

User Detection in Real-time Panoramic View through Image Synchronization using Multiple Camera in Cloud

Yung Fu Tan*, Mangal Sain**, Lee Byung Gook*

* Visual Content Department Dongseo University, Busan, South Korea

**Department of Information Engineering, Dongseo University, Busan, South Korea

Yungfu88@hotmail.com, mangalsain1@gmail.com, lbg@dongseo.ac.kr

Abstract— Nowadays to manage a big collection of image data is becoming a major issue for companies as well as for smart phone device. After panoramic view in camera, elongated fields of view technology and algorithms has been fully utilizes in industrial area, commercial systems, painting and mobile devices such as smart phones and portable computers. But still majority of panoramic view is only implemented to capture wide angle scenery or a three-dimensional model. In this paper a cloud computing service and its application for data synchronization with multiple cameras has been presented. This service has been implemented using image extracted data transferring without multiple devices for synchronization. For image data analysis, a dedicated application is developed with single capture device and each data will be sent to server for image processing computation. The goal of this research is to provide a better image data synchronization which perform a faster real-time panoramic view and analyze human tracking information precisely. In this system multiple cameras were used with image data synchronization to get panoramic view and performing human detection from each panoramic view frames.

The system is setup with multiple camera connected with few computer where each of them sending image direct to cloud server and one of machine will be retrieving all image contents directly from cloud server for information mapping, image stitching and computation part. For better performance different cloud server has been used to analyze data transferring size, speed, computation time with different features tracking method. Finally this paper concludes with a system which can be specially use in wide are view for multiple capture devices. This system can be used in various systems which can help in vehicle safety system as well as wide are human detection.

Index Terms— Cloud Computing, Human Tracking, Image Processing, information mapping and data synchronization

I. INTRODUCTION

After release of panoramic view in camera, this elongated fields of view technology and algorithms has not been fully utilizes in industrial area, commercial systems, painting and mobile devices such as smart phones and portable computers. Majority of the panoramic view was only implemented to capture wide angle scenery or a three-dimensional model such as some work by Christian Schönauer [1-2] to capture wide area motion tacking with 3d depth sensor.

On other side, Yu-Jin Hong, Jae-In Hwang, Sun-Bum Youn, Sang Chul Ahn, Hyung-Gon Kim, Heedong Ko [3] used interactive panorama video viewer with head algorithms. J. Segen and S. Kumar [4] uses a controlled background to localize the hand efficiently in real-time. Francois [5] presented head motion as a new input stream. Rony Ferzli, Ibrahim Khalife [6] show the importance and benefit of coupling cloud computing with mobile especially due to power limitation that mobile devices exhibit and included with head tracking.

In this paper a cloud computing service and its application for data synchronization with multiple cameras has been presented which is implemented using image extracted data transferring without multiple devices for synchronization. Especially for data analysis, a dedicated application is developed with single capture device and each data will be sent to server for image processing computation. To capture multiple pictures in same time frame and to get panoramic view and performing human detection from each panoramic view frames different cameras were used.

However, in this system, the first part is the combination of multiple photographic images together on the homogeneous virtual surface. The overlapped regions between images are match to each other, then images are warped onto the panorama surface using the estimated camera motion and geometric relation between the panorama surface and image coordinates inspired by those in [7], [8]. The outcome of panoramic images will generate wide scenes of view. It is sometime known as wide format photography which cannot be captured by single picture of usual cameras. Thus, panorama synthesis overcomes the limitations of view angles and resolutions in the usual cameras. [9] - [11].

On the others hand, real-time capturing with computation of panoramic images will send to cloud computing part for data transferring and retrieving. The cloud system become colloquial expression used to describe a variety of different types of computing concepts that involve a large number of computer that are connected through a real-time communication network. It provides a low-cost, highly accessible alternative to other traditional high-performance computing platforms. Cloud computing has many other benefits such as high availability, scalability, elasticity, and

free of maintenance. Majority of cloud are currently cloud database and data storing cloud [12].

