# Robust near-infrared structured light scanning for 3D human model reconstruction

Bo Fu and Ruigang Yang

University of Kentucky, 329 Rose Street, Lexington, USA

## ABSTRACT

In this paper we present a novel sensing system, robust Near-infrared Structured Light Scanning (NIRSL) for three-dimensional human model scanning application. Human model scanning due to its nature of various hair and dress appearance and body motion has long been a challenging task. Previous structured light scanning methods typically emitted visible coded light patterns onto static and opaque objects to establish correspondence between a projector and a camera for triangulation. In the success of these methods rely on scanning objects with proper reflective surface for visible light, such as plaster, light colored cloth. Whereas for human model scanning application, conventional methods suffer from low signal to noise ratio caused by low contrast of visible light over the human body. The proposed robust NIRSL, as implemented with the near infrared light, is capable of recovering those dark surfaces, such as hair, dark jeans and black shoes under visible illumination. Moreover, successful structured light scan relies on the assumption that the subject is static during scanning. Due to the nature of body motion, it is very time sensitive to keep this assumption in the case of human model scan. The proposed sensing system, by utilizing the new near-infrared capable high speed LightCrafter DLP projector, is robust to motion, provides accurate and high resolution three-dimensional point cloud, making our system more efficient and robust for human model reconstruction. Experimental results demonstrate that our system is effective and efficient to scan real human models with various dark hair, jeans and shoes, robust to human body motion and produces accurate and high resolution 3D point cloud.

**Keywords:** Near-infrared, Structured light, human body scan, 3D reconstruction

## 1. INTRODUCTION

Human modeling has been an active research area in both computer vision and computer graphics for a long time because of its various applications. In order to acquire the raw 3D measurement, several recently available low cost depth sensor have been developed. One of the most representative device is the Microsoft Kinect sensor. Many approaches have been proposed to generate 3D models based on these sensors. However, due to the relatively low-quality depth maps, more efforts have to be made for removing noise and outliers lays in raw depth maps. An alternative approach is derived from structured light technique which is capable of acquiring relatively high quality depth map of human because camera and projector can provide better resolution. The drawback of this method is caused by relatively low contrast over the dark surface of human, such as black shoes or dark clothes. Thus, this method can not correctly distinguish patterns. Moreover, both the Kinect and Structured Light requires the human to be static while scanning is in progress. This assumption is not user friendly and errors may be introduced due to human subject's unintentional motion, especially when the whole scan procedure takes a while.

In this paper, we present a near-infrared structured light scanning system that is aiming to improve the 3D model quality for human subject with dynamic motion. Our main idea is to combine the near-infrared light source, structured light scanning technique, and the high speed visual information capturing together. Structured light scanning, with high resolution camera, can yield high quality 3D model. The reflection of visible light over the dark area of the human body, such as black shoes and coat is unreliable and thus results in errors and artifacts of the 3D model. The near-infrared light, however, can provide enough contrast over these dark areas

Further author information: (Send correspondence to Bo Fu)
Bo Fu: E-mail: fu.bo@uky.edu, Telephone: 1 859 333 0722
Ruigang Yang: E-mail: ryang@cs.uky.edu, Telephone: 1 859 257 1257 ext 81282

due to its nature. Another advantage of using near-infrared light is its non intrusive nature, compare to standard structured light scanning which projects bright light and may be directly viewed by human subject. In order to minimize the error introduced by dynamic motion of human body, the time required for scanning should be as short as possible. Thus, high speed camera and projectors are employed in our system;

## 2. RELATED WORK

We review the related recent work in 3D human modeling and scanning techniques.

