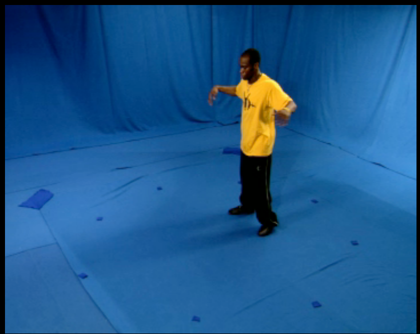
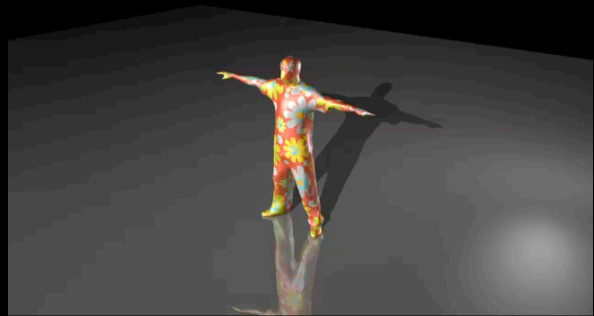
4D Performance Modelling & Animation

Adrian Hilton
Centre for Vision, Speech and Signal Processing
University of Surrey, UK

4D performance modelling & animation



Performance Capture



4D Model

Interactive Animation

Overview

Part I: Performance capture

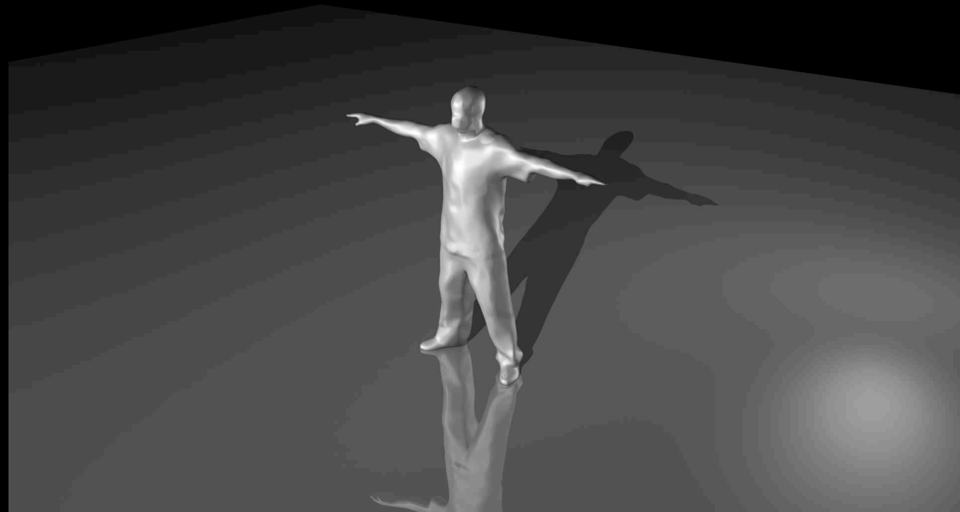
Part II: Structured representation

Part III: Interactive Animation

Future directions

I: Performance Capture

Multiple view video



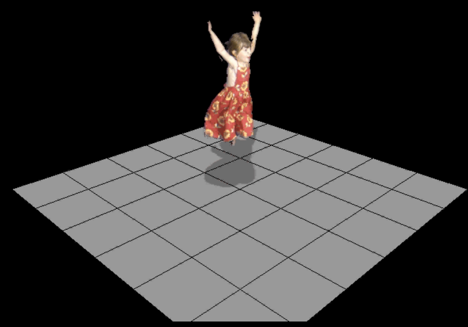
3D video

I:Video-based reconstruction of people

Model-free:

Free-viewpoint replay [Virtualized Reality - Kanade'96]
Photo-realistic rendering [Zitnick'04, Starck'05]

+ any scene
- Unstructured representation
- No change in movement

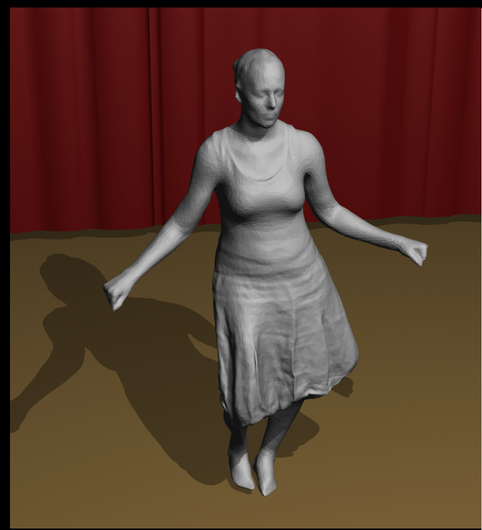


[Starck'05]

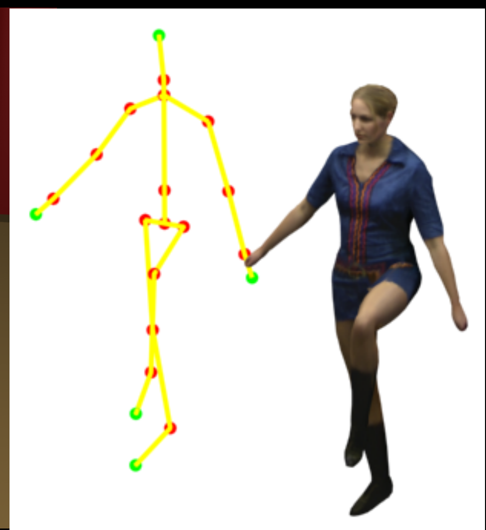
Model-based:

Single view [Hilton'00, Black'08, Gall'08]
Multi-view [Carranza'03, Starck'03]
High-detail [deAguiar'08]

+ Structured representation
+ User control of movement
- Fixed scene
- Limited visual quality

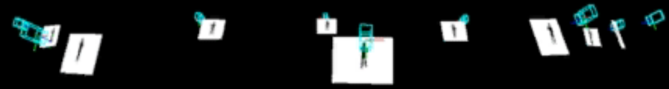


[de Aguiar'08]

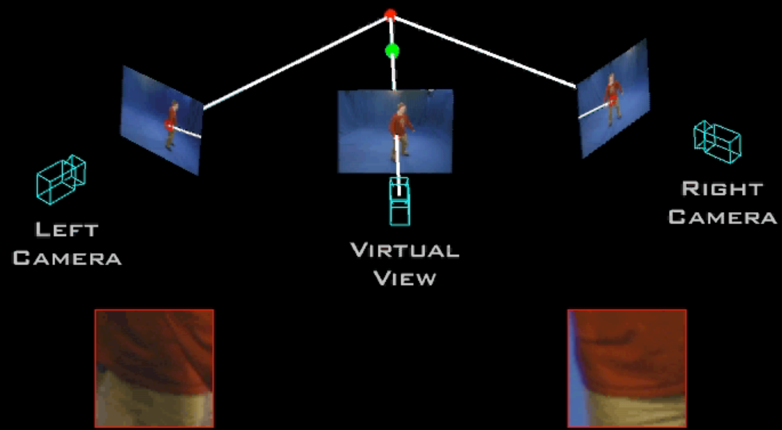


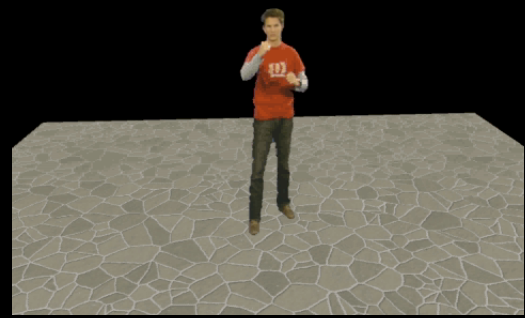
[Starck'03]

I: Multiple view reconstruction



Visual Hull

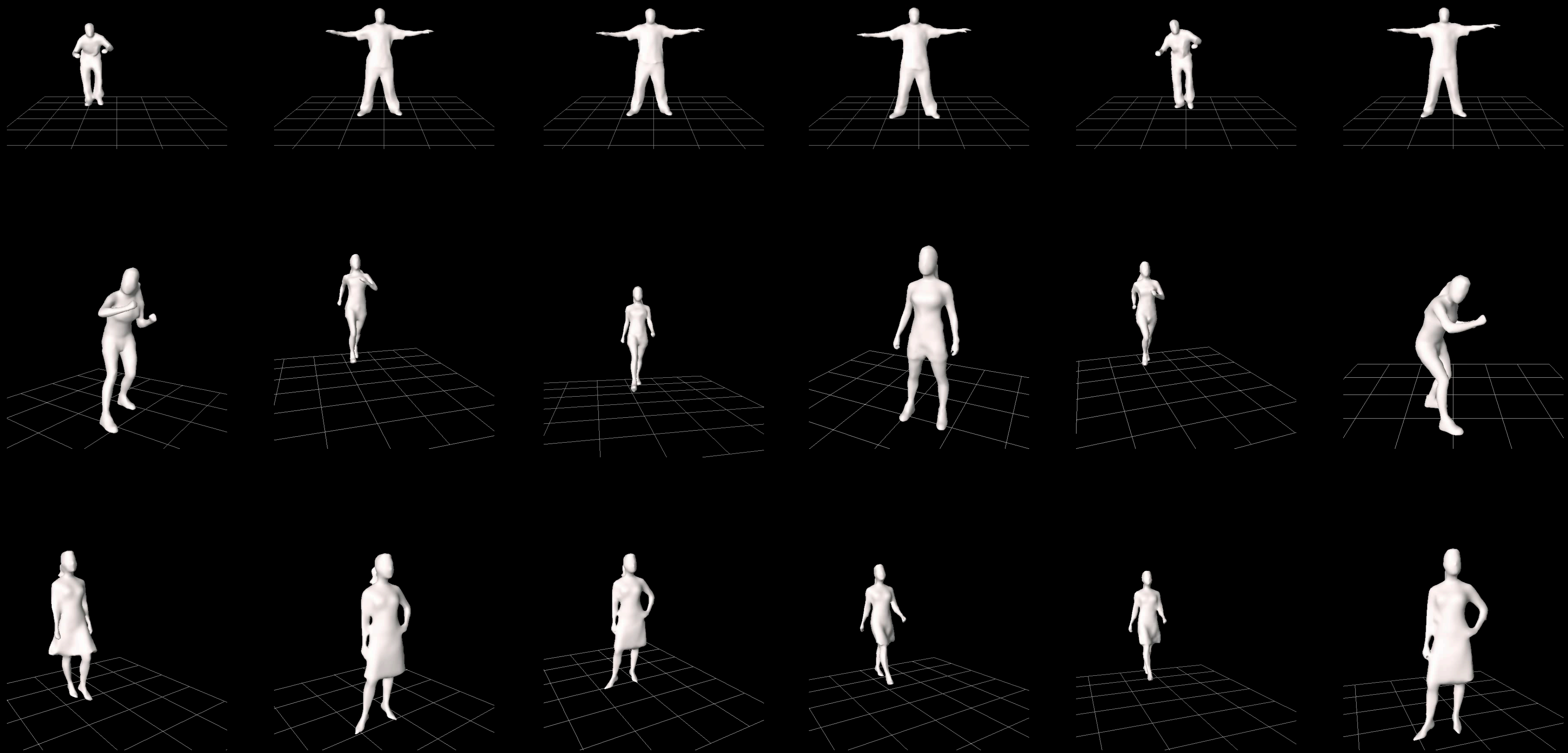




Stereo Refinement

[Starck et al. CVIU'08]

