Good Vibrations: A Modal Analysis Approach for Sequential Non-Rigid Structure from Motion

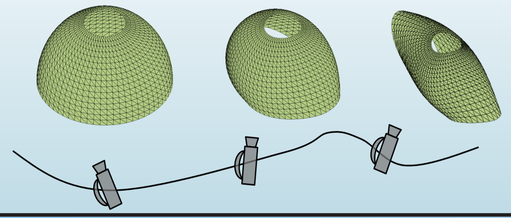
Antonio Agudo[1], Lourdes Agapito[2], Begoña Calvo[1] and J. M. M. Montiel[1]
[1]I3A-Universidad de Zaragoza, [2]University College London

European Research Council
Established by the European Commission

University College London
UCL
Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

Non-Rigid SfM

- 3D reconstruction of non-rigid objects from 2D temporal correspondences in a monocular image sequence.
- So far most approaches are batch.
- **Our Goal:** A sequential NRS*f*M method that is **real-time capable**.



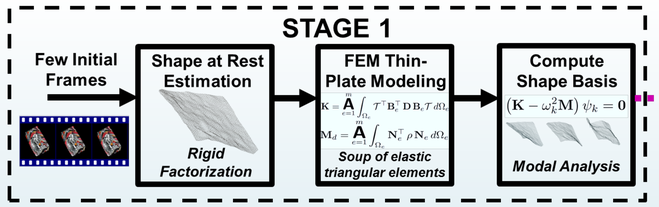
Our Contribution

- A **new shape basis** based on modal analysis with bending and stretching modes. No learning step.
- An **online solution** to NRS*f*M that estimates camera pose and deformable shape on a per-frame basis.

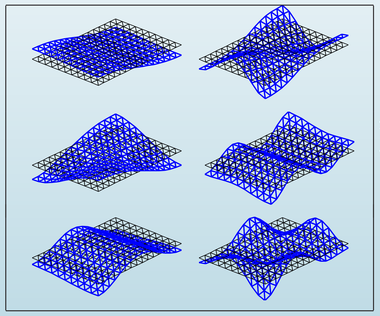
Our Approach

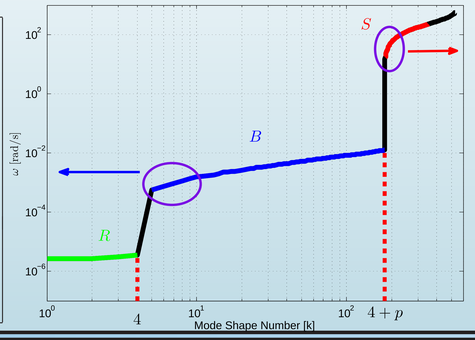
- **Stage 1:** Computation of the shape basis. Only an estimation of the 3D shape at rest is needed.
- **Stage 2:** Bundle Adjustment over a sliding window to optimize deformation coefficients and rigid pose.
- Suitable to model a wide variety of deformations: from inextensible to highly **extensible surfaces**.

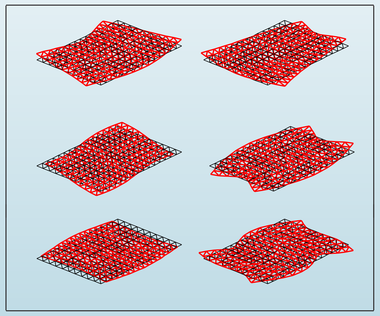
Stage 1: Physics-based Modal Shape Basis Computation



- Mode shapes are ordered by ***frequency spectrum***, i.e. according to the energy needed to excite each mode:
  - **Rigid modes (R):** Rigid body motions.
  - **Bending modes (B):** out-of-plane deformations. Interval $[5, p+4]$ $(quasi-isometric)$.
  - **Stretching modes (S):** in-plane extensible deformations. Interval $[p+5, 2p]$.



