

## Research Article

# Development for Multisensor and Virtual Simulator–Based Automatic Broadcast Shooting System

Wonjun Lee, Hyung-Jun Lim, and Mun Sang Kim 

*School of Integrated Technology, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea*

Correspondence should be addressed to Mun Sang Kim; [munsang@gist.ac.kr](mailto:munsang@gist.ac.kr)

Received 20 April 2022; Revised 24 May 2022; Accepted 13 June 2022; Published 16 July 2022

Academic Editor: Marco Rocchetti

Copyright © 2022 Wonjun Lee et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

To solve the limitations of complexity and repeatability of existing broadcast filming systems, a new broadcast filming system was developed. In particular, for Korean music broadcasts, the shooting sequence is stage and lighting installation, rehearsal, lighting effect production, and main shooting; however, this sequence is complex and involves multiple people. We developed an automatic shooting system that can produce the same effect as the sequence with a minimum number of people as the era of un-tact has emerged because of COVID-19. The developed system comprises a simulator. After developing a stage using the simulator, during rehearsal, dancers' movements are acquired using UWB and two-dimensional (2D) LiDAR sensors. By inserting acquired movement data in the developed stage, a camera effect is produced using a virtual camera installed in the developed simulator. The camera effect comprises pan, tilt, and zoom, and a camera director creates lightning effects while evaluating the movements of virtual dancers on the virtual stage. In this study, four cameras were used, three of which were used for camera pan, tilt, and zoom control, and the fourth was used as a fixed camera for a full shot. Video shooting is performed according to the pan, tilt, and zoom values of the three cameras and switcher data. Only the video of dancers recorded during rehearsal and that produced by the lighting director via the existing broadcast filming process is overlapped in the developed simulator to assess lighting effects. The lighting director assesses the overlapping video and then corrects parts that require to be corrected or emphasized. The abovementioned method produced better lighting effects optimized for music and choreography compared to existing lighting effect production methods. Finally, the performance and lighting effects of the developed simulator and system were confirmed by shooting using K-pop using the pan, tilt, and zoom control plan, switcher sequence, and lighting effects of the selected camera.

## 1. Introduction

In today's COVID-19 pandemic era, un-tact cultures have emerged to prevent the spread of the virus rather than human-to-human contact [1]. Consequently, un-tact cultures such as telecommuting, video conferencing, and online ordering have become entrenched in our lives [2]. These un-tact cultures have significantly influenced cultural lives, particularly sports and performances. Performances such as concerts, musicals, and plays have been performed non-face-to-face via online broadcasting. In the case of sports, games are played without spectators, but online broadcasts are made possible via various Internet platforms [3, 4]. Online performances and broadcasts have a significant impact on people's emotional care, particularly those who

are tired of the pandemic restrictions [5]. However, with the un-tact era, online performances can be viewed and enjoyed while working at home, and then access to these performances has increased because they can be viewed at a lower cost than offline performances [6]. However, technical aspects such as sound and emotion that can be perceived in the field are inevitably reduced. In the case of a camera, e.g., the audience cannot experience stage feelings because they rely on the installed camera to watch the performance. Furthermore, although the ticket price has reduced, the number of workers mobilized for both offline and online performances is barely reduced, resulting in economic loss to producers [7–9]. To compensate for these limitations, many broadcasting stations or research teams are developing and using automatic camera shooting systems such as

camera system, calibration, pan, tilt, and zoom control methods [10–15]. In particular, for reality entertainment and sports broadcasts such as the Olympics, cameras are attached to robots or control assistance systems [16, 17] capable of pan, tilt, and zoom to automatically take videos, thus replacing camera persons [18–24]. However, this shooting technology simply tracks a person or controls a camera using a remote control device such as a joystick, and there is a limit to performing or broadcasting using only this technology [25, 26]. Furthermore, because the above system simply controls a camera to be used in a performance, it is limited in broadcasting tasks requiring various broadcasting systems such as a switcher, music, lighting, and other systems. To compensate for these limitations, we developed an automatic filming system that can produce the same shooting results as existing offline performances or broadcasting systems using a minimum number of people in a pandemic era. The developed system can control and produce everything from rehearsal to sound, shooting, and lighting. The developed system comprises four parts, as shown in Figure 1.

First, there is a stage calibration part that creates a virtual stage using a simulator based on an actual stage. Second, there is a movement acquisition unit that acquires the movement of actors on stage through rehearsal. Third, there is a simulator unit that uses the obtained movement data to create a camera operation plan using the simulator. Fourth, the video of dancers taken during rehearsal is overlapped with the video shot only with the lighting effect on an empty stage, and this overlapped video is used to assess the lighting effect. Finally, our system comprises an integration control part that transmits and controls simulator data produced above to actual hardware. The system's abovementioned parts were developed to be controlled in one integrated control system program using UNITY. The developed system imitates the existing Korean music broadcasting sequence. The general order of stage equipment setting, rehearsal, lighting effect production, and the main shooting was followed. The song used for verification was a song called Red Velvet's Red Flavor among K-Pop, and amateur dancers were used for choreography.

The whole system is configured as shown in Figure 2, and all hardware is controlled through the simulator implemented as UNITY in the Main PC. First, after installing the camera, UWB, 2D LiDAR, and lighting, set the camera position and stage through the calibration process. After the hardware setup is complete, rehearsal proceeds. When conducting rehearsals, UWB and 2D LiDAR are used to obtain the movement lines and IDs of the dancers. Using the obtained movement data, the user creates a plan for pan, tilt, and zoom of the camera. When the camera control plan is completed, set up a switcher plan for screen switching. In addition, the user plans the lighting effect while watching the video recorded during the rehearsal. When the lighting effect plan is completed, only the lighting effect is filmed again on the empty stage, and then the rehearsal video and the lighting video are overlapped. The user corrects the lighting effect at the timing when correction or emphasis is needed while rechecking the overlapped image. Finally,

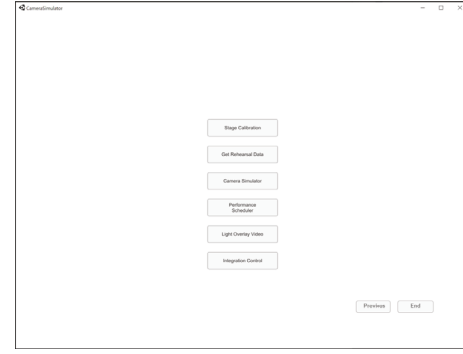


FIGURE 1: Composition of the developed integrated broadcasting control system to operate the existing shooting system in one program.

when the lighting effect plan is completed, the main shooting begins. The main PC controls not only the start signal but also all hardware necessary for the main shooting. When the main shooting starts, the 2D LiDAR checks the positions of the dancers and compensates if they are outside the set error range and controls the dancers to be located in the center of the screen as much as possible. The above contents are written in detail for each part below.

