Augmented Reality Game Interface Using Effective Face Detection Algorithm

Young Jae Lee, Geun Ho Lee
Dept. of Multimedia, Jeonju University
#45, Backma-Gil,Wansan-Gu,Jeonju,Jeonbul,560-759,Korea
leeyj@jj.ac.kr , ghlee@jj.ac.kr

Abstract

Viola and Jones Algorithmand Viola and Jones, the conventional face-detecting methods, have the advantages in that they could help fast and clear face recognition. But they also have their own drawbacks. With Viola and Jones, detection of slanted face is hard, and Viola and Jonescannot provide the precise recognition of faces since the detected area in faces is rather wide depending on the environment. To come with some kinds of solutions to these limitations, this paper propose the algorithm, which considers the position and size of the circle(the range of face recognition), detection intervals, and etc. The algorithm has been proven to be effective through a series of experiments. 3D AR game has been produced to utilize the algorithm as game interface in AR environment. Experiments in various game situations proved its effectiveness. Compared with the conventional methods, the algorithm was better in getting the desired information through face recognition, thus realizing interesting game interface. The algorithm could be used as the basis for making natural interaction with virtual object in augmented reality and also for developing interface in the smart environment.

Keywords: *Augmented Reality Game, Interface, face detecting, interaction*

1. Introduction

"Augmented Reality", which shows information added to the real world in real time, is rapidly advancing with the ever-developing IT technology. The technology of augmented reality is expected to generate add values in terms of increased convenience, experience, empathy, safety, and efficiency when connected to various industries. There is an increasing demand for convenience of various gadgets and information services in various spaces such as home, office and public areas. In this context, the mobile or game industry is best catching up with the current trend thanks to the related infrastructure established for the required hardware, software and wireless communication. Recently, a number of augmented reality solutions, Wikitude of Mobility, Layar and Sekai Camera, have appeared which provide information on visual targets by using smart phone.

There have also been efforts to improve sense of reality and immersion by using augmented reality in the fields of education and game. The augmented reality is most utilized in games to provide increased sense of reality and immersion which cannot be adequately provided online or through virtual space within the computer. The most common method is to use the real space reflected on camera lens as game space. Sony Entertainment has recently introduced Eye-Pet which uses the real space on camera lens as game space.[1]

Interface technology is the key to product's competitiveness in the areas where this kind of augmented reality is used. Human friendly, tangible user interfaces are leading the markets of smart phone and game as shown in Apple's I-phone, which uses multi-touch interface, Microsoft Kinect and SCE(Sony Computer Entertainment) Move, which use tangible interface based on recognition of movements. That is to say, intuitive and innovative interface is being welcomed in the market and more importantly, tangible interface with emotions is in the limelight.[2]

This paper proposes the new face-detecting method based on computer vision technology, which is applied for game interface to be proved for its efficacy. For the purpose of proving its effect, 3D game based on augmented reality has been developed by using webcam and Irrlicht. Its validity is tested by applying face-detecting algorithm to the game. Additionally, its interface with the virtual objects in the AR space is tested in various situations.

2. AR Game

AR game has the characteristic of using real space reflected on webcam as game space to provide added sense of reality and immersion which cannot be adequately provided only through virtual space in the computer. More excitement and immersion is provided by using human friendly tangible interface. Most of the games developed for PC use very limited hardware such as keyboard, mouse and joystick, which explains lessened sense of reality and limited utilization of space. Other interfaces using Haptic glove and HMD, also have limitations. Haptic glove, goggles and HMD gadgets are inconvenient in that they have to be worn on user's body. They are also hampering the movements of the users because of wires and their purchase decisions are not easy since they are very expensive.[3] If we could control the game with mere human movements, without keyboard manipulation or the need for attachment of the sensors, we could expect increased sense of reality and new human friendly interface.[4] The important factor in tangible 3D game based on augmented reality is the interaction between gamer's movements and virtual objects. In other words, how to effectively realize the events such as collision is the key. That's because it could determine the game levels and immersion factors, and also help the gamer enjoy the game while exercising.