Human modeling has attracted research interests for years. Weiss proposed the home 3D body scan[1] applies the SCAPE (Shape Completion and Animation for PEople)[2] approach to Kinect point cloud data and utilizes the silhouette constraints to make the fitting more robust to side views. This method reconstructs human shape from noisy monocular image and range data using a single inexpensive commodity sensor. The approach combines low-resolution image silhouettes with coarse range data to estimate a parametric model of the body. The full body multiple Kinect scanning system[3] captures a dense sequence of partial meshes while the subject standing still on a turntable. All the partial scans are registered together based on the error distribution approach.[4] 3D Self-Protraits[5] presents the first autonomous capture system for self-portraits modeling using a single Kinect. The user stands as still as possible during capture and rotate themselves with the same pose for a few scans from different views, their system stitches together the captured scans using multi-view non-rigid registration, and produces watertight final models. All above methods are based on raw data acquired by Kinect sensor which is inexpensive and well developed. The drawback of Kinect is its relatively low-quality depth maps which can not provide detail geometry of human body. Multiple overlapping depth maps have to be fused together to not only provide more coverage, but also reduce the noise and outliers in the raw depth maps. Therefore the data acquisition is limited to static objects(e.g., KinectFusion system[6]), or human in mostly static poses.[1,5]

Structured light scan, by its relative high-quality output is considered one of the most reliable techniques to recover object surfaces and has been introduced to human modeling. Zhang[7] proposed a color structured light technique for recovering shape from one or more images. The resulting approach is suitable for generating both high speed scans of moving objects when projecting a single stripe pattern and high-resolution scans of static scenes using a short sequence of time-shifted stripe patterns. Garcia[8] proposed a A low-cost system composed of multiple cameras and projectors positioned around a central performance area capturing the dynamic geometry of human. Lanman[9] proposed a system for rapidly acquiring complete 3-D surface models using a single orthographic structured light projector, a pair of planar mirrors, and one or more synchronized cameras. Using the mirrors, the structured light patterns illuminate the object from all sides (not just the side of the projector) and are able to observe the object from several vantage points simultaneously. These method are concentrate on object(could be applied on human body scan) 3D geometry reconstruction with visible light source. Given that the dark clothing is inevitable in daily life, near-infrared light illumination is a nature alternative for structured light scan.

There are three different families of invisible structured light: Imperceptible Structured Light (ISL), Infrared Structured Light (IRSL) and Filtered Structured Light (FSL). ISL sensors are composed by a unique light source and two cameras.[10,11] It is dedicated to combine the advantage of structured light (accurate correspondence and easy reconstruction) with advantage of classical vision. IRSL sensor use an infrared laser beam to generate invisible patterns. This method is widely used for 3D scanning, robot navigation, industrial production[12] etc. FSL filters the light source such that only infrared structured light passes. The pattern can be projected through a laser source or a video-projector.[13] When compared these three families of methods, IRSL requires mechanical scanning and ISL suffers from occupancy issue, made both of them not suitable for human body scan. FSL, however, enjoys the advantage of the high resolution and easy and free design of coded pattern. Thus our proposed near-infrared structured light scanning system is based on FSL.

## 3. SYSTEM PLATFORM

### 3.1 hardware layout

Based on the standard FSL structured light scanner setup, our system is composed by a high resolution CCD camera and a high speed near-infrared projector (illustrated in Fig. 1, Fig. 2).