## 2. Materials and Methods

**2.1. Stage Calibration and Human Data.** In general, the first step in music broadcasting is to set up the stage. The location of the camera and lighting is determined as per the size and shape of the stage, and so stage installation is performed first. In this study, lighting, cameras, and the stage size and shape were determined by a general director. The auditorium in the school served as the stage, and only the size of the stage section to be used was specified. The shape of the stage to be used for filming was rectangular.

First, install UWB anchors in a square shape at the corners of four sides of the stage, as shown in Figure 3(a). Then, install the camera in the desired location. Then, install the two-dimensional (2D) LiDAR sensor in the center of the stage. In this study, four cameras were used, three of which were combined with the pan tilt unit (PTU) to control the pan-tilt-zoom configuration. Because the remaining camera is a fixed camera for full shot, and no PTU is installed. After installing the cameras and 2D LiDAR sensor, as shown in Figure 4(a), measure the size of the stage using a laser range finder. After the stage measurement, the distance and angle from each vertex of the stage to all cameras and the 2D LiDAR sensor are measured and input into the simulator to simulate the virtual stage and cameras, as shown in Figure 3(b). The camera position was calculated using a simple trigonometric function and Euclidean distance equation using the stage coordinates and the angle between the stage and the cameras [27, 28].

After the completion of the camera and sensor installation, as well as the virtual stage and camera installation in the simulator are completed, dancing rehearsal is performed to acquire the dancers' movement. Two sensors were used to acquire the dancers' movement. The sensors used were a

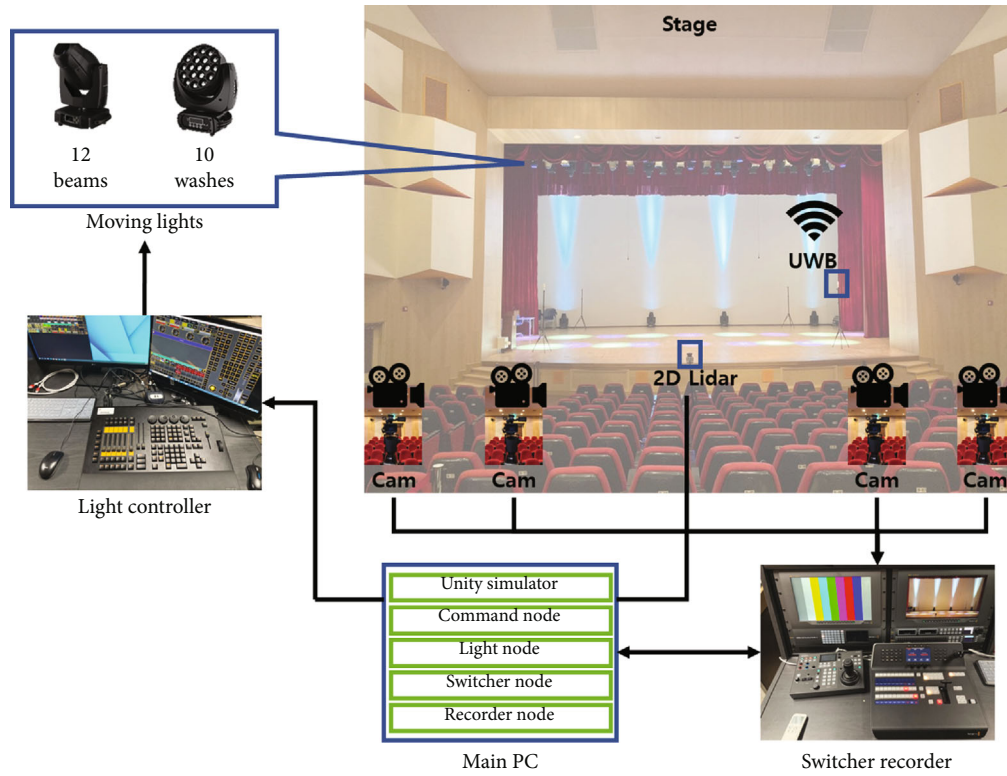


FIGURE 2: Overall system configuration and connection diagram.

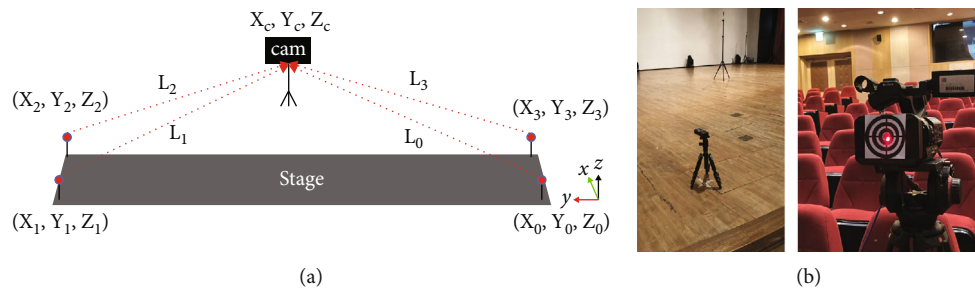


FIGURE 3: (a) Coordinate system diagram of laser measuring instrument and camera for calibration, (b) measuring distance and angle of camera using laser distance and angle meter. Installation of a laser meter on stage (left) and a laser meter measuring the distance and angle of camera (right).

UWB sensor and a 2D LiDAR sensor. For each sensor, Decawave's MDEK1001 Development Kit and SICK's LMS101-10000 model were used. The first sensor is a UWB sensor for identifying actors. Because each UWB sensor has a unique ID, it is optimal for identifying multiple actors on a stage. Furthermore, although the UWB sensor provides location data, such data are not used for obtaining the movement line in this study because the movement error is 10 cm [29, 30]. However, if there is a loss of position data obtained using the 2D LiDAR sensor, the location data were used for correction. Due to the limitations of UWB error, many studies are being conducted on the fusion of UWB and Lidar sensors [31–33]. A UWB sensor is attached to the shoulder of the vest; during rehearsal, actors wear the sensor-attached vest (Figure 4(b)) for identification while performing on stage. As the second sensor, a 2D LiDAR sen-

sor is used to acquire the movement of actors [34, 35]. A 2D LiDAR sensor is installed at the center of the stage to acquire the movement of actors on the stage. The 2D LiDAR sensor was installed 1 m above the stage floor such that actors could be scanned from the pelvis to the waist. Furthermore, when actors overlap back and forth or disappear from the scanning area, making it difficult to acquire actor movement data, the location data obtained using the UWB sensor are used to compensate for the actor movement data loss. Moreover, for position data compensation, the coordinate system of the 2D LiDAR sensor was synchronized with that of the UWB sensor. After acquiring the actor movement data, data filtering was performed using the curve fitting method, which is a data post-processing method. The final acquired movement data are input to the simulator and loaded on the already created virtual stage.





