

# Online 3D Reconstruction System Using Multi-view Smartphone Images

GIBEOM KIM<sup>O</sup>, Chung Hun Kang, Cheong Jun Young, In Kyu Park

Department of Information and Communication Engineering, INHA Univ.

gibeomkim@hotmail.com, chunghunk@gmail.com, cheongjunyoung@gmail.com, pik@inha.ac.kr

## Abstract

The purpose of this paper is to establish an online system reconstructing multi images photographed by smartphones in 3D form. The paper proposes the online 3D reconstruction system using web framework establishment and communication, camera calibration, editing photographed image information, SfM (Structure from Motion), and MVS (Multi-View Stereo) in order to reconstruct multi images photographed by smartphones in 3D form. Using the proposed system, multi images photographed by smartphones can be reconstructed in 3D form, and the user can control and check reconstructed objects in smartphones and PC through real time rendering. In addition, the process has been improved so that the process is not as complicated and enables easy use. As smartphone cameras and processes have improved, 3D form reconstruction using multi-images photographed by smartphones would be the new content in the market.

## 1. Introduction

Thanks to recent increase in 3D printer dissemination and services using 3D related technology, public interest of 3D images continues to increase significantly. Also, due to increase of high performance smartphone dissemination, users using 3D related contents by smartphones are expected to increase. Although the public interest of 3D related services has increased, difficulties follow in terms of the public creating their own 3D related contents.

The University of Washington has conducted a study of Photo Tourism[1] in which multi-angle images photographed by digital cameras are used to reconstruct structures. In addition, the study had a good result of creating 3D maps using multi-angle images of various structures. However, this method has its limit in use of images photographed by smartphones, and its shortcoming is that it is complicated and hard for the public to use.

This paper proposes the online 3D reconstruction system used to create 3D related contents using smartphone cameras, which is different from that of conventional complicated methods. The proposed system is subdivided into web framework establishment, web server communication, 3D reconstruction using server SfM (Structure from Motion)[1] and MVS (Multi-View Stereo)[2], and GUI (Graphic User Interface) used to control reconstructed objects in PC and smartphones.

The composition of this paper is as of the following. Chapter 2 describes components required for online 3D reconstruction system development such as the summary

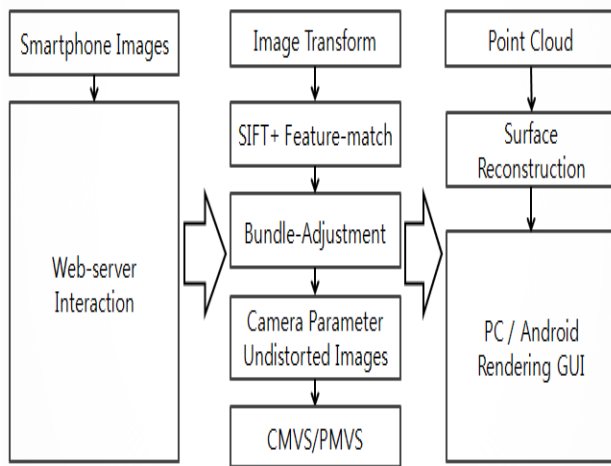
of online 3D reconstruction system, SfM, MVS, and user control GUI (Graphic User Interface). Chapter 3 conducts comparative tests and takes considerations of different conditions of input images using the proposed system. Finally, Chapter 4 draws a conclusion.

## 2. Online 3D Reconstruction System

### 2.1 Summary of Online 3D Reconstruction System

The block diagram of the proposed 3D reconstruction system using smartphones is shown in [Picture 1], and each system function is as of the following.

Multiple images photographed by smartphones are transferred to the server through web frame establishment and web server communication. Transferred images will perform camera calibration after image data conversion. Then camera location data and 3D point data are saved through SfM. Using extracted data such as undistorted images and camera location data, 3D Point Cloud will be created through CMVS (Cluster Views for Multi-View Stereo) and PMVS (Patch-Based Multi-View Stereo) [3]. The point based objects will be reconstructed to 3D form with smooth surfaces using the Surface Reconstruction [4] algorithm. The reconstructed 3D can be checked by the user with a smartphone or PC.

