

고해상도 어안렌즈 영상에서 움직임기반의 표준 화각 ROI 검출기법

류아침[†], 한규필^{**}

Motion-based ROI Extraction with a Standard Angle-of-View from High Resolution Fisheye Image

Ar-Chim Ryu[†], Kyu-Phil Han^{**}

ABSTRACT

In this paper, a motion-based ROI extraction algorithm from a high resolution fisheye image is proposed for multi-view monitoring systems. Lately fisheye cameras are widely used because of the wide angle-of-view and they basically provide a lens correction functionality as well as various viewing modes. However, since the distortion-free angle of conventional algorithms is quite narrow due to the severe distortion ratio, there are lots of unintentional dead areas and they require much computation time in finding undistorted coordinates. Thus, the proposed algorithm adopts an image decimation and a motion detection methods, that can extract the undistorted ROI image with a standard angle-of-view for the fast and intelligent surveillance system. In addition, a mesh-type ROI is presented to reduce the lens correction time, so that this independent ROI scheme can parallelize and maximize the processor's utilization.

Key words: ROI, Fisheye Lens, Distortion Correction, Standard Angle-of-view, Motion Detection

1. 서 론

현재 영상관련 산업에서는 고성능의 영상 미디어 기기가 많이 대중화되었고 특히 CCTV 관제시스템은 사후 처리용 녹화기가 아닌 실시간 능동 감시 시스템으로 발전하고 있으며, 더 나아가 지능형 영상분석 시스템으로 고도화되고 있는 추세이다[1]. 이러한 요구에 맞추어 CCTV용 카메라 제조업체는 8백만 화소(4K급) 이상의 고해상도 카메라를 개발보급하고 있으나, 일반적인 카메라 렌즈 특성상 감시 가능한 화각이 90도 이하에 머무르고 있어 감시 사각지대

가 존재하고 있으며, 사각영역을 줄이기 위해 다수의 카메라를 요구하는 단점이 있다[1].

이에 반해 어안렌즈 카메라는 180도 이상의 넓은 화각을 가지고 있어 감시 영역이 넓지만, 렌즈 굴절에 의해 심각한 방사형 왜곡을 가지는 문제가 있다.[2] Fig. 1에서와 같이 기존 어안렌즈를 장착한 카메라에서는 왜곡에 대한 보정기능을 제공하고 있지만, 여전히 불룩렌즈 왜곡이 존재하는 것을 확인할 수 있다[3]. 일반적으로 어안렌즈는 영상 획득 위치에 따라 통과하는 렌즈의 곡률이 크게 다르기 때문에 왜곡을 최소화하려면 가장자리 영역이 제거되고, 감

* Corresponding Author: Kyu-Phil Han, Address: (39177) 61 Daehak-ro, Gumi, Gyungbuk, Korea, TEL: +82-54-478-7525, FAX: +82-54-478-7539, E-mail: kphan@kumoh.ac.kr

Receipt date: Sep. 17, 2019, Revision date: Jan. 6, 2020
Approval date: Jan. 28, 2020

[†] Dept. of Computer Engineering, Kumoh National Institute of Technology, E-mail: axiuus@gmail.com

^{**} Dept. of Computer Engineering, Kumoh National Institute of Technology

* This research was supported by Kumoh National Institute of Technology(2017-104-060)



Fig. 1. 360-degree and calibrated views of a commercial fisheye lens camera[3].

시영역을 최대화하면 왜곡을 완전히 제거할 수 없다. 그래서 어안렌즈를 장착한 감시 시스템을 구성하고자 할 경우에는 영상의 왜곡 보정 계산으로 인한 시스템의 연산량 증가를 고려해야 한다. 보정에 소요되는 시간이 증가되면 입력되는 프레임이 누락되어 감시 성능이 떨어지게 된다. 또한, 다분할 및 파노라마 뷰 단일보정으로 인해 발생하는 사각지대의 발생을 최소화해야 한다. 그러므로 제한된 기법에서는 실시간 감시 효율을 높이기 위해서 고해상도 어안렌즈 카메라를 사용하면서 사각지대를 최소화하고 동시다발적 영역감지가 가능하며, 렌즈 왜곡이 보정된 표준화각의 관심영역(ROI: region of interest) 검출기법을 제안하고자 한다.

