한국산학기술학회논문지
Vol. 9, No. 1, pp. 100-107, 2008

Extracting and Tracing a Specified Object among Multiple Ones

Eun-Hwan Kim[1*], Dan-Song Han[2], Kwang-Hyoung Lee[3] and Moon-Seog Jun[4]

다중객체 환경에서 특정객체의 추출 및 추적에 관한 연구

김은환[1*], 한단송[2], 이광형[3], 전문석[4]

Abstract This paper proposes a real-time processing method to simplify the image input procedure while extracting and tracing a specified object among multiple ones. In order to extract an object in a specified area among multiple objects of indoor environment and tracing the extracted object continuously, it is verified through experiments that the information interchanged between cameras upwards and in front of it have effect on tracing a specified object continuously. The camera located upward transfers its x-axis data of the input image to the front camera so that the front camera can catch the area of object soon without computing the information of x-axis. The front camera can't resolve the problem of objects overlapping till they share information with the upward camera. The result of the experiment shows that the computation for tracing an object is simplified and the accuracy for extracting and tracing is upgraded.

Key Words : Object Extract, Object Trace, Real Security, Image Processing

요 약 본 논문에서는 다중객체 환경에서 특정한 객체를 추출과 추적을 하기 위하여 입력되는 영상의 처리과정을 간소화하여 실시간으로 수행할 수 있는 방법을 제시하였다. 실내환경의 많은 객체들 중 특정영역에 입력되는 객체를 추출하고 추출된 객체를 지속적으로 추적하기 위하여 상단의 카메라와 전면의 카메라의 상호정보교환을 통하여 특정 객체를 지속적으로 추적할 수 있음을 실험을 통하여 검증하였다. 상단의 카메라에서 입력받은 영상에서 x 축의 정보를 전면의 카메라에 전송하여 전면의 카메라는 객체의 x축 정보를 계산하지 않고 빠르게 객체영역을 획득할 수 있으며, 전면카메라에서는 객체들 간의 겹침현상을 해결할 수 없지만 상단의 카메라와 정보를 공유하므로 객체들 간의 겹침현상을 해결할 수 있었다. 실험결과 기존의 객체추적 시스템보다 객체를 추적하기위한 연산량이 줄어들었으며, 객체추출 및 추적의 정확성이 향상되었다.

1. Introduction

Recently, the research of computer vision systems has received growing attention to meet the development of information communication technology. As an area of vision systems, camera-used vision inspection is used for traffic control, surveillance of construction site and a shopping place, a place of unattended etc [1][2]. A human vision inspection is used in the past, as the development of vision inspection technology an surveillance system unattended is used recently [3].

Visual surveillance system analyzes the information of time and space of an image getting from camera to extract and analyze the moving object. Information (location, speed etc.) of a moving object can be extracted through analysis. Object identification, automatic classification and searching is possible via the analysis of color, shape and quality of a moving object [4]. There are three methods used in the camera surveillance, one camera extracting a moving object, several cameras in an area and the method of using several cameras in several areas [5].

Using multiple cameras for specified area monitoring needs to set up cameras to make images overlapped for tracing a moving object [6]. This method is not only cost

[1]Soongsil University Computer Institute
[2]Korea Information Security Agency
[3]Seoil College, Dept. of Internet Information.
[4]Soongsil Univ. Dept of Computing.
[*]Corresponding Author : Eunhwan Kim(ehkim@ssuci.ac.kr)

much but also have a low performance while processing input images captured by each camera. The method of using several cameras in several areas cost less while tracing an object in a wild area because it shares data with the cameras adjacent to it. It is difficult to decide which is research object while images moving from one camera to another. Moreover, the color distribution and light of the background image that inputs to each camera is different so it is not easy to tracing the same object. There are three main methods (such as, difference imaging, block conformance and using background imaging) for a real-time image processing to distinguish the background image and the input image, so as to recognize an extracted moving object.

Real-time object that extracting and tracing system needs to extract object correctly in a short time. The existed single object system has a disadvantage that it needs an environment of moving just in one place. For an input image that moving in several places and several objects, it can't extract and trace the object.

Leveling object tracing system is used for tracing multiple objects. In an environment of multiple objects, overlapping of objects happens. When overlapping decreased, it depends on the supposition of the former moving direction to decide which object get the leveling. Multiple object tracing system traces each object so that it has a high computation complexity. It is not proper for a real-time imaging security system.

