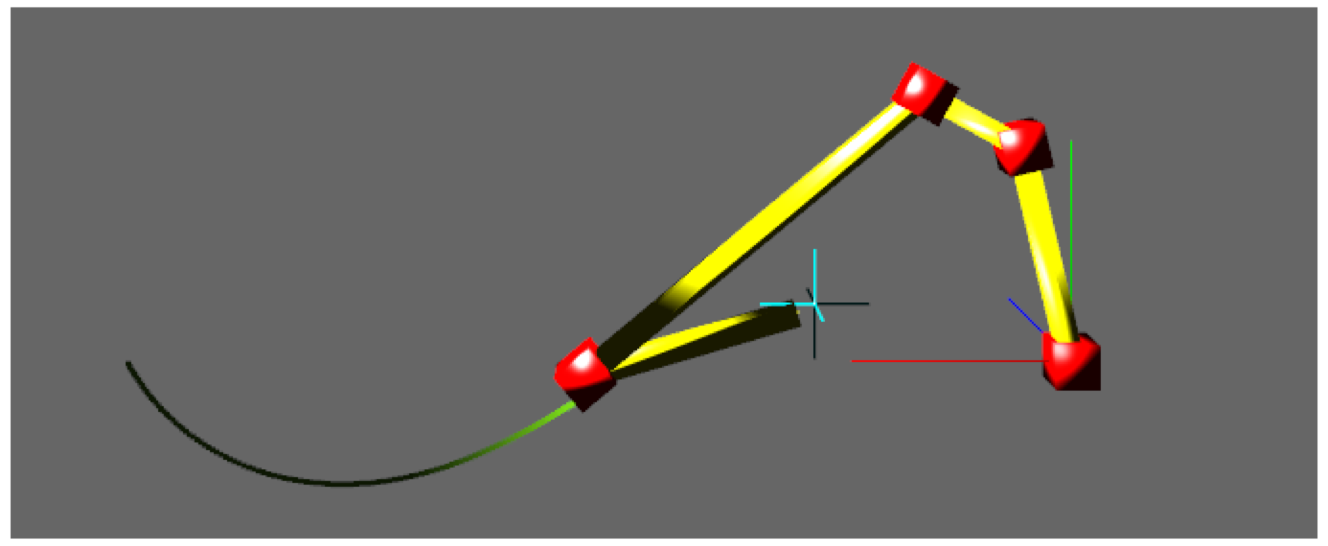
Doctor of Engineering (EngD) in Digital Entertainment, 1st Year