1: Studio Capture



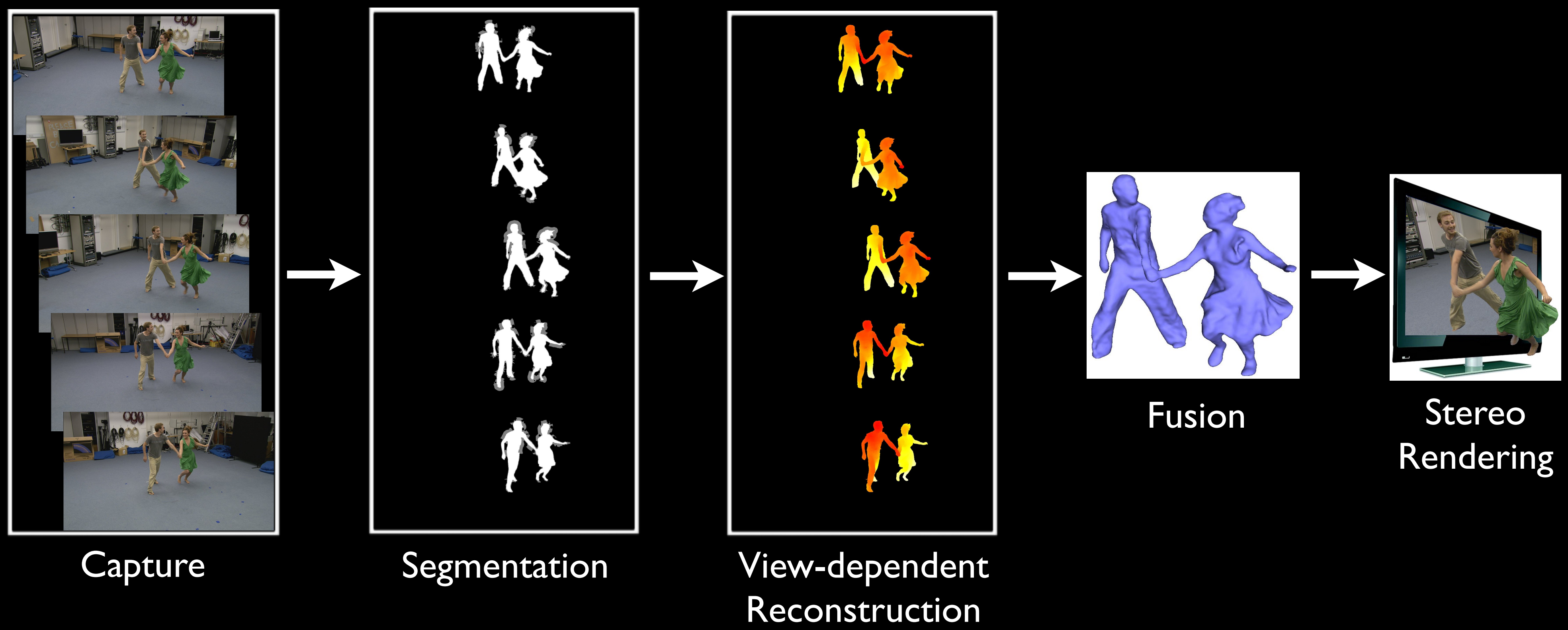
SurfCap 3D Video Database *[Starck et al. CGA'07]*

I: Outdoor Capture

Problems:

- large capture volume (film set/soccer pitch)
- uncontrolled environment
- non-uniform (moving) backgrounds
- less accurate camera calibration (moving cameras)

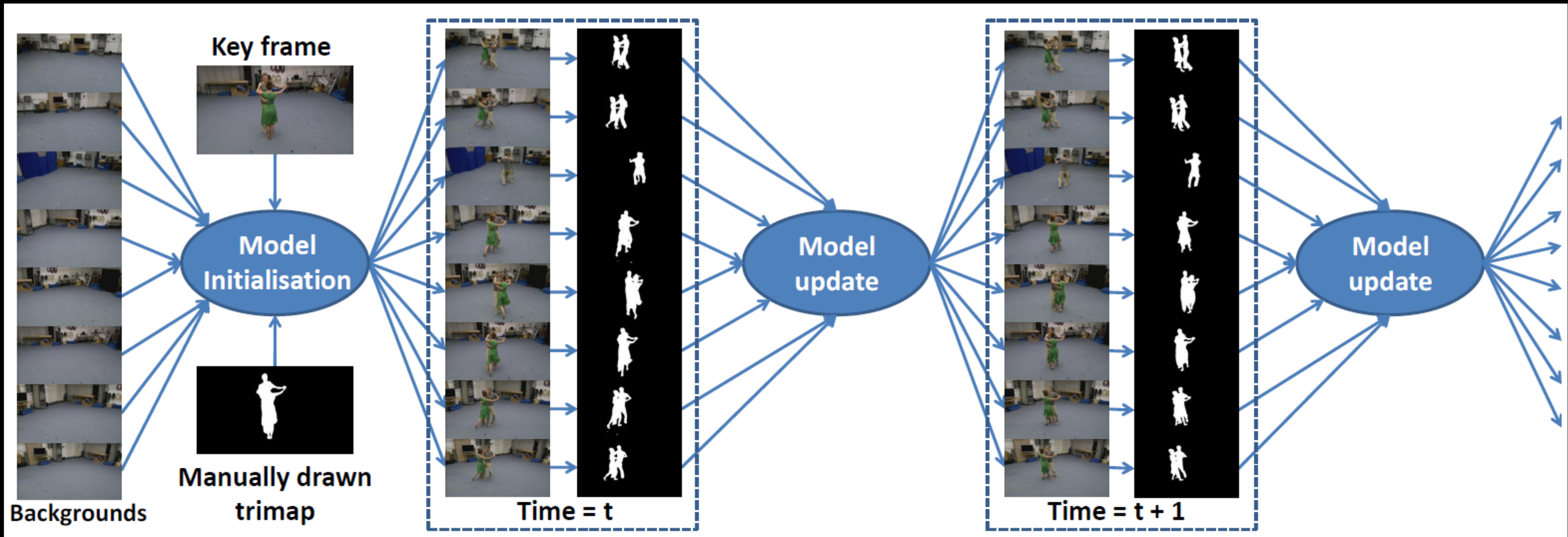
I: Robust Multi-view Reconstruction



I: Multi-view Segmentation

Bayesian inference to propagate trimap labels:

- exploit natural image matting
- manual trimap labels in single view (~every 100frames)
- spatio-temporal propagation of trimap labels
- local & global colour statistics + epipolar constraint



I: Multi-view Segmentation

Trimap label inference:

$$P(\mu_k, \Sigma_k \mid \boldsymbol{I_p}) = P(\boldsymbol{I_p} \mid \mu_k, \Sigma_k) P(\mu_k, \Sigma_k) / P(\boldsymbol{I_p})$$

label likelihood for kth colour model | model prior | pixel prior

Colour Models

- Local background

- Global background/foreground

- Local foreground estimated from initial foreground/background labelling

Natural image matting - closed form solution*[Levin et. al. CVPR'96]*

[Sarim et al. ICIP'10]

1: Multi-view segmentation

I: Robust Reconstrution

View-dependent joint reconstruction & segmentation

- pixel label estimation: depth *d*, layer *l*

d,l

- constraints

initial multi-view segmentation

initial depth (visual-hull)

[Guillemaut et al. IJCV'11]

I: Robust Reconstruction

Formulate as an energy minimisation problem:

$$E(l,d) = w_1 E_1(l) + w_2 E_2(l) + w_3 E_3(d) + w_4 E_4(l,d)$$

2D cues

3D cues

1. Foreground/Background colour:

$$E_1(l) = \sum_{\boldsymbol{p} \in \mathcal{P}} -\log P(\boldsymbol{I_p} | l_{\boldsymbol{p}})$$

3. Multi-view image matching:

$$E_3(d) = \sum_{\boldsymbol{p} \in \mathcal{P}} \mathrm{NCC}(\boldsymbol{p}, d_{\boldsymbol{p}})$$

2. Contrast

$$E_2(l) = \sum_{(\boldsymbol{p},\boldsymbol{q}) \in \mathcal{N}} e_2(\boldsymbol{p}, \boldsymbol{q}, l_{\boldsymbol{p}}, l_{\boldsymbol{q}})$$

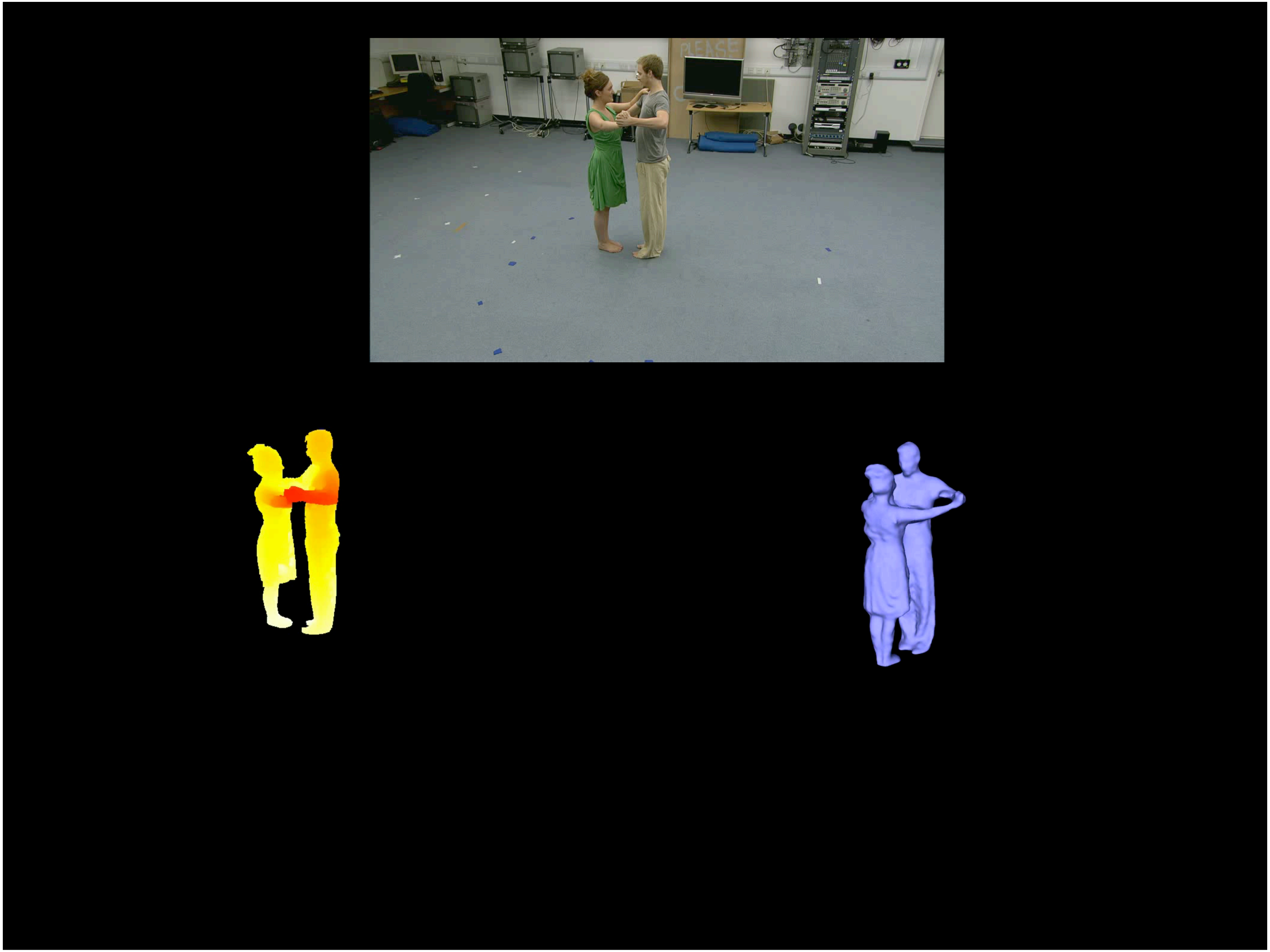
$$e_2(\boldsymbol{p}, \boldsymbol{q}, l_{\boldsymbol{p}}, l_{\boldsymbol{q}}) = \begin{cases} 0 & \text{if } l_{\boldsymbol{p}} = l_{\boldsymbol{q}}, \\ \exp(-\beta ||\boldsymbol{I_p} - \boldsymbol{I_q}||) & \text{otherwise.} \end{cases}$$

