ICVSS 2014
Sicily ~ 13-19 July
International Computer Vision Summer School

Exploiting Shading Cues in Kinect IR images for Geometry Refinement

Published in CVPR 2014

KAIST RCV

Phone: +82-42-350-5465
Email: gmchoe@rcv.kaist.ac.kr
Homepage: http://rcv.kaist.ac.kr/~gmchoe/

Gyeongmin Choe, Jaesik Park, Yu-Wing Tai and In So Kweon
Korea Advanced Institute of Science and Technology (KAIST), Republic of Korea

Motivation

Using IR Instead of color

Geometry from Low cost depth sensor

Images for geometry refinement

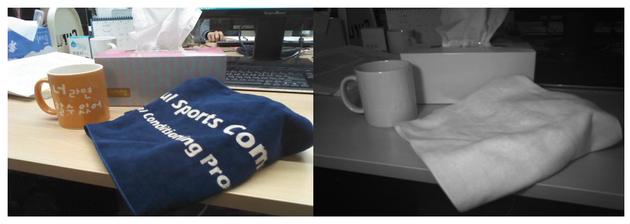
High-quality 3D geometry

Advantages of IR Image

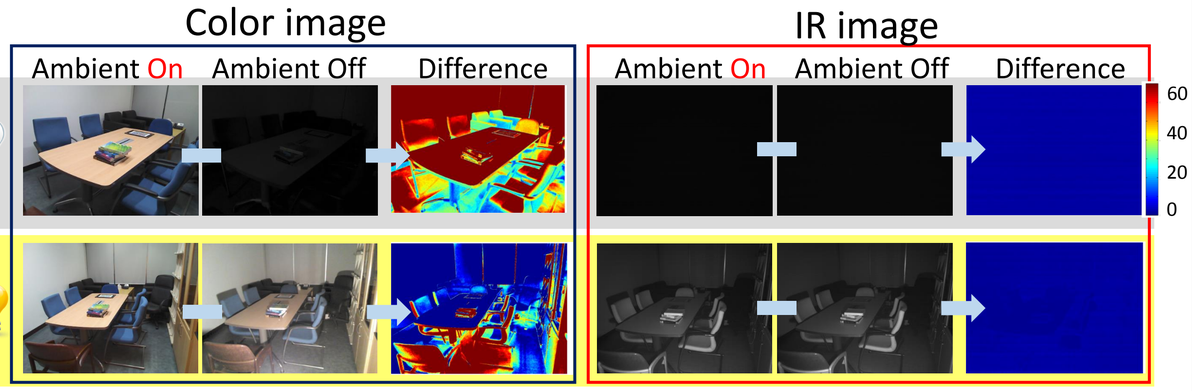
① Simple Setup

No need for color camera

② Less Albedo Variance

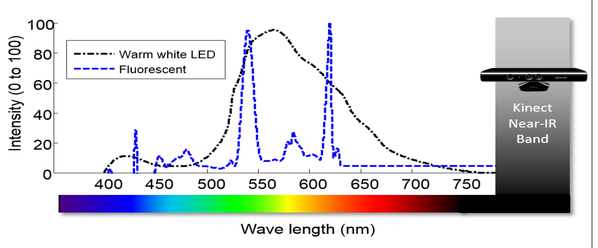


③ IR Invariability under Indoor Lightings



- Kinect IR camera filters out most undesired ambient light that makes our system robust to indoor illumination.
- Our method works well in both dark room and natural lighting environment.

Kinect Near-IR Band



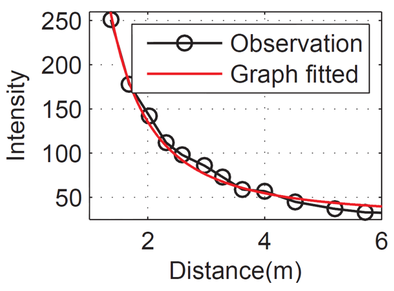
Our IR Light Model

$$I(u) = \frac{c\rho}{d^2}(n \cdot l) + I_{Ambient}$$

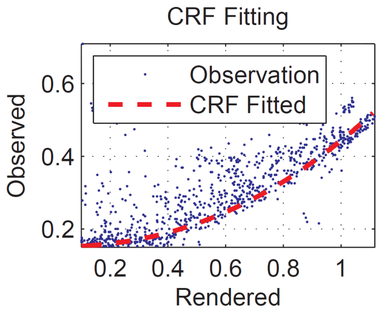
The beauty of IR

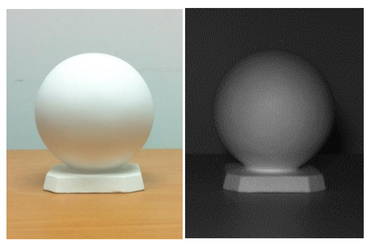
$I(u)$: pixel intensity,
ρ : albedo of surface, n : surface normal,
l : lighting direction, c is the global brightness
d : the distance between surface point and center of light source

Inverse Square Law



CRF Fitting





Color IR

- To radiometrically calibrate the IR shading images, we use intensity observations of a known geometry and determine the gamma function.