Centre for Digital Entertainment

Garoe Dorta Perez, Ieva Kazlauskaite and Richard Shaw   University of Bath

University of Bath

Inverse Kinematics

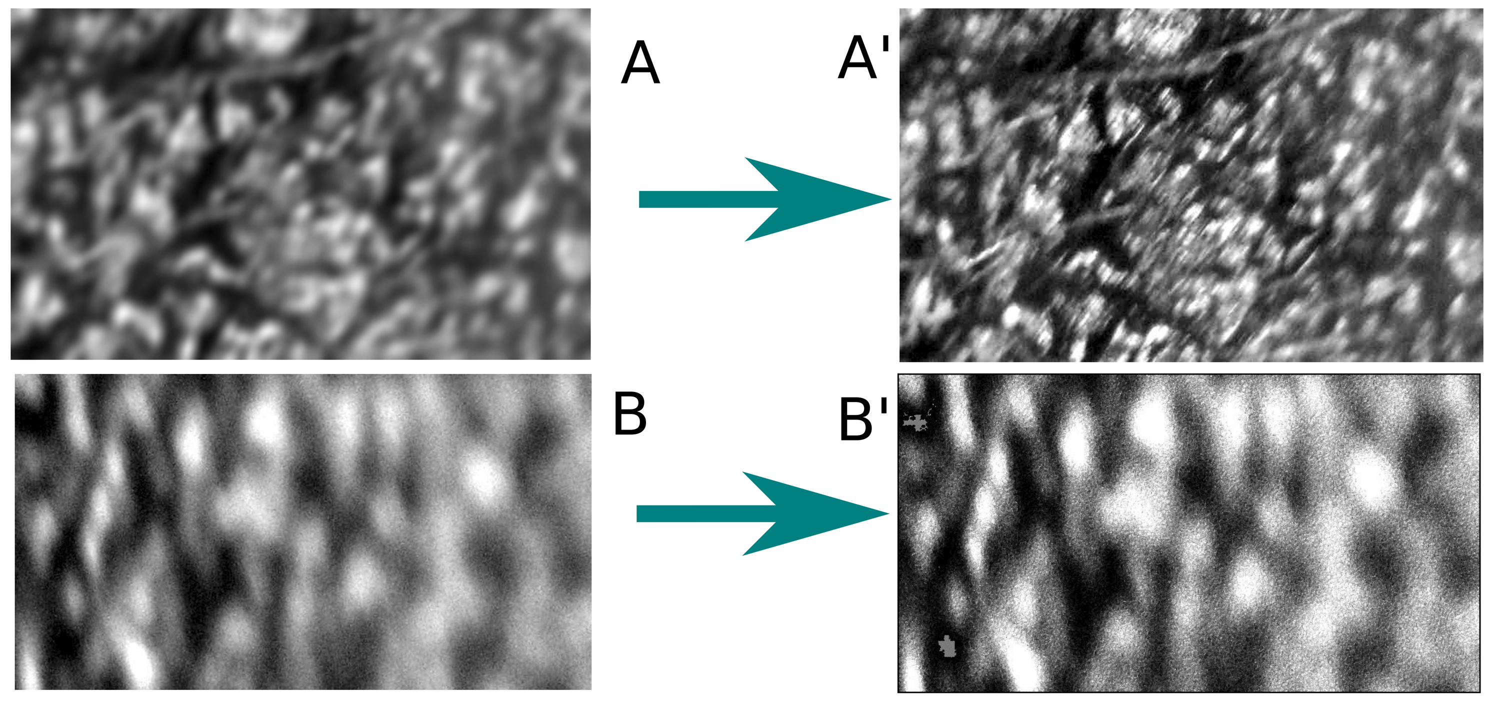


Inverse kinematics uses a known end-effector position $\mathbf{E}$ to calculate the required angles $\boldsymbol{\theta}$ between the parts which ensure that the object reaches the desired target position, such that

$$\mathbf{E} = f(\boldsymbol{\theta}) \rightarrow \boldsymbol{\theta} = f^{-1}(\mathbf{E})$$
$$\partial\mathbf{E} \approx J(\boldsymbol{\theta})\partial\boldsymbol{\theta} \rightarrow \partial\boldsymbol{\theta} \approx J^{-1}(\partial\mathbf{E}),$$

where f is the forward kinematics solver, J is the Jacobian matrix, and $J^{+} = (J^T J)^{-1} J^T$ is the pseudoinverse of J.

Skin Rendering



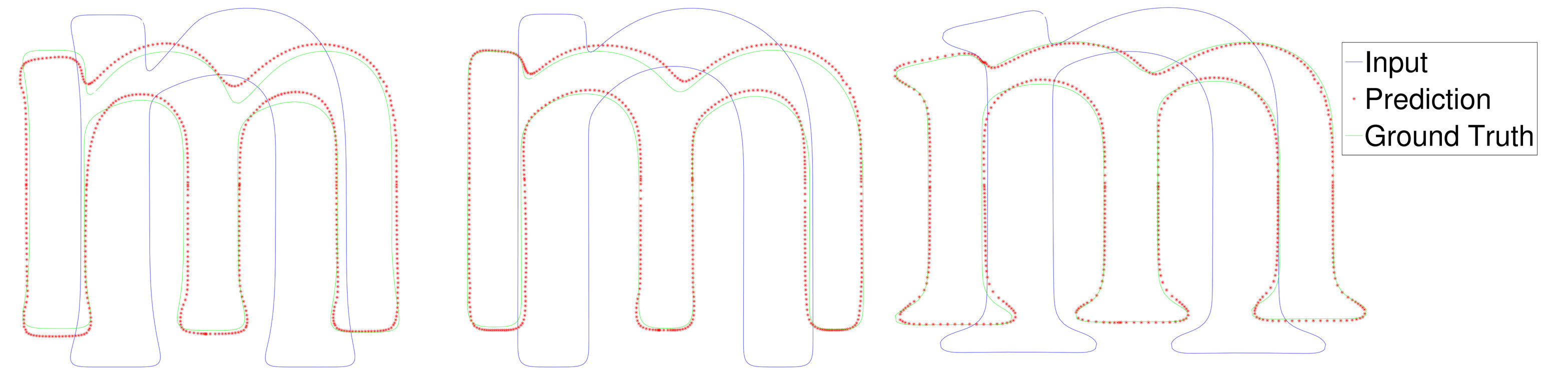
Hertzmann *et al.* [2] introduced a method to apply filters to images based on a best approximate match and a best coherence match pixel search. A high-resolution bump map B' can be synthesised from a lower resolution B and a pair of training samples A and A'.