3. System Configuration of the 3D AR Game

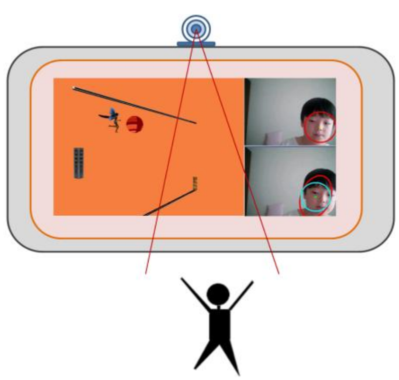

Figure 1. 3D Game Configuration Based on Augmented Reality

Figure 1 is the system map of AR game designed to realize the algorithm proposed in this paper. It is the first-person AR game in which the web camera is attached to PC and the gamer plays looking at the monitor. The real space made by webcam is used as game space. Results of face detection using the proposed method are applied to a particular object in the game and the player attacks and defends according to the scenario. In this game, gamer's moving parts are various and face and hands are to be used the most. However, considering all these moving parts in realizing the game will make the game very simple and boring. If we make a game where scoring occurs by recognizing a particular part of body, we could add degrees of difficulty and fun factors as well. Therefore, in the paper, we don't use the additional motion-capturing device and only use the web camera to get the image information of the gamer's face which will interact with the objects made in the virtual space as the game goes on.

4. Face Detection Algorithm

4.1 Viola and Jones Algorithm

With Viola and Jones Algorithm(Algorithm I), we generate classifier by using Haar-like Feature and cascade based on Boost technology, and search the results by using the Cascade structure. This algorithm is the Gray-based algorithm and constitutes the collection of very simple features to be studied. To extract these features, Haar-like Feature which uses the light-darkness distribution and has been proposed by Paul Viola et al[5-7], has been utilized as the feature collection for the purpose of face recognition. That is to say, integral image makes rapid calculation of face detection possible

4.1.1Haar-Like Feature : Haar-Like Feature is the face detecting method which is resistant to the surrounding factors and with which quick calculation is possible. This method utilizes the features of the objects in the face area as shown in Figure 4. Each feature is defined as Edge or Line, or Center-surround and etc. Collection of these features are defined as prototypes and through each experiment and the generally used set of prototypes are shown in Figure 2.

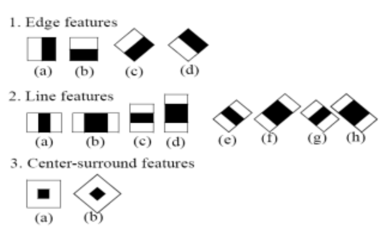

Figure2. Prototype Set of Haar-like Features

4.1.2 Integral Image : Integral image is similar to SAT(Summed Area Table) which is used for image processing. SAT minimizes the overlapping calculation and allows for speedy calculation.

(a) SAT (b) sum of pixels in the rectangle
Figure3. Integral image of the rectangle

SAT in Figure 3(a) is the sum of the pixel values within the rectangle area which includes the starting point (0,0) of the input image and the coordinate (x,y). Its calculations are shown in Formula(1) and Formula(2). In Figure 3(b), the rectangle is expressed as $r = (x, y, w, h, \alpha)$ when its coordinate for the upper left is (x,y), the width, w, the height, h and the rotating angle. The sum of the pixel values of r(the rectangle area) is expressed RecSum(r), which can be effectively calculated by using

$$SAT(x, y) = \sum_{x \le x, y \le y'} I(x', y') \quad (1)$$

$$SAT(x, y) = SAT(x, y-1) + SAT(x-1, y) + I(x, y) + SAT(x-1, y-1) \quad (2)$$

The case of rectangle RSAT which rotated at the angle 45°, shows the sum of all the pixels in the rectangle area can be calculated another formula.[7]

4.1.3AdaBoost and Cascade structure : Figure 4.shows cascade structure and its error rate has been proven to get exponentially closer to zero as the number of the weak classifier increases. That is to say, it is the method to generate the strong classifier with high detection performance through linear integration of the t number of weak classifiers. Haar-like which has been derived in the above becomes a weak classifier and it takes gradual approach from high-probability data in the studied data through probability distribution. Gradual approach is realized by the Cascade structure, which is similar to TLU(Threshold Logic Unit). The Cascade structure applies the high probability in the above and gradually applies low probability Haar-like. Each Sub-windows is applied to image and while going through classifier errors are eliminated in the process increasing speed as a result. If we control the number of the stages in the test, we can freely regulate the performance.