There are several reasons for research on specified object monitoring system in a category of real-time imaging security system. firstly, it is necessary to protect a specified object by security monitoring. In other words, preventing a rare and valuable thing from losing takes top priority. While applying the post-processing centralized algorithm which used in an existed object tracing system, it is not easy to recover the loss. So an efficient security monitoring policy is required to be pre-prevention. secondly, security system using camera requires an real-time environment. The existed imaging security system mainly analyzes the saved images when problems appear. If a security system for extracting and tracing a specified object has a real-time property, monitoring ahead can become possible. Thirdly, when an object to be investigated is shadowed by a specified object, it is hard to identify and trace. In other words, it is possible to identify and trace an object in the case of just one object existed or multiple objects distributed. In general, multiple objects appear overlapped. Few research work was done for an environment of multiple overlapped objects. Fourthly, security system for prevention monitoring requires a fast image processing. If a specified object is identified in a short time, the passing of an warning information in a prevention system will not spend much time. But it has a limitation to reduce time for identifying an object in the whole image.

2. Related Work

2.1 Object Tracing System

An object tracing system detects changes of an object in a sequence of time through analyzing the video in different point of view and convert tracing information to trace the location of an object. It consists of image input part, object extracting part, object tracing part which is shown in [Fig. 1].

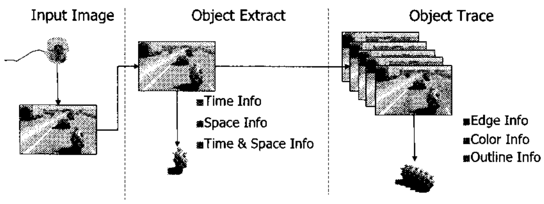


Fig. 1 Architecture of an object tracing system

Image input part input the monitoring scenes in the system by cameras. The image data get from the cameras is recorded as color, texture, shape and boundary.

Object extracting part analyzes the input image to identify a specified object. It requires time and space information for object extracting.

Object tracing part traces the moving track get from the object extracting part. In order to trace the movement of a specified object, information of color, edge, contour of an image updated by time is needed.

2.2 Object Tracing Algorithm

The image of an video is made up of frames. Moving detection is to predict such small changes which is happened from an frame to another and delete the duplication by time.

The principle for detecting movement is under the assumption that there is no change in pixels and just tracing the pixel values of current frame. We detect moving by the value change of the pixels and moving vector. It shows in [Fig. 2] [7].

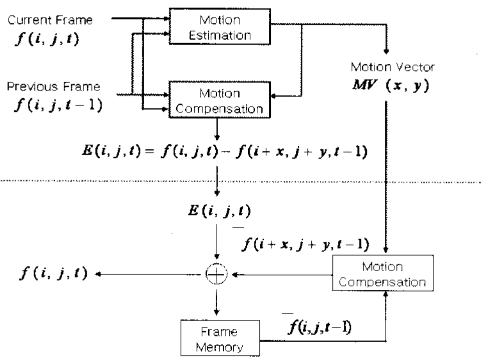


Fig. 2 Method for Extracting a moving object

There are two methods for object moving detection, that is, full search method and high-speed search method. In order to detect the movement of objects in an image, there are three methods, that is, difference imaging method, block conformation and full search method.

(1) Difference Imaging Method

Difference Imaging method detects the moving from the difference of current frame <Frame t-1> and the former frame <Frame t>. It consists of difference between two frames and the difference between the background and two frames.

The method of using the difference between two frames compares the difference value between the current frame and the former frame with the critical value of '0'. If the value is '0', it means the pixels don't have any change. If the value is more than '0', it means the pixels have been changed. The second method mentions above defines the background firstly and compares it with the current frame. The critical value describes the change of the pixels. It shows in [Fig. 3] and function(2.1)

$$Difference_{img}(i,j) = |f_t(i,j) - f_{t-1}(i,j)| \quad (2.1)$$

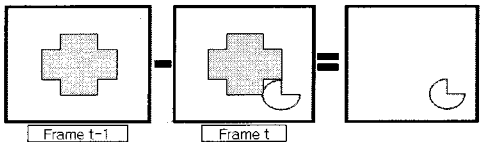


Fig. 3 Difference Imaging Method

(2) Block Conformation Method

Block conformation method [Fig. 4] is in a supposition that all the pixels in a block have the same moving vector and presume current frame from former frame in a unit of block.

