Simultaneous Pose and Non-Rigid Shape with Particle Dynamics

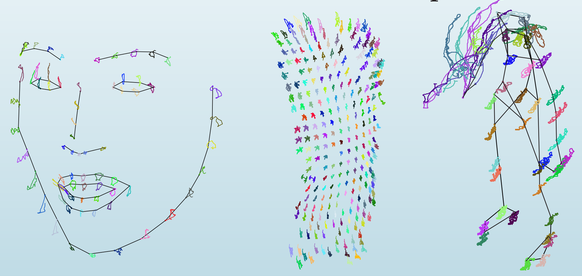
Antonio Agudo[1], Francesc Moreno-Noguer[2]
[1]I3A-Universidad de Zaragoza, [2]Institut de Robòtica i Informàtica Industrial (CSIC-UPC)

CVPR 2015 BOSTON JUNE 7-12

Institut de Robòtica i Informàtica Industrial
Instituto Universitario de Investigación en Ingeniería de Aragón
Universidad Zaragoza

Problem: Non-Rigid SfM

- 3D reconstruction of deformable objects from 2D temporal tracks in a monocular video.
- So far most approaches use global models and batch operation.
- *Our Goal:* A *sequential* NRS*f*M method based on *local* models that is **real-time capable**.



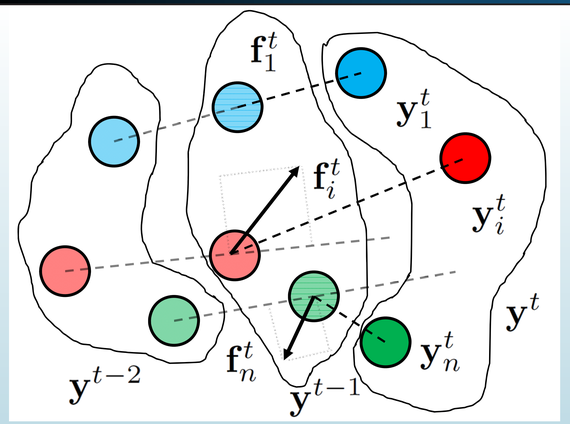
Contribution

- A piecewise model at **particle level** based on physical constraints to model the non-rigid deformation between consecutive frames.
- An **online solution** to NRS*f*M that estimates time-varying shape and camera motion.

Approach

- A dynamic object is modeled as a system of particles with force-perturbed motion models.
- Model parameters are estimated by Bundle Adjustment over a sliding window of images.
- Suitable to model a wide variety of deformations: from *articulated* to *non-rigid* motion even for **discontinuous surfaces**.

Physics-Inspired Motion Model



- We employ *Newton's second law of motion* $\mathbf{f}_i^t = m_i \mathbf{a}_i^t$ to constrain deformation per point. The acceleration at time t is approximated using second-order finite differences:

$$\mathbf{f}_i^t \approx m_i \left[\frac{\mathbf{y}_i^{t-2} - 2\mathbf{y}_i^{t-1} + \mathbf{y}_i^t}{(\Delta t)^2} \right]$$

- The 3D position of the particles at time t follows the dynamical model:

$$\mathbf{y}^t = \mathbf{f}^t + 2\mathbf{y}^{t-1} - \mathbf{y}^{t-2} = \mathbf{f}^t + \mathbf{d}^t$$

- The constant velocity model $\mathbf{d}^t$ is force-perturbed by $\mathbf{f}^t$ at each image.

Online Non-linear Optimization

- Orthographic camera model:

$$\mathbf{P}^t = [\mathbf{p}_1^t, \ldots, \mathbf{p}_n^t] = \mathbf{R}^t \mathbf{Y}^t + \mathbf{T}^t$$

- We model the non-rigid deformation of an object represented by a system of particles.

