IR Surface Reflectance Estimation and Material Type Recognition using Two-stream Net and Kinect Camera

SeokYeong Lee
Department of Computer Science and Engineering
Kyunghee University
shapin94@khu.ac.kr

HwaSup Lim
Center of Imaging Media Research
KIST
hslim@kist.re.kr

SangChul Ahn
Center of Imaging Media Research
KIST
asc@kist.re.kr

SeungKyu Lee
Department of Computer Science and Engineering
Kyunghee University
seungkyu@khu.ac.kr

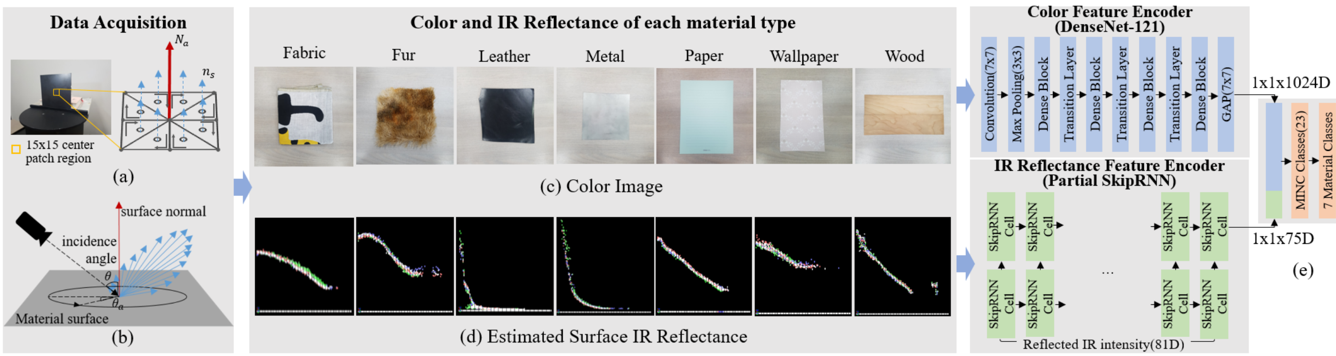

Figure 1: Proposed two-stream material type recognition framework. IR responses and color images are obtained using single Kinect camera in (a) and (b). IR Surface reflectance is estimated from the IR responses in (c) and (d). Color and reflectance features are individually trained and fused at the end of the network in (e).

ABSTRACT

Recently, material type recognition using color or light field camera has been studied. However, visual pattern based approaches for material type recognition without direct acquisition of surface reflectance show limited performance. In this work, we propose IR surface reflectance estimation using *off-the-shelf* ToF (Time-of-Flight) active sensor such as Kinect and perform surface material type recognition based on both color and reflectance clues. Two stream deep neural network consists of convolutional neural network encoding visual clue and recurrent neural network encoding reflectance characteristic is proposed for material classification. Estimated IR surface reflectance and material type recognition evaluation on our Color-IR Material Data set show promising performance compared to prior approaches.

CCS CONCEPTS

• **Computing methodologies** → **Supervised learning by classification**;

KEYWORDS

Material recognition, Surface reflectance

SIGGRAPH '19 Posters, July 28 - August 01, 2019, Los Angeles, CA, USA

ACM ISBN 978-1-4503-6314-3/19/07.
https://doi.org/10.1145/3306214.3338557

ACM Reference format:
SeokYeong Lee, HwaSup Lim, SangChul Ahn, and SeungKyu Lee. 2019. IR Surface Reflectance Estimation and Material Type Recognition using Two-stream Net and Kinect Camera. In *Proceedings of SIGGRAPH '19 Posters, Los Angeles, CA, USA, July 28 - August 01, 2019,* 2 pages.
https://doi.org/10.1145/3306214.3338557

1 INTRODUCTION

Material type recognition based on visual patterns have been studied in computer vision and graphics fields. Deep TEN[Zhang et al. 2017] shows state of the art material type recognition performance based on conventional color feature descriptors with deep neural network. However, some material types are not able to be well described only by visual appearance (ex. colored plastic and wooden objects). Angular imaging using light field camera [Wang et al. 2016] and mobile robot device collect angular visual variations, however, the devices are expensive and show limitation in practical applications. Kim et al. [Kim et al. 2018] propose material recognition method using surface roughness estimation and multi-directional scanning of target material surface using an active sensor. They have shown in laboratory settings that surface reflectance has potential in characterizing material types. In this paper, we propose surface reflectance estimation using *off-the-shelf* Kinect camera in practical environment for material type classification. Infrared modulation light of time-of-flight (TOF) camera observes reflected IR, that follows reflectional distribution of the surface with respect to emitted Infrared. A material type recognition method based on both color and surface reflectance is proposed. To this end, two stream deep neural network is constructed (a convolutional neural network

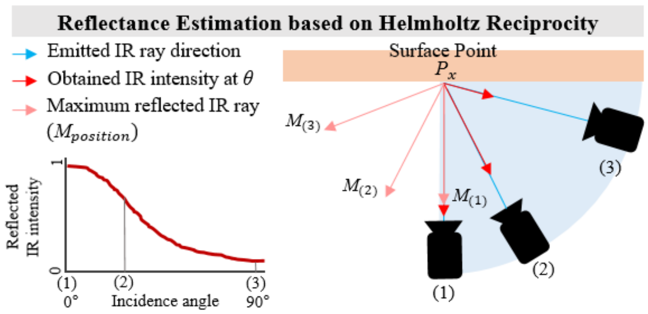

Figure 2: IR reflectance acquisition using Kinect. Based on Helmholtz reciprocity, reflected IR observations at moving Kinect on the right (with light source) corresponds to IR observations based on moving light source on the left side.

for color and a recurrent neural network for IR reflectance). For our classification, invisible Infrared surface reflectance is estimated and no photo-realistic rendering is involved.