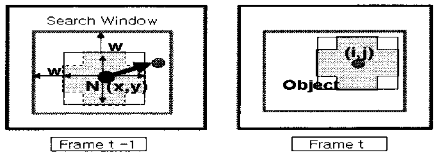


Fig. 4 Block Conformation Method

It divides the frame into blocks and finds the most similar block in the former frame. Finally, it presumes the current block according to the most similar one.

For a block which has a size of N×N, it can move W pixels up and down, left and right, so the area for searching data is (2W+N)×(2W+N). There are three standards for evaluating the agreement of blocks, as function (2.2), that is MSE(Mean of Squared Error), MAD(Mean of Absolute Difference) and SAD(Sum of Absolute Difference). SAD method has an advantage of calculation and implementation, so it is used popularly now[8][9].

$$MSE(x,y) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}[f(i,j,t) - f(i+x, j+y, t-1)]^2$$

$$MAD(x,y) = \frac{1}{N^2}\sum_{i=0}^{N}\sum_{j=0}^{N}|f(i,j,t) - f(i+x, j+y, t-1)|$$

$$SAD(x,y) = \sum_{i=0}^{N}\sum_{j=0}^{N}|f(i,j,t) - f(i+x, j+y, t-1)| \quad (2.2)$$

Block conformation algorithm consists of full search algorithm and high speed search algorithm. Full search algorithm can detect the moving accuracy. High speed algorithm compares part of the blocks, so that it has a high speed.

(3) Full Search Method

The common method for detecting the movement is the full search method. It compares current blocks of current frames with search area of former frames one by one. It has a high performance, simplified calculation and implementation and it is used on video compression popularly.

3. Proposed System Architecture

The object tracing system using multiple cameras is consisted of input part, processing part and output part as shown in [Fig. 5]. Input part consists of object monitoring camera(MC) which is located upward of the object and object tracing camera(TC) which is located at the front of the object. The input image is converted into digital data and transferred to the processing part. The processing part analyzes the image and tracing the specified object. The output part transfers an warning message when specified object appears and saves the image from TC to the memory.

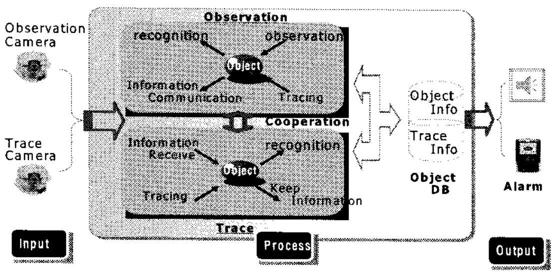


Fig. 5 Architecture of the Object Tracing System using Multiple Cameras

(1) Input Part

Input part transfers the images from MC and TC to the processing part. Because the images from MC and TC achieved from the same space, x-axis of MC is same with the x values of TC for a same specified object. Moreover, y-axis of MC which means the distribution of specified object is same to the z-axis of TC which stands for the distance of specified object. The information transfers from MC to TC is x value and y area which is same to the z value of TC.

(2) Processing Part

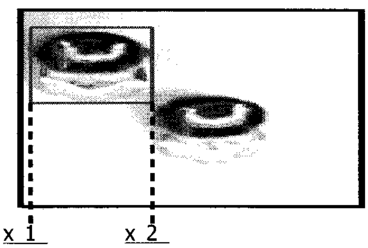
The processing part consists of monitoring part and tracing part. It detects whether the monitoring object is exist or not and the boundary processing of specified object, calculation of the minimum area, passing information form MC to TC, giving a warning message and tracing an object.

In order to extract the signature of the investigated object which is transferred from MC, information of object, such is, size, color and shape must be saved.

The confirmation of whether the investigated object is exist in real-time and the performing of the image which is transferred from MC is required. In order to detect whether a specified object enters a monitoring area, an area wider than the object is needed.

When a specified object coming into a monitoring area, it need to extract the boundary of the object according to part of the object that input to a monitoring part. The size of the object is defined according to the minimum value of x1, the maximum value of x2, the minimum value of y1 and the maximum value of y2. The information of x1 and x2 is sent to the tracing part. The middle value of y1 and y2, that is, mid(y1, y2) is sent to the tracing part to decide the distance of object.

The operation of TC starts when the specified object coming into MC. When TC is operated, warning message is sent to output part and the original image information of TC is saved. The value of x1 and x2 transfers from MC decides the x-axis value of the object. It is shown in [Fig. 6].