References

[1] M. Alexa, D. Cohen-Or, and D. Levin. As-rigid-as-possible shape interpolation. In *Proc. of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00, pages 157–164, NY, USA, 2000.

[2] A. Hertzmann, C.E. Jacobs, N. Oliver, B. Curless, and D.H. Salesin. Image analogies. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, pages 327–340, 2001.

Gaussian Processes

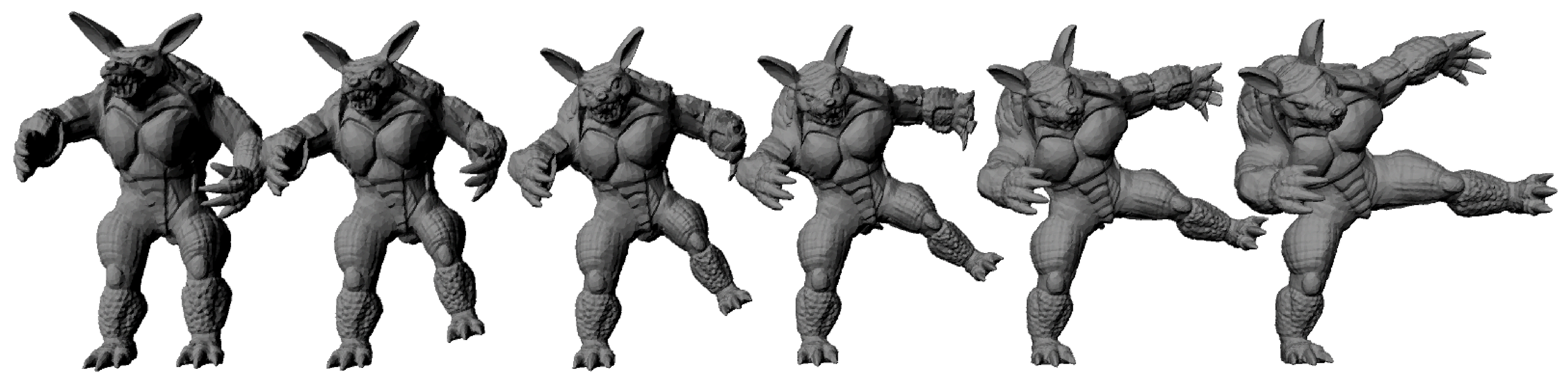


A Gaussian process is a random process that can be considered as an infinite-dimensional generalisation of the multivariate Gaussian distribution. The main assumption of Gaussian process regression is that our data can be represented as a sample from a multivariate normal distribution. A kernel function $k(x, x')$ must be chosen, and its parameters tuned to maximise the marginal likelihood $Pr(\boldsymbol{\theta}|\mathbf{X}, \mathbf{w})$

$$k(x, x') = \sigma_f^2 \exp\left[-(x - x')^2 \backslash 2\lambda^2\right]$$

Gaussian processes were used to predict the shape of fonts, given a training set. Even with very few training examples, the Gaussian process model gives a reasonable prediction of the shape of a font. The best results were achieved using an exponential kernel with optimised length scale and variance hyperparameters.

Shape Interpolation



Alexa *et al.* [1] introduced a transformation-based interpolation technique that aims to preserve the structure of the parts that are only translated or rotated between the two meshes. For each triangle, the transformation $\mathbf{A}$ is split into rotation and translation/shearing, both of which are interpolated linearly. The corresponding smooth transformation is estimated by minimising the error in Frobenius norm:

