3D Human Motion Capture from 2D Video using Cloud-Based CNNs

NVIDIA GPU Technology Conference
Imagine if Cameras Could See People

bodySLAM™ turns 2D video into 3D humans
Applications: AR Social Imagery & Filters
AR / VR Gaming
Sport Analytics
Eyes for Virtual Assistants
BodySLAM™: Deep Learning API

- **API**
  - **In:** 2D Video
    - fidelity/performance parameters
  - **Out:** Human Features
    - Joints: 21 joints including eyes and ears
    - Skeletons: 2D, 3D
    - Segmentation: background, individuals, limbs

- **Platforms**
  - CUDA
  - Metal Performance Shaders
  - openCL
Deep Learning / Computer Vision Secret Sauce
- 100+ layer CNN (Caffe)
- Synthetic Training data
- Human tracking

Game Tech
- 3D animation
- Joint Filtering
- Inverse Kinematics

Platforms
- NVIDIA GPU
- AWS
- MPS iOS
High Level Architecture
Training Pipeline
BodyFilter™: VFX filters built on top of BodySLAM™
Unity Game Engine for Training Pipeline and VFX App
Future Work

- Working with Customers
- Optimization
- Native Mobile Support
- Higher Fidelity
Questions? Let’s Start Augmenting Reality Together!