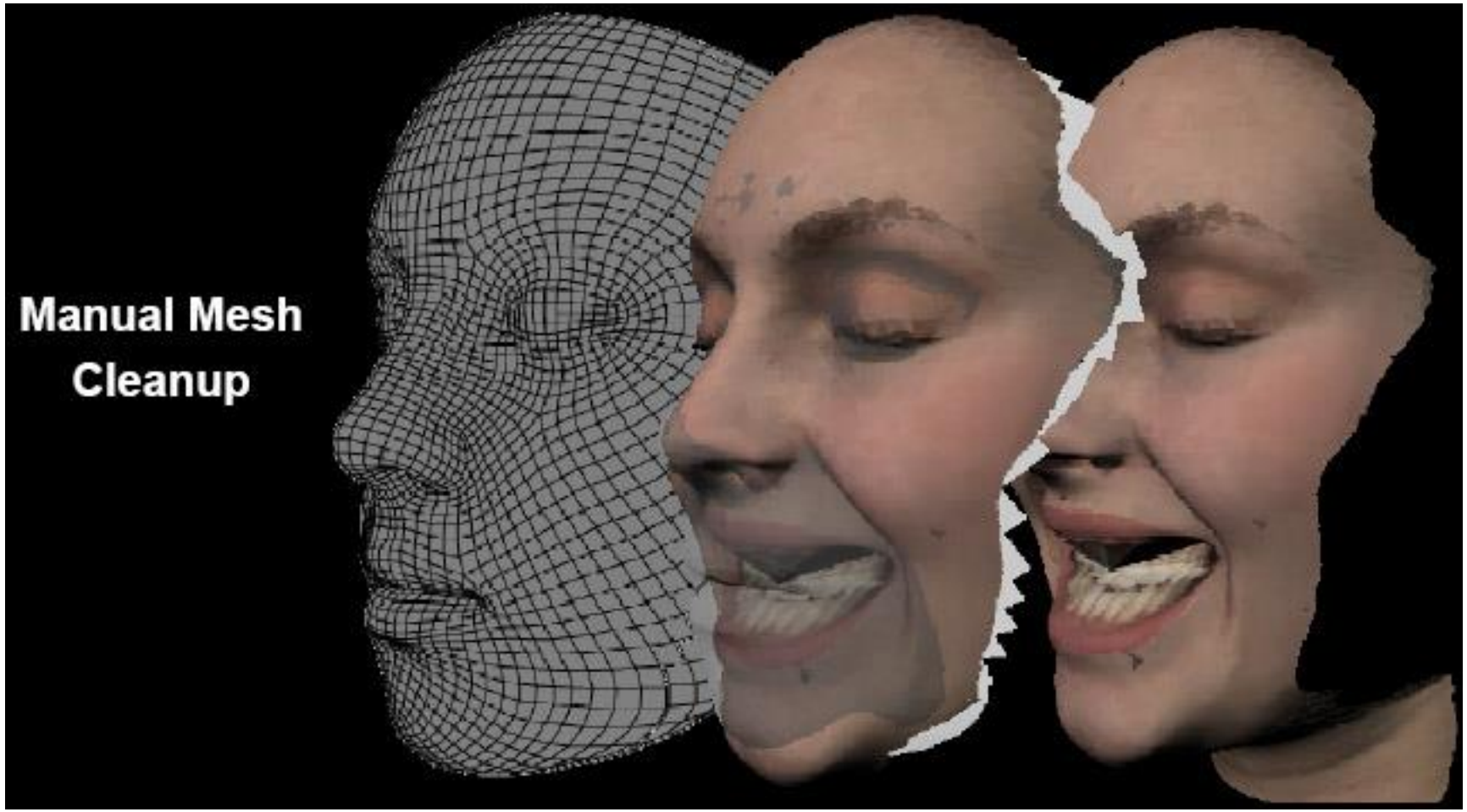
- 通过在Emily脸上设置标记的方法来解决这个问题！



为了增加可见性：黑点已经用显目的大黄点来表示

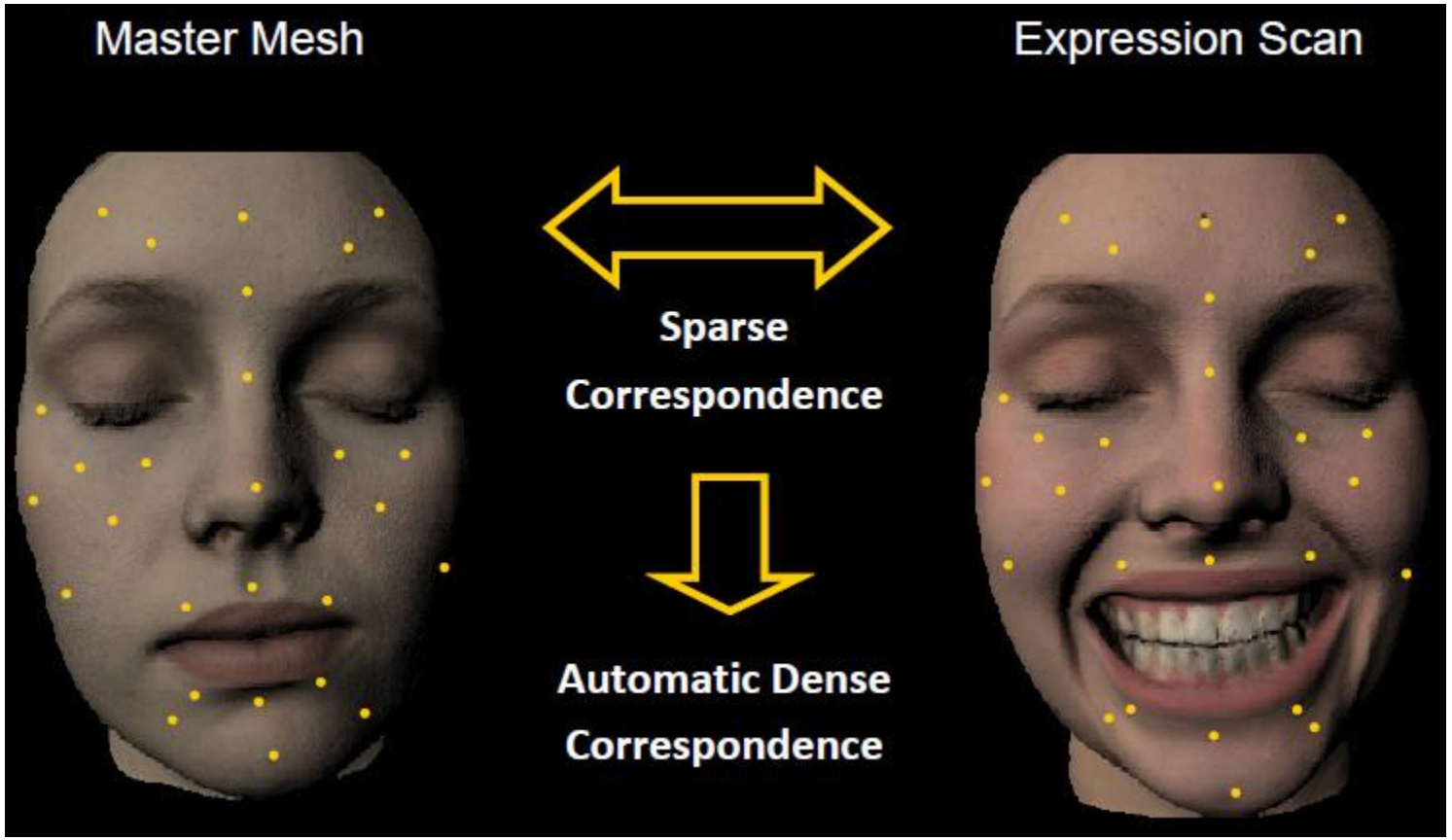
角色构建——Blend Shapes模型的构建

- 部分区域通过手工清除



角色构建——Blend Shapes模型的构建

- 扫描模型清洗完毕后，需要对所有顶点建立稠密对应关系
  - 标记点给出了Mesh之间的稀疏对应关系
  - 然后通过一个自主开发的方法建立稠密对应关系：基于三维空间位置+法向一致性+2D映射