(a)



(b)

FIGURE 4: Hardware settings used in this study, (a) A PTU for controlling pan and tilt and a 4K camera for video recording. An environment in which PTU+camera and 2D LiDAR sensor are installed in front of the stage to acquire movement data and correct movement errors of actors, (b) 22 lights and UWB anchor for identification installed on the stage. A UWB tag in the form of a vest.

**2.2. Simulator.** When the movement data of actors acquired via the virtual stage and rehearsal are loaded in the simulator, the editor or director creates a camera plan based on their experience. The camera plan comprises two parts. The first part is to control the camera's pan, tilt, and zoom configurations, and the second creates a switcher plan that selects the camera from which the video is transmitted among multiple cameras.

The camera plan generation part creates a camera plan by changing the pan, tilt, and magnification configurations of the camera as desired while watching the rehearsal video and the movements of actors on the virtual stage for each camera in the simulator. The camera plan was developed such that keyboard and joystick users can use our developed system conveniently. Because the movement data acquired during rehearsal is 2D data, it is difficult to create a tilt plan to create a camera plan by only simulation. To compensate for these limitations, as shown in Figure 5, the actual

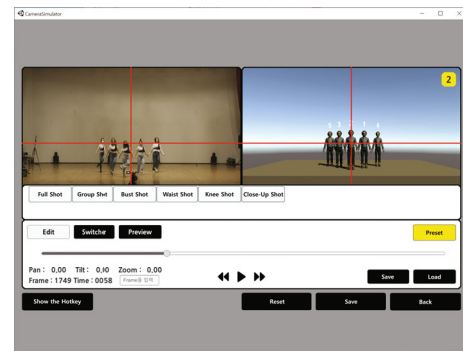


FIGURE 5: Camera and switcher plan authoring simulator to which rehearsal video and acquired movement data were applied on the virtual stage.

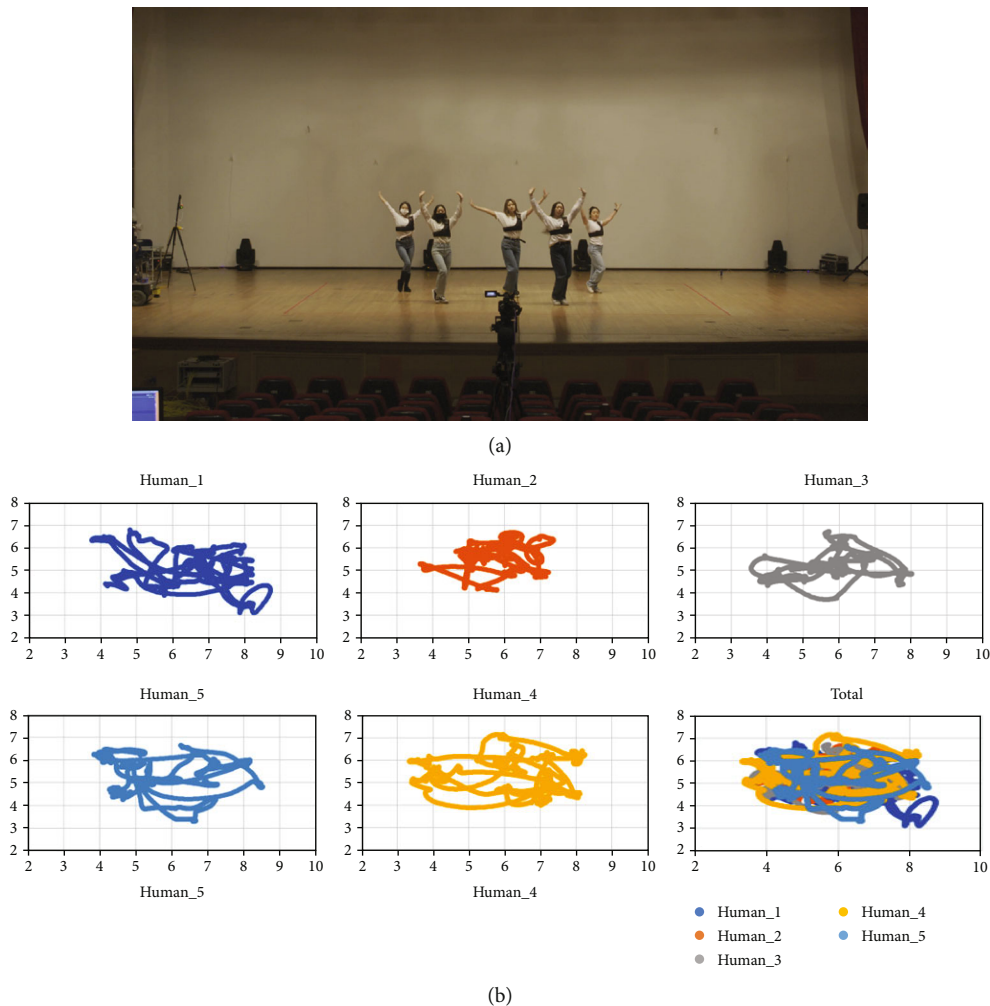


FIGURE 6: (a) Five dancers whose movement data are obtained during rehearsal. (b) Final obtained movement data of each dancer.

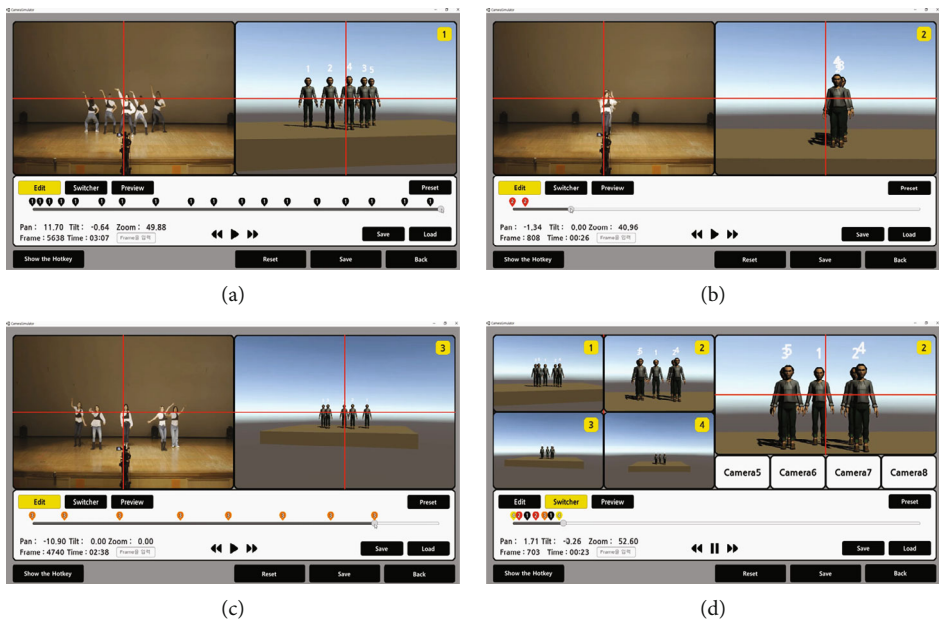


FIGURE 7: (a)–(c) A scene producing the pan, tilt, and magnification configuration of cameras No. 1–3 using the simulator. (d) A view of switching production, as shown in four cameras at the upper left corner.