The main goal of this research is to provide a better and faster method for image data synchronization through multiple real-time panoramic camera capture in the most efficient way with using cloud server included human tracking at the end results. The system separated into three parts where the first part is multiple camera capture with computation of panoramic view frames. Each panoramic frame is sending to cloud server. The second part will be retrieving all images from cloud server with computation of overall full panoramic view. Last part will be human tracking and position in the full panoramic view frames. Several methods has been analyzed and compare through during panoramic computation such as FAST Features Detection, SURF Features Detection, and Harris Corners Detection. Where else different cloud server has been tested and compared such as Google Drive, Dropbox, Microsoft Skydrive, Amazon Cloud Drive and Olleh Ucloud.

In this study, human tracking in real-time camera capture with wide angle view computation through cloud system is further implemented which enable the credibility of human detection, human interaction or vehicles safety system.

II. SYSTEM DESIGN

A. System Overview

Figure.1 shows the overall structure for this system. At first, we capture wide-angle image for each two HD webcam which is connected to PC.

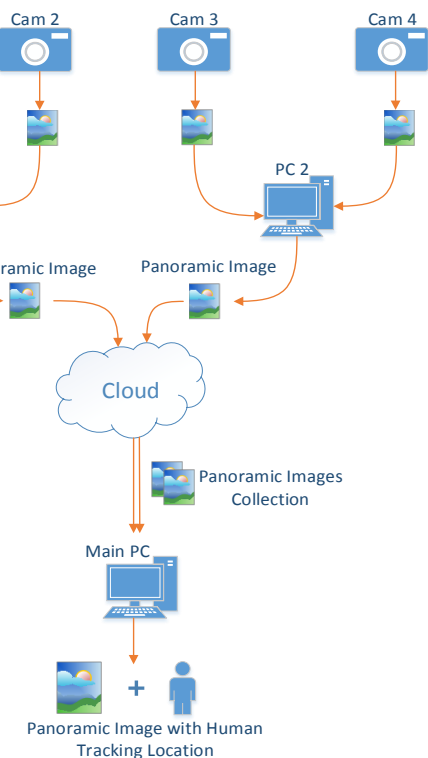


Figure 1. System Structure

These photos will be stitched and used as source of panoramic image. In order to send the panoramic image from each computer to same cloud server computed panoramic image will be label. Overall, a main PC will be functioning as result output PC to download each image from cloud server from time to time and stitching all panoramic images into one whole complete panoramic image. At the same time human location tracking result will be generated in the main PC as well.

B. Panoramic View

To generate a panoramic image, this system is computed in different feature detection methods which are FAST Features Detection [13-15], SURF Features Detection [16-19] and Harris Corners Detection [20-21] to compare and analyze the fastest and the most precise image stitching results. Each of the detection method will computed a homography that maps the relation in between images. Different features detection method will compute different type of mapping result and computation performance.

1. FAST Features Detection

FAST (Features from Accelerated Segment Test) features algorithm derives from definition of what constitutes a “corner”. As show in Figure 2, FAST method is based on image intensity around a putative feature point. The key point is gain by examining a circle of pixels centered at a candidate point. When an arc of contiguous points of length greater than 3/4 of circle perimeter is found in which all pixels significantly differ from the intensity of the center point, then a key point is declared. This algorithm result in very fast interest point detection and should be used when speed is a concern.



Figure 2. FAST Features Detection

2. SURF Features Detection

SURF (Speed up Robust Features) features as show in Figure 3, which is the scale-invariant features that has been introduced in computer vision. SURF not only scale-invariant features but they also benefit in efficiently computation. The

circles of detected key points change in size is proportional to the scale change.

