Shadow Matting and Compositing

Presented by Yunjun Zhang

How to write a paper for SIGGRAPH

- The applications are plentiful and the result is amazing.
- Idea could be simple but it really works and it is robust.
- The method is explained in good detail without too much fancy formulas.
Notations used in paper and this presentation

\[ f : \text{frames in a video sequence} \]
\[ C : \text{observed color image} \]
\[ L : \text{lit image } L = \max_f C_f \]
\[ S : \text{shadow image } S = \min_f C_f \]
\[ L', S', C' : \text{lit image, shadow image, and calculated color image for the target.} \]
\[ \beta : \text{it is the visibility of the light source} \]

Matting and Compositing
Contribution of this paper

- Composite the shadow without knowing the geometry and camera calibration.
- No blue screen is required for shadow matting.

Implementation

- Assumption
  - Single light source
  - Shadow should be cast on a flat ground in source video.

- Input and Output
  - Source Video (must be video to generate S and L)
  - Target Image
Implementation

- Estimating the shadow matte.
- Generate the deformed shadow matte from target image.
- Shadow compositing.

Estimating shadow matting

- Calculate shadow and lit images
- Calculate shadow matte for the selected image
- No pixel position can be covered by shadow all the time.

\[
\begin{align*}
\beta &= \frac{(C - S) \cdot (L - S)}{||L - S||^2} \\
L &\text{ lit image } L = \max_f C_f \\
S &\text{ shadow image } S = \min_f C_f
\end{align*}
\]
Estimating shadow matting
Estimating shadow deformations

- Basic idea:
  - Generate displacement for each pixel for the potential shadow area.

Target background Shadow Scan Video
Estimating shadow deformations

Algorithm

- For each directional scan s:
  - Find the first crossing time for each pixel by temporal analysis.
  - In each frame f, fit a shadow line
- For each pixel location p:
  - For each scan s, using the two nearest line equation’s parameters to generate the line equation for p
  - Compute the intersection point q of lines of all scans s
  - q-p is the displacement on point p.

Estimating shadow deformations

- Target background must has some region with planar surface matched to source background. This region is called reference plane.
Temporal Analysis

- For each Pixel position in each scan

$\text{(d) Temporal analysis}$

Fitting a line for each Frame in each Scan
Interpolating the lines for each pixel location

Get the displacement map for each Pixel position
Result

[Images of people and different textures]

Result
Discussion

- How to eliminate the planar background requirement?
- In video compositing, more interpolating methods are required.