(a) input image of MC (b) input image of TC

Fig. 6 x-axis of TC according to x1 and x2

In order to decide the value of y-axis from TC, the original value was achieved from MC by mid(y1, y2). It divides the image that input from MC and takes it as a standard when measure the height of the input object in

TC's image. As shown in [Fig. 7], y-axis of MC is divided into n equal parts. The size of the object that input from TC is decided on whether the most overlapped area is in the upward or downward. If the two objets are equal in size, the object in area ⓑ has a height of 59 pixels while the object in area ⓒ has a height of 96 pixels.

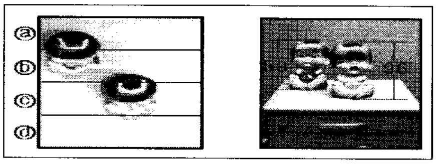
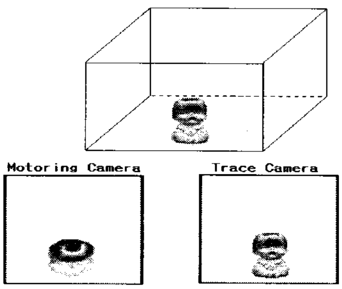


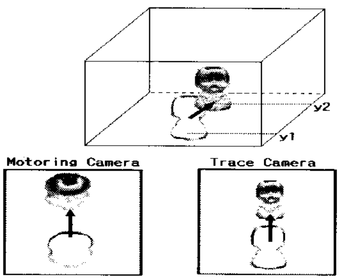
Fig. 7 Height of Objects in TC According to MC Areas

In order to calculate the height of the object, we use the value of mid(y1, y2) of MC as the start point of calculation, so the y-axis value of TC is calculated.

[Fig. 8] shows the mutual relations between y-axis of MC and z-axis of TC,



(a) A Specified Object located front



(b) Back moved images

Fig. 8 The Mutual Relation of MC's y-axis and TC's z-axis

(3) Output Part

When MC finds a specified object entering an area to be monitored, the output part transfers an information to TC that make it operates. An warning system is operated according to the missing risk of the object to be investigated. Moreover, when TC is operating, output part saves all the images of a specified object which is achieved in the period of tracing. The output part is made up of warning device and data saving device. A warning device is used to notify when a specified object enters a monitoring area. A data saving device is used to save the data of object tracing.

3.1 Proposed Algorithm of Object Tracing

The algorithm of the proposed object tracing using two cameras is made up of MC which is used to monitor the entering of an specified object and TC which is used to tracing the movement of the object. [Fig. 9]

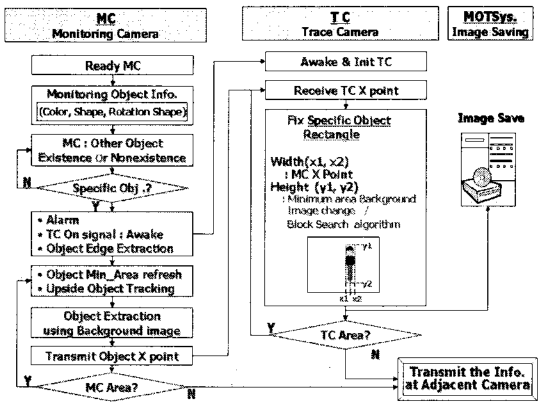


Fig. 9 Flowchart of the Proposed Algorithm

In the early period, MC has a real-time monitoring to detect whether a specified object enters the area. If a specified object enters, it operates TC and makes an architecture of multi-dimension for creating the object boundary and tracing the object in a real-time.

The information of objects in a monitoring area is input through MC, information of the specified object is achieved from color, shape and turnover information in a real-time. When a specified object enters the area, TC is operated and the minimum boundary of the specified object is transferred to TC.

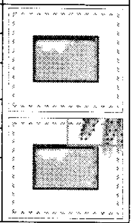
TC creates blocks which separate the background and the specified object for tracing according to the minimum boundary information of an specified object transferred from MC

4. Performance Evaluation

4.1 Calculation of Pixels for Object Recognition

Calculation of pixels for an object tracing system is defined as all of the pixels from the starting time that decides the object's boundary and detect whether the object enters. [Table 1] shows the calculation of pixels for object detection system. The calculation of pixels is made up of pixels used to decide whether an object enters the monitoring area and pixels used to get to know the shape of the object.