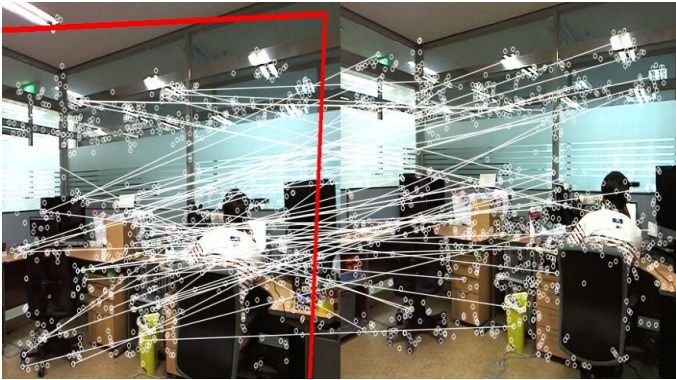


Figure 3. SURF Features Detection

To detect the features, the Hessian matrix is computed at each pixel. This matrix measures the local curvature of a function and is defined as:

$$H f(x, y) \equiv \begin{vmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{vmatrix}. \quad (1)$$

3. Harris Corners Detection

Harris looks at the average directional intensity change in small window around a putative interest point.

$$E(u, v) = \sum_{x,y} w(x, y) [I(x + u, y + v) - I(x, y)]^2 \quad (2)$$

The average intensity change can then be computed in all possible directions which lead to the definition of a corner as a point for which the average change is high in more than one direction. Harris is obtained with the direction of maximal average intensity change. If the average intensity change in the orthogonal direction is also high, then we will have a corner point as in Figure 3.1.



Figure 3.1 Harris Corners Detection

Once the interest point in Harris Corners has been detected, Accord.NET image stitching library has been use to perform Correlation matching, robust homography estimation and gradient blending. Correlation matching method will be measure and matching the interest point extracted from Harris Corners method. Given a set of correlation points there allow to pair the points in between two images as show in Figure 3.2.

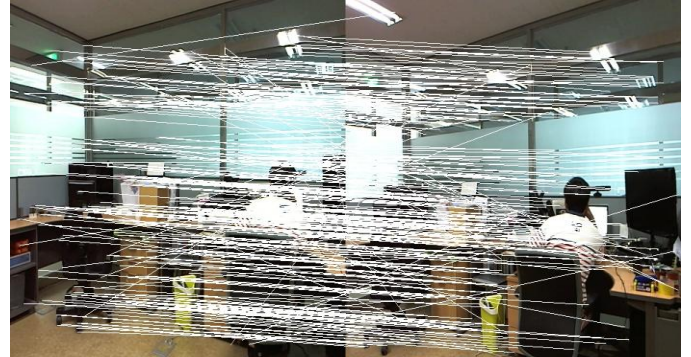


Figure 3.2. Correlation points Paring.

Next will be Robust Homography Estimation which using RANSAC for fitting a homography matrix. A robust estimator of RANSAC will generate a homography matrix from previous set of correlation feature points. RANSAC is actually an abbreviation for “RANDOM Sample Consensus”, which is an iterative method for robust parameter estimation to fit mathematical models from sets of observed data points which may contain outliers. RANSAC works by trying to fit several models using some of the point pairs and then checking if the models were able to relate most of the points as show in Figure 3.3 as below.

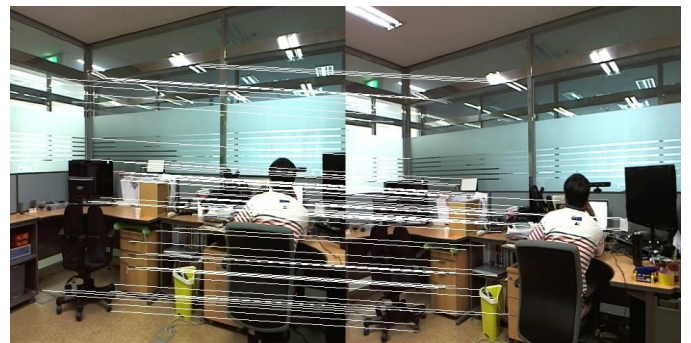


Figure 3.3. RANSAC estimator result.

