|  |  |
| --- | --- |
| Project | **Human Factor for Immersive Content Working Group**  <<http://sites.ieee.org/sagroups-3079/> **>** |
| Title | **Real-time** **User Tracking System Using Depth Camera** |
| DCN | **3079-21-0043-00-0002** |
| Date Submitted | **July 12, 2021** |
| Source(s) | **Sangkwon Peter Jeong** [ceo@joyfun.kr](mailto:ceo@joyfun.kr) **(JoyFun Inc.)**  **Gookhwan Lee** [**ghlee@joyfun.kr**](mailto:ghlee@joyfun.kr) **(JoyFun Inc.)**  **Jimmy Jaeyoung Jang** [**jimmyjanggg@joyfun.kr**](mailto:jimmyjanggg@joyfun.kr) **(JoyFun Inc.)**  **HyeonWoo Nam** [hwnam@dongduk.ac.kr](mailto:hwnam@dongduk.ac.kr) **(Dongduk Women’s University)** |
| Re: |  |
| Abstract | This contribution document defines input/output interface to interact and guide movement of user effectively when depth camera with gesture recognition function and beam project are synchronized, and content guiding and controlling user's movement is serviced through data exchange between depth camera and beam project. |
| Purpose | This contribution document is to define input/output interface interaction to develop and service content utilizing projection display and sensor based on gesture recognition using depth camera. |
| Notice | This document is offered as a basis for discussion and is not binding on the contributing individual(s) or organization(s). The material in this document is subject to change in form and content after further study. The contributor(s) reserve(s) the right to add, amend or withdraw material contained herein. |
| Release | The contributor grants a free, irrevocable license to the IEEE to incorporate material contained in this contribution, and any modifications thereof, in the creation of an IEEE Standards publication; to copyright in the IEEE’s name any IEEE Standards publication even though it may include portions of this contribution; and at the IEEE’s sole discretion to permit others to reproduce in whole or in part the resulting IEEE Standards publication. The contributor also acknowledges and accepts that IEEE 3079 may make this contribution public. |
| Patent Policy | The contributor is familiar with IEEE patent policy, as stated in [Section 6 of the IEEE-SA Standards Board bylaws](http://standards.ieee.org/guides/opman/sect6.html#6.3) <[http://standards.ieee.org/guides/bylaws/sect6-7.html#6](http://127.0.0.1:4664/cache?event_id=757737&schema_id=1&s=5X0vID10lu_E6yrIkWkNd4Wz2H8&q=hancock)> and in *Understanding Patent Issues During IEEE Standards Development* <http://standards.ieee.org/board/pat/faq.pdf> |

Real-time User Tracking System Using Depth Camera

1 Coverage

The technology tracks the user's location in real-time camera images, calculating where the user is within the overall image in 3D coordinates. After that, the 3D coordinates are mapped into the real-time RGB Camera to calculate where they are. By making the calculated area appear in a smaller area than the camera image, it is displayed in real-time user tracking and screen.

2 Term Definition

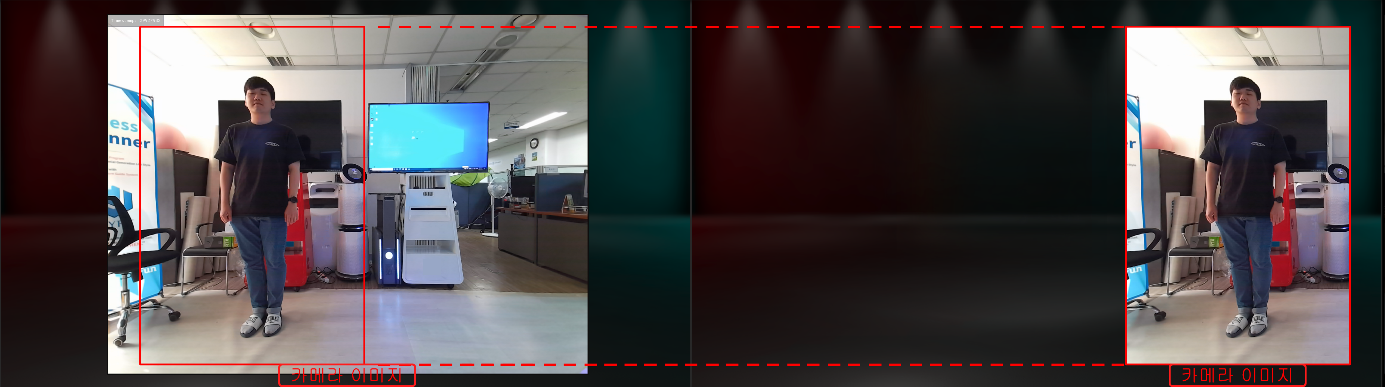
2.1 Skeleton

A line segment connects the major joint points of the human body extracted from the camera image.

