# Fast Tracking of Deformable Objects in Depth and Colour Video

Andreas Jordt
http://www.mip.informatik.uni-kiel.de
Reinhard Koch
http://www.mip.informatik.uni-kiel.de

Multimedia Information Processing
Department of Computer Science
Christian-Albrechts-University
Kiel, Germany

One challenge in computer vision is the joint reconstruction of deforming objects from colour and depth videos. So far, a lot of research has focused on deformation reconstruction based on colour images only, but as range cameras like the recently released Kinect become more and more common, the incorporation of depth information becomes feasible. In this article, a new method is introduced to track object deformation in depth and colour image data.
The tracking is done by translating, rotating, and deforming a prototype of an object such that it fits the depth and colour data best. The prototype can either be cut out from the first depth/colour frame of the input sequence or an already known textured geometry can be used. A NURBS [2] based deformation function allows to decouple the geometrical object complexity from the complexity of the deformation itself, providing a relatively low dimensional space to describe arbitrary 'realistic' deformations. This is done by first approximating the object surface using a standard NURBS function $\mathcal{N}$ and then registering every object vertex to the surface as depicted in figure 1.

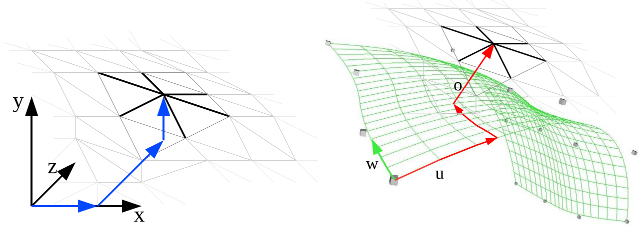


Figure 1: Left: a mesh vertex described by its native $(x,y,z)$ coordinates. Right: a vertex mesh described by the parameters $(u,w)$ and offset $o$.

For every vertex $v = (x,y,z)$ the closest NURBS surface point $\mathcal{N}(u,w)$ is determined, such that the surface normal directly points at $v$. Defining the signed distance between $v$ and $\mathcal{N}(u,w)$ along the surface normal as offset $o$, the vertex $v$ can be described completely by the parameters $(u,w,o)$ for a given NURBS surface function $\mathcal{N}$.
An object mesh, registered to a NURBS surface in such a way follows the translation and rotation of the NURBS surface as well as its deformation (see fig 2) when addressed via $(u,w,o)$.

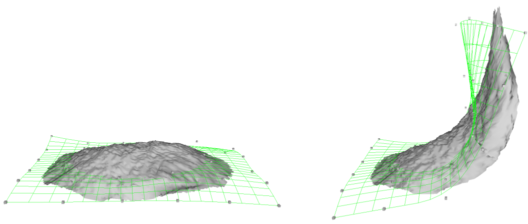

Figure 2: A triangle mesh registered to a NURBS function is deformed according to the NURBS surface it is registered to.

We assume to have a known, calibrated [3] acquisition setup, so synthesised depth and colour projections can be generated according to the real capturing devices. A mean square error sum over all vertices is calculated according to their fit of the synthesised pixels to the real input data. Additionally, a penalty is introduced to prevent the NURBS surface from shrinking or stretching too far.

The error function at hand provides a measure to determine the fit of a deformation to the image data. This error function is minimised for every frame separately using the global optimiser CMA-ES [1]. Figure 3 shows tracking results for a piece of flabby silicon.

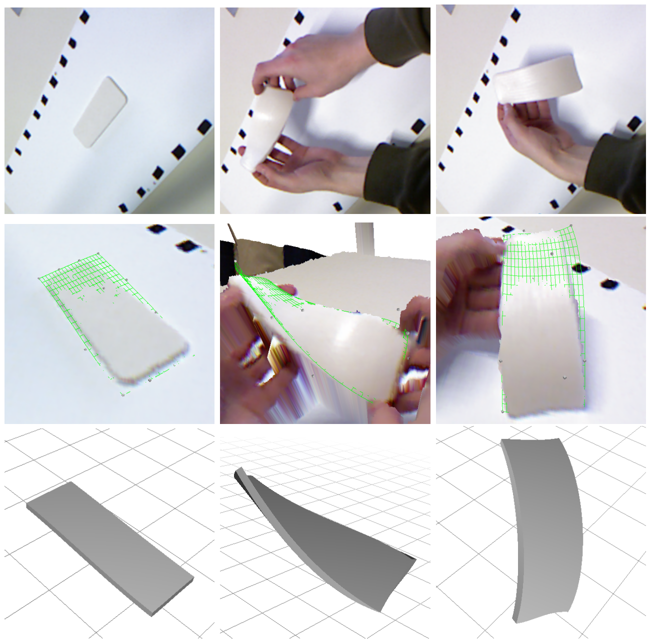

Figure 3: First row: Colour images from a deformation sequence recorded with a Kinect camera. Second row: The resulting coloured 3D mesh and the NURBS surface approximating the deformation of the tracked object. Third row: The deformation applied to a cube with a shape similar to the tracked object

Since the deformation information is completely decoupled from the object geometry, any other object can be registered to a NURBS function describing a deformation sequence. The bottom row of figure 3 shows a simple box model emulating the deformation of the tracked object from the input sequence.

Modelling the deformation tracking as an analysis by synthesis problem is robust, but usually computationally expensive, however, a set of optimisations is introduced, allowing a very fast calculation of the error function. The object vertices store the computationally expensive part of the NURBS evaluation in a 'NURBS cache', reducing the synthesis calculation to only few mathematical operations.
The small number of function evaluations CMA-ES needs to find the solution inside high dimensional search spaces leads to a system that is capable of tracking complex deformations of large objects (6000 triangles and more) with more than 6Hz on a common desktop machine.

[1] N. Hansen. The CMA evolution strategy: a comparing review. In J.A. Lozano, P. Larranaga, I. Inza, and E. Bengoetxea, editors, *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pages 75–102. Springer, 2006.

[2] Les Piegl and Wayne Tiller. *The NURBS book (2nd ed.)*. Springer-Verlag New York, Inc., New York, NY, USA, 1997. ISBN 3-540-61545-8.

[3] Ingo Schiller, Christian Beder, and Reinhard Koch. Calibration of a pmd camera using a planar calibration object together with a multi-camera setup. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume Vol. XXXVII. Part B3a, pages 297–302, Beijing, China, 2008. XXI. ISPRS Congress.