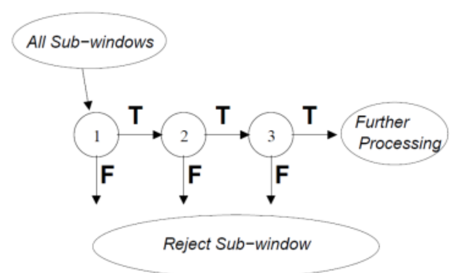

Figure 4.Schematic Depiction of the Detection Cascade.

4.2 CAMSHIFT Algorithm

CAMSHIFT algorithm(Algorithm II) [8-10] helps quickly calculate and track the position, rotating angle and even size by using the color information of the object, namely, the color probability distribution. So it allows for the benefits of being able to keep tracking the object in real time with the same minimal information about the search area, as long as another object does not overlap with the currently tracked object.

Another benefit of CAMSHIFT is that it is much smaller than the RGB color space with complex calculating, since it uses the Hue channel in the HSV color space. It is also relatively resistant to the color change since it ignores the extreme brightness or darkness thanks to the intrinsic characteristics of the color probability distribution model. But there is the prerequisite that it has to be supported by the algorithm which guarantees a certain amount of light because of the characteristics specific to the color information.[11]

Considering the time the whole process takes to process the complex algorithm and to eliminate noise real time in the area in which more than 30 images of frames have to be processed, CAMSHIFT algorithm is more convenient than Template, Active Contour Model, MeanShift, and etc.

But when we try to apply the exiting CAMSHIFT algorithm only, we have difficulties in automatically detecting and tracking the moving objects because we have to manually set up the ROI. CAMSHIFT algorithm tracks the object with the color value of the object, Hue and there can be many pixel values with the same Hue value. Therefore, if we do not control the area of the tracked object in an appropriate way, we can face the problem of the tracking area expanding unnecessarily.

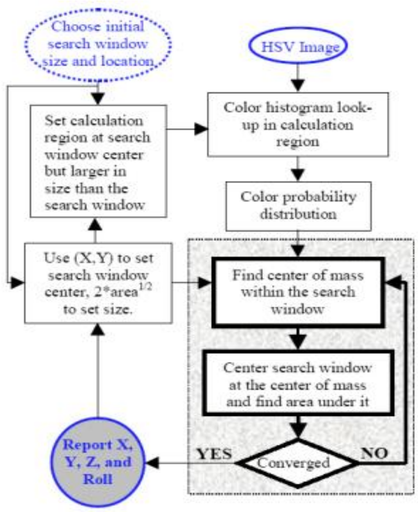

Figure5. The Flowchart of CAMSHIFT Algorithm

Figure 5.shows the flowchart of CAMSHIFT algorithm. As is shown in the figure, CAMSHIFT algorithm is the color-based, area-tracking algorithm which is based on the HSC color space.

4.5 The Proposed Algorithm(Algorithm III)

With only CAMSHIFT algorithm, we have to manually set the region of interest (ROI), making it difficult to automatically detect and track the moving object. CAMSHIFT algorithm tracks the object based on the object's hue, saturation and value, and there could be a wide range of pixel values with the same hue. So, there is a danger of having too wide a tracking area unless we don't limit the area for the tracked object in a proper way. To solve these problems, we analyze the face information gotten from Algorithm I to get the size and the center point of the face circle. CAMSHIFT algorithm is then used to get the inscribed circle in the calculated oval. Next, we finely modify hue value to get the new circle.

