PROBLEM STATEMENT
Given an uncalibrated monocular video where both rigid and non-rigid scenes can be observed.
We want to jointly and sequentially retrieve the self-calibration of the camera, the 3D non-rigid shape model, and the full camera trajectory.
We propose a Bayesian filtering approach based on a sum-of-Gaussians filter composed of a bank of non-rigid extended Kalman filters. Neither training data nor a calibration pattern are needed.

SUM OF GAUSSIAN (SOG) FILTER
- In our SoG filter, we approximate a probability density function as a sum of weighted multivariate Gaussians as:
  \[ p(x) = \sum_{g=1}^{G} \gamma_g \mathcal{N}(x; \mu_g, P_g) \]
- The SoG filter exploits a bank of Extended Kalman Filters (EKF). Every of them is estimated by means of a prediction-update strategy.
- The weight coefficients are updated every frame, removing those with low factor \( \gamma_g \).
- An overall mean \( \hat{x}_{k|k} \) and covariance \( P_{k|k} \) for the SoG filter can be considered for visualization:
  \[
  \hat{x}_{k|k} = \sum_{g=1}^{G} \gamma_g \hat{x}_{k|k}^g, \\
  P_{k|k} = \sum_{g=1}^{G} \gamma_g \left( P_{k|k}^g + \left[ \hat{x}_{k|k}^g - \hat{x}_{k|k} \right] \left[ \hat{x}_{k|k} - \hat{x}_{k|k}^g \right]^\top \right)
  \]

SELF-CALIBRATION NON-RIGID SOG
- The state of the camera is represented by an 18-dimensional vector, including calibration \((\alpha, \beta_x, \beta_y, k_x, k_y)\), camera pose \(r\) and orientation \(q\), and linear and angular \(\omega\) velocities:
  \[
  m_{k+1} = \begin{bmatrix}
  \alpha_{k+1} \\
  \beta_x_{k+1} \\
  \beta_y_{k+1} \\
  k_x_{k+1} \\
  k_y_{k+1} \\
  r_k + (v_x + \Delta v) \Delta t \\
  q_k \times q(v_{\omega} + \Delta \omega) \Delta t \\
  v_{\omega} + \Delta \omega \\
  \omega_{\omega} + \Delta \omega_{\omega}
  \end{bmatrix}
  \]
- The state of the scene is represented by a 3n-dimensional vector, with \(n\) the number of points.
- We use an elastic model based on finite elements, defining a compliance matrix \(C_0\) to relate the deformation of all points, and a vector of Gaussian acting forces \(\Delta f\). For every frame, the state function is:
  \[
  y_{k+1} = y_{k+1}(y_k, \Delta f) = y_k + C_0 \Delta f
  \]
- Our model can handle both inelastic and elastic materials, and it is computed every frame.
- A full perspective camera model is assumed.

TOTAL ESTIMATION FROM RGB VIDEO: ON-LINE CAMERA SELF-CALIBRATION, NON-RIGID SHAPE AND MOTION
ANTONIO AGUDO
INSTITUT DE ROBÒTICA I INFORMÀTICA INDUSTRIAL, CSIC-UPC, BARCELONA, SPAIN

CALIBRATION, MOTION AND NON-RIGID SHAPE RECONSTRUCTION

Non-Critical Motion Scenarios
Hand-held 320 × 240 camera/endooscope. Rigid and elastic sequences.

Critical Motion Scenarios
Hand-held 320 × 240 IEE1394 camera. Scene and/or calibration is not possible.

<table>
<thead>
<tr>
<th>Method</th>
<th>Tracking</th>
<th>Self-Calibration</th>
<th>Focus</th>
<th>Full</th>
<th>Batch</th>
<th>Sequential</th>
<th>Rigid</th>
<th>Non-Rigid</th>
<th>Geometric</th>
<th>Elastic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ours</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Critical Motion Sequences

Indoor and Loop Closing Sequences