4. Homography Matrix

The above methods are being used to extract feature points for homography computation, where the homography matrix is matching two images. A homography is a projective transformation, a kind of transformation used in projective geometry. It describes what happens to the perceived positions of observed objects when the point of view of observer changes. In others terms, a homography is an invertible transformation from the real projective plane to the projective plane that maps straight lines to straight lines.

By using homogeneous coordinates, which can represent a homography matrix as a 3x3 matrix with 8 degrees of freedom.

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \times \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} \tag{3}$$

Homogenous coordinates are very useful because they will allow the system to perform an image projective transformation by using only standard matrix multiplication, as shown by the equation and schematic diagrams above. Once, all the projected points have been computed. The original coordinate system is recovered by dividing each point by its homogenous scale parameter and then dropping the scale factor, which after division will be set at 1.

Result which following table showing is average of 30 set data with different execution time by different detection method.

TABLE 1 FEATURE DETECTION METHOD WITH EXECUTION TIME AND QUANTITY OF FEATURE POINTS RESULT

feature detection methods	Average Execution Time (millisecond)
FAST Feature Detection	449.667
SURF Feature Detection	334.033
Harris Corner Detection	1364.800

Homography matrix that has been computed which allow our system to perform Blending of two images from Accord.NET library. The system is using a linear gradient alpha blending from the center of one image to the other. Gradient blending works by simulating a gradual change in one image's alpha channel over the line which connects the centers of two images as show in Figure 4 and Figure 5.

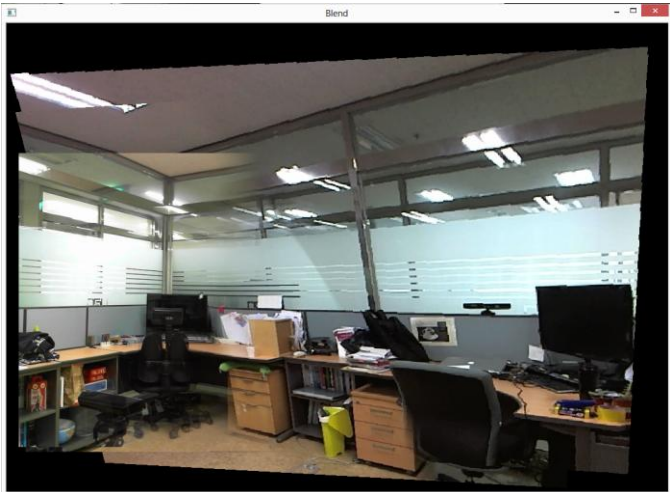


Figure 4. Image Blending

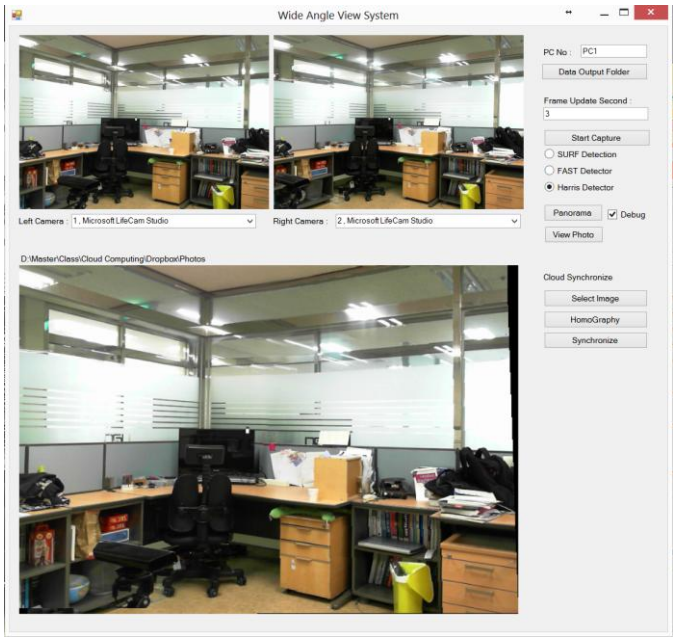


Figure 5. Image Blending. Top Left: Camera 1 Image, Top Right: Camera 2 Image, Bottom: Blending of two images.