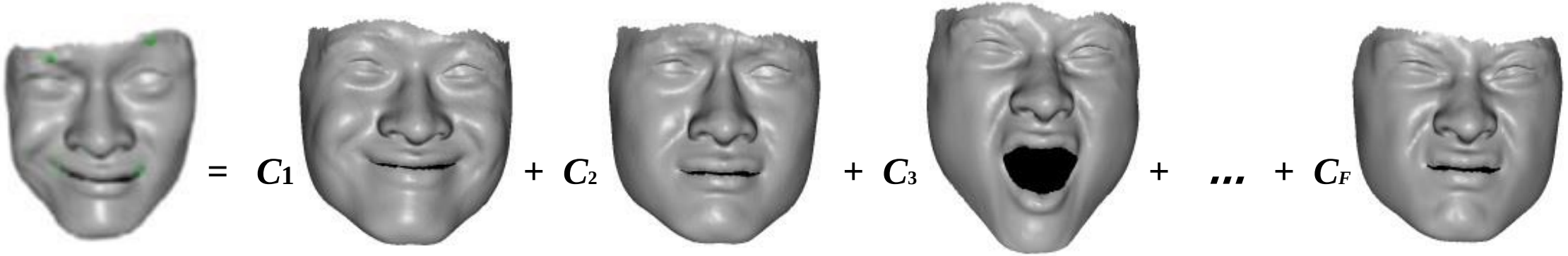
基于Blend Shapes的动画

- 关键姿态表情的形状——Blend Shapes
- 把这些基本的人脸Shape作为**表情基**

(a) (b) (c) (d) (e)

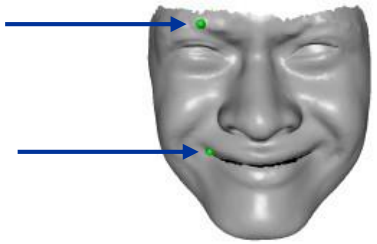
基于Blend Shapes的动画

- 根据插值这些Blend Shapes来得到任意的表情

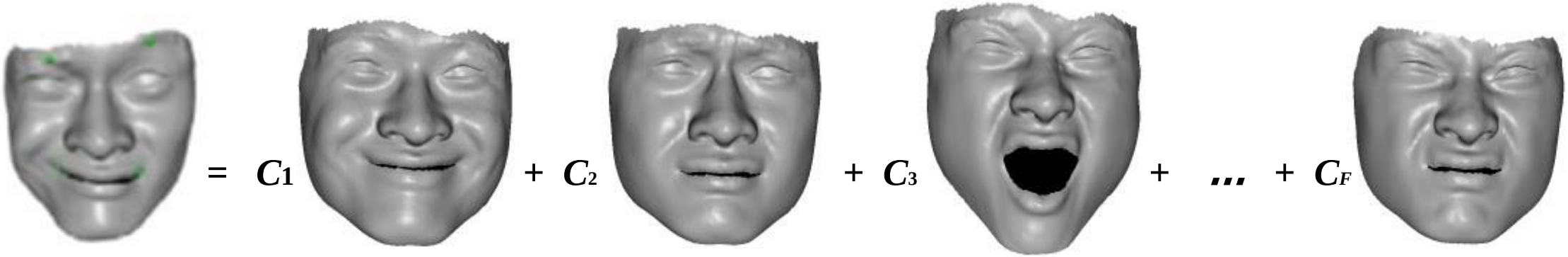


其中：$C_1 + C_2 + C_3 + \ldots + C_F = 1$

FACE IK



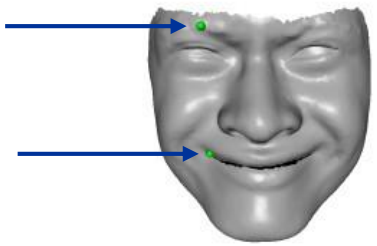
- 动画师更愿意控制脸部上的一些点（类似于IK），而不是权因子



- $\mathbf{S}_{i,f}$: **Blend Shapes**网格*f*的第*i*个顶点
- $\mathbf{S}_i$ ：最终人脸的第*i*个顶点（一个**Blend Shape**有*n*个点）

$$\mathbf{S}_i = C_1\mathbf{S}_{i,1} + C_2\mathbf{S}_{i,2} + \ldots + C_F\mathbf{S}_{i,F}, i = 1, \ldots n; \qquad \sum_{f=1}^{F} C_f = 1$$