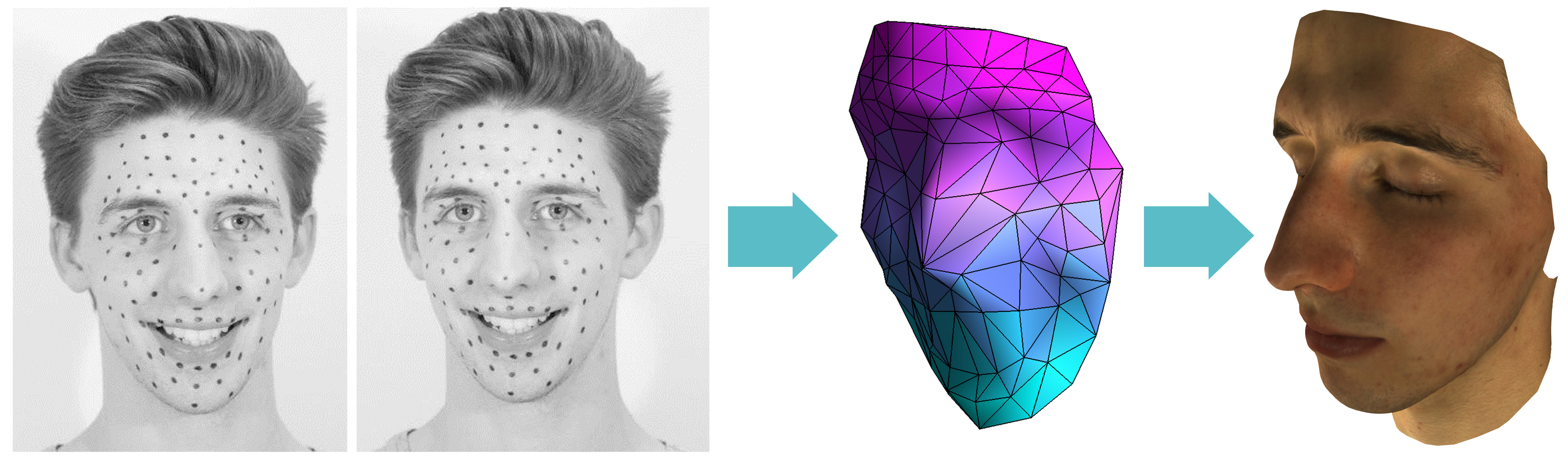
$$E_{V(t)} = \sum_{\Delta \in \mathcal{T}} ||\mathbf{A}_T(t) - \mathbf{B}_T(t)||_F^2,$$

where $V(t)$ are the intermediate positions of vertices, $\mathbf{A}$ is the ideal mapping, and $\mathbf{B}$ is the actual affine transformation.

Future Research

The students undertake an individual three-month summer project in a chosen research area before starting the industrial placement.

Performance-driven Facial Animation



A facial performance is captured using calibrated stereo cameras and tracking markers on the face. The marker positions produce a sparse 3D point cloud, which is in turn used to drive a high-resolution mesh. A new face $\mathbf{x}^*$ can be computed from the 'neutral-face' plus a weighted combination of blend-shape faces $\{\mathbf{x}_i\}_{i=1}^N$, where the required weights $\mathbf{w} = \{w_1 ... w_N\}$ are found through an optimisation procedure

$$\mathbf{x}^* = \sum_{i=1}^{N} w_i \mathbf{x}_i \quad \mathbf{w} = \arg\min_{\mathbf{w}} ||\mathbf{x}^* - \sum_{i=1}^{N} w_i \mathbf{x}_i||^2$$

The synchronised cameras are pre-calibrated using a checkerboard pattern and the intrinsic matrix $\mathbf{K}$ and external parameters $\mathbf{R}$ and $\mathbf{t}$ are recovered from the projection matrices $\mathbf{P}$ and $\mathbf{P}'$, where

$$\mathbf{u} = \mathbf{PX} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X}$$

The fundamental matrix $\mathbf{F}$ encompasses the intrinsic geometry between the two views and defines the epipolar constraint $\mathbf{u}'^T\mathbf{F}\mathbf{u} = 0$ from which we compute epipolar lines and rectify the stereo images to make the correspondence problem easier. Tracked facial points are constrained to lie on epipolar lines so that they intersect at points in 3D space. The captured sparse performance is then mapped to the detailed face mesh using the thin plate splines algorithm.

Contact Information

Web http://www.digital-entertainment.org
Garoe Dorta Perez g.dorta.perez@bath.ac.uk
Ieva Kazlauskaite ik359@bath.ac.uk
Richard Shaw richard.o.shaw@gmail.com