C. Cloud Computing

In cloud computing we separated images data into uploading part and downloading part. Panoramic images that being captured by each two computer will be named and uploading to cloud server folder for synchronization used.

For uploading part in our system as shown in Figure 1, stitching images of Cam 1 and Cam 2 for PC1 will be name as “ImgPC1.jpg” sending to Cloud server. At here, we were using SURF Features detection for homography matrix generation where this is the fastest feature detection method among the others shown in Table 1. Feature detection matrix only perform once and same homography will be reuse for next image stitching. Same goes to Cam1 and Cam2 in PC2 will be named as “ImgPC2.jpg”, where else same stitching and features detection method is same like in PC1 too.

Furthermore, in downloading part of our system, a Main PC will be operated to download all the images from cloud server and sorting the images according with images’ name. Image stitching with the same method that was used for image stitching method PC1 and PC2 will be used. Overall sorted images will generate a last panoramic view scene. At last, human tracking will be analyzed from overall Panoramic frame to extract an exact human location using OpenCV Human detection.

In our system we had tested and analyze our cloud system from different third party Cloud server such as Google Drive Dropbox, Microsoft Skydrive, Amazon Cloud Drive, Olleh Ucloud with 100kb to 200kb image file.

TABLE 2 CLOUD SERVER UPLOADING

Cloud Server	Uploading Time (s)
Google Drive	4
Dropbox	3
Microsoft SkyDrive	11
Amazon Cloud Drive	8
Olleh Ucloud	13

D. Human Detection

Human recognition [22] is an easy task for humans. Our system included human tracking that using EmguCV (OpenCV C#) Human Detection & Haar Classifier where referring the face recognition in [23]. EmguCV is an open source computer vision library which is cross platform .Net wrapper to OpenCV image processing library. EmguCV allow openCV function to be called from .Net. In our system each frame that gathers at main PC will be run through image stitching to generate overall view. In term of that, a human tracking algorithms will be functioning as human location tracking machine where will output the human location in the panoramic image as show in Figure 6 as below. This system able to perform a high accuracy location in images.

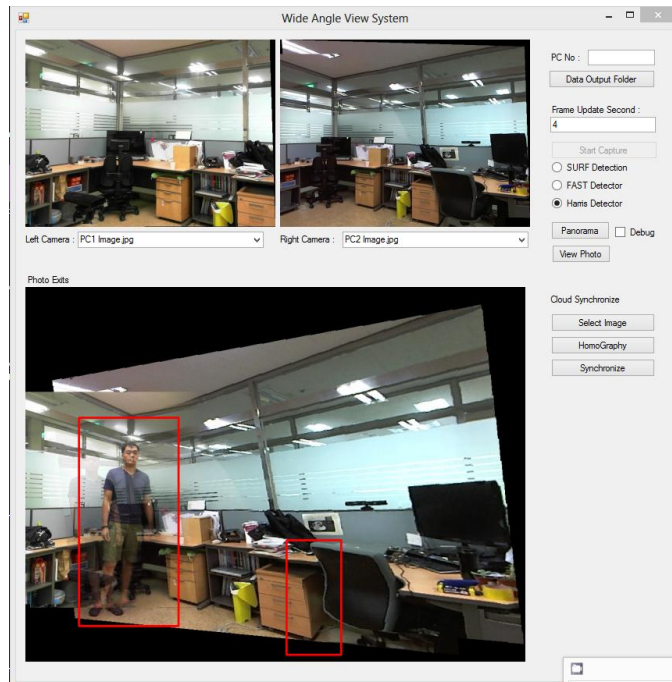


Figure 6. Human detection in Main PC Panoramic image.