4. Surface smoothness:

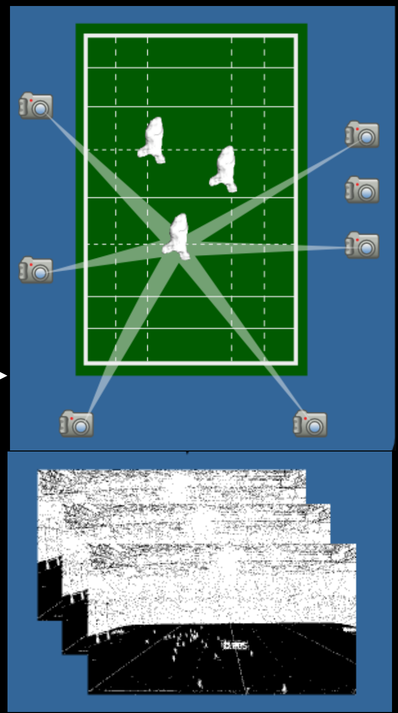
$$E_4(l,d) = \sum_{(\boldsymbol{p},\boldsymbol{q}) \in \mathcal{N}} D_{l_{\boldsymbol{p}}, l_{\boldsymbol{q}}}(d_{\boldsymbol{p}}, d_{\boldsymbol{q}})$$

Graph-cut solution using alpha expansion *[Boykov PAMI'01]*

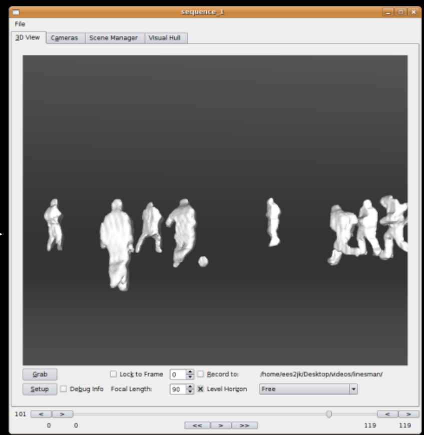
[Guillemaut et al. IJCV'11]


PLEASE

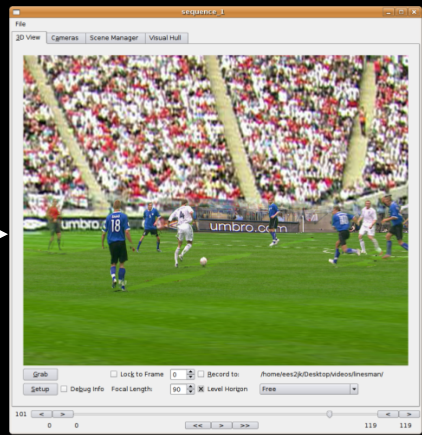
I: Outdoor Stadium Sports



2D matting
& calibration



Joint reconstruction
& segmentation



I: 3D video performance capture

Rugby dataset

Overview

Part I: Performance capture

Part II: Structured representation

Part III: Interactive Animation

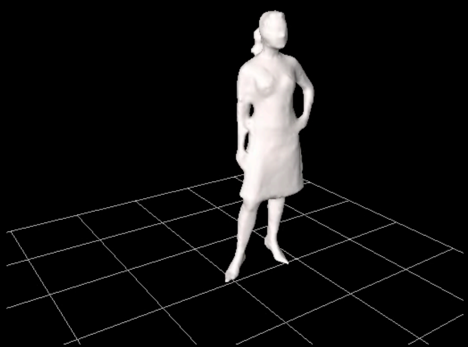
Future directions

II: Structured Representation

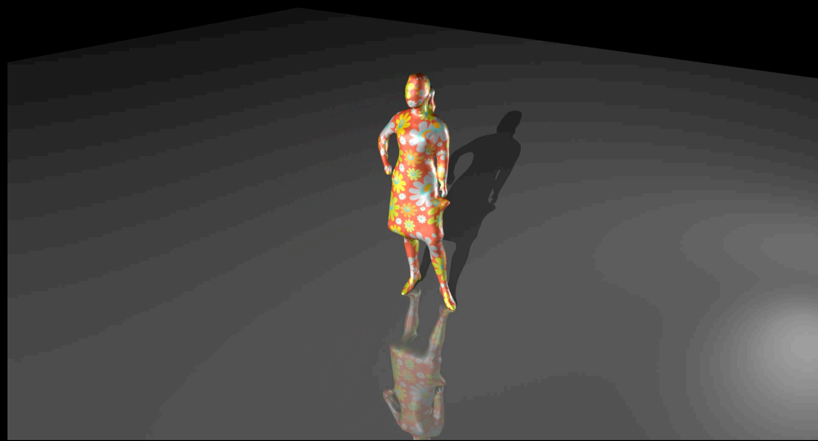
3D video performance capture

- unstructured mesh sequences
- no temporal correspondence

Goal: temporally coherent structure with correspondence

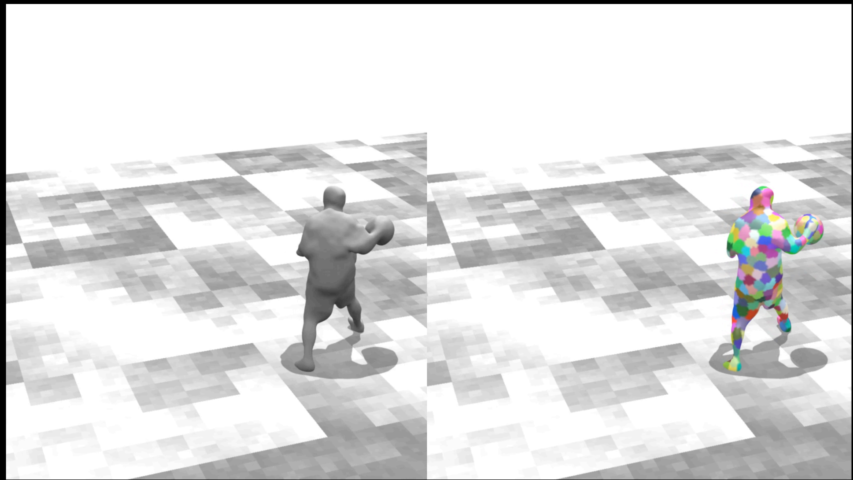


3D video

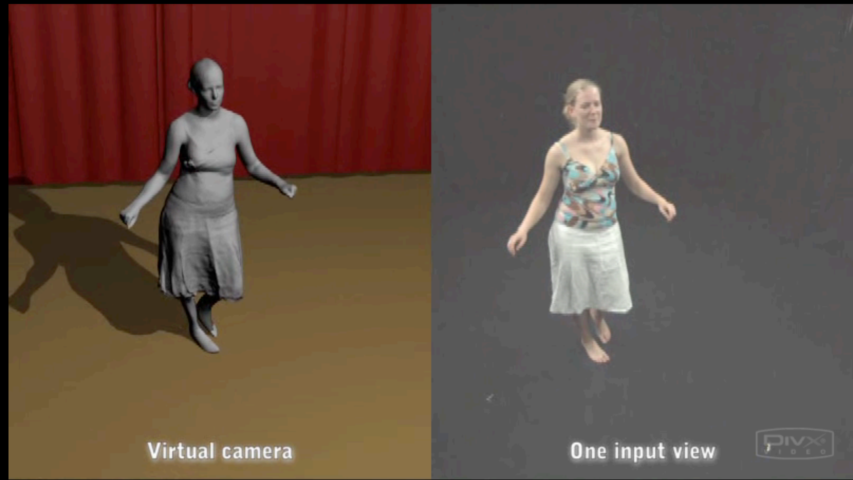


4D model

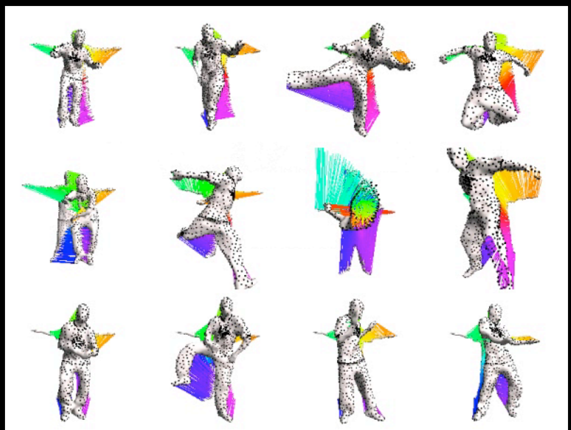
II: Non-rigid Surface Tracking



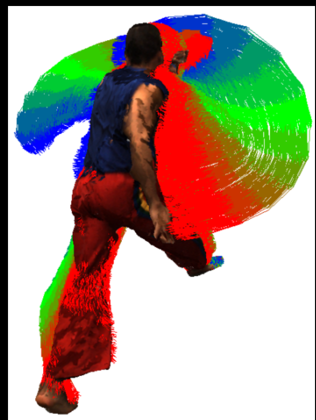
Cagniart et al. ECCV'10



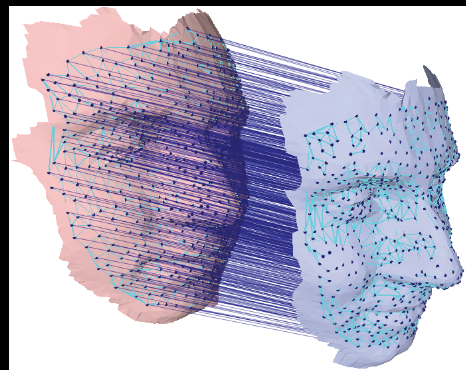
de Aguiar et al. SIGGRAPH'08



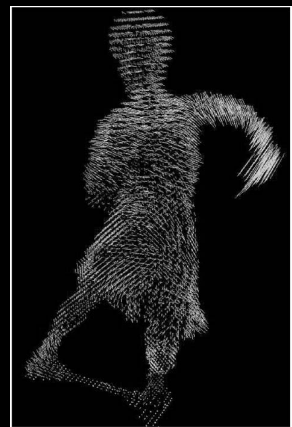
Starck et al. ICCV'07



Tung et al. CVPR'10



Zeng et al. CVPR'10



Vedula et al. PAMI'05

II: Global Non-rigid Alignment

3D video sequences

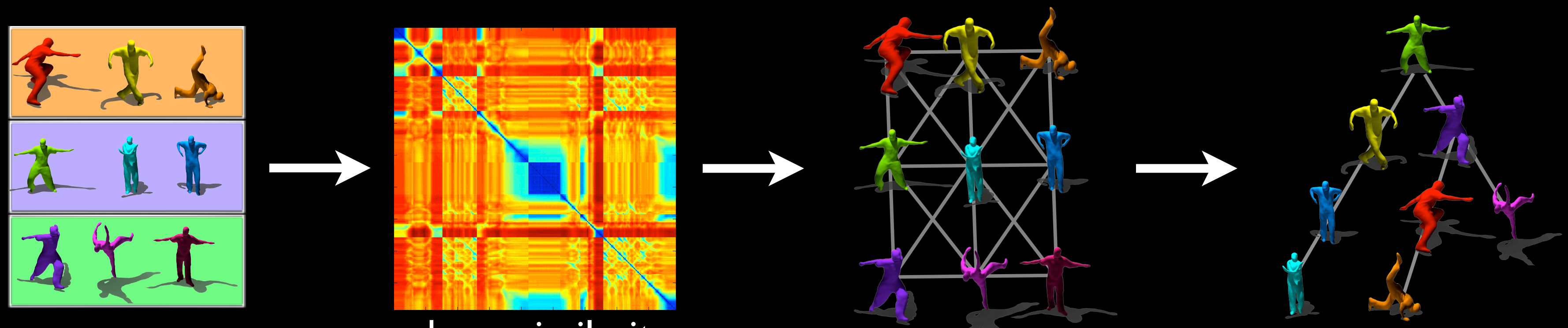
Shape similarity tree

4D model

II: Global Non-rigid Alignment

Shape similarity tree construction

- 3D shape similarity
- fully connected graph construction
- graph optimisation for shortest non-rigid alignment path