The global energy we define includes *image reprojection* error terms for all visible points within a temporal sliding window of 3 frames in addition to *spatial* and *temporal* smoothness priors. The optimization is solved using sparse Levenberg-Marquardt:

$$\mathcal{A}\left(\mathbf{R}^j, \mathbf{t}^j, \mathbf{F}^t\right) = \sum_{j=t-2}^{t} \sum_{\nu \in \mathcal{V}^j} \|\mathbf{p}_\nu^j - \mathbf{R}^j \mathbf{y}_\nu^j - \mathbf{t}^j\|_{\mathcal{F}}^2 + \alpha_e \sum_{e=1}^{n_e} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{d_e^{r\,2}}{2\sigma^2}\right) |d_e^r - d_e^t(\mathbf{F}^t)|$$
$$+\alpha_s \|\mathbf{Y}^t\left(\mathbf{F}^t\right) - \mathbf{Y}^{t-1}\|_{\mathcal{F}}^2 + \alpha_p \sum_{j=t-1}^{t} \|\mathbf{q}^j - \mathbf{q}^{j-1}\|_{\mathcal{F}}^2 + \alpha_t \sum_{j=t-1}^{t} \|\mathbf{t}^j - \mathbf{t}^{j-1}\|_{\mathcal{F}}^2$$

Conclusions

- Newton's second law of motion to model non-rigid deformations into bundle adjustment.
- Our method can handle different types of deformations and it can cope with missing data.

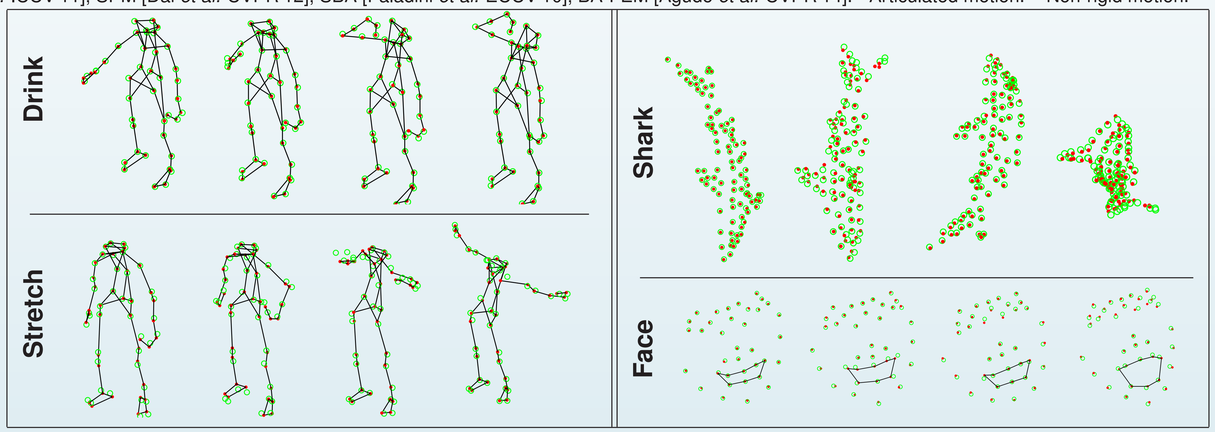
Future Work

- Generalization to full perspective cameras.
- To cope multiple non-rigid objects.
- Simultaneous feature tracking and outliers detection into a single process.