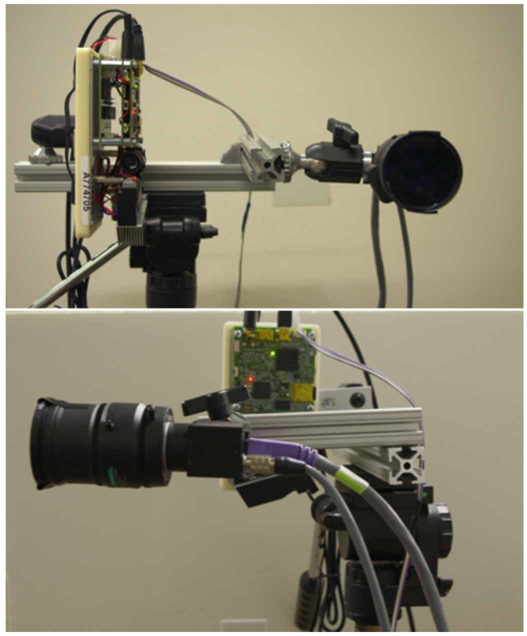

Figure 1. Near-infrared Structured Light Scanning system. Upper image: frontal view, lower image: side view. This system contains a set of high speed camera and a high speed near-infrared projector. Both of them are mounted in vertical direction to compensate width-heigth ratio of their field of view with human body's. The camera and projector are synchronize by GPIO in/out cable. Once the camera sends out trigger signal, camera itself starts capturing images; the projector, once received the income trigger signal, starts projecting preloaded patterns. The in/out time lag is neglectable.

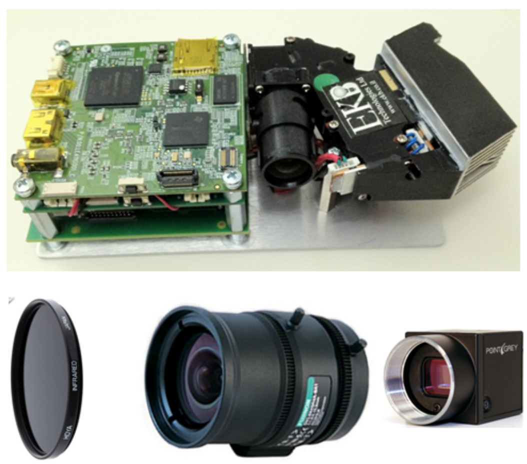

Figure 2. Components of our system. Upper image: E.K.B DLP LightCrafter MKII$^{TM}$ with native resolution of 608x684 and maximum binary pattern rate of 4kHz. Lower image: camera set. We are using Point Grey Flea3 CCD camera with native resolution of 1600x1200 and frame rate of 15Hz(note that this is only a demo image, another Flea3 camera with native resolution of 1280x1024 and maximum frame rate of 480Hz). Other components are a Fujinon DV3.8x4SR4A-1 zoom lens and a Hoya Infrared Glass Filter.

### 3.2 software introduction

We developed an fully functional structured light scanning system with GUI(Fig. 3). Our system is built upon standard structured light algorithm with subpixel accuracy. The average error is less than 1 millimeter. The projector is controlled by GUI provided by manufacture company.

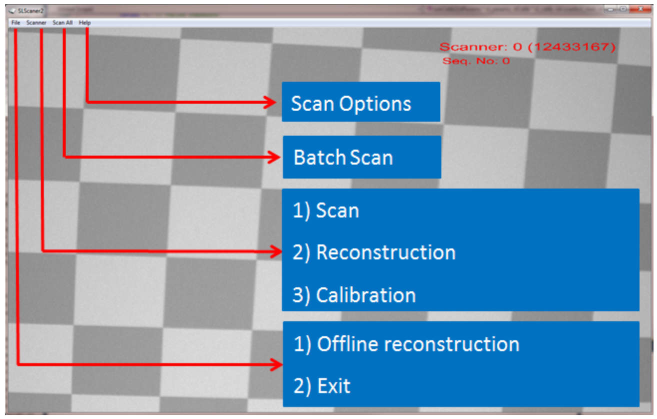


Figure 3. Near-infrared structured light scanning system GUI. This GUI is capable of control scan process, camera-projector calibration, and reconstruction for 3D geometry of object.

## 4. EXPERIMENT

In order to emphasizes the effectiveness of our scanning system, the experiment is conducted mainly on dark object of human surface, such as black shoe, dark hair, etc. Since we are concentrate on object, all of the experimental subject is static. Since our system is potentially to achieve a maximum scanning frame rate of 480Hz, dynamic motion of human body during the scanning is neglect able. We compare the performance of our near-infrared scanning system with conventional visible light scanning system and Kinect Fusion, which is widely used for human body reconstruction. Detail comparison is illustrated in Fig. 4. Based on the result, our system can generate more dense point cloud than Kinect Fusion method while keeps most of the detail 3D information than the conventional structured light scan.