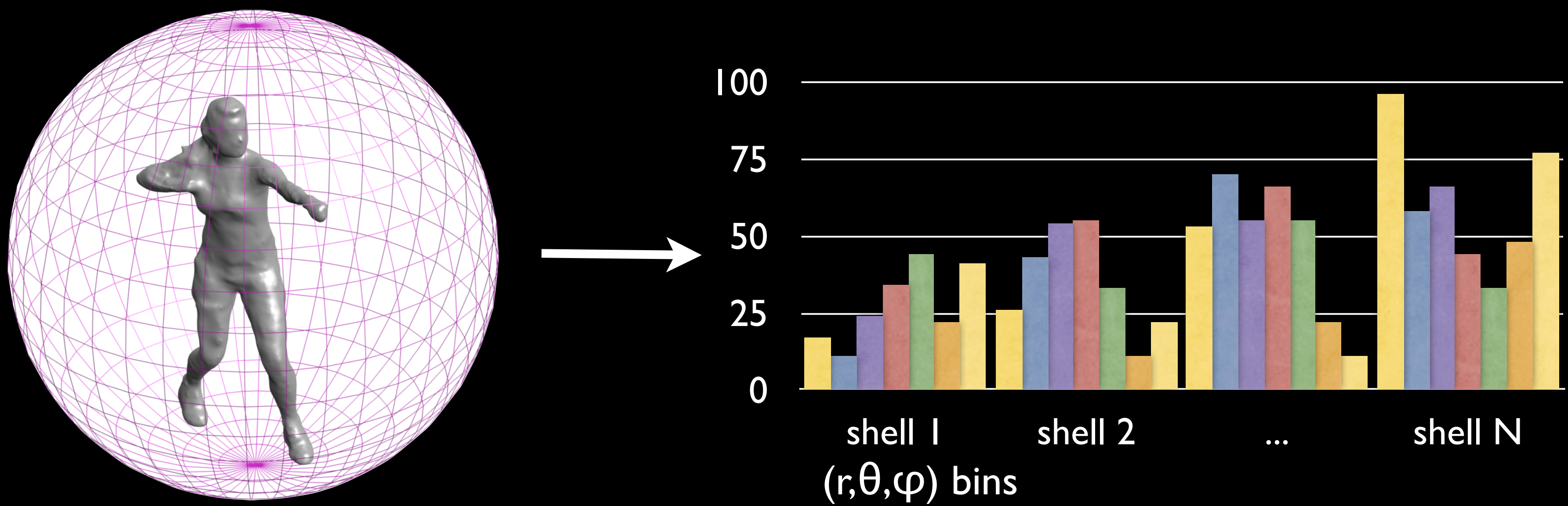


3D video sequences → shape similarity matrix → graph → Shape similarity tree

II: Global Non-rigid Alignment

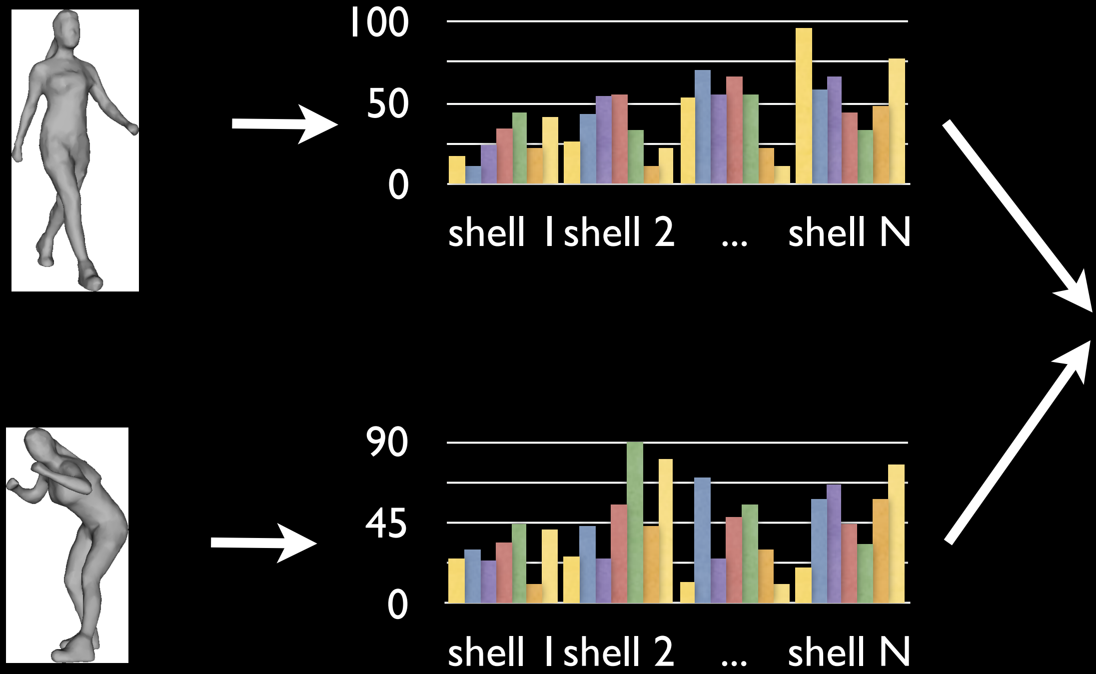
Shape similarity:

- spherical shape histogram *[Huang et al. IJCV'10]*



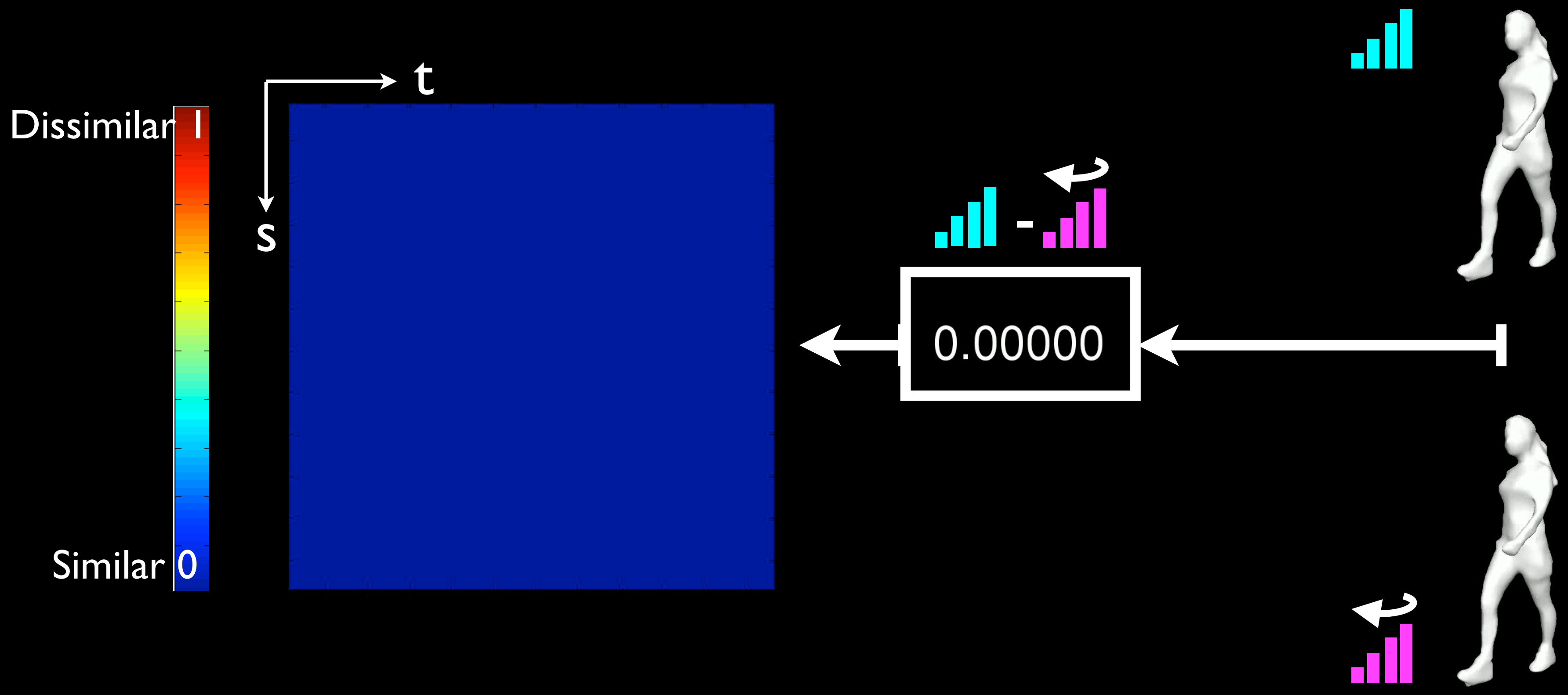
II: Global Non-rigid Alignment

Shape similarity:



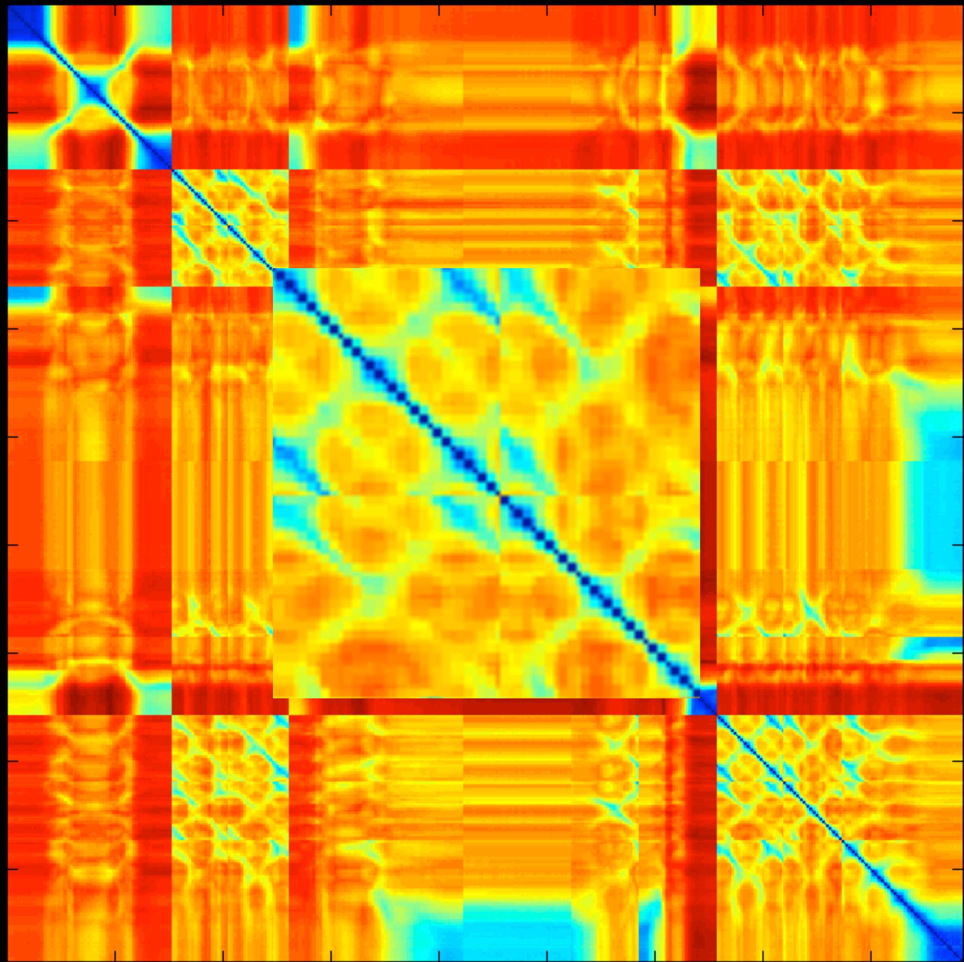
$$S(M_i, M_j) = \min_{\phi} \sum_{r=1}^{N} \left\| H_i(r) - H_j(r, \phi) \right\|^2$$

