Online Dense Non-Rigid 3D Shape and Camera Motion Recovery

Antonio Agudo[1]
aagudo@unizar.es
J. M. M. Montiel[1]
josemari@unizar.es
Lourdes Agapito[2]
l.agapito@cs.ucl.ac.uk
Begoña Calvo[1,3]
bcalvo@unizar.es

[1] Instituto de Investigación en Ingeniería de Aragón (I3A), Universidad de Zaragoza, Zaragoza, Spain.
[2] Department of Computer Science, University College London, London, United Kingdom.
[3] Centro de Investigación en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Zaragoza, Spain.

Recovering 3D reconstruction from 2D images of a deforming object is an inherently ill-posed problem and it usually requires prior knowledge on the scene structure. Most approaches model the non-rigid shape using a low-rank shape constraint [5, 7, 12] combined with additional priors such as temporal smoothness [4, 6, 12], smooth-time trajectories [3, 8], spatial smoothness [7, 12] and inextensibility constraints [13]. Although accurate results have been obtained in recent years, these approaches process all the frames in the sequence in batch manner after video acquisition, preventing them from *online* and *real-time* applications. While sequential rigid real-time solutions exist for a sparse set of salient points [9] and even per-pixel dense reconstruction [10], online estimation of non-rigid objects from a single camera based only on the measurements up to that moment remains a challenging problem. Only recently, sequential formulations have emerged using either sparse [1, 11] or dense correspondences [2].

In this paper, we propose a *sequential solution* to simultaneously recover camera motion and the 3D reconstruction of non-rigid objects from 2D point tracks in a monocular image sequence as the data arrives. We employ a *probabilistic linear subspace* to encode the non-rigid 3D shape at each frame where the shape basis is computed by modal analysis. Our contribution is to propose a new mode shape computation algorithm that makes possible the full extension of the method to dense shapes, and a sequential expectation maximization based algorithm to solve the latent variable problem providing both efficient and more accurate solutions with respect to state-of-the-art sequential methods. Our approach works in two stages: shape basis computation and online estimation.

In stage one, we estimate a shape at rest using a few initial frames, and then the surface is discretized by means of a soup of triangular finite elements where applying the continuum mechanics. The mode shapes can be computed by modal analysis solving an eigenvalue problem [2] obtaining two non-rigid families: bending and stretching modes. The first one is affordable to compute even for dense cases, but only it is valid for out-of-plane stretching deformations. However, to code shapes undergoing stretching deformations, stretching modes have to be included in the shape basis. Unfortunately, computing these mode shapes may become prohibitive –sometimes unfeasible– for some dense cases in terms of computational and memory requirements. It is our first contribution to propose a *growth of modes* (Fig. 1) to easily compute all frequency spectrum and to obtain the stretching modes at quite affordable cost. We compute the mode shapes on a down-sampled shape at rest, and then the sparse shape basis is grown back to dense exploiting the shape functions within the finite element.

In stage two, equipped with this low-rank deformable shape basis, it is our second contribution to propose an online expectation maximization based algorithm over a sliding temporal window of frames to estimate the model parameters as the data arrives. Since the basis weights are modeled with hierarchical priors, these can be marginalized out and we only optimize a small number of parameters per frame obtaining a low cost system that potentially runs in real-time.

We show successful non-rigid 3D reconstruction results on several challenging sequences from highly extensible to inextensible deformations. We also show the advantages of our approach w.r.t. competing sequential methods. Our approach is also valid from sparse to dense data, do not require any training data and can deal with missing data.

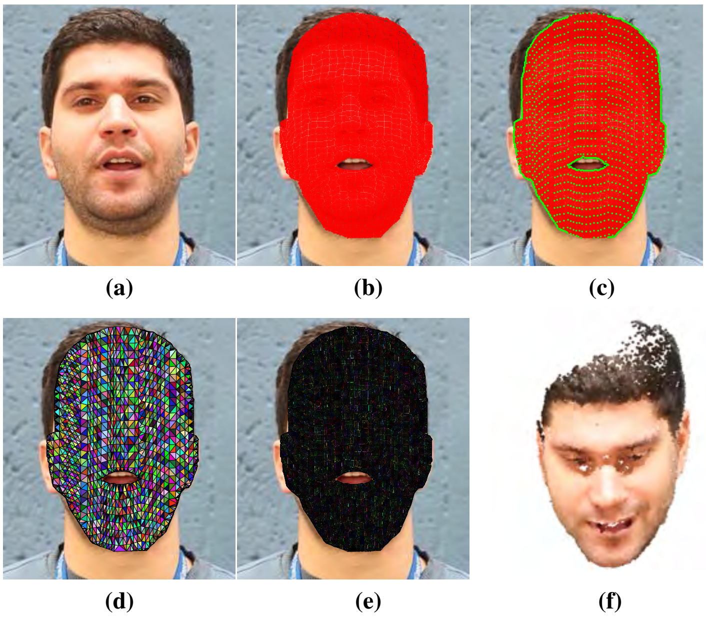

Figure 1: **Growth of modes for dense shapes. (a):** Reference image plane to compute optical flow. **(b):** Dense 2D tracking of p points. **(c):** Subsample of dense shape into q points (green points) with $q << p$. **(d):** Delaunay triangulation for sparse mesh. **(e):** Active search to match every point in the sparse mesh. **(f):** General viewpoint of the 3D reconstruction. Best viewed in color.

[1] A. Agudo, B. Calvo, and J. M. M. Montiel. Finite element based sequential bayesian non-rigid structure from motion. In *CVPR*, 2012.

[2] A. Agudo, L. Agapito, B. Calvo, and J. M. M. Montiel. Good vibrations: A modal analysis approach for sequential non-rigid structure from motion. In *CVPR*, 2014.

[3] I. Akhter, Y. Sheikh, S. Khan, and T. Kanade. Trajectory space: A dual representation for nonrigid structure from motion. *TPAMI*, 33 (7):1442–1456, 2011.

[4] A. Bartoli, V. Gay-Bellile, U. Castellani, J. Peyras, S. Olsen, and P. Sayd. Coarse-to-fine low-rank structure-from-motion. In *CVPR*, 2008.

[5] C. Bregler, A. Hertzmann, and H. Biermann. Recovering non-rigid 3D shape from image streams. In *CVPR*, 2000.

[6] A. Del Bue, X. Llado, and L. Agapito. Non-rigid metric shape and motion recovery from uncalibrated images using priors. In *CVPR*, 2006.

[7] R. Garg, A. Roussos, and L. Agapito. Dense variational reconstruction of non-rigid surfaces from monocular video. In *CVPR*, 2013.

[8] P. F. U. Gotardo and A. M. Martinez. Non-rigid structure from motion with complementary rank-3 spaces. In *CVPR*, 2011.

[9] G. Klein and D. W. Murray. Parallel tracking and mapping for small AR workspaces. In *ISMAR*, 2007.

[10] R. Newcome, S. Lovegrove, and A. J. Davison. DTAM: Dense tracking and mapping in real-time. In *ICCV*, 2011.

[11] M. Paladini, A. Bartoli, and L. Agapito. Sequential non rigid structure from motion with the 3D implicit low rank shape model. In *ECCV*, 2010.

[12] L. Torresani, A. Hertzmann, and C. Bregler. Nonrigid structure-from-motion: estimating shape and motion with hierarchical priors. *TPAMI*, 30(5):878–892, 2008.

[13] S. Vicente and L. Agapito. Soft inextensibility constraints for template-free non-rigid reconstruction. In *ECCV*, 2012.