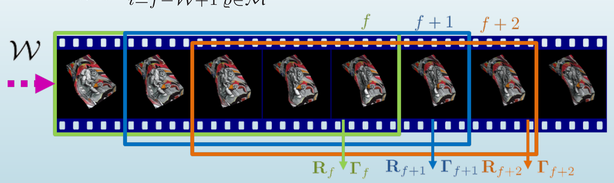




Stage 2: Online Sliding Window Optimization

Our energy includes image reprojection error terms for all visible points within a temporal sliding window of $\mathcal{W}$ frames in addition to temporal smoothness priors. The optimization is solved using sparse Levenberg-Marquardt:

$$\underset{\mathbf{R}_i, \mathbf{\Gamma}_i}{\arg\min} \sum_{i=f-\mathcal{W}+1}^{f} \sum_{\varrho \in \mathcal{M}} \| \mathbf{W}_{i\varrho} - \mathbf{R}_i \left(\mathbf{S}_\varrho + \left(\mathbf{I}_3 \otimes \mathbf{\Gamma}_i^\top \right) \mathbf{\Psi}_\varrho \right) \|_{\mathcal{F}}^2 + \lambda_q \sum_{i=f-\mathcal{W}+1}^{f} \| \mathbf{q}_i - \mathbf{q}_{i-1} \|_{\mathcal{F}}^2 + \lambda_\gamma \sum_{i=f-\mathcal{W}+1}^{f} \| \mathbf{\Gamma}_i^\top - \mathbf{\Gamma}_{i-1}^\top \|_{\mathcal{F}}^2$$



- Non-rigid 3D displacement is approximated as a **linear combination** with Γ_i weight coefficients.
- Orthographic camera model:

$$\mathbf{W}_f = \begin{bmatrix} u_{f1} & u_{f2} & \dots & u_{fp} \\ v_{f1} & v_{f2} & \dots & v_{fp} \end{bmatrix} = \Pi \mathbf{Q}_f (\mathbf{S}_f + \mathbf{T}_f)$$

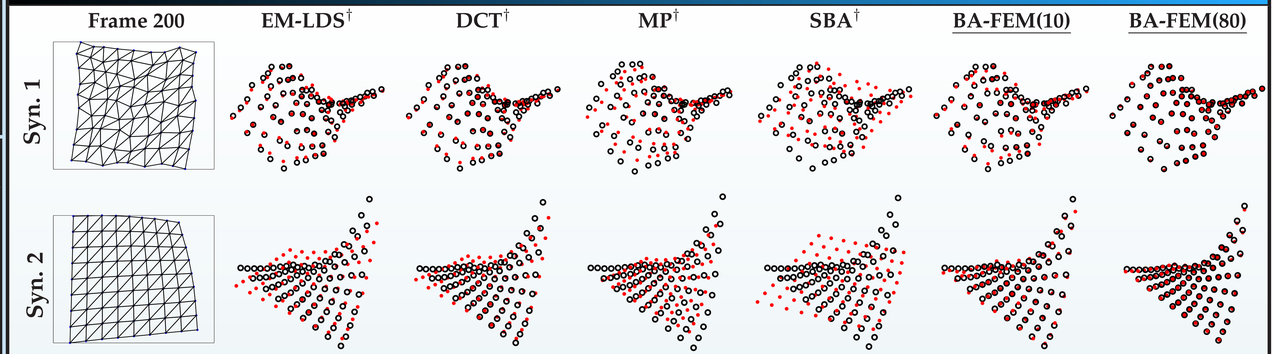
Conclusions

- Real-time capable, accurate and scalable solution.
- Our method can handle both sparse and dense meshes and can cope with structured missing data.
- No additional inextensibility constraints.

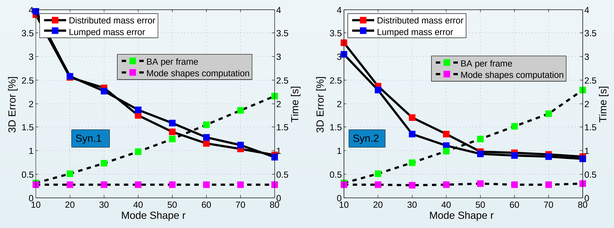
Future Work

- Incorporate feature tracking and outlier detection to provide a unified framework.
- To exploit biomechanical priors.
- Application to Minimally Invasive Surgery.