Frame to Frame Comparison



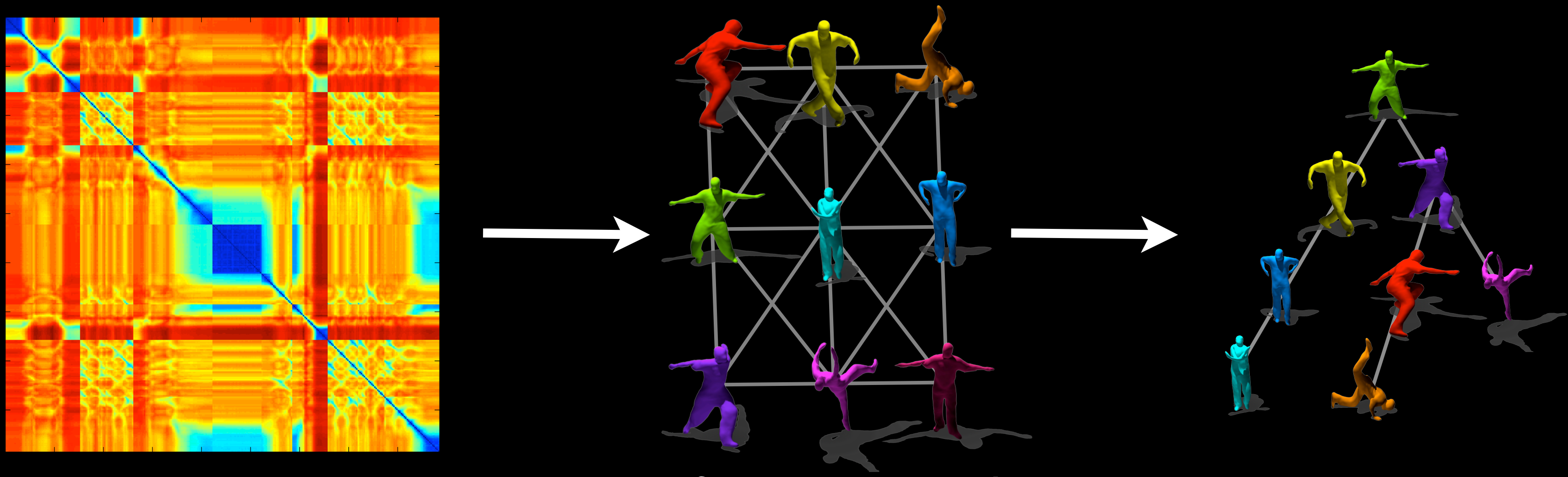
Similarity Matrix

20 motions



Compute similarity matrix for all frames (420x420 frames)

II: Global Non-rigid Alignment



Optimise for shortest path in fully connected graph:

$$\arg\min_{P \in \mathbb{Q}} \left(\sum_{\forall (i,j) \in P} S(M_i, M_j) \right)$$

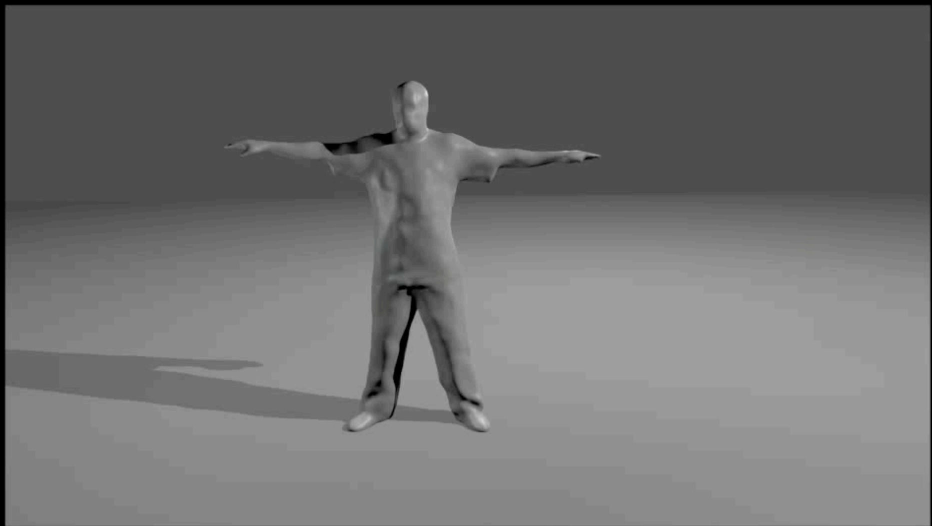
sum of similarities S() for all edges in p

Solution: minimum spanning tree

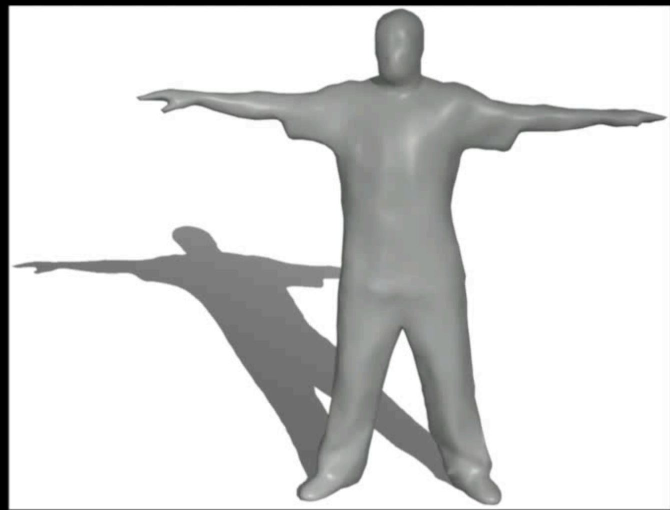
[Budd et al. 3DIMPVT'11,Huang CVPR'11]

II: Global non-rigid alignment

Shape Tree Construction



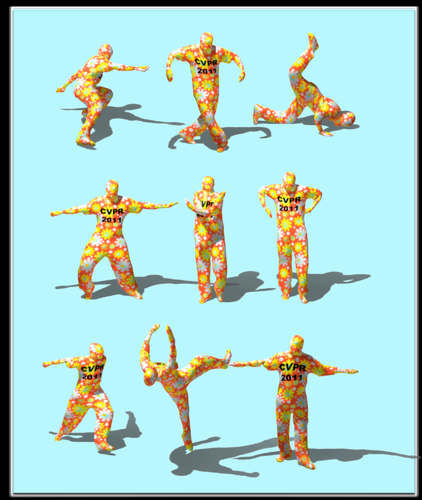
Original Reconstruction



Shape Tree Building

II: Global non-rigid alignment

Shape similarity



4D model

Non-rigid alignment:

- shape similarity tree gives minimum non-rigid deformation for alignment of all frames
- Laplacian deformation framework
- Geometric & photo-metric feature constraints

II: Global non-rigid aligment

Laplacian deformation framework *[Sorkine CGF'06]:*

$$Lx = \delta$$

mesh Laplacian — vertex positions — differential coordinates

Laplacian represents mesh shape & connectivity

Energy minimisation:

$$\arg\min_{X} \left\| Lx - \delta(x_0) \right\|^2 + \left\| W(x - x_C) \right\|^2$$

original position — constraints

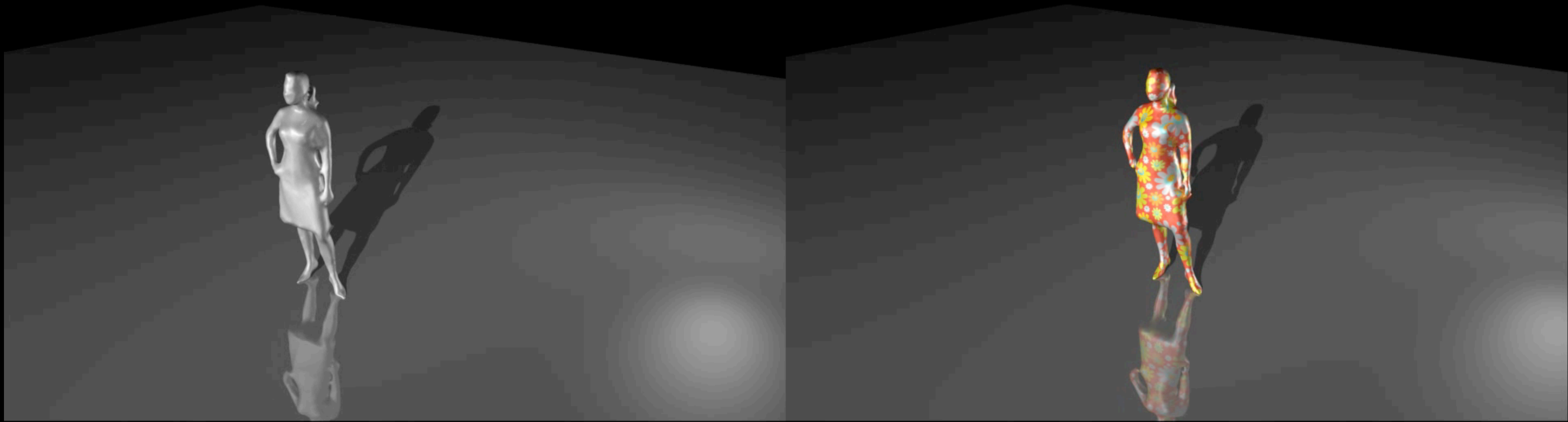
constraints given by geometric & photometric correspondence

Globally Aligned Sequence Database

Original Reconstruction

Temporally Consistent

Fashion1-Global Alignment



Original Reconstruction

Temporally Consistent

Overview

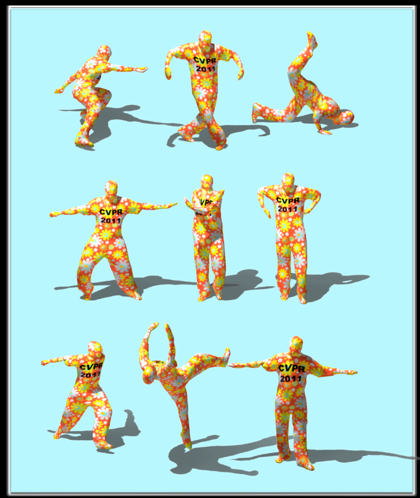
Part I: Performance capture

Part II: Structured representation

Part III: Interactive Animation

Future directions

III: Interactive Animation



4D model

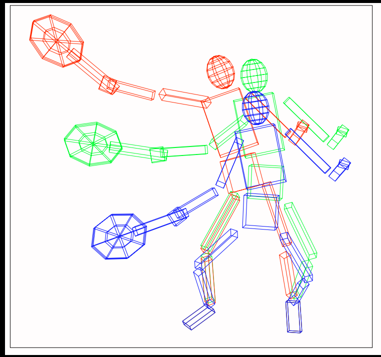
Interactive Animation

Interactive control of character animation:

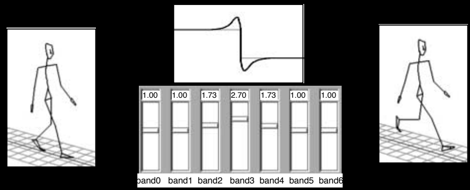
- editing motion
- high-level parameterisation of motion
- transitions between motions

Skeletal Character Animation

Motion Editing

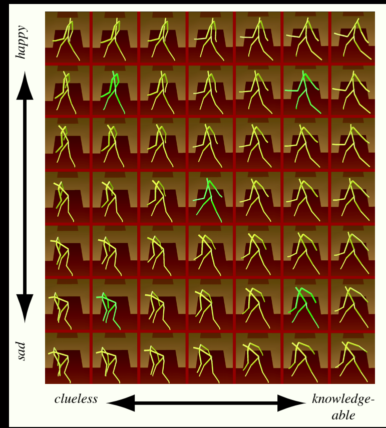


Witkin'95



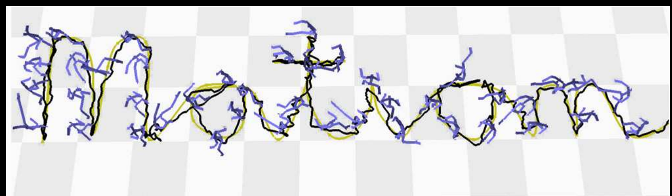
Brundelin'95

Motion Parameterisation



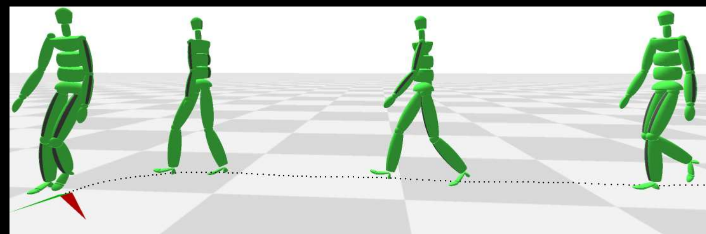
Rose'98

Motion Graphs



Kovar&Gleicher'02

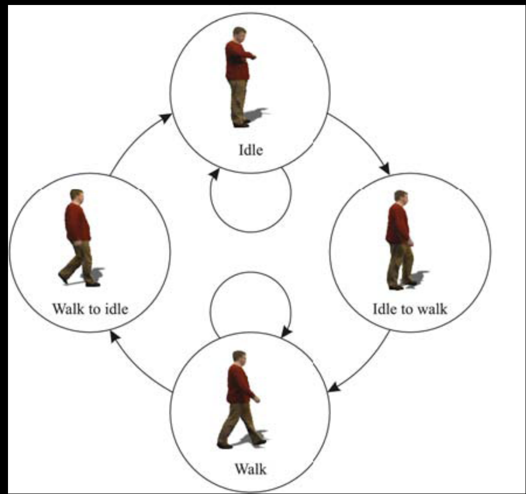
Arikan&Forsyth'02



Heck & Gleicher'07

III: 3D video concatenation *[Huang et al. CVPR'09]*

surface motion graph representation



key-frame animation

key start

key end

III: 3D video concatenation *[Huang et al. CVPR'09]*

Synthesized 3D Character Animation

III: 4D motion editing

Goal: Interactive editing of 4D models

Space-time key-frame editing

- Laplacian deformation framework
- learnt 4D motion space

[Tejera et al. CVMP'11]

III: 4D motion editing

Laplacian deformation in learnt motion space *r*:

$$\arg\min_{x,r} \left\| Lx - \delta(r) \right\|^2 + \left\| W(x - x_C) \right\|^2$$

learnt 4D space

user specified constraints

surface deformation constrained to learnt space to preserve anatomical structure

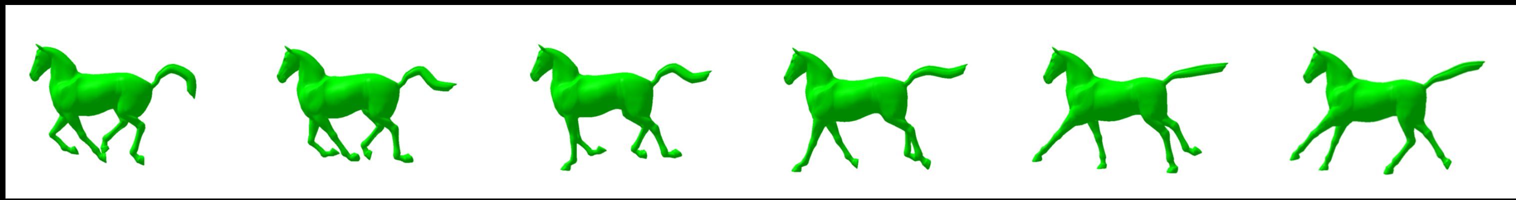
$$\delta(r) = \bar{\delta} + \sum_{k=1}^{L} r_k e_k$$

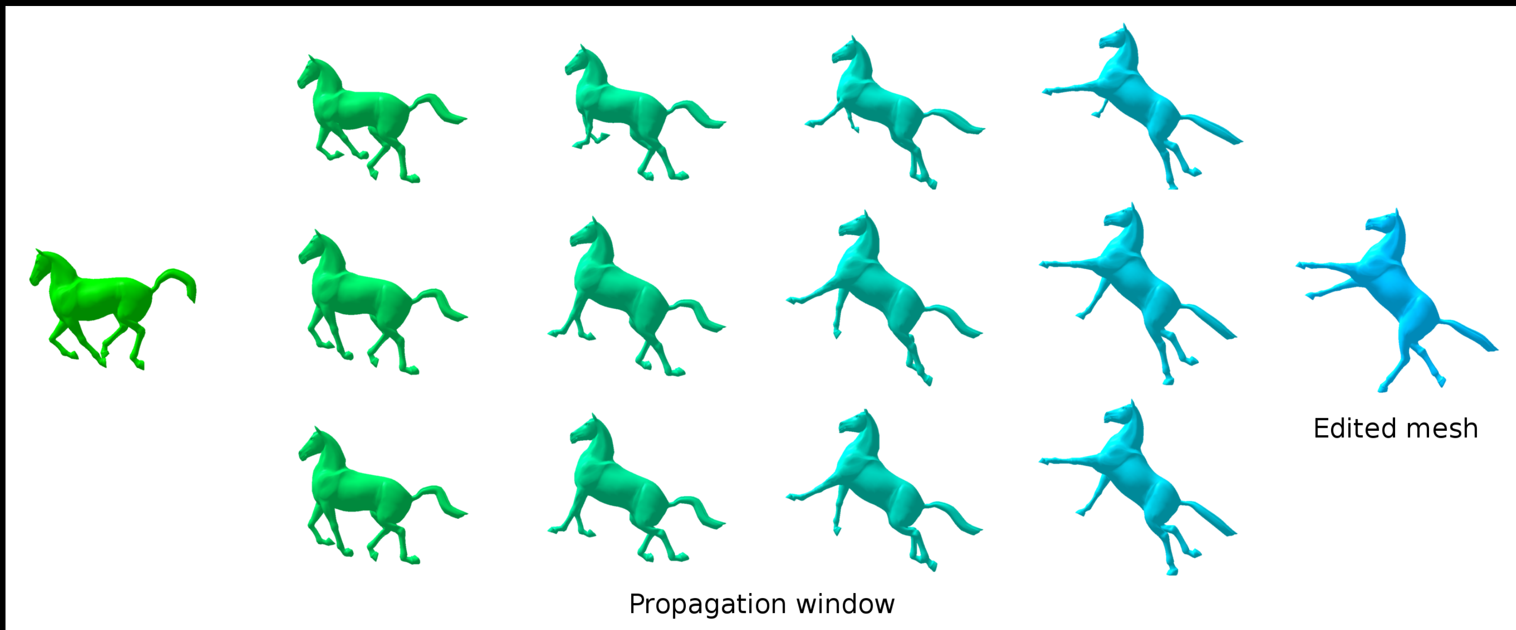
learnt 4D basis
in differential coordinates

[Tejera et al. CVMP'11]

III: 4D motion editing

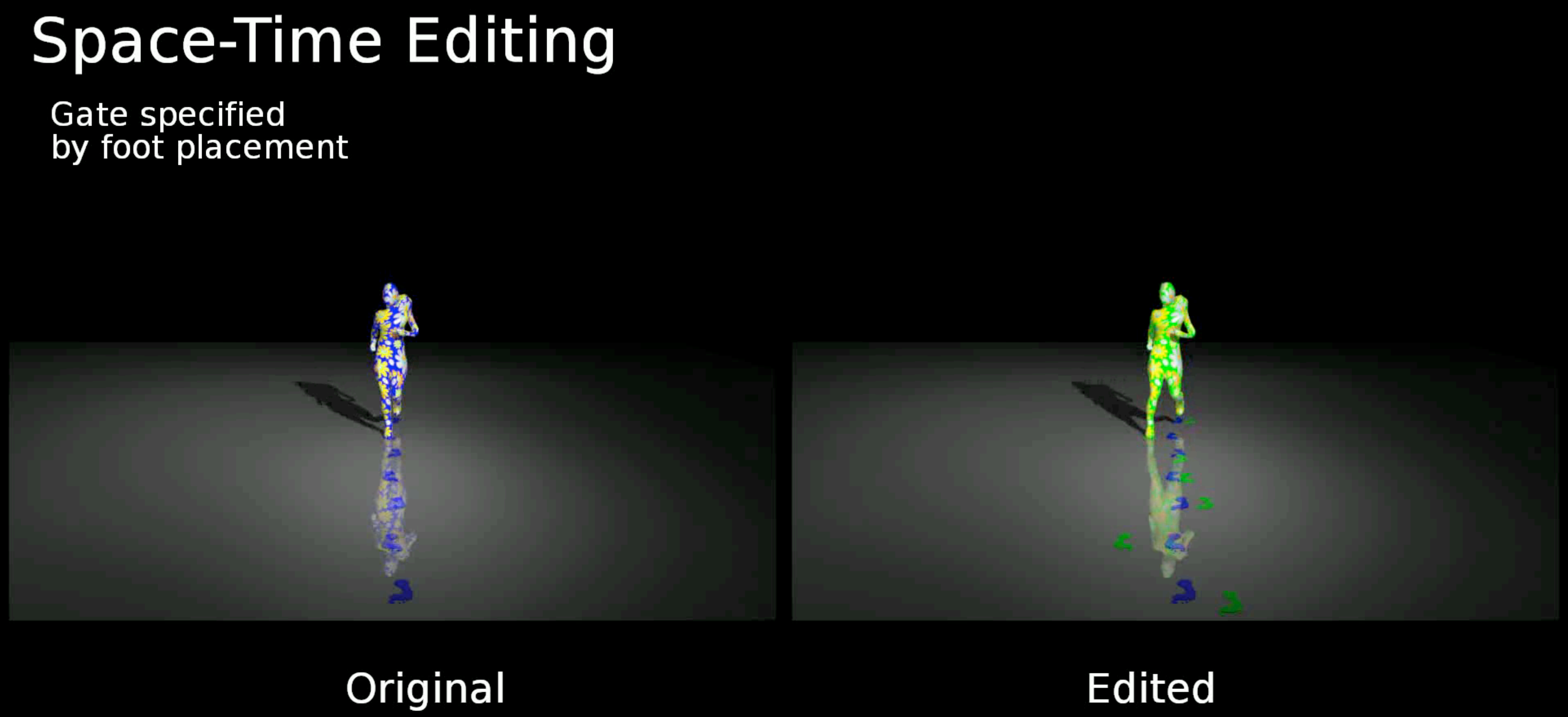
Key-frame edits propagated over space-time window