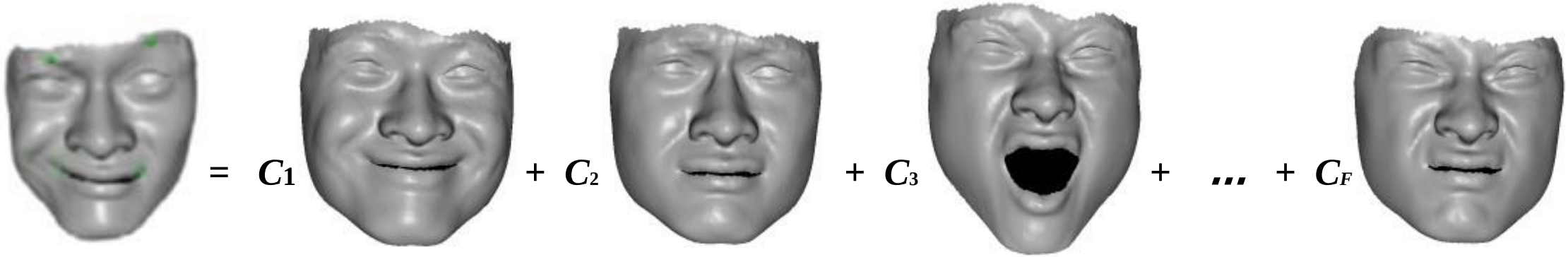
FACE IK



- 我们希望约束L个点（由动画师控制），使得顶点l的位置为$\mathbf{P}_l$，

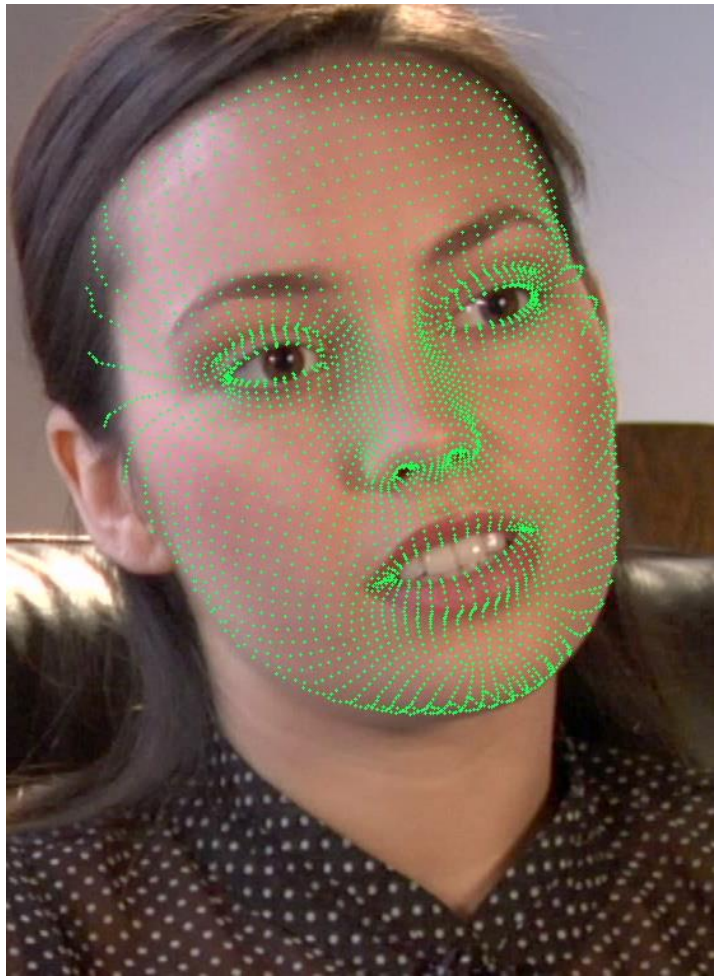
$$\sum_{f=1}^{F} C_f \mathbf{S}_{l,f} = \mathbf{P}_l, l = 1, \ldots, L, \text{ 其中} \sum_{f=1}^{F} C_f = 1$$