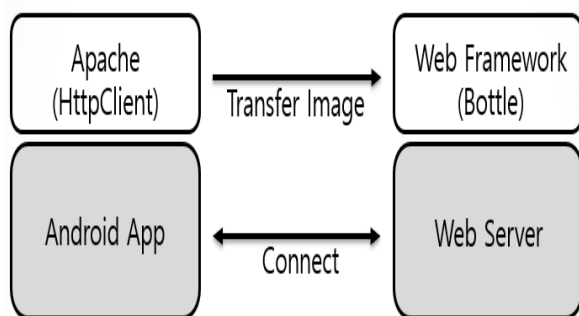


Picture 1. The Block Diagram of 3D Reconstruction System

## 2.2 Smartphone and Web Server Communication

The web framework establishment and server communication processes of this paper are shown through [Picture 2].

The web framework is established using Bottle, Python Web Framework, and the web server transfer multipart communication is conducted using HttpClient.



Picture 2. Smartphone and web server communication

## 2.3 SfM and MVS Technology & User Control GUI

SfM technology dealt in this paper extracts features from multiple 2D images and obtains connected features using information of images and camera data for 3D reconstruction.

MVS technology creates 3D Point Cloud using camera data obtained from various images.

Using obtained data from SfM and MVS, users can create a GUI environment to view reconstructed 3D images in PC and smartphones.

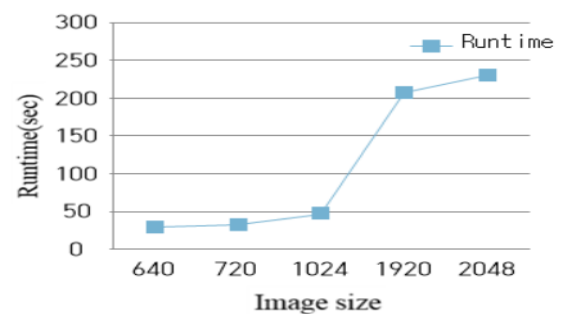
## 3. Comparison Test of Online 3D Reconstruction System Run Time

For the test environment, smartphone API 19: 4.4 (KitKat) Galaxy A7-SM-A7005, computer-1.7GHz Intel(R) Core(TM) i5 CPU, 8192MB RAM, Intel(R) HDGraphics 4000 for Window 7 Cygwin and Virtual Box Ubuntu 4.04.3, Bundler[5], CMVS/PMVS, EXIV2, OpenCV, OpenGL, OpenGL ES, C language and Shell programing are utilized.

### 3.1 Comparing Run Time Depending on Image Resolution

The test was conducted by inputting different maximum values of width and height including 640, 720, 1024, 1920, and 2048. The number of the images was set to 20.

[Picture 3] is a graph comparing the run time of extracting 3D points of each image according to the resolution. 1024 was decided as the optimal resolution for performing 3D reconstruction because the test revealed 3D reconstruction was performed without increasing too much run time.

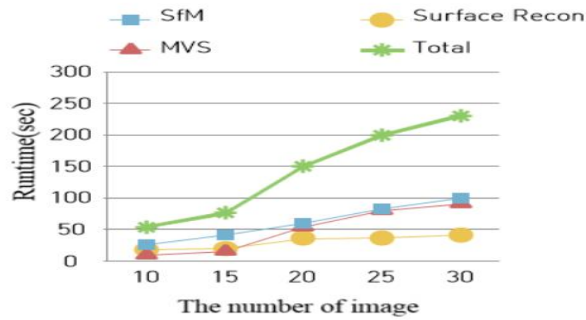


Picture 3. Comparing the run time in SfM depending on resolution

### 3.2 Comparing Run Time Depending on Input Image Numbers

The test was conducted by change the number of images transferred from smartphones to web server to 10, 15, 20, 25, and 30. The maximum resolution was set to 1024.