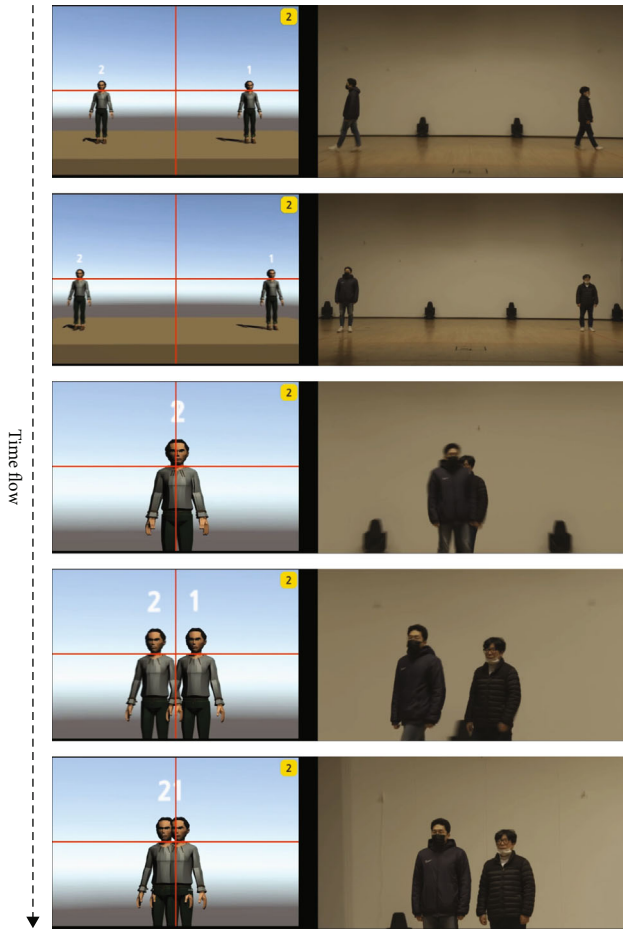


FIGURE 8: Magnification value that changes as per the distance between two people and the tracking result as per the movement of two people (left: simulator, right: actual video).

rehearsal video was played in sync with the simulator stage. While watching the actual rehearsal video, the director will be able to assess when to tilt the camera and create a tilt plan. A preset function was developed for the convenience of the producer's camera plan. Preset functions include close-up, bust shot, waist shot, knee shot, and full shot that track one person. Each shot tracks the actor selected by the creator, and the camera's pan, tilt, and zoom can be changed as per the type of shot. In particular, if two actors are selected to shoot through the group shot preset, a scene can be created in which the camera's pan, tilt, and zoom change as per the movement of the two actors. After creating each camera's pan, tilt, and zoom plans, a camera switcher plan for video transmission is finally created. The producer selects the scene in which the video will be finally transmitted in units of time or frame while watching the camera plan simulation of four cameras. The producer can switch cameras at any time or frame, and it is possible to modify the switcher plan. Finally, the pan, tilt, zoom, and switcher values of the planned cameras are transmitted to the Unity program. The produced data is loaded from the Unity-based broadcasting control system and prepared to control the camera, PTU, lighting control panel, switcher (Blackmagic ATEM 1 M/E Advanced Panel and Blackmagic ATEM

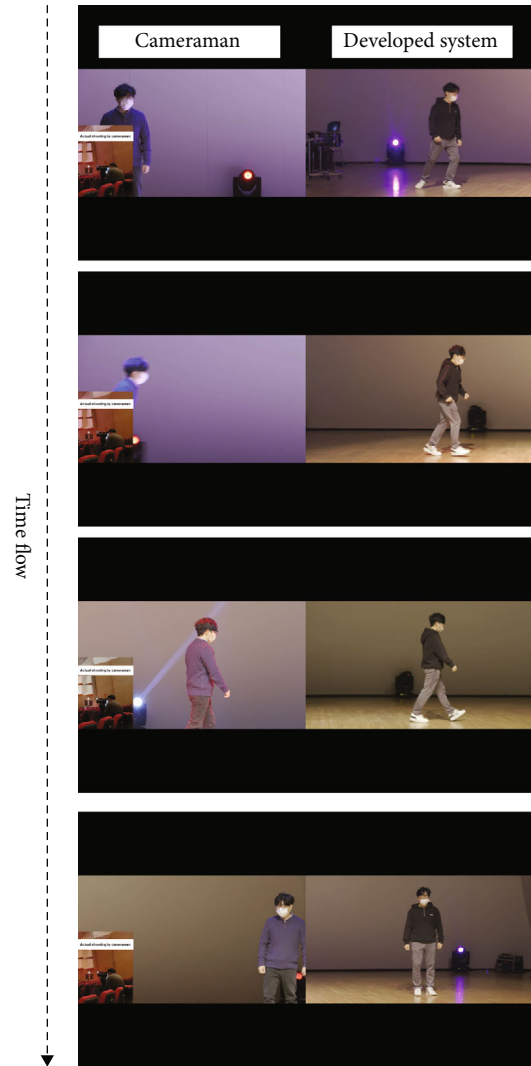


FIGURE 9: Comparison of video shot by an actual camera operator and that using the developed system.

1 M/E Production Studio 4k), and recording equipment (Blackmagic Duplicator 4k and Blackmagic HyperDeck Studio 4k Pro) at the same time.

**2.3. Camera Control.** The pan, tilt, and magnification values of each camera produced using the simulator are transmitted to the camera for PTU and magnification control for each pan and tilt control as per the command of the primary program. FLIR's D-100E product was used for PTU, and Sony's Z-280 was used for the cameras. PTU was fixed and seated on a tripod, and the camera was fixedly installed on PTU. The movement of PTU is based on the camera's pan and tilt values obtained using the simulator. However, the above-mentioned values are the movement data obtained through the existing rehearsal. Therefore, when the primary performance is filmed using the abovementioned data, an error occurs between the obtained values and actual filmed image unless the actor behaves exactly as in the rehearsal. To solve the abovementioned limitation, actor position data obtained using the 2D LiDAR sensor were used for compensation.