$$Circle\ 1: x_1^2 + y_1^2 = r_1^2 / \eta_1 \quad (1)$$

$$Circle\ 2: (x_2 - r_1)^2 + (y_2 - r_2)^2 = r_3^2 \quad (2)$$

$$Circle\ 3: x_3^2 + y_3^2 = r_5^2 / \eta_2 \quad (3)$$

$$\sqrt{(r_3+x_1)^2+(r_4+y_1)^2} \le r_{Dist} \quad (4)$$

$$((r_5^2/\eta_2)/(r_1^2/\eta_1))^2 \le E_{radius} \quad (5)$$

$$D_{t2}-D_{t1} \le D_{tref} \quad (6)$$

$$D_{t4}-D_{t3} \le D_{tref2} \quad (7)$$

Formula(1) shows the size of the face-detection circle created using algorithm I. We divided the radius by $\eta_1(1.0 \le \eta_1 \le 2.0)$ because it is the improved the formula to be able to more precisely detect the face by narrowing the face-detection area. Formula(2) is the oval formula generated using CAMSHIFT algorithm. Formula(3) shows the circle inscribed inside the oval in Formula(2). As we see in Formula(3), we have improved the formula to get more precise face recognition by using $\eta_2(1.0 \le \eta_2 \le 2.0)$ in circle 3. Formula(4) shows the distance between the center points of the Formula(1) circle and the Formula(3) circle. In this case, when the distance is less than r_{Dist}, proper recognition of face is possible and in case the distance is bigger, the face recognition is either possible in only one algorithm or in error, to be represented by algorithm I and CAMSHIFT algorithm respectively. So the case of failed face recognition in algorithm I is to be ruled out as an exception.

Formula(5) shows the subduplicate ratio of the radius sizes of the two circles. The center points are within the pseudo range, but the sizes of the radiuses are different, showing the distorted face recognition as a result. In this case, as in the case above, we simultaneously perform the two algorithms to get the stable radius value(time gap and size of the radius) to perform the face recognition. Formula(6) and (7) show the improved time-delaying method to be able to react to the abrupt environment or condition changes by maintaining the current value for the time being, in the case where the environment is not stable(light, noise, background color and etc.) and where the circle for recognizing the face changes so quickly in its position and size. Values for $r_{Dist}, E_{radius}, D_{tref}, D_{tref2}$ can be set properly to fit the environment through experiments. The makeup for the proposed improved algorithm is as follows. Figure 6 shows the flowchart of the proposed algorithm.

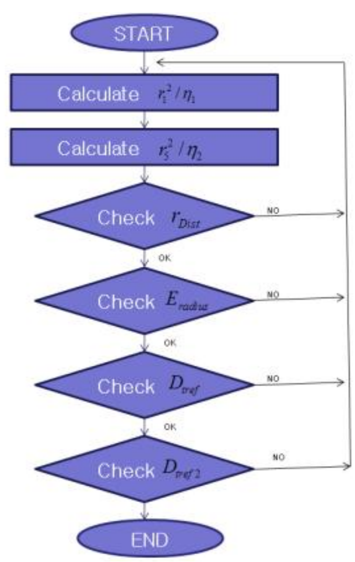

Figure 6. The Flowchart of the Proposed Algorithm

5. Experiments for the Proposed Algorithm

5.1 Experiment 1

(a) (b) (c)

Figure7. Result Image of Experiment 1

Experiment 3.1 shows the result of the test on the conventional face-detecting algorithms and the proposed algorithm. Figure 7(a) shows the result of face recognition using algorithm I. Figure 7(b) is the result of using CAMSHIFT algorithm which shows wide oval area recognized including face and neck areas. Figure 7(c) is the result of using the proposed algorithm, which shows more precise and smaller face area than in the case of Figure 7(b) since it considered the inscribed circle in Figure(b) and the circle size in Figure 7(a) to recalculate the factor.