Table 1. Calculation of Pixels in a MC Monitoring Area

calculation pixels / # of test	object entering decision total(240)	minimum pixel	maximum pixel	object monitoring area
1~100	147.3 pixel	40	236	
101~200	142.8 pixel	43	239	
201~300	147.9 pixel	40	238	
301~400	136.7 pixel	40	237	
401~500	144.3 pixel	40	239	
average	143.8 pixel	40.6	237.8	

As shown on [Table 1], the proposed system do not mentions about the whole entering image, but investigates the specified object that enters the monitoring area.

The minimum amount is 40 pixels and the maximum is 239 pixels. If the pixels in a monitoring area is 50 × 40 = 2000, we do not check each pixels but combine every 5 pixels as a group to be investigated. That saves a lot of time.

Object recognizing calculation of pixels at TC is the pixels to recognize the overall object after transferred from the x-axis of MC. The result of the experiment is shown on [Table 2].

The average number of pixels in an object area describes how much area does the object occupied. The minimum number of calculation of pixels is the minimum pixels that has been used to decide the size of the object. The maximum number of the pixels is the maximum pixels to be used.

Because for TC, the value of x1 and x2 is transferred from MC, the information of x-axis doesn't need to be calculated, the number of pixels used for calculation is decreased. The minimum calculation pixels can be got when object is far from the camera and the maximum calculation pixels can be got when object is near to the camera.

Table 2. Object Recognizing Calculation Pixels at TC

caculation pixels / # of test	average pixels in the object area	minimum pixels	maximum pixels	TC object areas
1~10	2,550.96	1,350	5,538	
11~20	2,890.68	1,352	5,543	
21~30	3,506.45	2,480	4,613	
31~40	3,513.35	1,919	5,210	
41~50	2,755.12	1,153	4,896	
Average	3,043.00	1,651	5,160	

4.2 Object Extracting Accuracy

The object extracting accuracy is measured by having different time that object enters the monitoring area and recognizing the overall size of the object.

The result of the experiment is that if the object enters the monitoring area fast, other area will be recognized as a part of object incorrectly. Extracting continuously can decrease the incorrect area. [Table 3] shows the pixels that recognized as a part of object incorrectly in different moving speed.

Table 3. Object Extracting Accuracy

calculation pixels / # of test	average pixels of object area	pixels of the error area	object extraction ratio[%]	image of the extracted object
1~10	2,029	101	95.25	
11~20	1,843	22	98.82	
21~30	2,323	62	97.40	
31~40	3,070	42	98.65	
41~50	1,889	70	96.42	
Average	2,230.8	59.4	97.40	

The pixels of the object area are the number of pixels that object occupied when it is static. The error area pixels are the number of pixels that recognized as a part of object incorrectly. The object extraction ratio is the pixels of an object area to pixels of an error area. In practise, object area is matching to the extract object area as an average value of 97.4%.

4.3 Object Tracing Accuracy

In order to measure the object tracing accuracy of the proposed system, we have the experiments in a different moving speed. It also compares with the objects except to the investigated object.

As shown on [Table 4], it is failed if the overlapped part of object move away from the area of MC and TC. In this case, the specified object is traced continuously but for TC it is not easy to decide the location of the object.

The speed is calculated from the extracting time of a specified object in a image to the time that object lost. [Table 4] shows the object extraction including time information. The object extraction time and speed of the object extraction in [Table 4] is from the system time.

Table 4. Object Tracking Accuracy

item # of times	MC tracing	TC tracing	time [sec]	item # of times	MC tracing	TC tracing	time [sec]
1	O	O	3.915	11	O	O	3.972
2	O	O	4.165	12	O	O	4.157
3	O	O	4.513	13	O	O	5.197
4	O	O	5.127	14	O	O	3.127
5	O	O	3.057	15	O	O	6.054
6	O	O	5.972	16	O	X	3.175
7	O	O	3.509	17	O	O	4.195
8	O	O	4.167	18	O	O	4.718
9	O	O	4.756	19	O	O	4.008
10	O	X	3.854	20	O	O	5.175

O : succeed, X : failed

5. Conclusions

This paper proposes an object tracing system which is made up of two cameras called MC and TC. It is different from the traditional methods which mainly use the front image. In this research, MC is used to monitor an object from upward and decreased the calculation of pixels by passing the information of MC's x-axis values to TC. In this paper, we have the experiments only in the monitoring area and doesn't mention about the whole input image. That decreased the calculation of pixels. The mutual relation of MC and TC is also used to decrease the calculation of pixels. Moreover, for an environment of multiply overlapping objects, continuous tracing can't monitor the image of TC, but in the image of MC it is possible. When eliminate the overlapping in the image of TC, it is possible to trace continuously in TC by using the information transferred from MC.