3 Real-time tracking system using success information

3.1 The concept of a real-time user tracking system

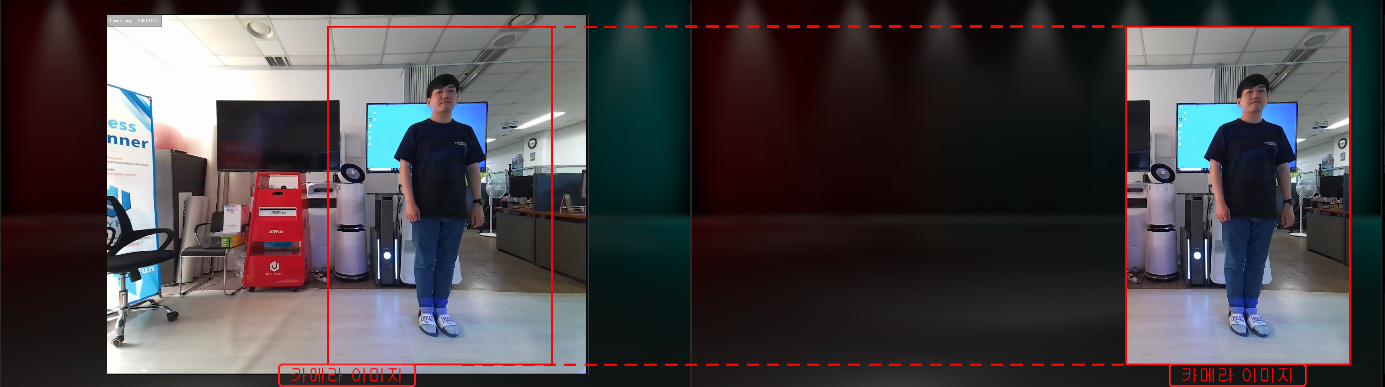
The resolution of the high-performance RGB camera is 2048×1536, but on the other hand, if there is not enough space to express the resolution, it is inappropriate to express the user because the user looks too small if the entire screen is reduced. In this case, only a specific area is expressed so that only a part of the entire screen is optimized for the corresponding resolution.



(Figure 1 a) Show around you when standing on the left



(Figure 1 b) Show around you when standing in the center

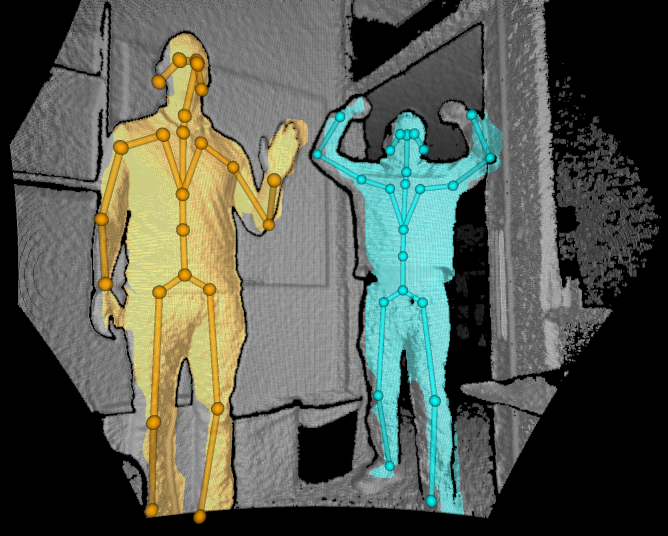


(Figure 1 c) If you stand on the right side of the real-time user tracking system

As shown above (Figure 1), when the user moves, the purpose is to track the user's location and display it in a certain size on the right side.

3.2 Depth camera-based user skeleton information extraction

To find out where the user is from the RGB camera, we need to extract the skeleton information. As shown in (Figure 2), there is a method of extracting skeletal information with a Kinect Depth camera, etc. By mapping the extracted skeleton information to the RGB camera image, it is possible to find out where the user is currently located.

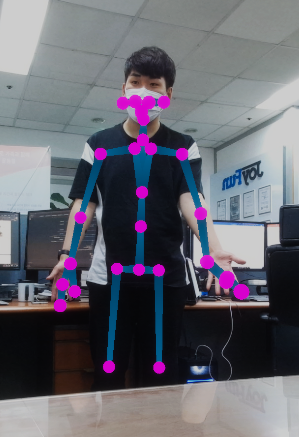


(Figure 2) Kinect Depth camera-based user skeleton information

[Source] https://docs.microsoft.com/en-us/azure/kinect-dk/media/quickstarts/samples-simple3dviewer.png

3.3 Mapping skeleton information to RGB cameras

By mapping the extracted skeleton information to the RGB camera image, it is possible to find out where the user is currently located. With this location, it is possible to specify in which part of the entire RGB Camera area the user is located, and to cut out and display a part of the area around the user (tracking target).



(Figure 3) RGB Camera user skeleton information mapping

3.4 Improved bouncing when tracking every frame in real time

When the user is tracked in real time and the display area is updated, a real-time user tracking system is completed. When tracking every frame in real time, the motion is not constant, so screen flickering occurs. To alleviate this problem, the user's central coordinates were set as the target point in real time and moved little by little toward the target point over time.

3.5 Example of real-time user tracking system using skeletal information

If you look at the example in use in (Fig. 4), you can see that the user moves to the right, starting from the far left. RGB Camera is in a fixed state, and when only the user starts from the left and moves gradually to the right, it can be seen that the image is displayed by tracking the user in a small fixed area.



(Figure 4) Real-time user tracking system screenshot