5.2 Experiment 2.

(a) (b) (c)

Figure8. Result Image of Experiment 2

Figure 8.shows that the results of face detection in the conventional algorithms and the proposed algorithmIn Experiment 2, the face is more slanted than in Experiment 1, which cannot be recognized by using Algorithm I in Figure 8(a). CAMSHIFT Algorithm in Figure 8(b) provides face recognition even in the case of slanted face but the demerit is the recognized areas is too big. In Figure 8(c), the proposed algorithm not only recognized the slanted face but also recognized the precise and narrow image of the face.

5.3 Experiment 3.

In the Experiment3, the proposed algorithm has been proven to be effective. In the Experiment 4, the proposed algorithm is applied on 3D AR Game, which has been produced by using Irrlicht Engine, openCV and Visual C++. The background of the game is based on 800* 600 sized Maya mode (Figure 9(e)). The main character is Farie(Figure 9(a)) and cube thin node and spare thin node functions of Irrlicht Engine have been used to realize the three dimensional attacking two bars and a ball. (Figure 9(b)(c)(d)) To add to the sense of reality, the ball has been given a spin and the collision algorithm for detecting collision with the bars has been applied the square-collision method for rapid detection. Background music and effect sound have been added to help immersion of the gamer and the collision has been realized by using unique sounds considering the condition of the ball spin.

The game has been realized through game interface using face-detection information. There could be one or two players who can make a score by attacking from left to right or from right to left. The attacker is Faerie who attacks the other party who can also defend by using the bar. In the process, the ball movements and the face movements have been mapped 1:1 to use the detected face information.

(a) Faerie (b) Ball (c) Bar1 (d) Bar 2 (e) Game scene

Figure 9. Game Design Image for Experiment 3

5.3.1 Experiment 3 : Experiment 3 is the crisis situation in which Faerie[11] attacks with the ball from left to right as Figure10(a). To defend, the gamer determines the defense position and moves the bar in the left to defend the attack from Faerie. In Figure10(b), the gamer's face is recognized by using algorithm I and the face cannot be moved to the place in reaction to the moved position of the ball. In Figure10(c), the proposed algorithm III has been used to detect the face and face can be detected through the bar spot reacting to the moved place of the ball. In Figure10(d), Faerie's attack has been defended by moving the bar by using face. The ball is in the process of movement from right to left by Faerie. Figure10(e) is the stage right after the face moved a little to move the bar from Figure10(b). The circle of recognized face area has become bigger than the original face and the center of the face is a little off the face. But the Figure10(f) where the proposed algorithm has been used, shows better recognition of the face and the center of the face is also positioned properly. Therefore, we could know that the proposed algorithm provides more precise recognition of the face and is more adaptable to the changing environment.

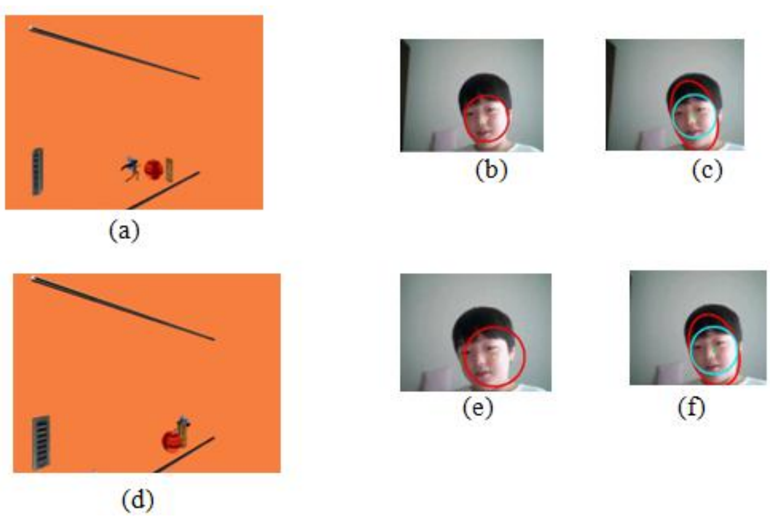

Figure10.Input and Result Image of Experiment 3

5.3.2 Experiment 4 : In Experiment 4, Faerie is attacking from left to right as in Figure11(a) and the gamer is positioned in the right bottom area as shown in Figure 11(b). In Figure 11(b) and Figure 11(c), face is recognized by using algorithms I and III respectively and the movement to the next stage is possible. If the face is turned to defend the ball in this situation, there is no recognizing the face if we apply algorithm I as in Figure 11(e). But if we use algorithm III, proper recognition of face is possible as in Figure 11(f) and we can use the face information to defend the attack as in Figure 11(d) and turn the attacking direction from right to left by using reaction. Effective counterattack is possible if we use Algorithm III as the game interface in various situations happening in 3D AR games.