III. REAL-TIME PANORAMIC USER DETECTOR CLOUD SYSTEM

Real-time camera capture has been widely used and implemented in human tracking broadcast system [24-25]. As in our system, the real-time panoramic system conducts precise and high resolution images stitching. Then system blends each frame before sending and after retrieving from cloud server. Different methods of features extraction enable the most natural images being compute in most efficient time. Real time

images will be uploading and downloading through cloud server agency over internet. Cloud server agency like Dropbox, Amazon Cloud, Microsoft Skydrive and Olleh Ucloud has been analyzed with the server synchronization speed. Dropbox, which has been tested as the most efficient cloud server, with the benefit of cloud agency, maintenance and cost for server will be highly reduce.

Through the experiment and data analyzing, we use SURF Detection where having 334.033 millisecond of fastest execution time for image stitching homography generation. Beside we used dropbox as the fastest image data cloud transferring tools which was 3 second per frame for uploading time. Though this experiments, hardware requirements and internet speed also will bring different result for overall cam captured images.

Instead of having a good security system providing by cloud server agency, Cloud system enable to work portable over wifi or land network which we can setup our system at any location all over the world as long as connected to internet. Our system is added and tested with Human tracking of overall panoramic images where enable tracking an exact human location in wide area view images which can be further improve and implement the human tracking system over wide area traffic secure system, wide area football match system, 360 degree 3D view system and wide area filming system.

IV. CONCLUSION

The main goal of this research is to provide a better and faster method for image data synchronization through multiple real-time panoramic camera capture in the most efficient way with using cloud server included human tracking at the end results. Consequently, the analyzed data of the result of this system which enable precise feature detection for panoramic image generation. Beside, multiple capture devices sharing data information with better computation speed and user mapping location precisely in a wide area view. These systems enable the credibility of wide area human detection, traffic road safety system, human interaction or vehicles safety system.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MKE/KEIT. [10041682, Development of high definition 3D image processing technologies using advanced integral imaging with improved depth range]

REFERENCES

- [1] Christian Schönauer, Hannes Kaufmann, "Wide Area Motion Tracking Using Consumer Hardware", ACM Advances in Computer Entertainment Technology conference (ACE 2011), Lisbon, Portugal, 08.11.2011.
- [2] Christian Schönauer, Hannes Kaufmann, "Wide Area Motion Tracking Using Consumer Hardware", Youtube link, <https://www.youtube.com/watch?v=qWSay6Cc840>, 22 April 2013.
- [3] Yu-Jin Hong, Jae-In Hwang, Sun-Bum Youn, Sang Chul Ahn, Hyung-Gon Kim, Heedong Ko, "Interactive panorama video viewer with head tracking algorithms", Human-Centric