References

[1] G. L. Foresti, "A Real-Time System for Video Surveillance of Unattended Outdoor Environments", IEEE Trans, on Circuit and Systems for Video Tech., vol. 8, no. 6, pp. 142-145, 1998.

[2] D. M. Gavrila, "Pedestrian detection from a moving vehicle", Proc. Eur. Conf. Comp. Vis., Vol. 2, pp. 37-49, 2000.

[3] G.D. Hager and P.N. Belhumeur, "Effecient Region Tracking with Parametric Models of Geometry an Illumination", IEEE TPAMI, vol, 20, no. 10, pp. 1020-1039, 1998

[4] G. P. Stein, A. Shashua, "A Robust Method for Computing Vehicle Ego-motion", In IEEE Intelligent Vehicles Symposium(IV2000), 2000.

[5] D. Beymer and K. Konolige, "Real-Time Tracking of Multiple People using Stereo", In IEEE Frame Rate Workshop, 1999.

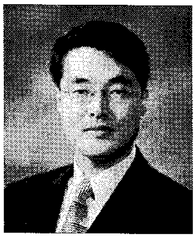
[6] Z. Duric, F. Li, Y. Sun. and H. Wechsler. "Using Normal Flow for Detection and Tracking of Limbs in Color Images". In Proc. International Conference on Pattern Recognition, Quebec City, Canada, August 2002.

[7] Yen, Ching Chang and Shyang Chang, "A fast estimation algorithm on the hurst parameter of discrete-time fractional brownian motion," IEEE Trans on Signal Processing, Vol. 50, No. 03 pp. 0554~0559, March 2002.

[8] Markus Flier and Bernd Girod, "Video coding with motion- compensated lifted wavelet transforms," Signal Processing: Image Communication, Vol. 19, No. 7 , pp. 561~575, August 2004.

[9] Viet Anh Nguyen and Yap-Peng Tan, "Fast block-based motion estimation using integral frames," IEEE on Signal Processing Letters, Vol. 11, No. 9, pp. 744~747, September 2004.

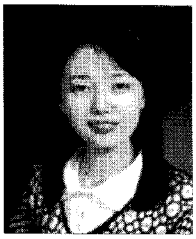
Eun-Hwan Kim [Regular Member]



- Feb. 1990 : Dept. of Computer Science Soongsil University(B.S.)
- Aug. 1997 : Dept. of Computing, Soongsil University(M.S.)
- Feb. 2003 : Dept. of Computing, Soongsil University(Ph.D.)
- Sep. 1997 ~ current : Professor with Dept. of Internet Communication in Soongsil University Computer Institute.

<research area>
Multimedia Communication, Network Security, Ubiquitous,

Dan-Song Han [Regular Member]



- Sep. 1997 : Dept. of Computing & Electronic, Yanbian Univ. of Science & Technology, China(B.S.)
- Aug. 2000 : Dept. of Computing, Soongsil University(M.S.)
- Sep. 2000 ~ current : Dept. of Computing, Soongsil Univ.(Ph.D. Candidate)
- Aug. 2005 ~ current : Junior Researcher, Korea Information Security Agency

<research area>
Information Security, Botnet response, honeynet, RFID

Kwang-Hyoung Lee [Permanent Member]

- Feb. 1998 : Dept. of Computer Engineering, Kwang Ju University.(B.S.)
- Feb. 2002 : Dept. of Computing, Soongsil University.(M.S.)
- Feb. 2005 : Dept. of Computing, Soongsil University.(Ph.D.)
- Mar. 2005 ~ current : Professor with Dept. of Internet Information, Seoil College.

<research area>
Multimedia data search , RFID, Ubiquitous

Moon-Seog Jun [Regular Member]

- Feb. 1980 : Dept. of Computer Science, Soongsil Univ.(B.S.)
- Feb. 1986 : Dept. of Computer Science, Univ. of Maryland(M.S.)
- Feb. 1989 : Dept. of Computer Science, Univ. of Maryland(Ph.D.)
- Mar. 1991 ~ current : Professor with Dept. of Computing in Soongsil University.

<research area>
Network Security, Computer Algorithm, Parallel Processing, Cryptology.