# Video Registration: Key Challenges

Richard Szeliski
Microsoft Research

# Key Challenges

1. Mosaics and panoramas
2. *Object-based segmentation (MPEG-4)*
3. *Engineering Support Data*
4. *3-D registration*
5. Wide base-line stereo
6. *Correlation*
7. *Video registration vs. visual tracking*
8. Role of image features
9. Next most important problems
10. *Most successful solutions and approaches*

# Mosaics and Panoramas

- Accomplishments:
  - Fast “direct” methods
  - Feature-based approaches for large motion
  - True “3D” alignment (*spherical mosaics*)
  - Focal length (*f*) estimation (*self-calibration*)
  - Moving objects (*deghosting*)
  - Exposure compensation

# Mosaics and Panoramas


+

+

+ ……
+

+


# Mosaics and Panoramas

- Removing moving objects (*deghosting*)

# Mosaics and Panoramas

- Removing moving objects (*deghosting*)

# Mosaics and Panoramas

- Exposure compensation

# Mosaics and Panoramas

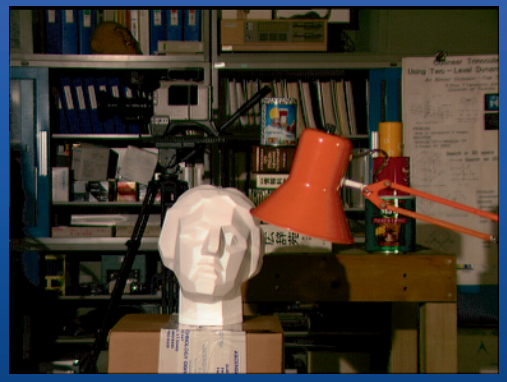
- Exposure compensation *and* deghosting

# Mosaics and Panoramas

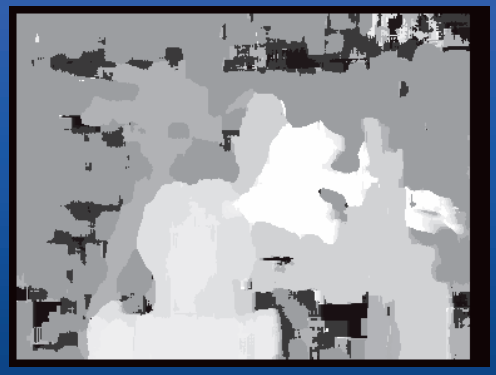
- Challenges:
  - dealing with parallax
  - dealing with more complex motions
  - large motions

# Wide-baseline stereo

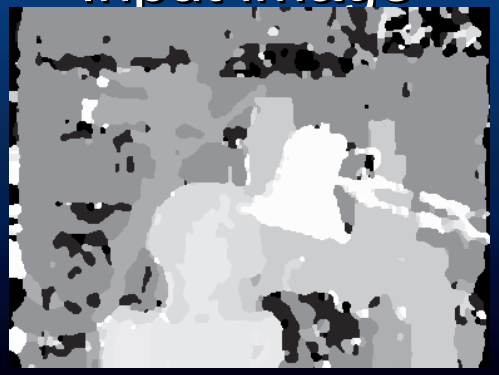
- *Is wide base-line stereo solved now?*