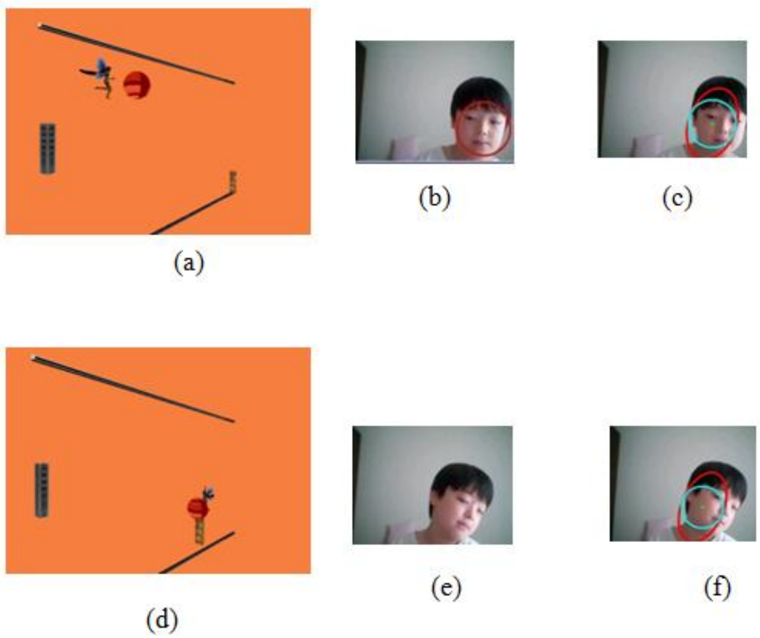

Figure11. Input and Result Image of Experiment 4

5.3.3 Experiment 5 : As in Figure 12(a), Faerie is attacking from the left toward the right. To fight it, the gamer's face is positioned in mid bottom as inFigure 12(c). In this case, since the gamer's face is downward, if we apply algorithm I, there is a danger that we might not recognize the face as in Figure 12(c), whereby we don't have the information to defend the attack from Faerie. But if we apply Algorithm III, proper recognition of the face, as shown in Figure 12(d), is possible and we can use the face information to counterattack as in Figure 12(b).Figure 12(e) shows attack result scene using face as a AR game interface. We could verify the effective response by using the proposed Algorithm III in various game situations in the produced 3D AR game.