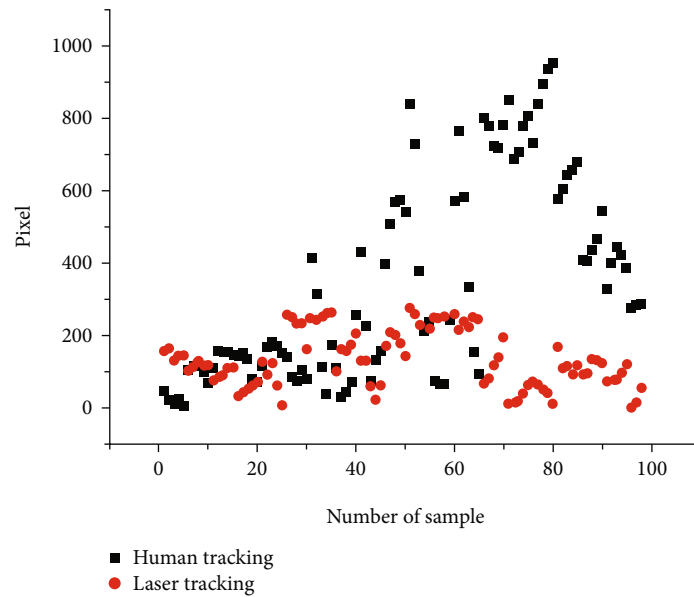


FIGURE 10: A graph comparing videos taken by camera operator and developed system.

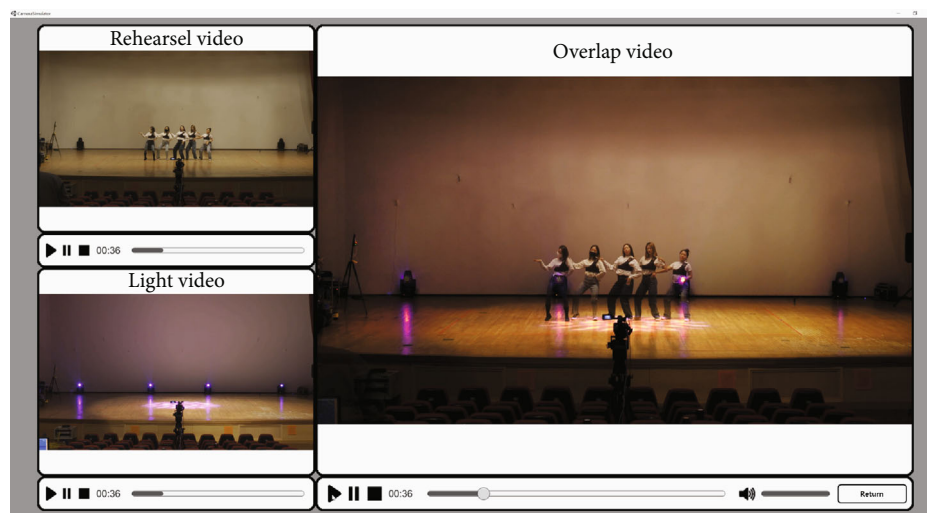


FIGURE 11: Simulator technology that can evaluate lighting effects by overlapping rehearsal and simulated videos.

When an error occurs by comparing the already acquired movement data with the current 2D LiDAR sensor data, the current 2D LiDAR sensor data are used to minimize the error of the captured image.

**2.4. Light Control.** In existing music broadcasting programs, the lighting director produces lighting effects by listening to music, and the lighting effects produced are based on his or her experience. Because this method creates lighting effects only through the director's experience and listening to music, these effects do not consider the movement of actors. To overcome this limitation, the lighting production method of the developed system first produces lighting effects while watching the video filmed during rehearsal. Second, the lighting effect produced is played on an empty stage and video. Third, lighting effects with dancers are simultaneously

assessed by overlapping the rehearsal and lighting videos in the simulator. After the assessment, the lighting director can repeatedly modify the parts that require to be corrected such that lighting effects can be created in more detail than with the existing lighting production method.

### 3. Results and Discussion

**3.1. Stage Calibration and Human Data.** The calculated stage and camera positions were compared with the actual values, which were measured using a laser angle and distance measuring device. The position of the camera from the origin of the stage was measured as X, Y, and Z coordinate values using a laser measuring device. The comparison results demonstrate that only an average error of 3–5 cm occurred; thus, the abovementioned value is not unsuitable for imaging. Five





FIGURE 12: A captured image of a lighting effect comparison video modified using the existing lighting effect and overlapping technology.

dancers danced to the music of Red Velvet's Red Flavor, one of K-pop, and the dancers' movement data were acquired. The music is 3 min and 25 s long, and while dancing to the music, the five dancers wore UWB-attached vests, and their identification and movement lines were obtained using the 2D LiDAR sensor. The movement data comprise an ID and X and Y coordinates for each dancer and are stored at a cycle of 30 Hz. The stored movement data are filtered using a data post-processing method, i.e., curve fitting, to realize cleaner movement data, as shown in Figure 6.

**3.2. Simulator.** The movement data acquired during the rehearsal using the UWB and 2D LiDAR sensors and the calculated positions of the stage and cameras were transferred to the virtual stage of the developed simulator. In this study, four cameras were used, three of which were used to control the camera pan, tilt, and magnification configurations, and the fourth was installed farthest from the center of the stage and used as a fixed camera for the full shot. Directly in the simulator, the director plans by watching the characters on the simulator move the pan, tilt, and zoom of each camera in chronological order in the simulator as

shown in Figures 7(a)–7(c). Because the camera used for the full shot does not control the camera pan and tilt, the director sets the magnification value to the desired view by setting only the magnification.

When the pan, tilt, and magnification plans of all cameras are completed, the last switcher task is performed. The switcher task determines which four cameras to use as final output images over time, as shown in Figure 7(d). Finally, when the switcher task is completed, data, including the pan, tilt, and magnification values for each camera and the switcher values, are extracted, and these data are then used for camera control. The video in Figure S1 shows the camera top view of three cameras, the plan of the switcher, and the actual comparison video taken on the basis of the simulator and the simulator. Furthermore, in the simulator, a preset function was developed such that a user could easily move the cameras. The developed preset functions comprise close-up, bust shot, waist shot, knee shot, full shot tracking one person, and group shot tracking group. In particular, the group shot is a preset function created to emphasize the advantages of the developed system. It was developed to zoom in or out in response to the movements of the two



TABLE 1: Questionnaire for user evaluation of simulators and automatic shooting systems.