[Picture 4] is a graph comparing the run time of SfM, MVS, and Surface Reconstruction depending on input number of images. The maximum number of input images was decided as 20 in this paper, considering image numbers, run time, and reconstruction accuracy through the test.



Picture 4. Comparing the run time depending on input image numbers in SfM, MVS, and Surface Reconstruction stage

### 3.3 Test Results and Consideration

According to the test results through the proposed system in this paper, reconstructed 3D objects were able to be confirmed using PC and smartphones as shown in [Picture 5] and [Picture 6].



Picture 5. 3D object implemented by PC GUI



Picture 6. 3D object implemented by smartphone GUI

The structure of the proposed system was connected as one, and it took approximately 5 minutes to perform 3D reconstruction of 20 images or less. This system has its advantage as the user can easily create a 3D reconstruction object.

### 4. Conclusion

The conventional methods of reconstructing images photographed by smartphones to 3D form are very complicated and hard for the public to use as well as program accessibility and the method of use. Therefore, this paper has established one integrated system of algorithms and technologies performing 3D reconstruction.

Through the proposed system, the public would be able to easily create photographed object images to 3D contents using smartphones. Furthermore, the advancements of various industries can be promoted using 3D related technologies.

### References

- [1] Noah Snavely, Steven M. Seitz, Richard Szeliski, "Photo tourism: Exploring photo collections in 3D," ACM Trans. Graphics, vol.25, no.3, pp. 834-846, 2006.
- [2] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein and Richard Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 519-528, 2006.
- [3] Yasutaka Furukawa, Jean Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1362-1376, 2010.
- [4] Michael Kazhdan, Hugues Hoppe, "Screen Poisson Surface Reconstruction," ACM Trans. Graphics, vol. 32, no. 3, 2013.
- [5] Noah Snavely, Bundler: Structure from Motion (SfM) for Unordered Image Collections, <http://www.cs.cornell.edu/~snavely/bundler/> (Search Date: 2015. 09. 11)

# 스마트폰으로 촬영된 다중 영상을 이용한 온라인 3D 형상 복원 시스템

김기범<sup>○</sup>, 강충현, 정준영, 박인규

인하대학교 정보통신 공학부

gibeomkim@hotmail.com, chunghunk@gmail.com, cheongjunyoung@gmail.com, pik@inha.ac.kr

## 요 약

본 연구의 목적은 스마트폰으로 촬영한 다중 영상을 3D 형상 복원시키기 위한 온라인 시스템의 구축이다. 스마트폰으로 촬영한 다중 영상을 3D 형상으로 복원하기 위한 웹 프레임워크 구축 및 통신, 카메라 캘리브레이션, 촬영한 영상의 정보 수정, SfM (Structure from Motion), MVS (Multi-View Stereo)를 이용한 온라인에서의 3D 형상 복원 시스템을 제안한다. 제안한 시스템을 이용하여 스마트폰으로 촬영한 다중 영상을 3D 형상으로 복원하고 복원된 객체를 스마트폰 및 PC 에서의 실시간 렌더링을 통해 사용자가 조작하며 확인할 수 있다. 또한 그 과정이 전보다 복잡하지 않고 사용자가 쉽게 이용할 수 있도록 개선되었다. 스마트폰 카메라와 프로세스의 성능이 향상됨에 따라 스마트폰에서 촬영된 다중 영상을 이용한 3D 형상 복원이 앞으로의 시장에서 새로운 콘텐츠로 이용될 것이다.