[Tejera et al. 2011]

III: 4D motion editing



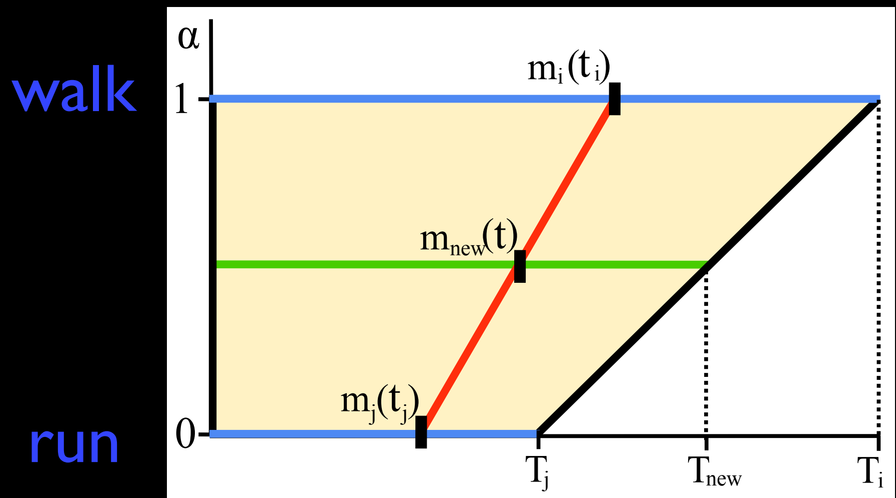
[Tejera et al. 2011]

III: 4D motion parameterisation

High-level real-time motion control

- parameters: walk speed/direction, jump height etc
- combine multiple skeletal sequences *[Rose'98]*

solution: mesh sequence blending ie walk/run



$$speed = f(\alpha)$$

$$\alpha = f^{-1}(speed) = h(speed)$$

III: 4D motion parameterisation

Mesh sequence blending

(I) temporal alignment: dynamic time warp

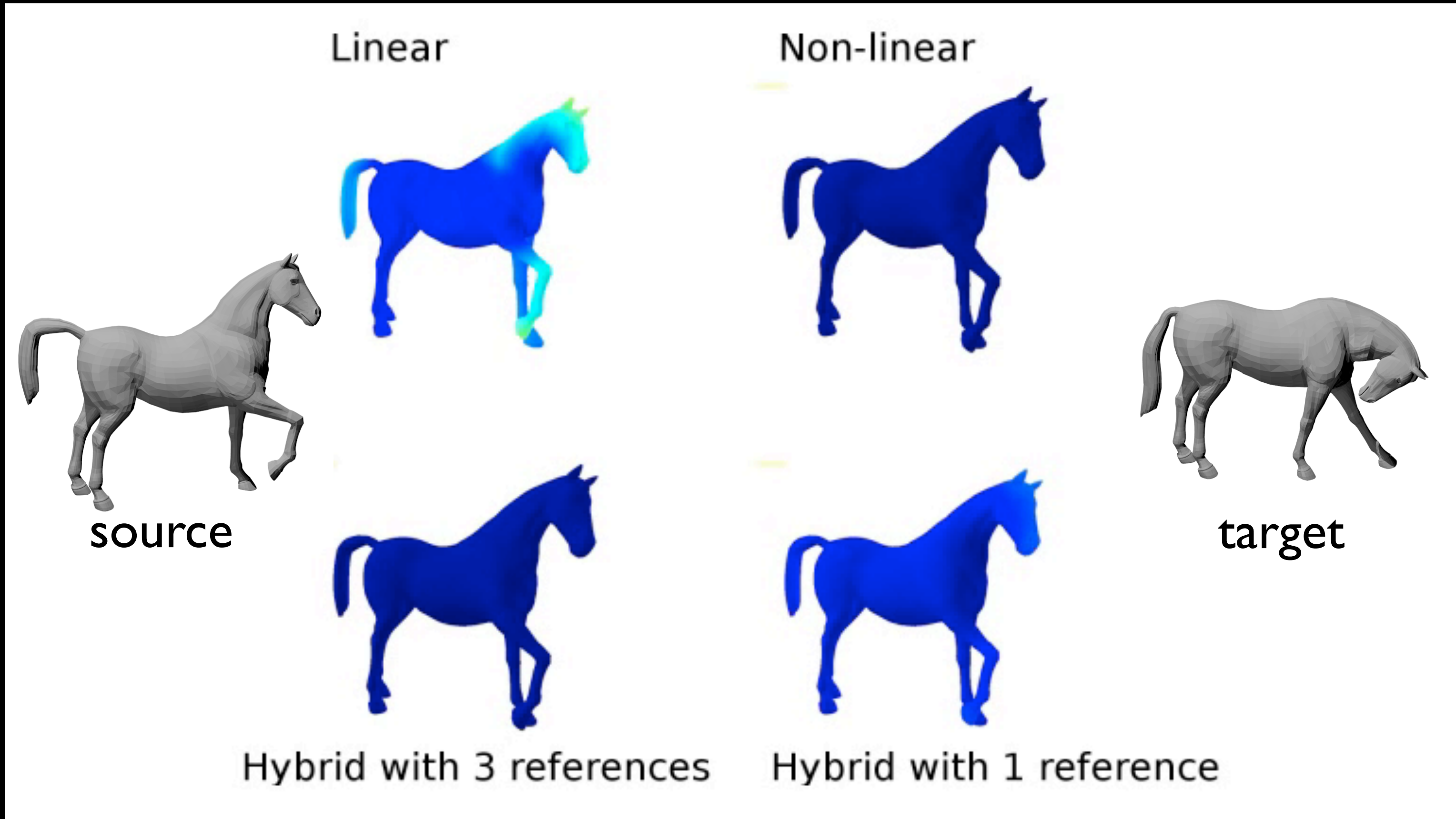
(II) blend corresponding frames

non-linear blending (Laplacian): ~100ms/frame

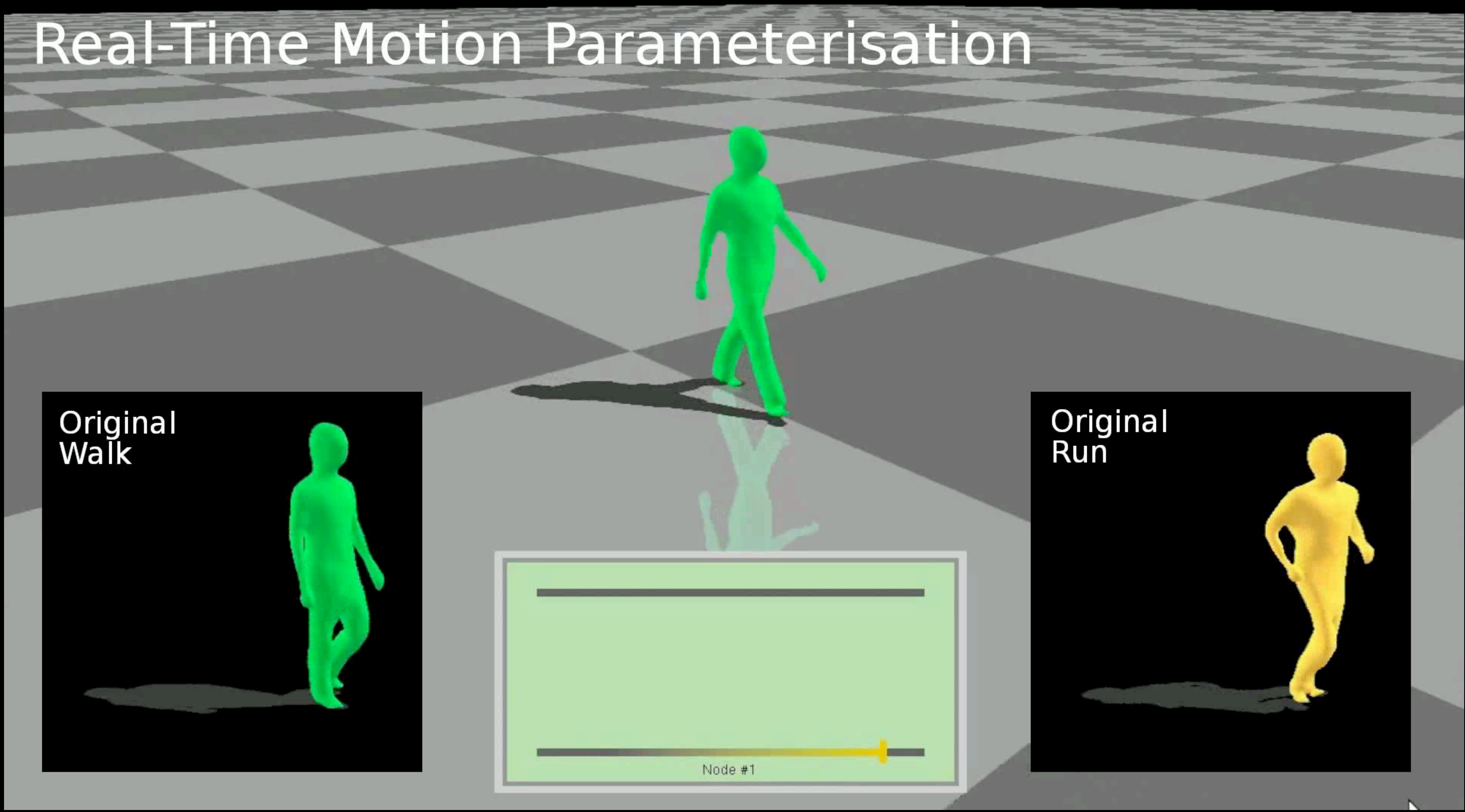
linear blending: <8ms/frame but unrealistic

solution: hybrid non-linear blending 10ms/frame

[Casas et al. MIG'11]



[Casas et al. MIG'2011]



[Casas et al. 2011]