- 采用最小二乘法求解系数$\mathbf{C}_f$

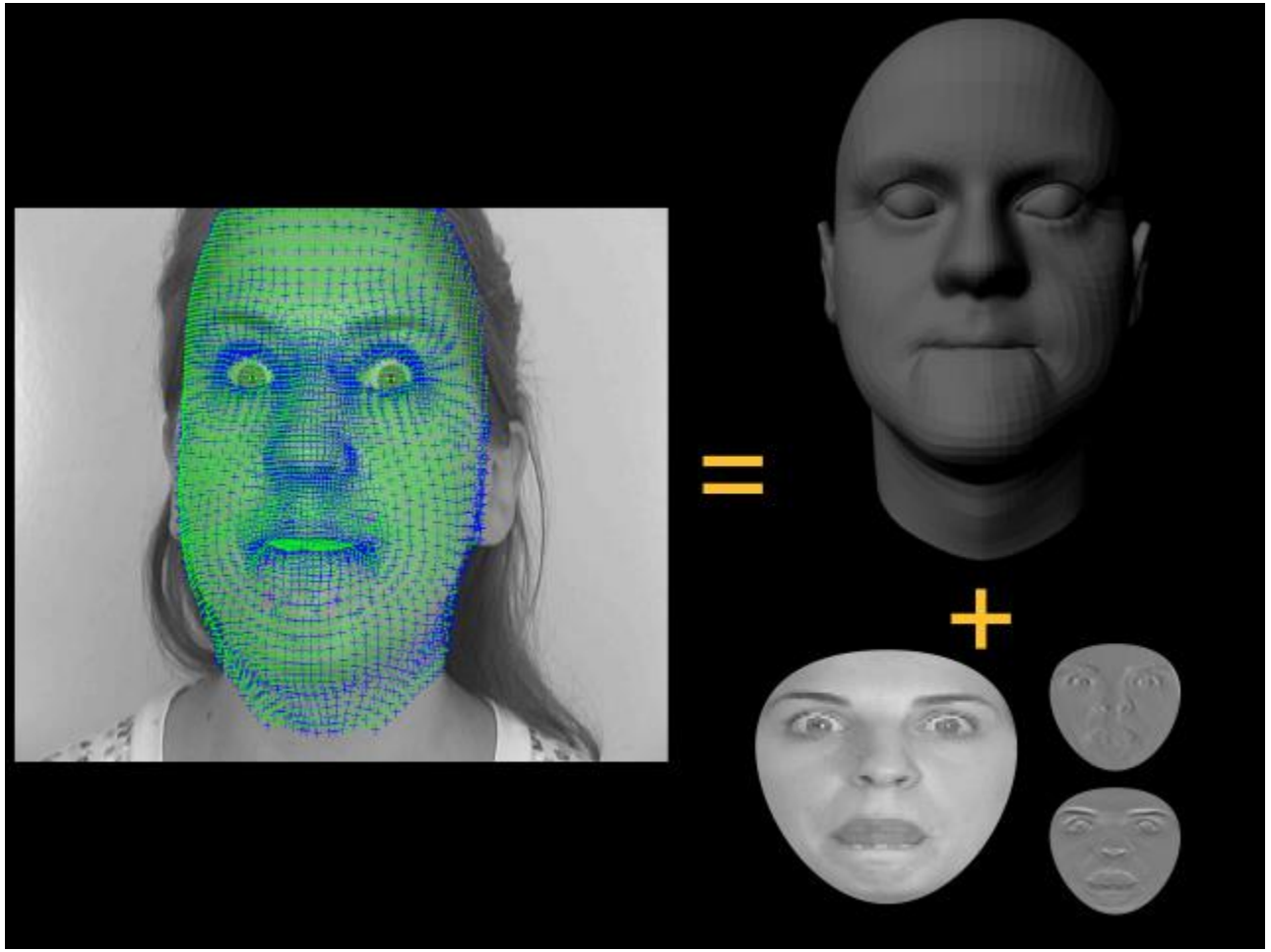


表演分析(Performance Analysis)

- 在捕获演员的表演后，采用计算机视觉的方法进行分析
  - 跟踪脸部特征
  - 把角色模型拟合到数据
  - 逐个像素分析



表演分析(Performance Analysis)