Input image



Sum Abs Diff



Mean field

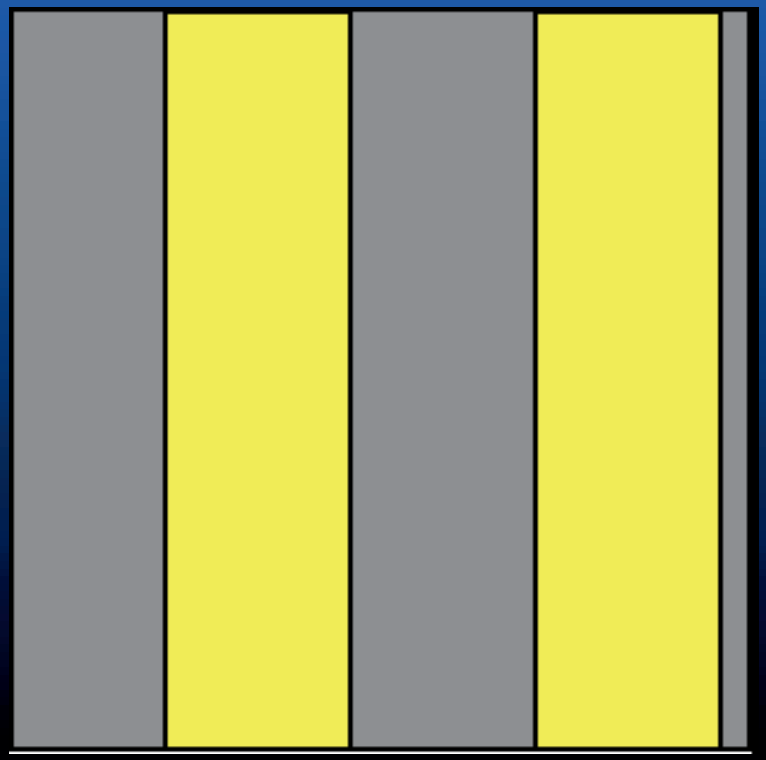
Graph cuts

# Wide-baseline stereo

- What about *really* wide baselines?

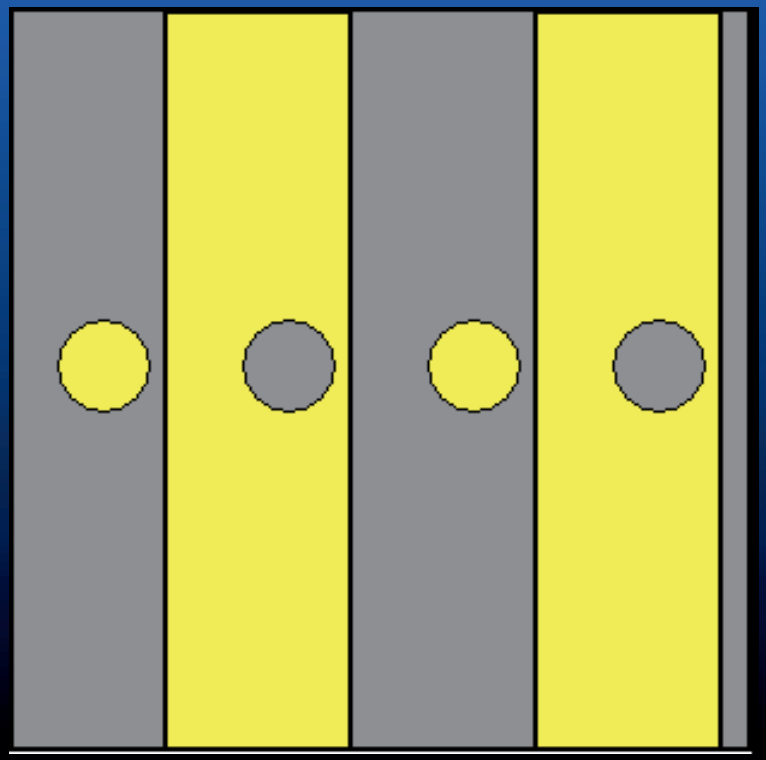
# Wide-baseline stereo

- What about *untextured regions*?



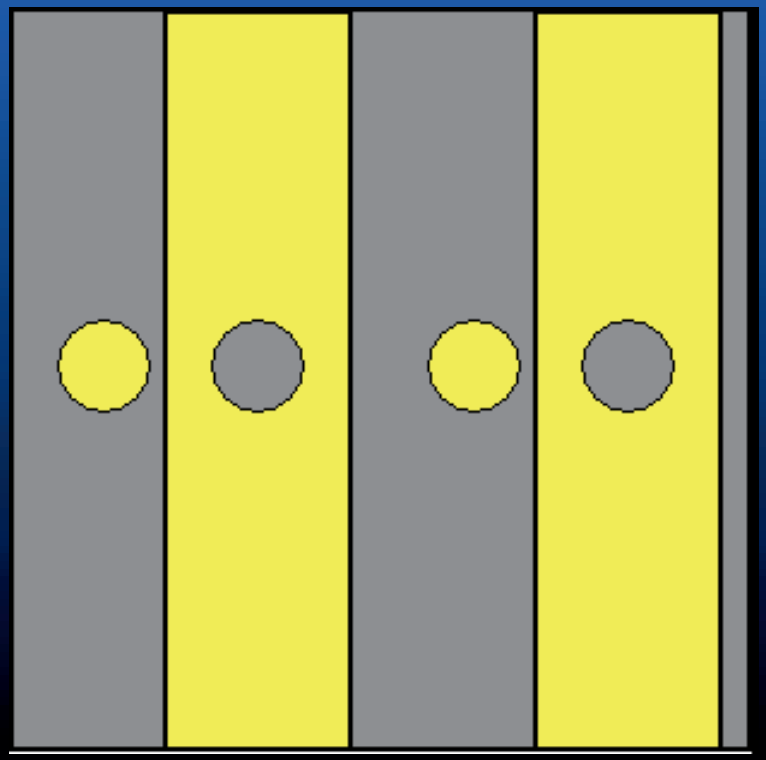
# Wide-baseline stereo

- What about *untextured regions*?