## 5. CONCLUSION

In this paper we present a robust Near-infrared Structured Light (NIRSL) Scanning for three-dimensional human model scanning. The proposed robust NIRSL, as implemented with the near infrared light, is capable of recovering those dark surfaces under visible illumination. Successful structured light scan relies on the assumption that the subject is static during scanning. Due to non-intentional body motion, it is very time sensitive to keep this assumption. The proposed sensing system, by utilizing the near-infrared high speed projector, is robust to motion, provides accurate and high resolution three-dimensional point cloud, making our system more efficient and robust for human model reconstruction. Experimental results demonstrate that our system is effective and efficient than conventional structured light scan and in some cased better than the popular Kinect Fusion based scanning technique.

## ACKNOWLEDGMENTS

Thanks to my adviser Dr.Ruigang Yang and special thanks to E.K.B technologies ltd for their generous technique support on the LightCrafter projector.

## REFERENCES

[1] Weiss, A., Hirshberg, D., and Black, M. J., "Home 3d body scans from noisy image and range data," in [*Computer Vision (ICCV), 2011 IEEE International Conference on*], 1951–1958, IEEE (2011).

[2] Anguelov, D., Srinivasan, P., Koller, D., Thrun, S., Rodgers, J., and Davis, J., "Scape: shape completion and animation of people," in [*ACM Transactions on Graphics (TOG)*], **24**(3), 408–416, ACM (2005).

[3] Tong, J., Zhou, J., Liu, L., Pan, Z., and Yan, H., "Scanning 3d full human bodies using kinects," *Visualization and Computer Graphics, IEEE Transactions on* **18**(4), 643–650 (2012).