- Computing (HumanCom), 2010 3rd International Conference. 11-13 Aug. 2010.
- [4] J. Segen, S. Kumar, "Shadow gestures: 3D hand pose estimation using a single camera", Proc. of the Computer Vision and Pattern Recognition Conference, CVPR99, v. 1: 485, 1999.
 - [5] F. B&ard, "The perceptual window: Head motion as a new input stream", Proc. of the IFIP Conference on Human-Computer Interaction, INTERACT99, pp. 238-244, 1999.
 - [6] Rony Ferzli, Ibrahim Khalife, "Mobile Cloud Computing Educational Tool For Image/Video Processing", Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop (DSP/SPE), 2011 IEEE, 4-7 Jan. 2011.
 - [7] R. Szeliski, "Image alignment and stitching: A tutorial," Technical Report MSR-TR-2004-92, Microsoft Research.
 - [8] R. Szeliski, "Video mosaics for virtual environments," *IEEE Computer Graphics and Applications*, vol. 16, no. 2, pp. 22-30, 1996.
 - [9] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International Journal of Computer Vision*, vol. 74, no. 1, 2007.
 - [10] S. J. Ha, H. I. Koo, S. H. Lee, N. I. Cho, and S. K. Kim, "Panorama mosaic optimization for mobile camera system," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 4, pp. 1217-1225, 2007.
 - [11] S. J. Ha, S. H. Lee, N. I. Cho, S. K. Kim, and B. J. Son, "Embedded panoramic mosaic system using auto-shot interface," *IEEE Transactions on Consumer Electronics*, vol. 54, no. 1, pp. 16-24, 2008.
 - [12] Deka Ganesh Chandra, Ravi Prakash, Swati Lamdharia, "A Study on Cloud Database", 2012 Fourth International Conference on Computational Intelligence and Communication Networks, 2012.
 - [13] Robert Laganiere, OpenCV 2 Computer Vision Application Programming Cook Book, "Detecting and Matching Interest Points", pages 191-212, May 2011
 - [14] Edward Rosten, Reid Porter, and Tom Drummond, "Faster and better: a machine learning approach to corner detection", Pattern Analysis and Machine Intelligence, IEEE, Jan. 2010.
 - [15] EmguCV, "FAST Features Detector in CSharp", http://www.emgu.com/wiki/index.php/FAST_feature_detector_in_CSharp, 21st July 2013
 - [16] Pan Jie, Chen Wenjie, Peng Wenhui, "A new moving objects detection method based on improved SURF algorithm", Control and Decision Conference (CCDC), 25-27 May 2013.
 - [17] Luo Juan, Oubong Gwun, "SURF applied in panorama image stitching", Image Processing Theory Tools and Applications (IPTA), 2010 2nd International Conference, 7-10 July 2010.
 - [18] Shahzad Ali, Mutawarra Hussain, "Panoramic Image Construction using Feature based Registration Methods", Multitopic Conference (INMIC), 13-15 Dec. 2012.
 - [19] EmguCV, "SURF feature detector in CSharp", http://www.emgu.com/wiki/index.php/FAST_feature_detector_in_CSharp, 21st July 2013.
 - [20] C. Harris, and M. Stephens, "A Combined Corner and Edge Detector", Alvey Vision Conference, page 147--151. (1988).
 - [21] Jianbo Shi, Carlo Tomasi, "Good Features to Track", Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94. 1994 IEEE Computer Society Conference, 21-23 Jun 1994.
 - [22] Elad Ben-Israel, "Tracking of Humans Using Masked Histograms and Mean Shift", Efi Arazi School of Computer Science · The Interdisciplinary Center Herzliya · March 2007.
 - [23] EmguCV, "Face detection in Csharp", http://www.emgu.com/wiki/index.php/Face_detection, 21st July 2013.
 - [24] Rene Kaiser, Marcus Thaler, Andreas Kriechbaum, Hannes Fassold, Werner Bailer, Jakob Rosner, "Real-time Person Tracking in High-resolution Panoramic Video for Automated Broadcast Production", Visual Media Production (CVMP), 2011 Conference, 16-17 Nov. 2011.
 - [25] Beom Su Kim, Sang Hwa Lee, Nam Ik Cho, "Real-time panorama canvas of natural images", Consumer Electronics, IEEE Transactions on (Volume:57, Issue: 4), November 2011.



Yung Fu Tan is a master student majoring in visual contents at Dongseo University, South Korea, where he has start his research here since August 2013. Prior to this he graduated with a B.S (Hons) in Software Engineering from Multimedia University, Malaysia. His recently work has focused on hand pose and hand skeleton estimation for depth sensor camera.



Mangal Sain is an assistant professor in Department of Information Engineering at Dongseo University, Busan Republic of Korea. He received his Ph.D. majoring in Ubiquitous Information technology from Dongseo University, Busan, Korea in 2011. He finishes his master in 2003 from India. During 2003-2007, he joined BSES Ltd and Altivolus InfoTech as a software engineer and Sr. Software Engineer respectively. His research interests are Wireless Sensor Networks, Ubiquitous Healthcare, Embedded Systems, Middleware, Cloud Computing and Cloud Middleware.



Professor Lee Byung Gook is a professor in Visual Contents department at Dongseo University, Busan, South Korea. Prior to this, he graduated with B.S in Mathematics from Yonsei University, M.S in Applied Mathematics from KAIST and Ph.D in Applied Mathematics from KAIST. Currently, his research interest is in development if high-definition 3D image processing technologies using advanced integral imaging with image processing technologies using advanced integral imaging with improved depth range.