## 1. 서론

최근 3D 프린터의 보급 및 3D 관련 기술을 활용한 서비스가 증가하면서 3D 영상에 대한 대중들의 관심은 계속해서 크게 증가하고 있다. 또한 고성능 스마트폰의 보급률이 증가하면서 스마트폰을 이용한 3D 관련 콘텐츠를 이용하는 사용자가 증가할 것으로 예상된다. 이렇게 대중들의 3D 관련 서비스 관심도가 증가했지만 일반인들의 3D 관련 콘텐츠를 직접 제작하는 일에는 많은 어려움이 따른다.

University of Washington 은 디지털 카메라로 촬영된 다 각도의 영상들을 이용하여 건축물들을 복원하는 Photo Tourism[1]의 연구를 수행하였다. 게다가 여러 각도의 다양한 건축물 영상을 이용하여 3D 지도를 그려내는 좋은 연구 성과를 낳았다. 그러나 이 방법은 스마트폰으로 촬영된 영상들에는 한계가 있고, 일반 사용자가 사용하기에는 복잡하고 어려운 단점이 있다.

본 논문에서는 스마트폰 카메라를 이용하여 3D 관련 콘텐츠를 제작하기 위해 기존의 복잡한 방법과는 달리 일반 사용자가 쉽게 이용할 수 있는 온라인 3D 형상 복원 시스템을 제안한다. 제안한 시스템은 웹 프레임워크 구축 및 웹 서버 통신, 서버에서 SfM (Structure from Motion)[1]과 MVS (Multi-View Stereo)[2]를 이용한 3D 복원, 복원된 개체를 PC 및 스마트폰을 통해 사용자 조작할 수 있는 GUI (Graphic User Interface)로 나누어 진다.

본 논문의 구성을 살펴보면 다음과 같다. 2 장에

서는 온라인 3D 형상 복원 시스템 개발에 필요한 요소들으로써 온라인 3D 형상 복원 시스템 개발의 개요와 SfM, MVS 및 사용자 조작 GUI (Graphic User Interface)에 대해 설명한다. 3 장에서는 제안한 시스템을 이용하여 입력 영상들의 조건에 따른 실행 시간 비교 실험 및 고찰을 한다. 그리고 4 장에서는 결론을 맺는다.

## 2. 온라인 3D 형상 복원 시스템

### 2.1 온라인 3D 형상 복원 시스템 의 개요

본 논문에서 제안하는 스마트폰을 이용한 3D 형상 복원 시스템의 블록 구성도는 [그림 1]에 도시하였으며, 각 시스템의 역할은 다음과 같다.

웹 프레임워크 구축 및 웹 서버 통신을 통해 스마트폰에서 촬영한 여러 장의 영상을 서버로 전송한다. 전송된 영상들은 이미지 데이터 변환 후 카메라 캘리브레이션 과정을 수행한다. 그 후 SfM 을 통해 카메라 위치 정보와 3D 포인트 정보를 저장한다. 추출한 정보들을 활용하여 얻은 Undistorted 이미지와 카메라 위치 정보를 이용하여 CMVS (Cluster Views for Multi-View Stereo) 와 PMVS (Patch-Based Multi-View Stereo) [3]를 통해 3D Point Cloud 를 형성해 낸다. 포인트 기반의 3D 물체는 Surface Reconstruction[4] 알고리즘을 통해 매끄러운 표면을

가진 3D 형상으로 복원시킨다. 복원된 3D 는 스마트폰과 PC 에서 사용자가 확인 할 수 있다.

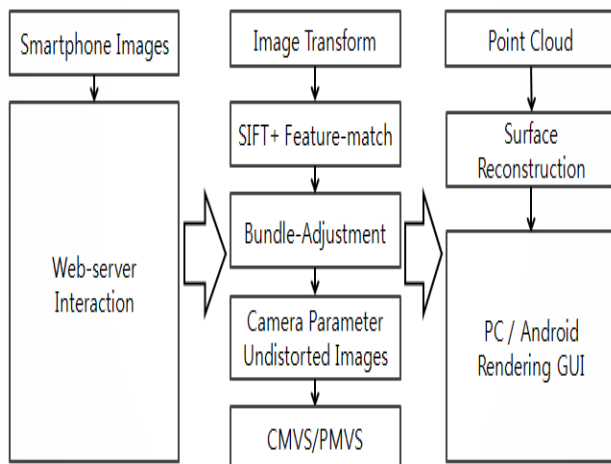


그림 1. 3D 형상 복원 시스템 블록 구성도

## 2.2 스마트폰과 웹 서버 통신

본 논문에서 웹 프레임워크 구축 및 서버 통신 과정은 [그림 2]를 통해서 나타낸다.

