Motion Denoising
with Application to Time-lapse Photography

Michael Rubinstein
MIT CSAIL

Ce Liu
Microsoft Research NE

Peter Sand

Fredo Durand
MIT

Bill Freeman
MIT
Time-lapse Videos

Construction

Natural phenomena

Medical

Biological/Botanical
For Personal Use Too!

9 months

7 years

16 years

http://www.danhanna.com/aging_project/p.html

Source: YouTube
“Stylized Jerkiness”
Motion Denoising

World

Time-lapse

Motion denoising
Motion Denoising
Time-lapse in Vision/Graphics Research

• Video summarization (video $\rightarrow$ time-lapse)

[Bennett and McMillan 2007]

[Pritch et al. 2008]

• Time-lapse editing

[Original]

[Without shadows]

[Sunkavalli et al. 2007]
Motion Denoising is Challenging!

- Naïve low-pass (temporal) filtering
  - Pixels of different objects are averaged

- Smoothing motion trajectories
  - Motion estimation in time-lapse videos is hard!
    * Motion discontinuities
    * Color inconsistencies
Formulation

- **Key idea:** long-term events in videos can be statistically explained within some local spatiotemporal support, while short-term events are more distinctive
  - Assumption: world is smooth
  - Short-term variation = *noise*, long-term variation = *signal*

- Our algorithm **reshuffles** the pixels in both space and time to maintain long-term events in the video, while removing short-term noisy motions
Formulation

\[ E(w) = \sum_p |I(p + w(p)) - I(p)| \]
\[ + \alpha \sum_{p,r \in N_t(p)} \|I(p + w(p)) - I(r + w(r))\|^2 \]
\[ + \gamma \sum_{p,q \in N(p)} \lambda_{pq} |w(p) - w(q)| \]

Fidelity (to input)
Temporal coherence (of the result)
Regularization (of the warp)

\[ p = (x, y, t) \]
\[ I - \text{input video, } I(p + w(p)) - \text{output video} \]
\[ N_t(p) - \text{Temporal neighbors of } p, N(p) - \text{Spatiotemporal neighbors of } p \]

\[ w(p) \in \{ (\delta_x, \delta_y, \delta_t) : |\delta_x| \leq \Delta_s, |\delta_y| \leq \Delta_s, |\delta_t| \leq \Delta_t \} - \text{displacement field} \]

\[ \lambda_{pq} = \exp(-\beta \|I(p) - I(q)\|^2), \quad \beta = (2 \langle \|I(p) - I(q)\|^2 \rangle)^{-1} \]
Optimization

- Optimized discretely on a 3D MRF
  - Nodes represent pixels
  - State space of each pixel = volume of possible spatiotemporal shifts

- Complicated (huge!) inference problem
  - E.g. $500^3$ nodes, $10^3$ states per node
  - Optimize using Loopy Belief Propagation
Optimization

- Potential functions
  - Message structure stored on disk; read and write message chunks on need

\[ \psi_p(w(p)) = |I(p + w(p)) - I(p)| \]

\[ \psi_{pr}^t(w(p), w(r)) = \alpha \|I(p + w(p)) - I(r + w(r))\|^2 + \gamma \lambda_{pr} |w(p) - w(r)| \]

\[ \psi_{pq}^t(w(p), w(q)) = \gamma \lambda_{pq} |w(p) - w(q)| \]

- message passing
  - Linear in state space + Pre-compute
  - Quadratic in state space (non convex)
  - But can be computed in linear time (distance transforms)
Multi-scale Processing

• Spatiotemporal video pyramid
  – Smooth spatially
  – Sample temporally

• Displacements in the coarse level used as centers for the search volume in the finer level
Results

Source

Result

Spatial Displacement

Temporal Displacement
Comparing with Other Optimization Techniques

<table>
<thead>
<tr>
<th>ICM</th>
<th>GCUT</th>
<th>LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="ICM Image" /></td>
<td><img src="image2" alt="GCUT Image" /></td>
<td><img src="image3" alt="LBP Image" /></td>
</tr>
</tbody>
</table>

Spatial Displacement
Results

Source  Result

Spatial Displacement  Temporal Displacement

future

past
Results
Comparison with Naïve Temporal Filtering
Support Size

Figure 7. Zoom-in on the rightmost plant in the sprouts sequence in four consecutive frames shows that enlarging the search volume used by the algorithm can greatly improve the results. “Large support” corresponds to a $31 \times 31 \times 5$ search volume, while “small support” is the $7 \times 7 \times 5$ volume we used in our experiments.
Motion-scale Decomposition
Motion-scale Decomposition
Other Scenarios
Future Work

• User-controlled motion scales
  – Not necessarily binary decomposition into long-term and short-term

• Modify the time-lapse capturing process to help post-processing
  – E.g. use short videos instead of still images and find best “path” through the video

• Explore motion-denoising with time-lapse from other domains
  – Embryos research, satellite imagery
Thank you!

http://csail.mit.edu/mrub/timelapse