2. 관련 기술 및 연구

2.1 렌즈 왜곡과 보정

이상적인 핀홀(pin-hole) 카메라 모델에서는 단순한 기하학적 투영을 통해 2차원 영상생성이 가능하지만, 실제 카메라에서는 사용된 렌즈의 왜곡 및 카메라 내부(intrinsic) 파라미터[4] 등의 요인에 의한 영향을 받게 되므로, 영상을 이용한 정합 및 인식 분야에서는 렌즈의 왜곡 보정이 성능에 크게 영향을 준다[5]. 렌즈 왜곡 보정은 보정 패턴을 촬영하여[6] 특징점의 이동거리로부터 왜곡을 보정하는 방법이 보편적이다. 일반적으로 볼록렌즈를 많이 사용하기 때문에 방사왜곡(radial distortion)중 술통왜곡(barrel distortion)이 대표적이며, 영상 중심에서부터 주변으로 갈수록 왜곡이 심해져 n 차 보정계수(distortion coefficient) k_n 의 항을 Eq. (1)에서와 같이 늘려가면서 왜곡을 보정한다. Eq. (1)은 짝수차수 다항식모델(even-order polynomial model)로써 가장 많이 활용되는 렌즈왜곡모델이며[7], 여기서 x, y 는 왜곡되어 획득된 영상좌표이고 $x_{corrected}, y_{corrected}$ 는 왜곡이 보

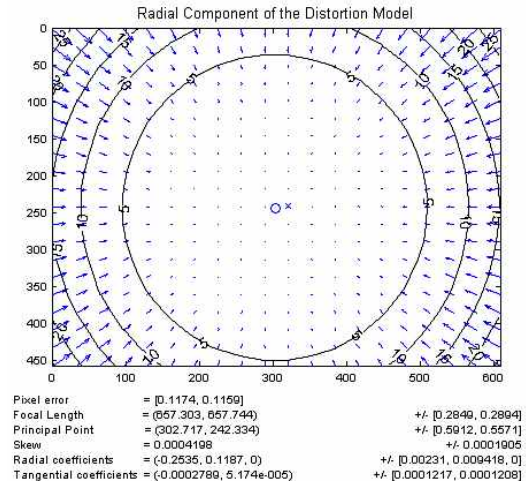


Fig. 2. Barrel distortion.

정된 좌표이며, $r = \sqrt{x^2 + y^2}$ 이다.

$$\begin{aligned} x_{corrected} &= x(1 + k_1r^2 + k_2r^4 + k_3r^6 \dots) \\ y_{corrected} &= y(1 + k_1r^2 + k_2r^4 + k_3r^6 \dots) \end{aligned} \quad (1)$$

Fig. 2는 술통왜곡의 예로써 파란색 선은 왜곡된 점과 보정된 점을 연결한 것으로 위치별 왜곡을 알 수 있으며, 가장자리로 갈수록 렌즈 곡률에 의한 왜곡이 심한 것을 볼 수 있다.

2.2 다해상도 영상과 움직임 예측

영상에서 변화를 검출하는 방법 중 가장 단순하면서 효과적인 기법은 Fig. 3에서처럼 프레임간의 차(frame difference) 영상을 이용하는 것이다[8]. 차 영상을 활용할 경우 잡음이나, 조명의 변화에 민감한 단점이 있기 때문에 잡음 제거와 adaptive threshold 등과 같은 처리를 필요로 한다. 고해상도 영상일 경우에는 고속 처리를 위해 이미지의 크기를 단계적으로 축소시켜 역 피라미드(inverse pyramid) 형태의

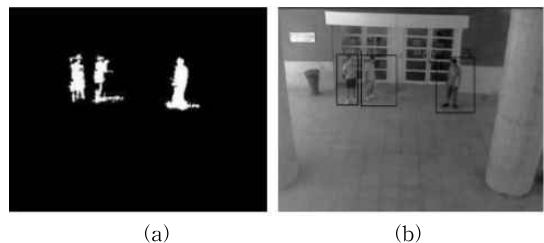


Fig. 3. An object detection using frame difference images. (a) Difference areas (b) Detected objects.