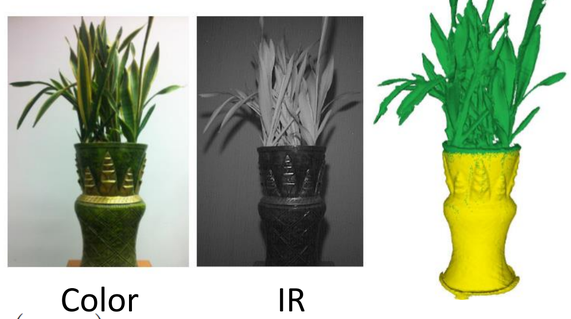
Multiple Albedo handling

Single albedo

$$c\rho = \frac{1}{Z}\sum_{i=1}^{N}\sum_{\substack{m=1,\\ \mathbf{u}_{i,m}\in V_i}}^{M}\frac{d_{i,m}^2}{\mathbf{n}_{i,m}\cdot\mathbf{l}_{i,m}}I_m(\mathbf{u}_{i,m})$$

Multiple albedo

$$c\rho_i = \frac{1}{Z_i}\sum_{\mathbf{x}_j\in N(\mathbf{x}_i)}\sum_{\substack{m=1,\\ \mathbf{u}_{i,m}\in V_i}}^{M}\frac{d_{j,m}^2}{\mathbf{n}_{j,m}\cdot\mathbf{l}_{j,m}}I_m(\mathbf{u}_{j,m})$$

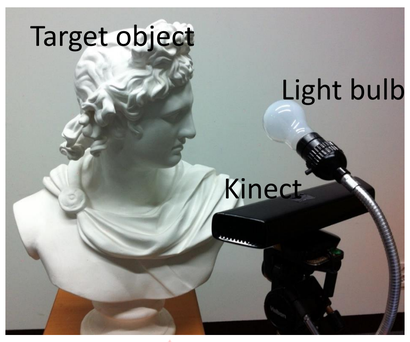


Color IR

Grouped IR albedo

K-means clustering for albedo grouping. 4-dimensional feature $[x\ y\ z\ c\rho]$ is used.

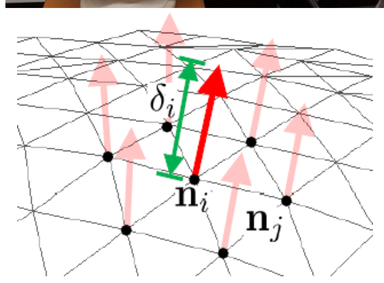
Capturing Setup & Mesh Optimization



input shading images

Projected mesh vertices

Depth map derived from a Kinect fusion mesh



Update $\boldsymbol{x_i}$ along with the normal direction $\boldsymbol{n_i}$

$$\mathbf{x}_i^t = \mathbf{x}_i^{t-1} + \delta_{i,t}\mathbf{n}_i^{t-1}$$

Minimizing the energy function below.

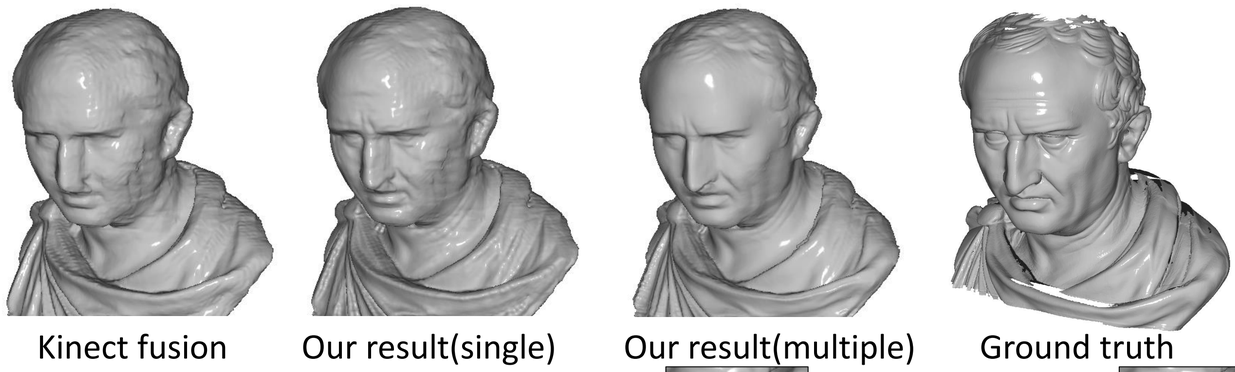
$$\arg\min_{\delta} E_p(\boldsymbol{\delta}) + E_s(\boldsymbol{\delta}),$$

$$E_p(\boldsymbol{\delta}) = \sum_{i=1}^{p}\sum_{k\in V_i} w_{i,k}\left(i_{i,k} - c\rho\frac{\mathbf{n}_{i,k}(\delta_{i,k})\cdot\mathbf{l}_{i,k}}{d_{i,k}^2}\right)^2$$

$$E_s(\boldsymbol{\delta}) = \sum_{i=1}^{p}\sum_{j\in N_i}\lambda_1(\delta_i-\delta_j)^2 + \sum_{i=1}^{p}\lambda_2(\delta_i)^2,$$

i is shading intensity, δ is displacement of vertices, $w_{i,k} = \boldsymbol{n_{i,k}}\cdot\boldsymbol{l_{i,k}}$
λ_1 and λ_1 are the weights for local smoothness and regularizer term respectively.

Experimental Results



Kinect fusion | Our result(single) | Our result(multiple) | Ground truth