Experimental Results



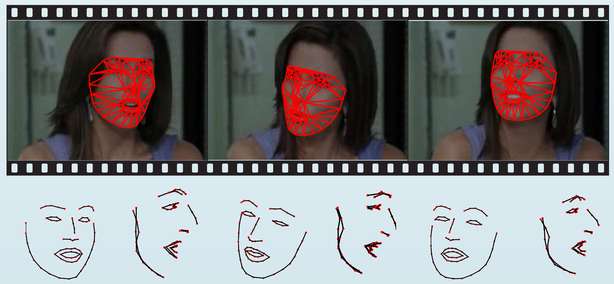
SYNTHETIC SEQUENCES



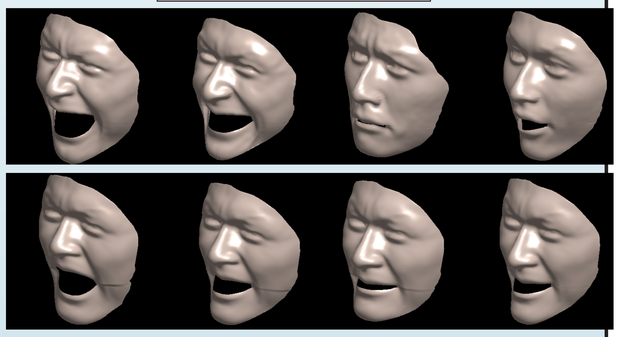
e_{3D}%	Syn. 1	Syn. 2	B†	S†
EM-LDS†	11.12(2)	10.92(2)	✓	
DCT†	9.25(2)	11.81(5)	✓	
MP†	12.42(2)	18.84(2)	✓	
SBA†	14.03(16)	20.90(8)		✓
BA-FEM	**3.89(10)**	**3.04(10)**		✓
BA-FEM	**0.86(80)**	**0.82(80)**		✓

† EM-LDS [Torresani *et al.* PAMI'08], DCT [Akhter *et al.* PAMI'11], MP [Paladini *et al.* CVPR'09], SBA [Paladini *et al.* ECCV'10]. B: Batch algorithm and S: Sequential algorithm.

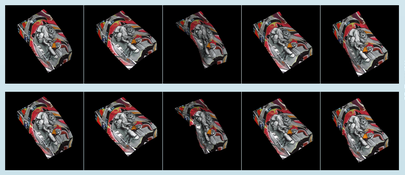
ACTRESS SEQUENCE

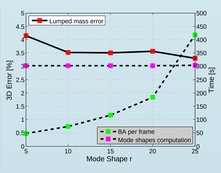


DENSE FACE SEQUENCE

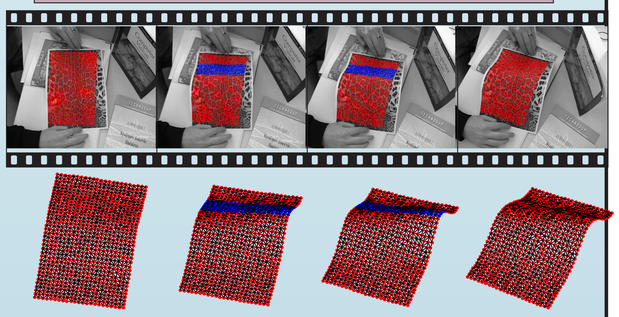


DENSE FLAG SEQUENCE





PAPER BENDING SEQUENCE with STRUCTURED OCCLUSION



	Size		Run-time (sec)		Error
Seq.	p‡	r‡	MSC‡	BA‡	e_{3D}%
Syn. 1	81	10-80	0.3	0.3-2.2	3.89-0.86
Syn. 2	81	10-80	0.3	0.3-2.2	3.04-0.82
Actress	68	10	0.4	0.3	-
Bending	828	10	60	0.85	-
Flag	9,622	5-25	300	47-416	4.14-3.29
Face	28,887	5	1540	60	4.64

‡ MSC: Modes Shape Computation, BA: Bundle Adjustment time, p and r number of points and modes respectively.

Acknowledgements: This work was partly funded by the MINECO projects DIP2012-32168 and DPI2011-27939-C02-01; by the ERC Starting Grant agreement 204871-HUMANIS; and by a scholarship FPU12/04886.