The following is a questionnaire for user evaluation of “simulator and automatic shooting system”	Strongly disagree	Very disagree	Disagree	Neutral	Agree	Very agree	Strongly agree
The simulator helped create the performance	1	2	3	4	5	6	7
The information provided by the simulator is accurate and reliable	1	2	3	4	5	6	7
After using the simulator was more helpful in writing the camera plan than before	1	2	3	4	5	6	7
The virtual environment provided by the simulator did not have a sense of heterogeneity in use	1	2	3	4	5	6	7
If the automatic shooting system and simulator are activated, it will be helpful for stage authoring	1	2	3	4	5	6	7
The movements of the dummy in the simulator were sufficient to grasp and utilize the stage contents	1	2	3	4	5	6	7
The simulator seems to be easy for even first-time users	1	2	3	4	5	6	7
The camera control plan using the simulator gave satisfaction	1	2	3	4	5	6	7
Simulator and automatic shooting system can be recommended to people	1	2	3	4	5	6	7
I am willing to use the simulator and automatic shooting system again	1	2	3	4	5	6	7
The simulator provided an experience that could replace the performance of an actual performer	1	2	3	4	5	6	7
The video made with the simulator was natural, and there was no sense of difference with the video taken by a real person	1	2	3	4	5	6	7
The shooting method using the simulator is more convenient than the existing shooting method	1	2	3	4	5	6	7
The simulator UI is convenient to use	1	2	3	4	5	6	7
The simulator UI responded quickly	1	2	3	4	5	6	7
The simulator UI worked as intended	1	2	3	4	5	6	7
Other comments							
Affiliation	Career		Responsibilities				

selected dancers and to enable tracking. As shown in Figure 8, the magnification value that changes as the distance between two people changes can be evaluated using the group preset function. Moreover, it was developed such that the camera automatically tracks the movement of two people. The video in Figure S2 shows the simulator work scene using presets, the simulation video made using group shots among the presets, and the comparison video actually shot.

As future work, we plan to add more various presets and camera shooting techniques by reflecting the experiences of camera directors. In addition, we think that users should consider the system environment and interface, etc., that are easier and more immersive for users to create camera plans and stages using the simulator. In particular, if the diegetic design, which has been widely used in VR or games, is used, it is expected that it will become a more realistic simulator because users can give the effect of composing a camera plan and stage in an actual studio [36, 37].

**3.3. Camera Control.** The pan-tilt-magnification value of each camera according to time was created in the simulator operates as per the start signal of the integrated control system. The pan-tilt-magnification configuration of the camera is controlled on the basis of the pan-tilt-magnification data

created in the simulator. However, the position of an actor on stage during rehearsal and that in the actual performance cannot exactly match. Therefore, an error is bound to occur in the image taken using only the simulator data. To solve this limitation, in this study, the position error of the actor was corrected using 2D LiDAR sensor data during this performance. All cameras were calibrated to be positioned at the center of the actor using data acquired using the 2D LiDAR sensor. To confirm the result of the correction, the movement data in the left and right directions on the stage were obtained. One camera was made to follow the person in the simulator using the acquired movement data. The person was tracked using the camera plan produced in the simulator and 2D LiDAR sensor correction. Furthermore, the performance of a camera operator with more than five years of experience was compared with the existing system using the same camera to capture the same movement at the same location, as shown in Figure 9. In Figure S3, in more detail, the camera operator and the developed system show a comparison video recorded by tracking a person moving on the same course.

For quantitative comparison of the tracking degree of the captured videos, it was determined whether the human was well positioned in the center of the screen. As an evaluation

TABLE 2: Questionnaire for user evaluation of light overlay.

The following is a questionnaire for user evaluation of "light overlay system"	Strongly disagree	Very disagree	Disagree	Neutral	Agree	Very agree	Strongly agree
The use of video overlays was helpful for editing and supplementing lighting authoring	1	2	3	4	5	6	7
The information provided by the video overlay is accurate and reliable	1	2	3	4	5	6	7
Using video overlays was helpful for lighting authoring	1	2	3	4	5	6	7
Using video overlays allowed me to focus on lighting authoring	1	2	3	4	5	6	7
If video overlays are enabled, it will be helpful for lighting authoring	1	2	3	4	5	6	7
The video provided by the video overlay was sufficient to understand and utilize the content of the provided video	1	2	3	4	5	6	7
Video overlay seems to be easy even for first-time users	1	2	3	4	5	6	7
Video overlays gave me satisfaction with lighting authoring	1	2	3	4	5	6	7
You can recommend video overlays to people	1	2	3	4	5	6	7
I am willing to use video overlay again	1	2	3	4	5	6	7
The video provided as a video overlay provided an experience of seeing the performance of real performers	1	2	3	4	5	6	7
The video provided by the overlay was natural, and there was no sense of heterogeneity	1	2	3	4	5	6	7
The video overlay image provided an effect and environment similar to the real stage	1	2	3	4	5	6	7
The video overlay UI is convenient to use	1	2	3	4	5	6	7
The video overlay UI responded quickly	1	2	3	4	5	6	7
The video overlay UI worked as intended	1	2	3	4	5	6	7
Other comments							
Affiliation	Career		Responsibilities				

method, the entire image was captured one per 100 frames to determine how far the center of the body deviated from the center of the screen. From the evaluation results, as shown in Figure 10, the tracking performance of the developed system outperformed the camera operator by 36%. Pixel is defined as the distance between the center of a person's body and the center pixel of the person's image. Even after evaluating the entire video, the video obtained by the camera operator frequently appeared to deviate from the center of the screen and sometimes even disappeared from the screen. However, in the developed system, a person is stably centered on the screen, and there is no case where a person disappears from the screen.

**3.4. Light Control.** In the existing music broadcasting lighting effect production, only the lighting director creates the lighting effect based on his/her experience and perception while listening to music. Because this lighting effect does not consider the movement and choreography of singers or actors, it is insignificant. In particular, on a solo stage or when it should be emphasized, it is difficult to create a lighting effect by simply listening to music. To solve this limitation, a lighting effect simulator was developed. The developed lighting effect simulator first records videos of dancers during rehearsal. Second, while watching and listening to the videos and music of the dancers filmed dur-

ing the rehearsal, lighting effects are created. Only after the lighting director has created the lighting effects are they recorded in the same manner. At this time, the camera used during the rehearsal should have the same position, angle, and magnification values as the camera used during the rehearsal. Only the video of the recorded dancers and the lighting effect are used to overlap the two videos using the overlapping technology of the developed simulator, as shown in Figure 11.

The lighting director watches the overlapping video to ensure that the dancers, music, and lighting are in good harmony. Then, the lighting effect in the areas that require correction or emphasis is corrected. The modified lighting effect is re-shot and evaluated by overlapping with the rehearsal videos of the existing dancers. The abovementioned series of processes can be repeated as per the requirements of the lighting director.