蓝色的点表示角色的三维结构，而绿色的点表示从图像上跟踪出来的点

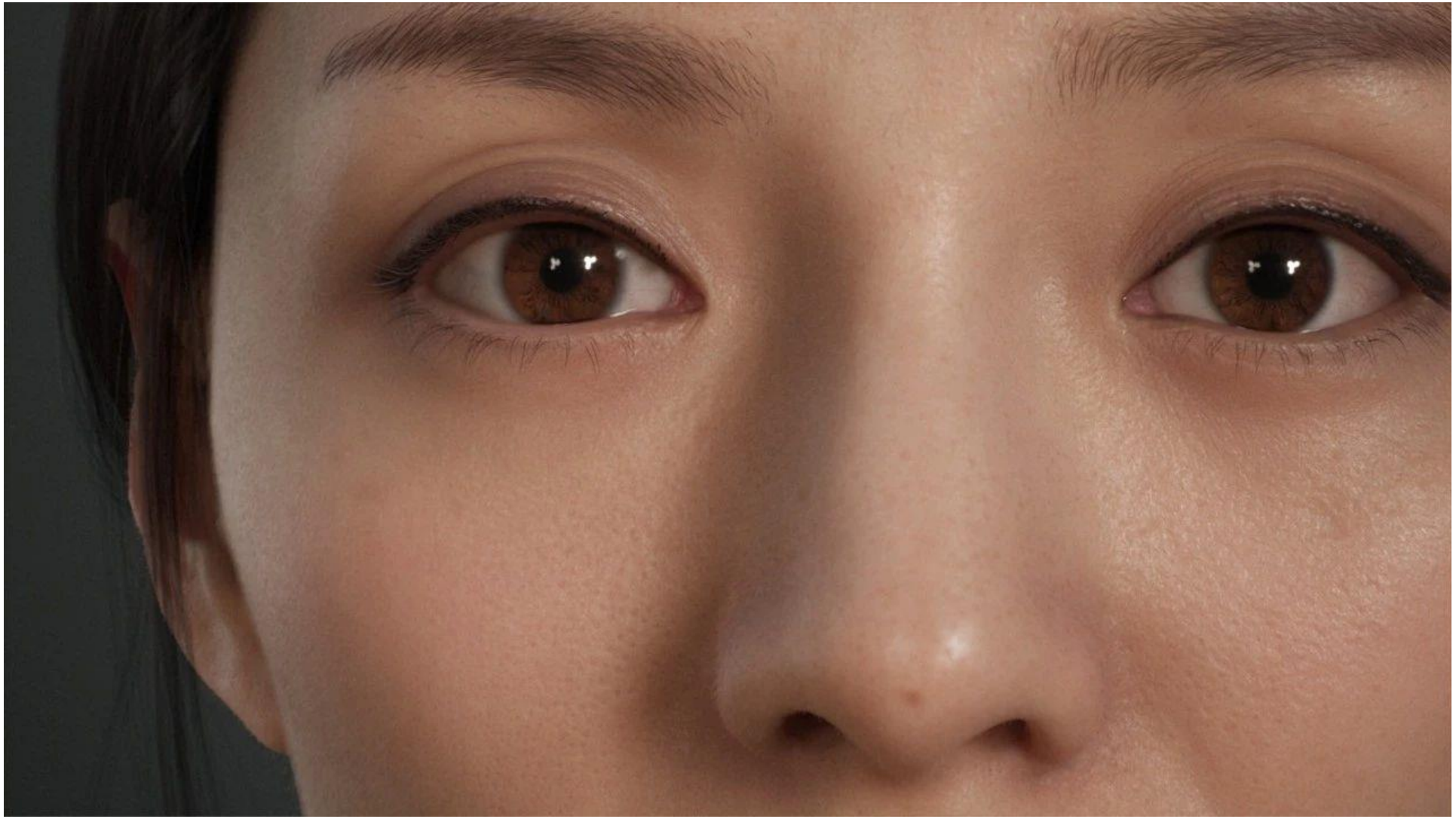
Calibration-Based Animation

该技术已经用于布拉德·皮特主演的电影《返老还童》
(The Curious Case of Benjamin Button)