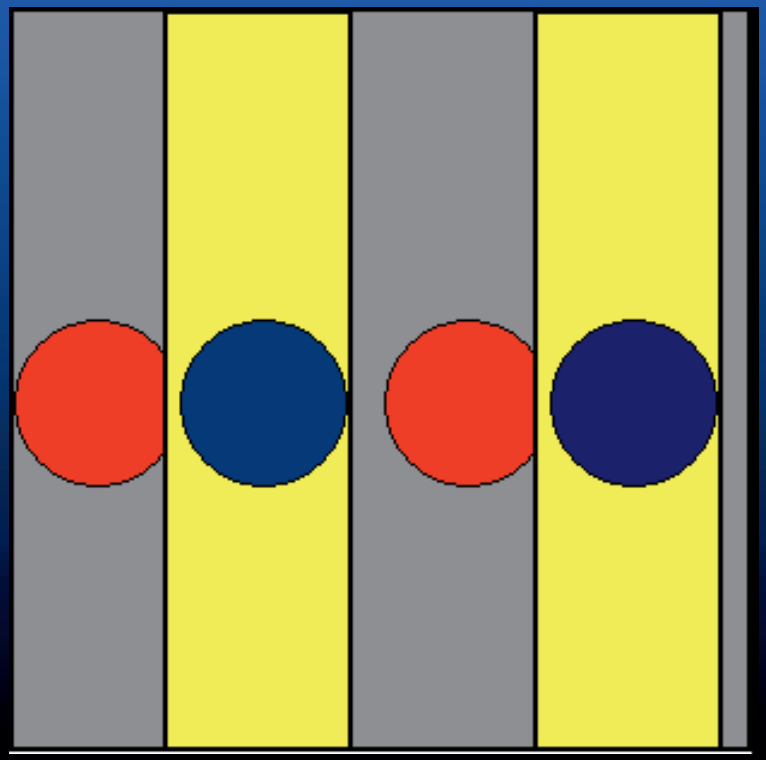
# Wide-baseline stereo

- What about *untextured regions*?



# Wide-baseline stereo

- What about *untextured regions*?



# Wide-baseline stereo

- What is it being *used* for?
  - view interpolation
  - view extrapolation
  - object removal / insertion
  - video compression
- Desired solution depends on application

# Role of image features

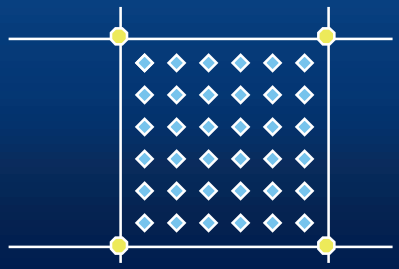
- Needed to establish original epipolar geometry [but see Hannah's direct methods]
- Once epipolar geometry is known, can use linear features or direct methods
- Useful for long-range motion: efficiency and robustness
- Features may vary in appearance [nice recent work by Schmid and Lowe]

# Role of image features

- Not statistically optimal:
  1. Weighting by feature certainty (doable)
  2. Not using all of the pixels
- Patch-based alignment [Shum & Szeliski]

$$\mathbf{A} \approx \sum_j \mathbf{J}_j \mathbf{A}_j \mathbf{J}_j^T \quad \text{with} \quad \mathbf{A}_j = \sum_{i \in \mathcal{P}_j} \mathbf{g}_i \mathbf{g}_i^T$$

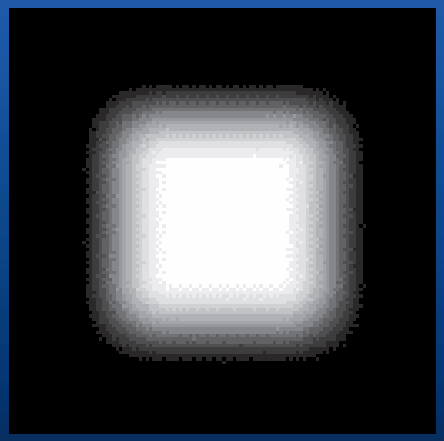
$$\mathbf{b} \approx \sum_j \mathbf{J}_j \mathbf{b}_j \quad \text{with} \quad \mathbf{b}_j = \sum_{i \in \mathcal{P}_j} e_i \mathbf{g}_i.$$



- Spline-based registration [Szeliski & Coughlan]

# Role of image features

- Can your feature tracker track this?