다해상도 분석 작업을 수행하는 것이 보편적이다[9]. 피라미드 형태로 축소된 영상들은 원본에 비해 매우 작은 크기이므로 전체 영상의 변화를 쉽게 관측할 수 있고 decimation과정에서 Gaussian blurring을 수행하므로 잡음에 강한 특징을 가지기 때문에, 최근 고해상도 동영상을 활용하는 분석시스템에서는 역 피라미드 기법에 의한 저해상도 변환과 프레임 차이를 이용한 움직임 검출 방식이 많이 활용되고 있다[8].

3. 제안한 방법

기존 어안렌즈 감시 카메라는 왜곡이 보정되지 않은 360도 뷰나 전체영상을 표준 화각에 맞게 왜곡 보정을 수행한 후 디스플레이하는 기능(다분할/파노라마 모드)을 제공하고 있다. 360도 무보정 영상을 이용할 경우에는 영상획득이 가장 빠르지만, 왜곡 보정을 사용자가 수행해야하는 부담이 있기 때문에 보통 제조사가 제공하는 카메라 내장 보정기능을 주로 이용한다. 이 경우 카메라 내부에서 하드웨어적으로 일괄 보정되어 전송되며, Fig. 4(b)에서 보는 것과 같이 보정 후 검은 영역이 없는 최소 사각형으로 출력하기 때문에 가장자리 부분이 많이 제거된다[10].

또한, 대부분의 어안렌즈 카메라는 보정된 영상의 임의 영역을 선택적으로 접근할 수 없기 때문에 영상 전체를 전송받아야 하는 문제도 있다. 그 결과, 제공되는 시스템의 하드웨어 및 네트워크 시스템의 성능이 낮은 경우 많은 부하가 걸릴 수밖에 없는 형태이다. 그러므로, 제안하는 방식은 Fig. 5에 나타난 것과 같이 영상획득이 가장 빠른 360도 고해상도 영상을 입력받은 후, 취급이 편리하고 잡음이 제거된 저해상도 영상으로 변환한다. 그리고 배경이 고려된 움직임 검출을 통해 변화가 발생된 부분만 정확한 왜곡보정 연산을 수행하여 일반 카메라로 촬영한 것과 같은

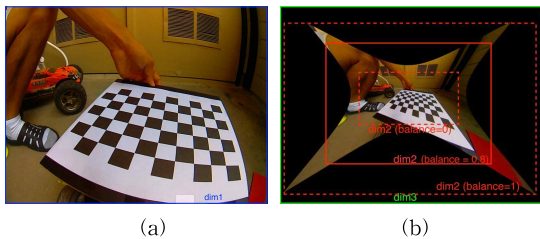


Fig. 4. Conventional lens calibration. (a) Input image (b) Compensated image.

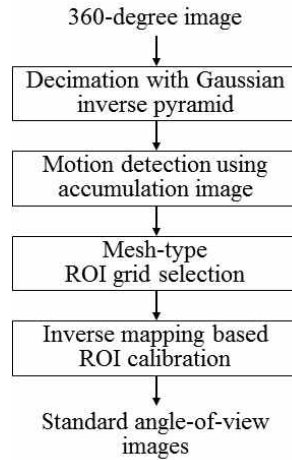


Fig. 5. Block diagram of the proposed algorithm.

표준화각의 영상을 빠르게 추출하는 방식을 제안한다. 또한, 왜곡보정의 병렬화가 용이하도록 동일크기의 격자기반 영상 보정 방식을 채택한다.

3.1 영상 축소

감시 시스템에서 주요 관심 객체는 사람이며, 대부분의 증강현실 및 인식 시스템에서 사람을 인식하는 최소 객체의 폭을 HD(2백만 화소)이하 급 영상에서 32-화소 이상으로 제한하는 것이 보편화되어있기 때문에, 제안한 기법에서 사용하는 9백만 화소의 원본영상기준으로 40-화소 이상의 차이 값을 의미 있는 움직임이라고 설정하였다. 이미지 피라미드를 통해 스케일을 낮출 때 한 옥타브 당 1/2크기로 Fig. 6과 같이 축소되며, 역 피라미드 방식은 가우시안 블

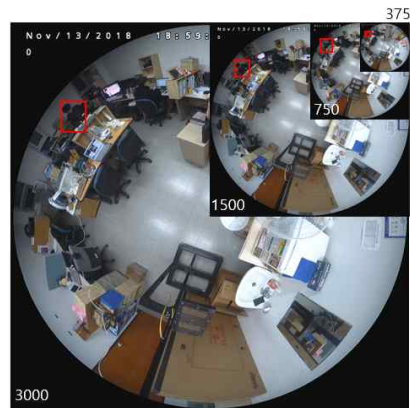


Fig. 6. Image decimation using the inverse pyramid.