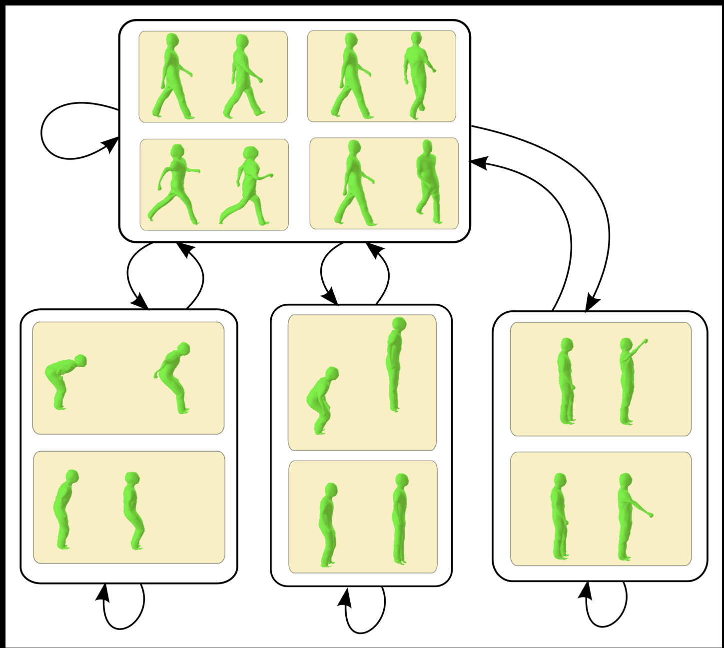
III: Interactive Animation

interactive motion transitions

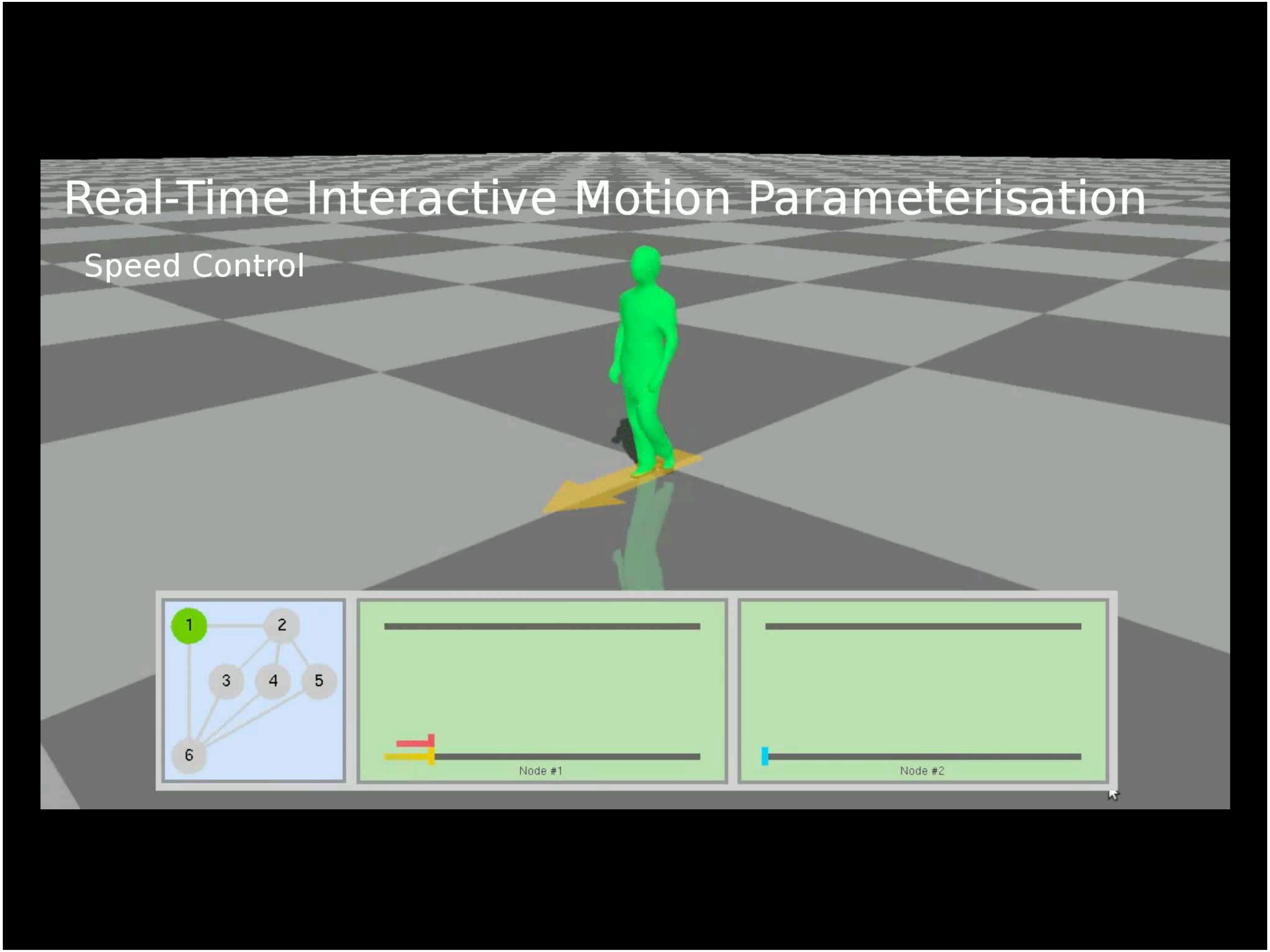
- skeletal motion graph *[Gleicher'02,Arikan'02]*

solution: 4D parametric motion graph

- real-time transitions using shape similarity



[Casas et al. 2011]


Real-Time Interactive Motion Parameterisation
Speed Control
1
2
3
4
5
6
Node #1
Node #2

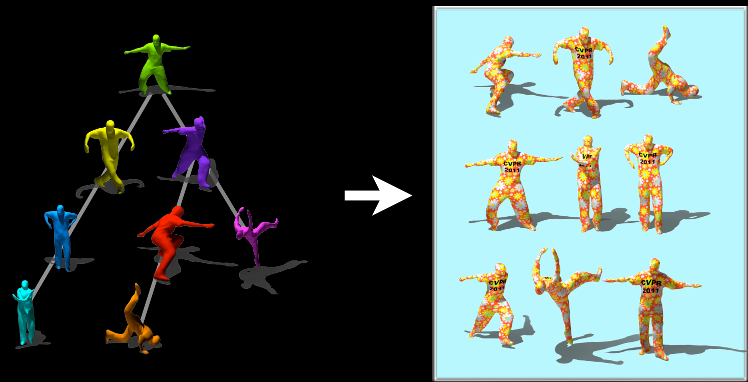
Summary

Part I: Performance capture

- 3D video capture indoor/outdoor
- joint segmentation & reconstruction

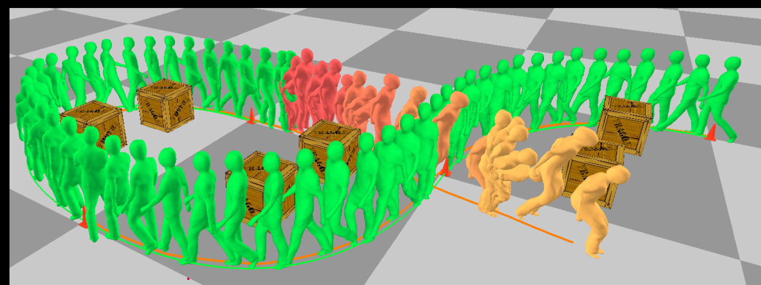
Part II: Structured representation

- Global non-rigid alignment
- shape similarity tree
- 4D models



Part III: Interactive Animation

- 4D motion parametrisation
- 4D parametric motion graphs



Future Challenges

Part I: Performance Capture

- general dynamic scenes

Part II: Structured Representation

- Accurate alignment of non-rigid surface detail

Part III: Interactive Animation

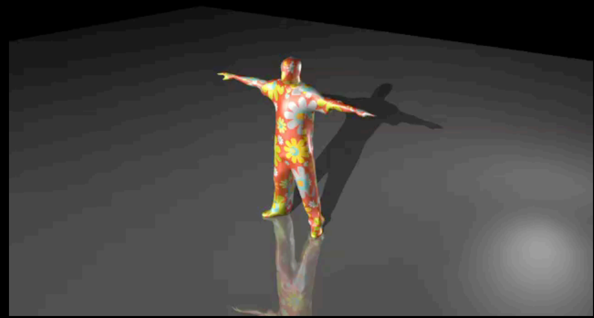
- characters with highly dynamic clothing/hair
- photo-realistic appearance

Other Applications

- performance analysis
- real-time remoteinteraction

4D Performance Modelling & Animation

Adrian Hilton
a.hilton@surrey.ac.uk



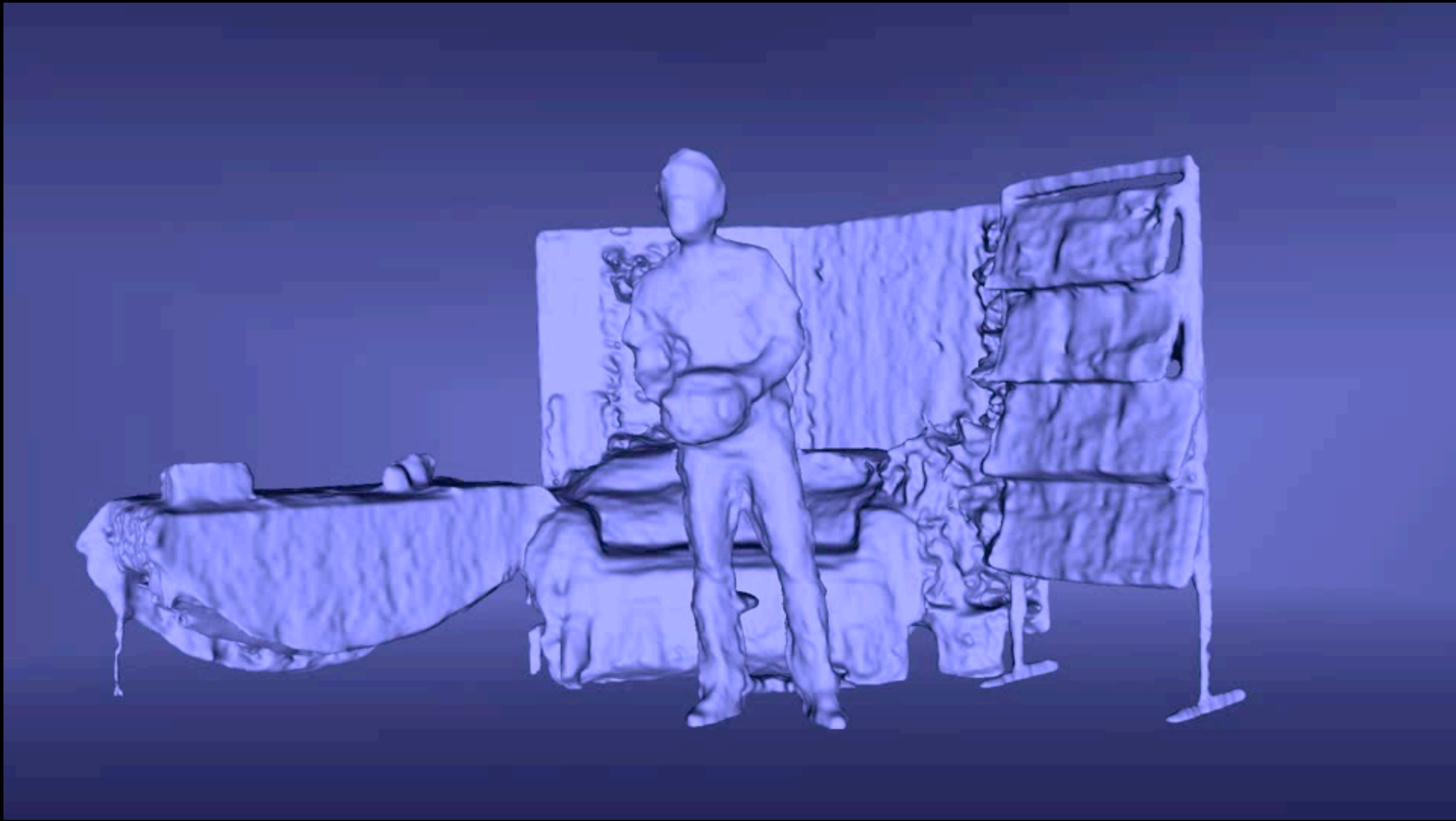
Researchers: Jon Starck, Jean-Yves Guillemaut, Peng Huang, Joe Kilner, Hansung Kim, Takeshi Takai, Evren Imre, Chris Budd, Dan Casas, Margara Tejera, Martin Klaudiny, Sarim Muhammad, Peter Stroia-Willams, Mykyta Fastovets, Alexandros Neophytou

Collaborators: BBC, Framestore, The Foundry, Sony, Vicon, DNeg, Filmlight, Bodymetrics, BUF, QuanticDream

Funding: Royal Society, EPSRC, EU ICT, industry


Blue Peter
Blue Peter
Blue Peter
Blue Peter


BMVC
BIG
GUIDE




Huang, Casas & Hilton
2011