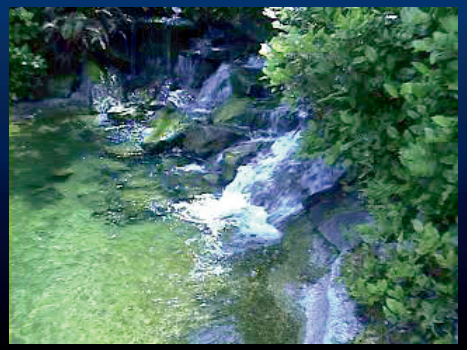
- Sometime direct methods track the *only* data in the sequence

# Next most important problems

1. Sub-pixel accurate registration

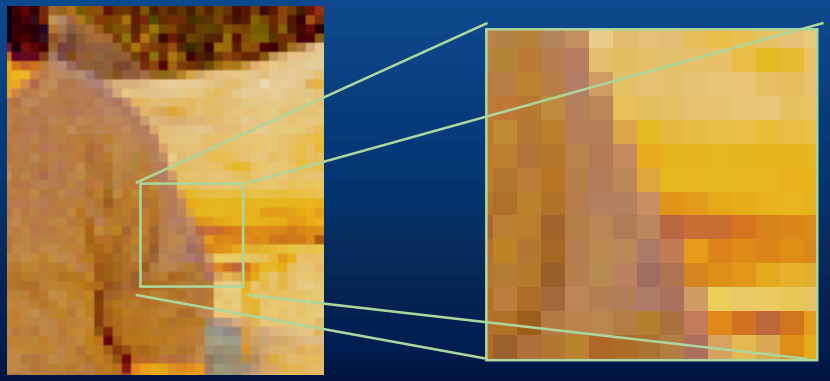
2. Transparency, reflections and specularities

3. Non-rigid motion



# Sub-pixel accurate registration

- Problems at and near occlusions
- Incorrect color extraction, no partial occupancy in (*mixed*) border pixels



# Layered Stereo

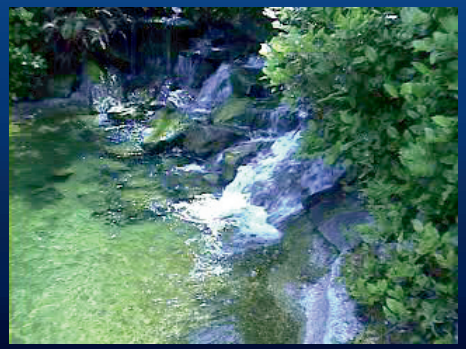
- Layers with *alpha* (opacity)