DEMO

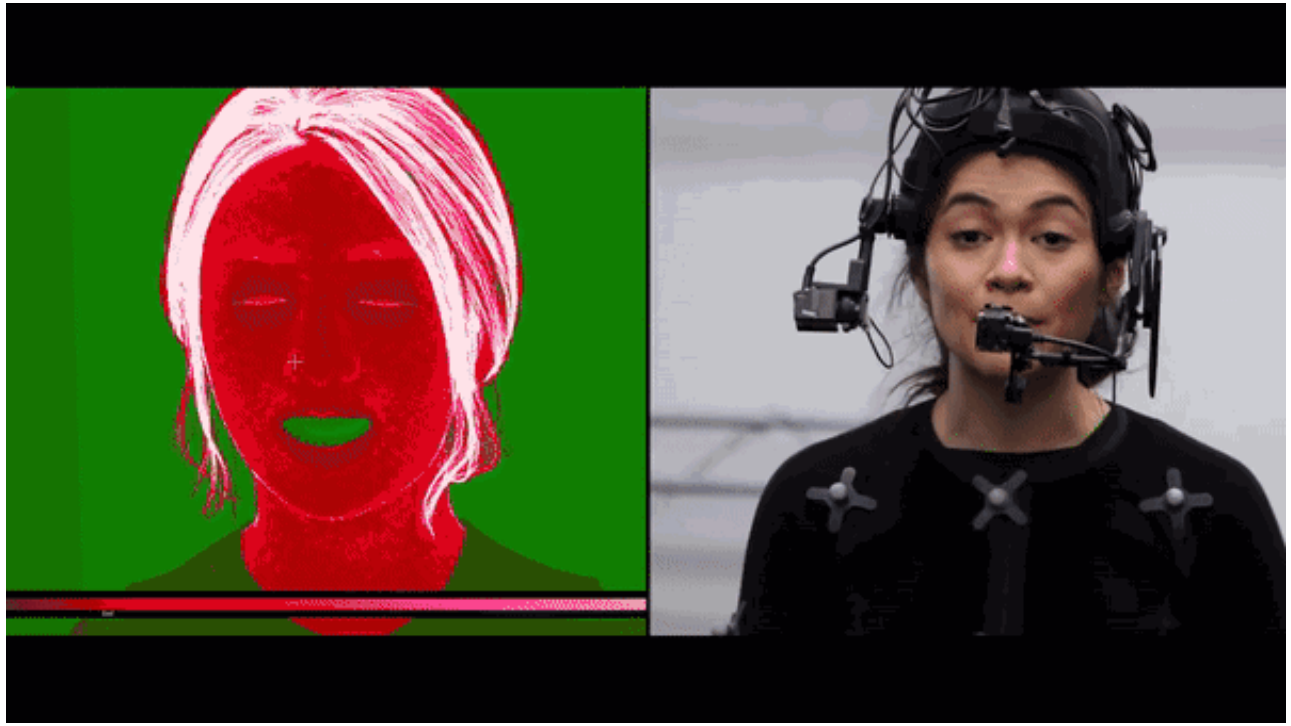
Unreal的实时角色"Siren"简直逆天了！

Unreal的实时角色"Siren"简直逆天了！



Unreal的实时角色"Siren"简直逆天了！

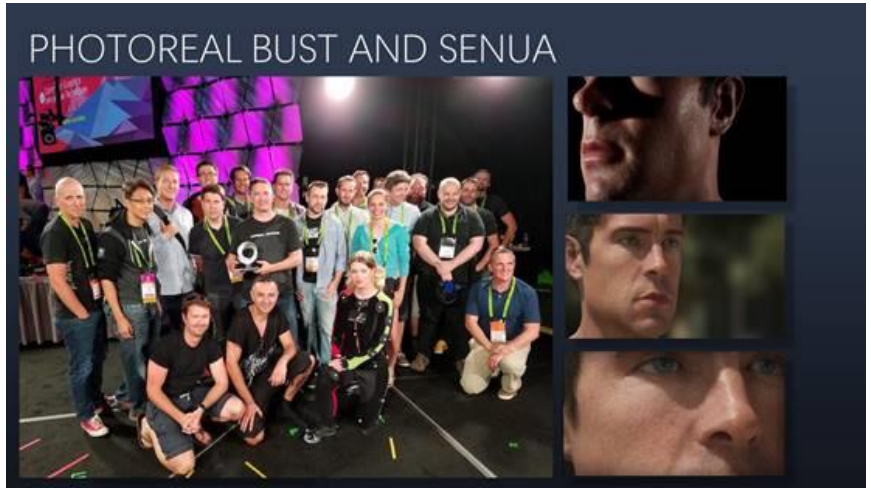
- Siren这位小姐姐的容貌身形，也不是凭空虚构出来。而是一位来自中国的演员：姜冰洁。
- GDC现场：一个真实的女性人类（英国演员Alexa Lee）在现场侃侃而谈，她的表情、动作、声音，被实时的赋予身后大屏幕上的女性数字人类Siren。