Python Web Framework 인 Bottle 을 이용하여 웹 프레임워크를 구축하고 HttpClient 을 이용하여 웹 서버 전송 Multipart 통신을 진행한다.

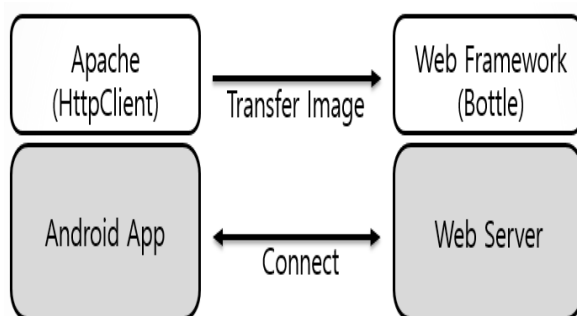


그림 2. 스마트폰과 웹 서버 통신

## 2.3 SfM 과 MVS 기술 및 사용자 조작 GUI

본 논문에서 다루는 SfM 기술은 주어진 여러 개의 2D 이미지들에서 특징 점을 추출하여 3D 형상 복원을 구현하기 위해 이미지 정보와 카메라 정보들을 이용하여 연결된 특징 점들을 얻어내는 기술이다.

MVS 기술은 주어진 여러 개의 이미지로부터 얻어낸 카메라 정보들을 이용하여 3D 형상의 Point

Cloud 를 만들어내는 기술이다.

SfM 및 MVS 통해 취득한 정보를 이용하여 사용자가 PC 및 스마트폰에서 복원된 3D 형상을 보기 위한 GUI 환경을 구축한다.

## 3. 온라인 3D 형상 복원 시스템의 실행시간 비교 실험

본 논문에서 제안한 시스템의 실험 환경은 스마트폰-API 19: 4.4 (KitKat) Galaxy A7-SM-A7005, 컴퓨터-1.7GHz Intel(R) Core(TM) i5 CPU, 8192MB RAM, Intel(R) HDGraphics 4000 의 Window 7 Cygwin 및 Virtual Box Ubuntu 4.04.3, Bundler[5], CMVS/PMVS, EXIV2, OpenCV, OpenGL, OpenGL ES, C 언어 및 쉘 프로그래밍을 이용하였다.

### 3.1 이미지 해상도 크기에 따른 실행 시간 비교

본 실험에서는 입력 이미지의 너비와 높이 중 최대 값을 다음과 같이 640, 720, 1024, 1920, 2048 으로 바꿔주면서 실험을 진행하였다. 이미지 개수는 20 개로 고정하였다.

[그림 3]은 각 입력 이미지의 해상도 크기에 따라서 3D 포인트를 추출하는 실행 시간을 비교하는 그래프이다. 실험을 통해 실행 시간 증가 폭이 크지 않으면서 3D 형상 복원을 수행할 수 있는 크기인 1024 를 최적의 크기로 결정한다.