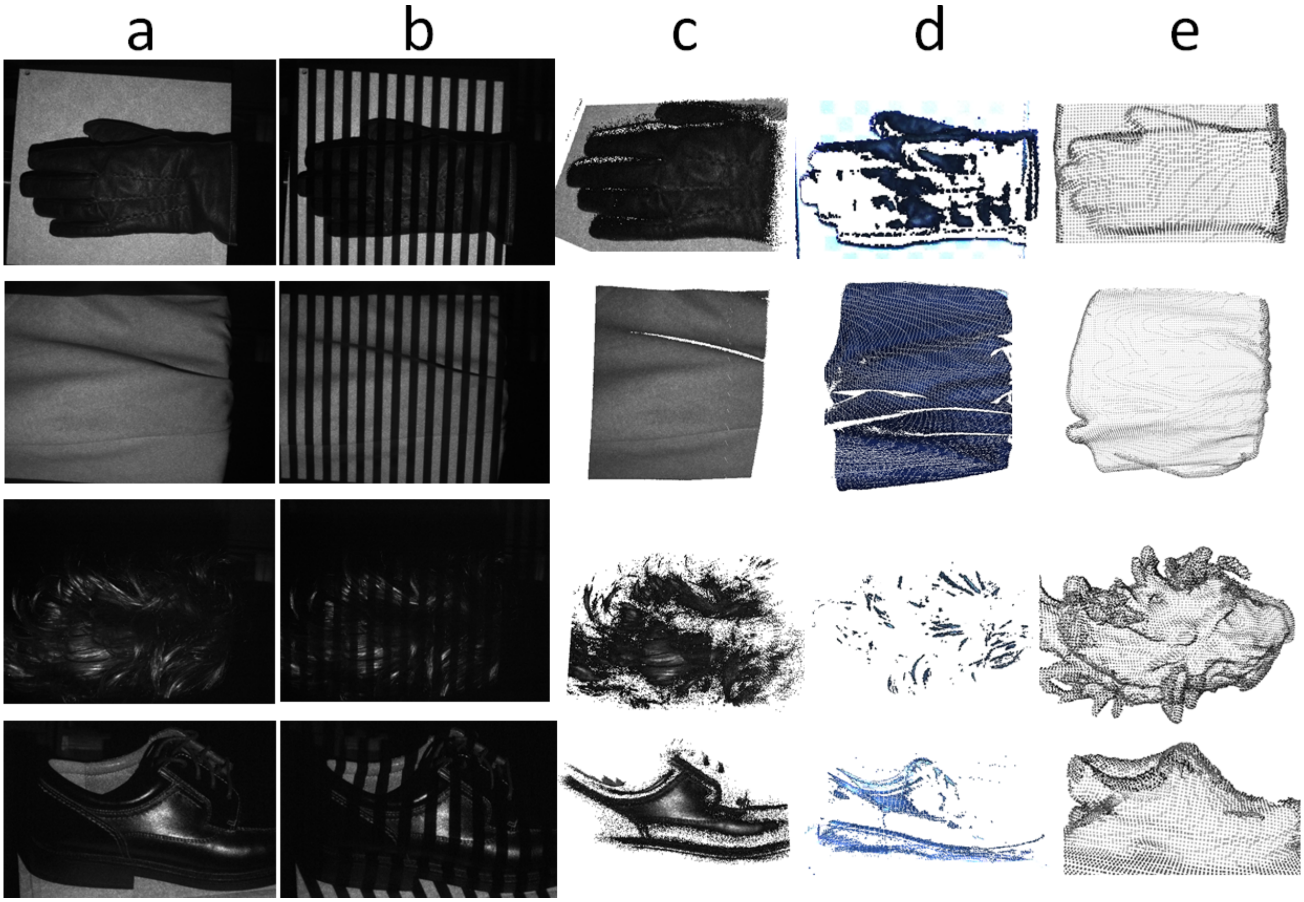


Figure 4. Experiment Result. Column a: raw texture image of the object; Column b: a sample image of the object with binary pattern projected on it during scanning with our system; Column c: reconstructed 3D piont cloud using our system; Column d: 3D point cloud using conventional structured light scanning system; Column e: mesh model generate by Kinect Fusion. In order to make these three method comparable, the object is placed at same distance away from the sensor. Row 1: black glove; Row 2: dark blue shorts (note that its color texture under near-infrared light appears more lighter; Row 3: dark or grey hair; Row 4: black shoes.

[4] Sharp, G. C., Lee, S. W., and Wehe, D. K., "Multiview registration of 3d scenes by minimizing error between coordinate frames," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **26**(8), 1037–1050 (2004).

[5] Li, H., Vouga, E., Gudym, A., Luo, L., Barron, J. T., and Gusev, G., "3d self-portraits," *ACM Transactions on Graphics (TOG)* **32**(6), 187 (2013).

[6] Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., et al., "Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera," in [*Proceedings of the 24th annual ACM symposium on User interface software and technology*], 559–568, ACM (2011).

[7] Zhang, L., Curless, B., and Seitz, S., "Rapid shape acquisition using color structured light and multi-pass dynamic programming," in [*3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on*], 24–36 (2002).

[8] Garcia, R. R. and Zakhor, A., "Multi-view dynamic geometry capture using structured light," in [*SPIE Newsroom DOI: 10.1117/2.1201301.004724*],

[9] Lanman, D., Crispell, D., and Taubin, G., "Surround structured lighting: 3-d scanning with orthographic illumination," *Computer Vision and Image Understanding* **113**(11), 1107–1117 (2009).

[10] M.Livingston, "Vision-based tracking with dynamic structured light for video see-through augmented reality," *PhD Thesis* (1998).

[11] R.Raskar, G. Welch, M. A. L. H., "The office of the future: A unified approach to image-based modeling and spatially immersive displays," in [*Proceedings of SIGGRAPH*], ACM (1998).

[12] S.Boverie, M.Devy, F., "3d perception for new airbag generation," in [*15th IFAC World Congress on Automatic Control*], (2002).

[13] Wei, B., Gao, J., Li, K., Fan, Y., Gao, X., and Gao, B., "Indoor mobile robot obstacle detection based on linear structured light vision system," in [*Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference on*], 834–839, IEEE (2009).