2 PROPOSED METHOD

Data Acquisition: To minimize the cost of material classification, we have three assumptions avoiding exhaustive scanning of the prior method [Kim et al. 2018]. First, we assume that observed materials have optically isotropic surfaces. Although proposed method has no limitation in dealing with an-isotropic surfaces, this assumption allows us to reduce scan range dramatically. Secondly, following Helmholtz reciprocity, the observation at moving camera on one side (with attached light source varying incidence angle) coincides with the observation by light source at opposite side (see Figure 2). Lastly, we ignore Fresnel effect. Based on the assumptions, we observe reflected IR signals from one side of 0° to 90° incidence angles (θ). 15x15 patch of target surface and corresponding observations are averaged to obtain robust and clean data. Estimated reflectance is treated as 1-dimensional sequence of IR intensities. As the estimated IR reflectance shown in Figure 1. (d), observations near 90° are discarded because they are very noisy due to weak reflect IR signal from inclined surface.

Material Recognition: Our network incorporates two types of deep neural networks for two different modality of input data. DenseNet-121[Huang et al. 2016] is adapted for color feature encoding and SkipRNN [Campos et al. 2017] is modified and incorporated for IR reflectance feature extraction. First, pre-trained DenseNet is fine-tuned separately for material classification (using MINC-2500[Bell et al. 2015] dataset). Material type recognition on MINC-2500 test set shows competitive performance over state of the art material recognition method [Zhang et al. 2017] and converges faster than it (Table 1). And then, our two stream net (DenseNet + partial SkipRNN) is constructed and trained with our Color-IR Material data set. *Update gate*(u_t) of original SkipRNN[Campos et al. 2017] determining the update-state of LSTM cell is replaced by our new *Partial input gate*(p_t) that discards invalid input value of estimated IR reflectance as follows.

$$s_t = p_t \cdot S(s_{t-1}, x_t) + (1 - p_t) \cdot s_{t-1} \quad (1)$$

Table 1: Color only material type recognition accuracy on MINC-2500[Bell et al. 2015] data set.

Model	Data	Accuracy	Epoch
AlexNet (Bell et al.)	MINC-2500	76.0 ± 0.2%	N/A
Deep TEN * (*SOTA*)(Zhang et al.)	MINC-2500	*multi : **81.3%** *single : 80.6%	35
DenseNet-121	MINC-2500	**81.13%**	35
DenseNet-169	MINC-2500	78.64%	>35

Table 2: Material type recognition accuracy on our Color-IR Material Data set (Mean and Std.)

Model	SkipRNN	Partial SkipRNN	Two-stream Net (SkipRNN)	Two-stream Net (Partial SkipRNN)
	$62.67_{\pm 5.5}$	$64.67_{\pm 1.8}$	$\mathbf{74.67}_{\pm 3.0}$	$\mathbf{76.00}_{\pm 4.9}$

where s_t is state value at time t resulted from state transition model S defined in SkipRNN[Campos et al. 2017] and $p_t = u_t$ if $x_t \geq 0.05$, $p_t = 0$ otherwise.

3 EXPERIMENTAL RESULTS

Total 116 material samples (color + IR reflectance) of 7 material types are collected for quantitative evaluation. Material types such as fur and metal show unique distribution (figure 1. (d)). On the other hand, specular material such as metal and leather show similar distribution despite clear separation in their colors. Extracted IR reflectance characterizes surface type (such as its roughness) and color pattern adds additional information on the visual characteristic of each material type. For robust and material specific feature extraction, proposed reflectance features are trained for 1000 epochs using our partial SkipRNN. DenseNet for color is also fine-tuned using MINC-2500[Bell et al. 2015] for 35 epochs. Extracted color and IR reflectance features are concatenated at the end of our two stream nets and trained for 1000 epochs. Material recognition result of our method is summarized in Table 2. Two-stream network performs better than color-based network.

ACKNOWLEDGMENTS

This work was supported by the Global Frontier RD Program on Human-centered Interaction for Coexistence funded by the National Research Foundation of Korea grant funded by the Korean Government (MSIT) (2010-0029752)

REFERENCES

Sean Bell, Paul Upchurch, Noah Snavely, and Kavita Bala. 2015. Material Recognition in the Wild With the Materials in Context Database. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015).

Victor Campos, Brendan Jou, Xavier Giró i Nieto, Jordi Torres, and Shih-Fu Chang. 2017. Skip RNN: Learning to Skip State Updates in Recurrent Neural Networks. *CoRR* abs/1708.06834 (2017). arXiv:1708.06834 http://arxiv.org/abs/1708.06834

Gao Huang, Zhuang Liu, and Kilian Q. Weinberger. 2016. Densely Connected Convolutional Networks. *CoRR* abs/1608.06993 (2016). arXiv:1608.06993

J. Kim, H. Lim, S. C. Ahn, and S. Lee. 2018. RGBD Camera Based Material Recognition via Surface Roughness Estimation. (March 2018), 1963–1971. https://doi.org/10.1109/WACV.2018.00217

Ting-Chun Wang, Jun-Yan Zhu, Hiroaki Ebi, Manmohan Chandraker, Alexei A. Efros, and Ravi Ramamoorthi. 2016. A 4D Light-Field Dataset and CNN Architectures for Material Recognition. *CoRR* abs/1608.06985 (2016). arXiv:1608.06985

Hang Zhang, Jia Xue, and Kristin Dana. 2017. Deep TEN: Texture Encoding Network. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.