# Results: Michael and Lee

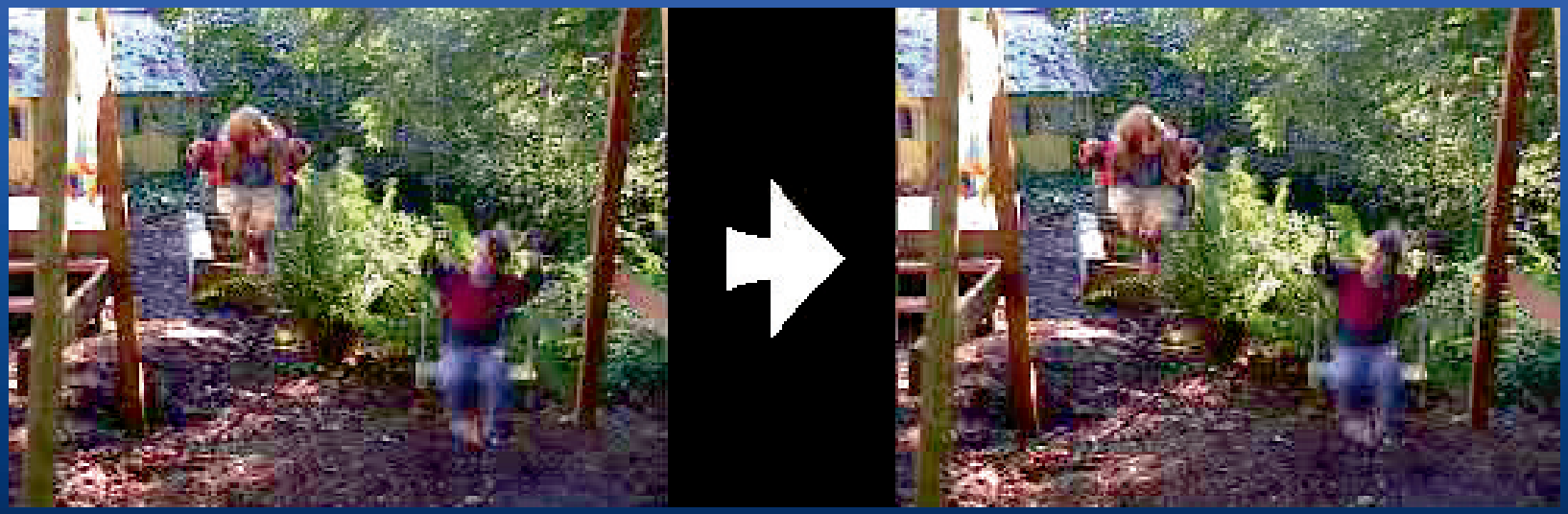
# Results: Anne and books

# Non-rigid motion

- Multiple moving objects (segmentation)
- Articulated and “soft” motion
- Video textures (quasi-random or quasi-periodic)



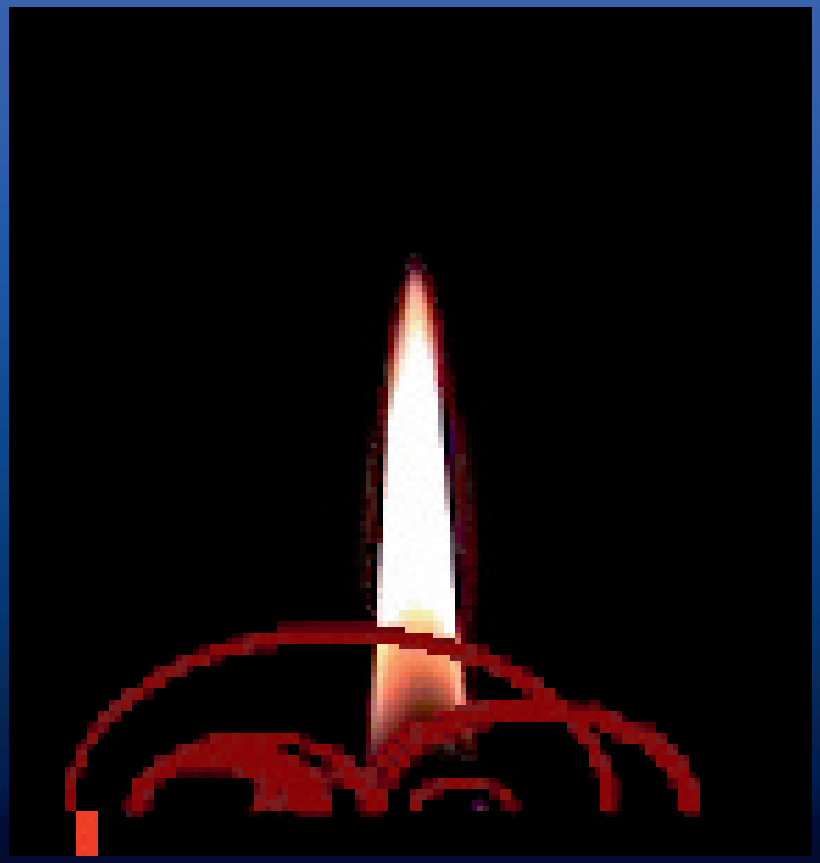
# VideoTextures



video clip

video texture

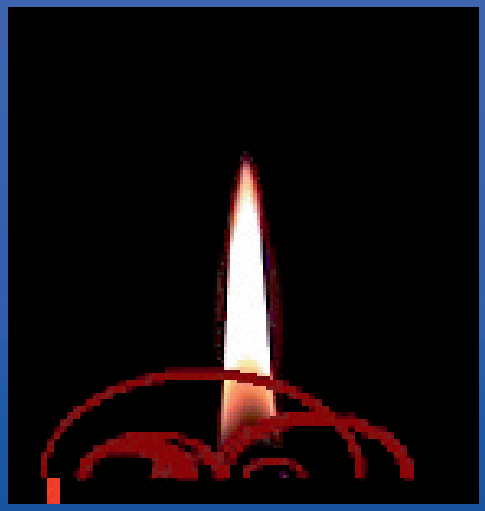
# Video Textures



How do we find good transitions?

# Complete animation

# Summary



- Video clips → Video Textures
  - discover Markov structure
  - preserve dynamics
  - disguise visual discontinuities
  - separate regions
  - user input
  - create video-based animations
- Example of *Video-Based Rendering*

# Next most important problems

1. Sub-pixel accurate registration

2. Transparency, reflections and specularities

3. Non-rigid motion