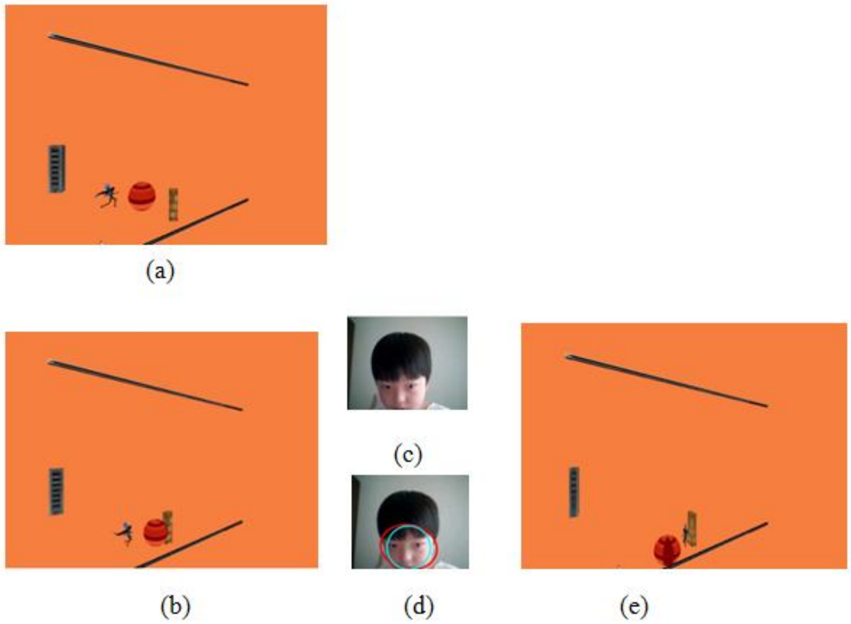

Figure12. Input and Result Image of Experiment 5

5. Conclusion

The conventional face-detecting methods, Viola and Jones(Algorithm I) and CAMSHIFT Algorithm(Algorithm II) are effective in detecting the face rapidly and clearly. But Algorithm I shows a problem in recognizing slanted faces. Algorithm II has a problem in precisely detecting face since its recognizing area is relatively wide depending on environment. To solve these problems, the paper presents the new algorithm(Algorithm III) by considering the position and size of the circle(the area of face-recognition) and detection intervals. The experiment has proven the efficacy of the new algorithm. The 3D AR game has been produced to prove its utilization as game interface in AR environment and its efficacy has been proven through experiments in various game situations. Whereas the conventional methods have limitations in detecting face depending on environment and the position of gamer's face, the proposed algorithm has been proven to be successful in various situations, thereby realizing exciting game interface.

References

[1] SERI Management Note, March.11 2010, Book No. 46, The future change which will be brought about by augmented reality

[2] SERI Economy Focus, Oct. 27, 2008 No. 216, Interface technology which enhances the value of the product

[3] Kim, Ky Young et al ,"ARPushPush : Augmented Reality Game in Indoor Environment", KHCI, pp354--359,(2005)

[4] Handheld Augmented Reality Game System Using Dynamic Environment Kang, Won Hyung, "Handheld Augmented Reality Game System Using Dynamic Environment ", KAIST, 2007, Thesis for Master's Degree

[5] P. Viola, M. Jones, "Rapid object detection using a boosted cascade of simple features", *IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawaii, USA, 2001

[6] http://cafe.naver.com/opencv.cafe?iframe_url=/ArticleRead.nhn%3Farticleid=1328

[7] An extended set of Haar-like feature for rapid object detection / Lienhart, R and Maydt, J /International

Conference on Image Processing (2002)

[8] J. G. Allen, R.YD Xu, J. S. Jin "Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces," *Pan-Sydney Area Workshop on Visual Information Processing*, pp.3-7, 2003

[9] MunHo Lee, SeonhWhanJeong, Computer vision programming using OpenCV, HongReung Press, (2007)

[10] Young Jae Lee and Yong Jae Lee,"Interface of Augmented Reality Game Using Face Tracking and Its' Application to Advertising",Security-Enriched Urban Computing and Smart Grid Communications in Computer and Information Science, 2010, Volume 78, 614-620,

[11] http://irrlicht.sourceforge.net/

Authors

Young Jae Lee received the B.S. degree in ElectronicEngineering from Chungnam National University in 1984,the M.S. degree in Electronic Engineering from Yonsei

University in 1994, and the Ph.D. in Electronic Engineering from Kyung Hee University in 2000. He is presently an Associate Professor in the College of Culture and Creative Industry at Jeonju University. Hisresearch interests include Smart Media, Computer Vision, AR(Augmented reality), and Computer Games.

Geun Ho Lee received the B.S. degree in Electrical Engineering from Chonbuk National University in 1994, the M.S. degree in Electrical Engineering from Chonbuk National University in 1997, and the Ph.D. in Electrical Engineering from Chonbuk National University in 2000. He is presently an Associate Professor in the College of Culture and Creative Industry at Jeonju University. Hisresearch interests include Digital signal processing, Smart TV, Digital Logic System and Multimedia.