If the lighting effect is corrected using the lighting overlapping technology, many effects can be obtained, as shown in Figure 12. Because the positions of the dancers are not known, the lighting effect was produced broadly as shown in the pictures before the overlap of Figures 12(a) to 12(d). However, since the newly created lighting effect after the overlap can know the positions of the dancers, it can be confirmed that more appropriate lighting effects are produced by focusing the lighting on the positions where the dancers

are standing. As described above, unlike the existing lighting effect method that relies solely on music, lighting effects are produced while watching the positions and choreography of dancers using the overlapping technology, enabling the production of more lively and colorful lighting effects. The video in Figure S4 shows a comparison video before and after overlapping.

**3.5. User Evaluation.** When the performance shooting was tested, 5 camera directors and 5 lighting directors were invited. Each director is a director with at least 5 to 12 years of experience in each field. All the directors watched all the filming in the order set in the simulator, actually dealt with the simulator, and conducted the survey below (Tables 1 and 2). As for the survey, two surveys were prepared for camera supervision and lighting supervision.

Each of the two surveys consisted of 16 items, with a total score of 112 points. The camera directors did an overall evaluation of the simulator and automatic shooting system, and the lighting director did an overall evaluation of the lighting overlay. First of all, the evaluation result of the five camera directors recorded an average of 101.8 points out of 112 points. Overall, I gave high marks to the UI and system configuration. However, they felt that the image of the virtual stage consisting of a dancer and a rectangle expressed on the virtual stage gave a sense of different from reality in making camera plans, so most of them gave lower scores than other items. As an additional opinion, it was said that the camera's pan, tilt, and zoom were in sync with the simulator and the actual camera, so it was convenient to write a camera plan. In addition, it was evaluated that it would be more effective not only for performances but also for areas with a fixed order such as news or home shopping.

The lighting director gave an average of 108.2 points out of 112 points. Overall, I was very satisfied with the lighting overlay technology. As an additional opinion, the existing lighting effect production process only listens to music, and the lighting director imagines and makes the lighting effect. However, in the case of production using the lighting overlay technology, it was evaluated that the work was easy and appropriate effects could be produced because lighting effects were made while viewing the video acquired through rehearsal. In addition, it was evaluated as an advantage to be able to experience the actual stage in advance using the lighting overlay effect.

## 4. Conclusions

A new system was developed to compensate for the limitations of manual methods of rehearsal video recording, camera control, switcher control, and lighting effect production for broadcast shooting. The video in Figure S5 was taken using the finally developed system. When shooting using this developed system, shooting errors can be reduced by up to one-fifth compared to the existing workforce while maintaining or improving video quality. The developed system is based on the production order of music broadcasting in Korea. The developed system handles the

pan, tilt, and magnification control plan of cameras and the switcher control plan for screen switching using the virtual stage based on the working table in the developed simulator after acquiring the movements of dancers during rehearsal. The abovementioned production process iteratively allows a lighting director to easily add and modify an optimal camera and switcher control plan. For lighting effect production, only the rehearsal video of dancers recorded during rehearsal and the video produced in advance by the lighting director were overlapped. The lighting director corrects the parts that require to be corrected through the reconfirmation procedure of the overlapped video, and only the lighting video is re-recorded and overlapped with the existing rehearsal video to evaluate the modified lighting effect. The abovementioned operations can be easily repeated as desired by the lighting director to produce the best lighting effects for the music. Based on the authored data, when shooting this performance, all cameras and switcher lighting are automatically directed and filmed. Because the control plan of all cameras is created based on the rehearsal during automatic video shooting, the positions of dancers may change during the primary performance. When this occurs, 2D LiDAR sensor position data are used for real-time correction to place the actors in the center of the current camera screen. Finally, the performance of the developed system was confirmed by five amateur dancers performing the song "Red Flavor" by Red Velvet among K-pop. In the future, we plan to test and use the developed system in various fields such as news or talk shows; we will add a preset function to acquire high-quality videos. Moreover, we are attempting to overcome several shortcomings limiting online performances by increasing the number of cameras of various types to convey as much realism of the stage as possible.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research is supported by the Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2020, and this work was supported by the GIST Research Project grant funded by the GIST in 2022.

## Supplementary Materials

S1. See Supplemental Material Video 1 for further explanation of Figure 7. S2. See Supplemental Material Video 2 for further explanation of Figure 8. S3. See Supplemental Material Video 3 for further explanation of Figures 9 and 10. S4. See Supplemental Material Video 4 for further explanation of Figure 12. S5. See Supplemental Material Video 5 for

further explanation of developed system. (*Supplementary Materials*)

