Super-Resolution Keyframe Fusion for 3D Modeling with High-Quality Textures

Robert Maier, Jörg Stückler, Daniel Cremers

Computer Vision Group

Technische Universität München

Motivation

Given:

Low-res. (LR) RGB-D frames (640 x 480)

Accurate geometric reconstruction

Goal:

High-quality texture from low-cost RGB-D sensors

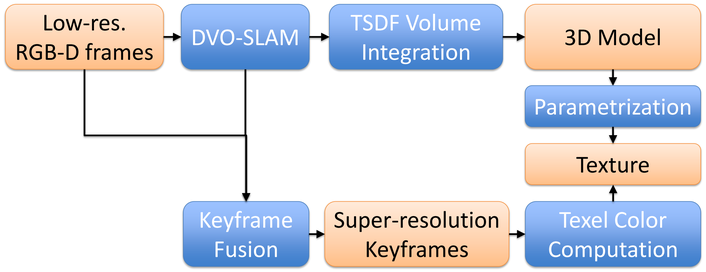
State-of-the-art:
- Vertex colors: limited resolution
- Texture mapping: good quality, slow

➔ Gap in prior work

Contributions

Our approach:
- Super-resolution (SR) keyframe fusion and deblurring
- Texture mapping using SR keyframes (weighted median)

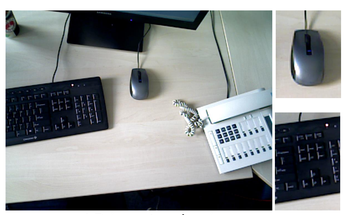
System Overview



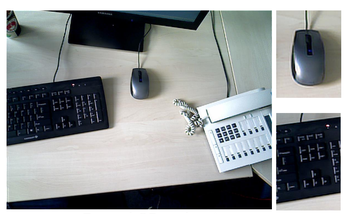
Keyframe Fusion

Fuse LR input RGB-D frames into high-res. RGB-D keyframes

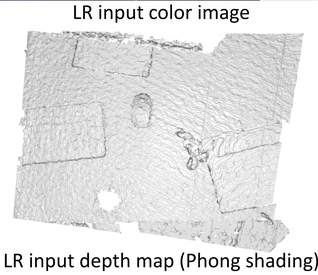
- **Depth fusion:**
  - Warp LR depth maps into keyframe (using relative poses)
  - Upsample and fuse depth using weighted averaging
- **Color fusion:**
  - Deconvolution: Wiener Filter on LR input images
  - Warp keyframe depth to input images for color lookup
  - Fuse colors using weighted median



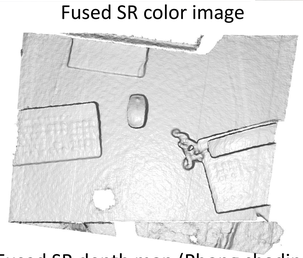
LR input color image



Fused SR color image



LR input depth map (Phong shading)



Fused SR depth map (Phong shading)

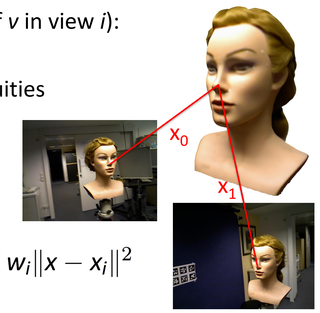
Vertex Color Computation

- Mesh vertex v, input views C_i (blur measure b_i), camera poses ξ_i
- Compute observations (x_i is obs. of v in view i):

$$x_i = \mathcal{C}_i(\pi(v, \xi_i))$$

- Discard x_i close to depth discontinuities
- Weights of obs. x_i:

$$w_i = \frac{\cos(\theta) * b_i}{d^2}$$



- Compute vertex color x:
  - Weighted mean: $\arg\min_x \sum_{x_i} w_i \|x - x_i\|^2$
  - Weighted median: $\arg\min_x \sum_{x_i} w_i \|x - x_i\|$

Unweighted Mean

Weighted Mean

Weighted Median

High-Quality Texture Mapping using SR Keyframes

- **Texture Parametrization:**
  - One-to-one mapping between 3D mesh and 2D texture
  - Least Squares Conformal Maps (Levy et al., ACM ToG 2002)
- **Texel color computation:**
  - Compute **3D vertex for 2D texel** (based on enclosing triangle using barycentric coordinates)
  - Compute **color from SR keyframes** analogous to per-vertex recoloring scheme (weighted median)

Runtime Evaluation

Datasets:

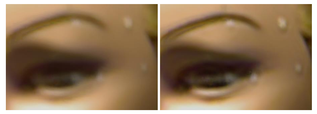
	face	*phone*	*keyboard*
# RGB-D frames	512	1359	642
# vertices (original)	159583	82942	155842
# triangles (original)	319176	165888	311686
# triangles (decimated)	40000	40000	40000

Runtimes:

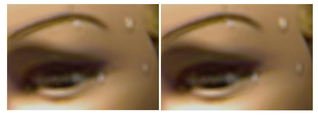
		face		*phone*		*keyboard*	
	s	t [s]	fps	t [s]	fps	t [s]	fps
Texture Mapping		91.5	5.6	330.8	4.1	128.8	5.0
Keyframe Fusion	2	57.5	8.9	222.0	6.1	72.1	8.9
SR Texture Mapping	2	18.7	2.8	50.7	2.7	18.8	3.5
Keyframe Fusion	4	100.9	5.1	362.8	2.2	214.9	3.0
SR Texture Mapping	4	26.4	2.0	58.2	1.4	42.6	1.5

(Standard desktop PC with Intel Core i7-2600 CPU with 3.40GHz and 8GB RAM)

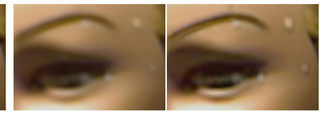
Qualitative Results



Without vs. with deconvolution

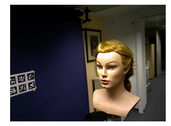


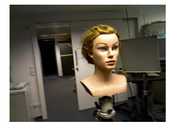
Keyframe dimensions 1280 x 960 vs. 2560 x 1920



With LR input frames vs. with SR keyframes

Face





RGB input images

Vertex colors

Our approach

Phone

RGB input images

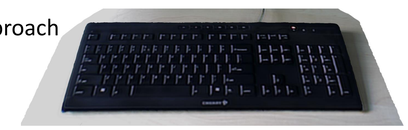
Vertex colors

Our approach

Keyboard

Vertex colors

Our approach



RGB input images

Conclusion

- **Robust** and **efficient** method for high-quality texture mapping in RGB-D-based 3D reconstruction
  - Fuse low-quality color images into **SR keyframes**
  - **Map** high-quality keyframes onto 3D model **texture** using **weighted median** scheme
- Experimental results:
  - **Increased photo-realism** of reconstructed 3D models
  - Very **efficient** and **practical** post-processing step (runtimes within a few minutes)