THE GOAL

• Add 3D objects to any video
  • Effortlessly
  • User-friendly
  • Accurate
  • Naturally
MOTIVATION

• Professional tools are powerful, but complicated

• Amateur tools are too limited
  • Simple effects
  • Sequencing

• Tedious work is usually required
  • Repeated manual interaction with many frames
RELATED WORK

- Mosaic-based video editing [Rav-Acha et al., 2008]
- 2.5D video editing using a mesh [Chen et al., 2011]
- 3D modeling from video [van den Hengel et al., 2007]
PROPOSED METHOD: OVERVIEW
VIDEO REPRESENTATION

- Use mosaic from [Rav-Acha et al., 2008]
  - Models a video as a single image
  - Uses the “most frontal” view of each part
  - Produces warping transforms
  - Handles occlusions

- Segment the object (semi-automatically) if needed

\[ E = E_{image} + \lambda_w E_{smooth} + \lambda_b E_{occlusion} \]
“IMAGE” OF NORMALS FORMATION

• Cloud of 3D surface normals as a world
• Camera projects 3D normals to frames
  • 3D location $\rightarrow$ 2D pixel
  • 3D normal vectors $\rightarrow$ “color”
• Projection and occlusions identical to image
• The “world” is identical for all the cameras
MOSAIC OF NORMALS

- Cost function as in the mosaic representation

\[ E = E_{data}^{image} + \lambda_w E_{smooth} + \lambda_b E_{occlusion} \]

- No new smoothness terms
- Similar optimization process
MATCHMOVING

- Adding to mosaic
  - By click – drag – rotate

- By brushing

- Camera motion is calculated automatically
COMPOSITING

• Given an input frame...
• The inlay is transformed into frame’s coordinate system...
• And added to the input frame.
• Blur is applied to make it look consistent with the video
EXPERIMENTS: DATA

- Free videos
- Moving object or camera
- Rigid scenes
- Non-rigid objects
Video Inlays: A System for User-Friendly Matchmove
CONCLUSIONS

• Video representation using mosaic and normals
  • Interact with image
  • Automatic object orientation

• As system with
  • User-friendly matchmoving
  • Basic compositing
Thank you for your attention