## References

- [1] S. Y. Bae and P.-J. Chang, "The effect of coronavirus disease-19 (COVID-19) risk perception on behavioural intention towards 'untact' tourism in South Korea during the first wave of the pandemic (March 2020)," *Current Issues in Tourism*, vol. 24, no. 7, pp. 1017–1035, 2021.
- [2] R. Shahabi Sorman Abadi, Z. Ghaderi, C. M. Hall, M. Soltaninasab, and A. Hossein Qezelbash, "COVID-19 and the travel behavior of xenophobic tourists," *Journal of Policy Research in Tourism, Leisure and Events*, pp. 1–23, 2021.
- [3] Y.-H. Kwon, Y.-K. Cheung, and B.-W. Ahn, "Effect of outdoor sports participants on leisure facilitation, recreation specialization, and leisure satisfaction: yacht and golf participants," *International Journal of Environmental Research and Public Health*, vol. 18, no. 15, article 8128, 2021.
- [4] Y. Fan, S. Ouyang, B. Zhang, and W. Wang, "The Sustainability of Online Concert and Live Concert," *E3S Web of Conferences*, vol. 308, 2021.
- [5] X. Xie, Z. Zang, and J. M. Ponzoa, "The information impact of network media, the psychological reaction to the COVID-19 pandemic, and online knowledge acquisition: evidence from Chinese college students," *Journal of Innovation & Knowledge*, vol. 5, no. 4, pp. 297–305, 2020.
- [6] F. Solihah, "Fans' satisfaction on watching virtual concert during COVID-19 pandemic," in *3rd Jogjakarta Communication Conference (JCC 2021)*, pp. 154–156, Atlantis Press, Netherlands, 2021.
- [7] O. M. Hylland, "Tales of temporary disruption: digital adaptations in the first 100 days of the cultural Covid lockdown," *Poetics*, vol. 90, article 101602, 2022.
- [8] S. Bilici and A. Yildirim, "Economic reflections of the transformation of social distance concept in everyday life after COVID-19 to virtual contact," *Kafkas Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, vol. 12, no. 23, pp. 554–591, 2021.
- [9] B. Gilbert, "Online Learning Revealing the Benefits and Challenges," 2015.
- [10] S. Zhang, Z. Wei, J. Nie, L. Huang, S. Wang, and Z. Li, "A review on human activity recognition using vision-based method," *Journal of Healthcare Engineering*, vol. 2017, Article ID 3090343, 31 pages, 2017.
- [11] L. W. He, M. F. Cohen, and D. H. Salesin, "The virtual cinematographer: a paradigm for automatic real-time camera control and directing," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 217–224, United States, August 1996.
- [12] J. Yang, Z. Tang, Z. Pei, and X. Song, "A novel motion-intelligence-based control algorithm for object tracking by controlling PAN-tilt automatically," *Mathematical Problems in Engineering*, vol. 2019, Article ID 9602460, 11 pages, 2019.
- [13] M. Kano, H. Okubo, M. Takahashi, K. Ikeya, K. Hisatomi, and T. Mishina, "Accurate and practical calibration of multiple pan-tilt-zoom cameras for live broadcasts," *IEEE Access*, vol. 8, pp. 153993–154006, 2020.
- [14] M. Pollefeys, "PTZ camera network calibration from moving people in sports broadcasts," in *2012 IEEE Workshop on the Applications of Computer Vision (WACV)*, pp. 25–32, Breckenridge, CO, USA, January 2012.
- [15] A. Girbau Xalabarder, "Sports Broadcasting and Multiple Object Tracking with Deep Learning Methods," 2021.
- [16] H. Wang, R. Li, Y. Gao, C. Cao, L. Ge, and X. Xie, "Target recognition and localization of mobile robot with monocular PTZ camera," *Journal of Robotics*, vol. 2019, Article ID 8789725, 12 pages, 2019.
- [17] P. Deshpande, O. Gaikwad, M. Jain, K. Choksi, P. V. Kasambe, and S. S. Rathod, "Automating Camera Rover for Sports Broadcast," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1–7, Kharagpur, India, July 2021.
- [18] H. P. Chou, J. M. Wang, C. S. Fuh, S. C. Lin, and S. W. Chen, "Automated Lecture Recording System," in *2010 international conference on system science and engineering*, pp. 167–172, United States, July 2010.
- [19] T. Tsuda, M. Okuda, K. Mutou, and Y. Nishida, "Automatic Tracking Camera System Utilizing the Position of Faces in the Shot Image," in *2006 9th International Conference on Control, Automation, Robotics and Vision*, pp. 1–6, Singapore, December 2006.
- [20] Y. Ariki, S. Kubota, and M. Kumano, "Automatic Production System of Soccer Sports Video by Digital Camera Work Based on Situation Recognition," in *Eighth IEEE International Symposium on Multimedia*, pp. 851–860, San Diego, CA, USA, December 2006.
- [21] Z. Byers, M. Dixon, W. D. Smart, and C. M. Grimm, "Say cheese! Experiences with a robot photographer," *AI Magazine*, vol. 25, no. 3, pp. 37–37, 2004.
- [22] P. Carr, M. Mistry, and I. Matthews, "Hybrid robotic/virtual pan-tilt-zoom cameras for autonomous event recording," in *Proceedings of the 21st ACM international conference on Multimedia*, pp. 193–202, United States, October 2013.
- [23] F. Chen and C. De Vleeschouwer, "Autonomous Production of Basketball Videos from Multi-Sensored Data with Personalized Viewpoints," in *2009 10th Workshop on Image Analysis for Multimedia Interactive Services*, pp. 81–84, London, UK, May 2009.
- [24] S. Daigo and S. Ozawa, "Automatic pan control system for broadcasting ball games based on audience's face direction," in *Proceedings of the 12th annual ACM international conference on Multimedia*, pp. 444–447, United States, October 2004.
- [25] K. Yachi, T. Wada, and T. Matsuyama, "Human Head Tracking Using Adaptive Appearance Models with a Fixed-Viewpoint Pan-Tilt-Zoom Camera," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 150–155, Grenoble, France, March 2000.
- [26] U. Khair, Y. D. Lestari, A. Perdana, and A. Budiman, "CCTV camera controller simulation using Arduino Uno and joystick," *Journal of Physics: Conference Series*, vol. 1361, article 012086, 2019.
- [27] S. Xu and K. Doğançay, "Optimal Sensor Deployment for 3D AOA Target Localization," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2544–2548, South Brisbane, QLD, Australia, April 2015.
- [28] J. Fonseca, A. Baptista, M. Martins, and J. P. N. Torres, "Distance measurement systems using lasers and their applications," *Applied Physics Research*, vol. 9, no. 4, pp. 33–43, 2017.
- [29] I. Guvenç, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *2007*



- IEEE Wireless Communications and Networking Conference*, pp. 1571–1576, Hong Kong, China, March 2007.
- [30] L. Taponneco, A. A. D'Amico, and U. Mengali, "Joint TOA and AOA estimation for UWB localization applications," *IEEE Transactions on Wireless Communications*, vol. 10, no. 7, pp. 2207–2217, 2011.
- [31] J. Feng, L. Wang, J. Li, Y. Xu, S. Bi, and T. Shen, "Novel LiDAR-Assisted UWB Positioning Compensation for Indoor Robot Localization," in *2021 International Conference on Advanced Mechatronic Systems (ICAMEchS)*, pp. 215–219, Tokyo, Japan, December 2021.
- [32] C. Zhang, X. Ma, and P. Qin, "LiDAR-IMU-UWB-based collaborative localization," *World Electric Vehicle Journal*, vol. 13, no. 2, p. 32, 2022.
- [33] Z. Chen, A. Xu, X. Sui et al., "Improved-UWB/LiDAR-SLAM tightly coupled positioning system with NLOS identification using a LiDAR point cloud in GNSS-denied environments," *Remote Sensing*, vol. 14, no. 6, p. 1380, 2022.
- [34] Á. M. Guerrero-Higueras, C. Álvarez-Aparicio, M. C. Calvo Olivera et al., "Tracking people in a mobile robot from 2d lidar scans using full convolutional neural networks for security in cluttered environments," *Frontiers in Neurorobotics*, vol. 12, article 85, 2019.
- [35] W. Wang, X. Chang, J. Yang, and G. Xu, "LiDAR-based dense pedestrian detection and tracking," *Applied Sciences*, vol. 12, no. 4, article 1799, 2022.
- [36] P. Salomoni, C. Prandi, M. Roccetti, L. Casanova, L. Marchetti, and G. Marfia, "Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift," *Journal on Multimodal User Interfaces*, vol. 11, no. 2, pp. 173–184, 2017.
- [37] Q. Marre, L. Caroux, and J.-C. Sakdavong, "Video game interfaces and diegesis: the impact on experts and novices' performance and experience in virtual reality," *International Journal of Human-Computer Interaction*, vol. 37, no. 12, pp. 1089–1103, 2021.