Unreal的实时角色"Siren"简直逆天了！

实时数字人Siren渲染技术揭秘（腾讯Next工作室）

- https://gameinstitute.qq.com/course/detail/10130
- 这一整套技术的前身是在SIGGRAPH 2016的Realtime live环节展示的游戏Hellblade，Epic凭借此项突破拿下Realtime live大奖，这个奖项相当于图形学界的“诺贝尔奖”，获得此殊荣足以说明这一整套解决方案是虚拟数字人方面做得最领先的。



我们跟腾讯Next工作室合作做的数字人

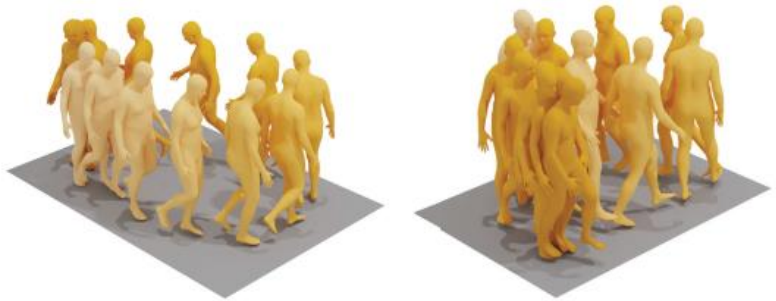
Qinjie Xiao, Hanyuan Zhang, Zhaorui Zhang, Yiqian Wu , Luyuan Wang, **Xiaogang Jin**, Xin Jiang, Yong-Liang Yang, Tianjia Shao, Kun Zhou: **EyelashNet: a dataset and a baseline method for eyelash matting.** ACM Trans. Graph. 40(6): 217:1-217:17 (2021)

英伟达：推出Metaverse概念创作平台Ominiverse。黄仁勋虚拟主题演讲

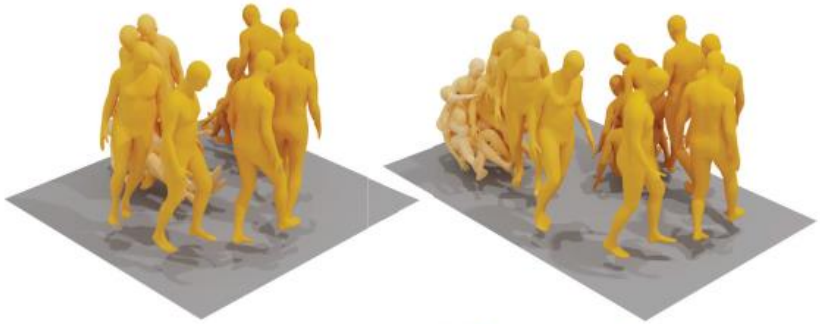
总结

- Motion Capture数据处理
- Motion Graph
- 人脸表情动画
- The Digital Emily Project

最近的研究热点：Text2Motion



a person steps forward with left leg walking in a circle counterclockwise motion swinging arms and returns back.



a man gets up from the ground, walks around in a circle and then sits down on the ground again.

a person runs forward, then turns around and walks in the opposite direction while swinging their right arm.

a man postures his arms like holding a dance partner and dances the waltz from the right hand side to the left hand sid

"A person punches in a manner consistent with martial arts."

最近的研究热点：TEXT2HUMAN

最近的研究热点：Text2Portrait

"A boy, solo, male focus, black eyes, pink hair, wide shot, day, teen"

Yiqian Wu, Hao Xu, Xiangjun Tang, Xien Chen, Siyu Tang, Zhebin Zhang, Chen Li, Xiaogang Jin, "Portrait3D: Text-Guided High-Quality 3D Portrait Generation Using Pyramid Representation and GANs Prior", ACM Transactions on Graphics (Proc. Siggraph 2024), 2024, 43(4): Article 45.

最近的研究热点：Text2Portrait

The End