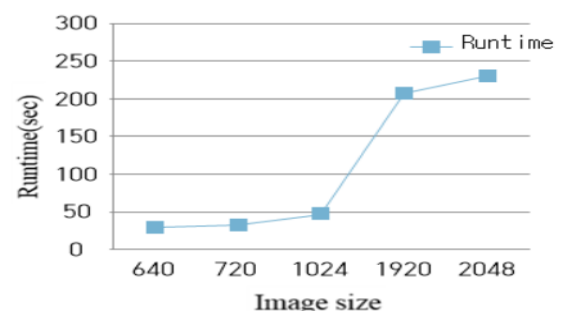


그림 3. SfM 에서 이미지 크기에 따른 실행 시간 비교

### 3.2 입력 이미지 개수에 따른 실행시간 비교

본 실험에서는 스마트폰에서 웹 서버로 전송된 이미지의 개수를 10, 15, 20, 25, 30 개로 변경 하면서 실험을 진행하였다. 이미지의 해상도의 최대값 크기는 1024 로 고정했다.

[그림 4]은 입력 이미지의 개수에 따라 SfM, MVS, Surface Reconstruction 의 실행 시간을 비교한 것을 그래프로 나타내었다. 실험을 통해 입력 이미지의 개수와 실행시간 및 복원의 정확도를 고려하



여 본 논문에서는 최적의 입력 이미지 개수를 20 개로 정하였다.

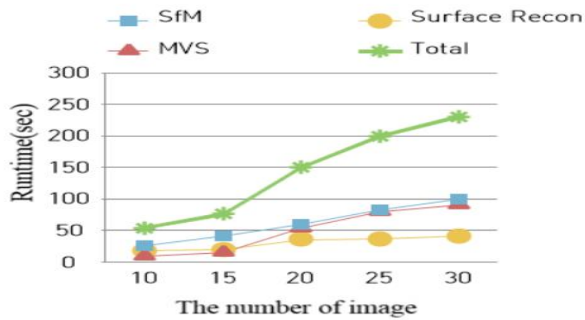


그림 4. SfM, MVS, Surface Reconstruction 단계에서 입력 이미지 개수에 따른 실행 시간 비교

### 3.3 실험 결과 및 고찰

본 연구에서 구축된 시스템을 통해 실험을 한 결과 [그림 5], [그림 6]과 같이 PC 및 스마트폰에서 복원된 3D 객체를 확인 할 수 있었다.



그림 5. PC GUI를 통해 구현된 3D 객체



그림 6. 스마트폰 GUI 를 통해 구현된 3D 객체

제안한 시스템은 구조가 하나로 연결되어 20 개 이내의 이미지로 3D 복원을 실행하는데 약 5 분 정도의 실행 시간이 걸렸다. 이 시스템은 사용자가 3D 복원 객체를 만들기에 편리한 이점이 있다.

### 4. 결론

기존의 방법들을 이용하여 스마트폰으로 촬영된 이미지를 3D 형상으로 복원을 하는 방법은 일반 사용자가 이용하기에 프로그램 접근 용이성 및 사용 방법이 매우 까다롭고 어렵다. 이에 본 연구에서는 3D 형상 복원을 진행하는 알고리즘 및 기술들을 하나로 통합한 시스템을 구축하였다.

본 연구 결과를 통해 일반인 사용자가 스마트폰을 이용해서 직접 촬영한 사물 이미지를 3D 콘텐츠로 쉽게 제작할 수 있을 것이다. 나아가 3D 관련 기술을 활용한 다양한 산업 발전을 도모할 수 있을 것이다.

### 참고문헌

- [1] Noah Snavely, Steven M. Seitz, Richard Szeliski, "Photo tourism: Exploring photo collections in 3D," ACM Trans. Graphics, vol.25, no.3, pp. 834-846, 2006.
- [2] Steven M. Seitz, Brian Curless, James Diebel, Daniel Scharstein and Richard Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 519-528, 2006.
- [3] Yasutaka Furukawa, Jean Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 32, no. 8, pp. 1362-1376, 2010.
- [4] Michael Kazhdan, Hugues Hoppe, "Screend Poisson Surface Reconstruction," ACM Trans. Graphics, vol. 32, no. 3, 2013.
- [5] Noah Snavely, Bundler: Structure from Motion (SfM) for Unordered Image Collections, <http://www.cs.cornell.edu/~snavely/bundler/> (검색 일자: 2015. 09. 11)