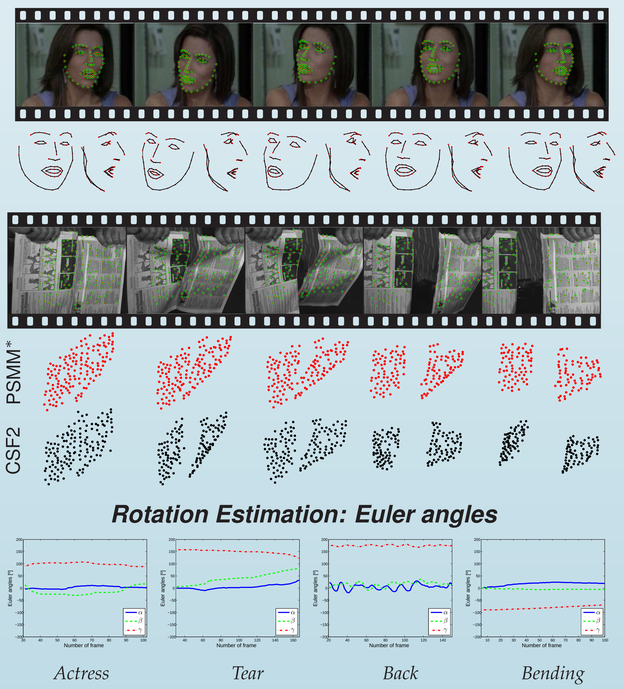
Experimental Evaluation

Seq. \ Met.	*Batch Methods* EM-PPCA†	MP†	PTA†	CSF2†	KSTA†	SPM†	*Sequential Methods* SBA†	BAFEM†	PSMM
Drink‡	5.56(5)	4.14(6)	1.38(13)	1.14(6)	**0.94(12)**	1.60(12)	11.25(12)	-	**1.93**
Stretch‡	13.72(15)	8.13(5)	3.85(8)	2.46(8)	2.00(7)	**1.86(11)**	17.61(20)	-	**5.76**
Yoga‡	11.89(14)	12.98(8)	2.42(8)	1.84(7)	2.12(7)	**1.65(10)**	15.84(20)	-	**6.65**
Shark+	1.82(2)	9.34(23)	5.91(6)	1.09(5)	**1.03(3)**	6.29(2)	8.81(5)	-	**6.99**
Jacky+	**1.80(5)**	2.74(5)	2.69(3)	1.93(5)	2.12(4)	1.82(7)	2.90(16)	3.43(15)	**2.80**
Face+	7.30(9)	3.77(7)	5.79(2)	6.34(5)	6.14(8)	**2.67(9)**	6.92(27)	6.89(2)	**4.49**

† EM-PPCA [Torresani *et al.* PAMI'08], MP [Paladini *et al.* CVPR'09], PTA [Akhter *et al.* PAMI'11], CSF2 [Gotardo *et al.* CVPR'11], KSTA [Gotardo *et al.* ICCV'11], SPM [Dai *et al.* CVPR'12], SBA [Paladini *et al.* ECCV'10], BA-FEM [Agudo *et al.* CVPR'14]. ‡Articulated motion. +Non-rigid motion.

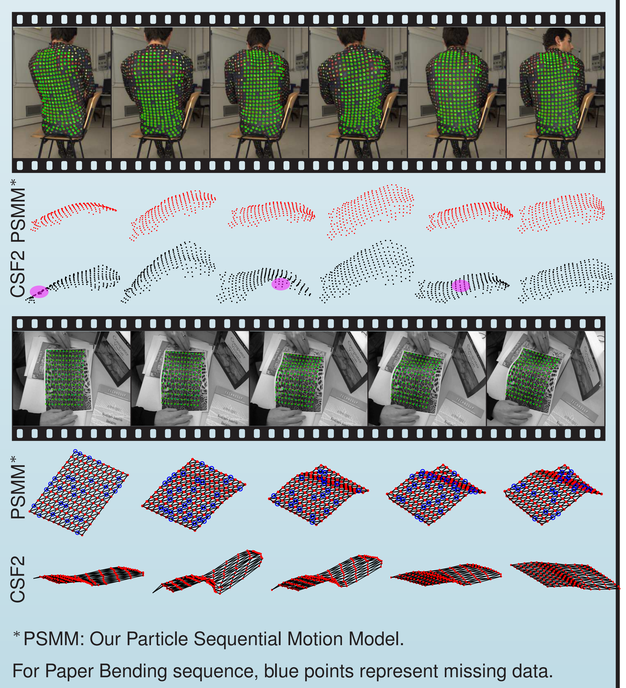


Real Video Sequences



Rotation Estimation: Euler angles

Actress *Tear* *Back* *Bending*



*PSMM: Our Particle Sequential Motion Model.

For Paper Bending sequence, blue points represent missing data.