러링을 통해서 LPF(저역통과여파기) 처리되기 때문에 저해상도에서는 작은 변화와 잡음이 제거되고 유의미한 큰 변화를 검출할 수 있다.

3.2 누적 영상을 이용한 움직임 검출

프레임의 차 영상을 이용한 전통적인 움직임 검출은 프레임 간 시간 간격에 비해 움직이는 객체가 너무 느릴 경우 움직임을 제대로 검출하지 못하고 잡음에 민감한 단점을 가지고 있기 때문에, 이들 문제를 해결하려면 많은 비용이 소요된다. 또한, 배경성분을 제거하고 변화를 검출하면 일반적으로 성능을 높일 수 있지만, 배경정보 갱신이 복잡하다는 단점도 있다 [11]. 그러므로, 제안한 방법에서는 단순하면서 좋은 성능의 움직임 검출을 위해서, 일정시간의 입력 프레임을 누적한 이동평균(moving average) 영상을 배경성분으로 설정하여 사용하였으며, 제안한 기법에서는 과거 5개 프레임을 묶어 누적하였다. Fig. 7은 움직임 검출 블록도를 보여주며, Fig. 8은 기존 차영상과 누적평균영상을 이용한 차 영상 결과를 보여준다. 여기서, 누적 평균을 활용하는 경우의 움직임 검

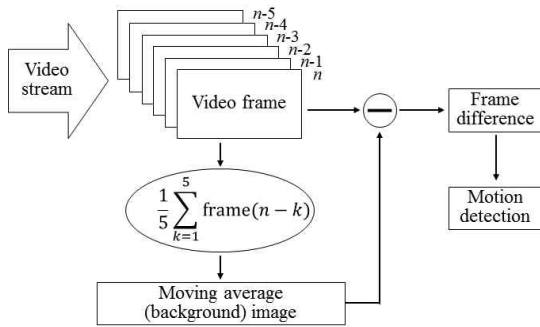


Fig. 7. Block diagram for the motion detection.

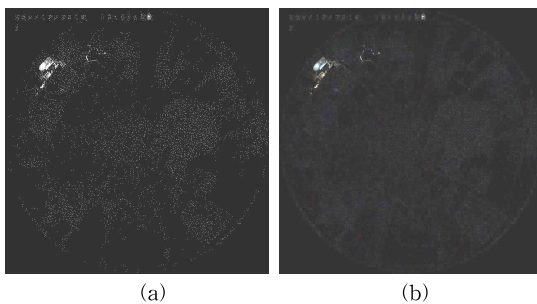


Fig. 8. Two frame difference images. (a) Conventional (b) Using the moving average.



Fig. 9. Conventional ROI selection. (a) Selected area (b) The calibrated image.

출 결과가 우수함을 알 수 있다.

3.3 격자형 ROI

기존 어안렌즈 카메라의 왜곡 보정 기능은 전체 영역에 대한 보정 이외에도 Fig. 9에서처럼 사용자 입력에 따라 임의의 위치에 대해 보정할 수 있다. 그러나 관심영역 변경이 이루어질 경우 새 영역에 대한 좌표계산의 변경뿐만 아니라 위치에 따라 다른 비선형적 렌즈곡률 변화를 적용한 보정이 이루어져야 하므로 속도가 느리다. 이러한 단점을 해결하기 위해 제안한 방법에서는 Fig. 10에서와 같이 미리 $N \times N$ 격자를 설정한 다음, 각 고정된 격자에 대한 보정 좌표를 초기에 계산해 놓음으로써 움직임이 발생한 격자의 보정좌표를 계산없이 참조하도록 하였고, 각 격자간 처리는 독립적으로 수행될 수 있기 때문에 효율적인 병렬화가 가능하도록 할 수 있다. 격자의 수는 사용자가 설정할 수 있으며, 세밀한 감지가 필요하거나 병렬 코어의 개수에 따라 조절가능하다.

3.4 왜곡 보정

제안한 기법은 Fig. 11과 같이 어안렌즈에 적합한 equidistant 투영모델의 f-theta 사상 함수를 사용한 왜곡보정 기법을 적용하였다[12]. 여기서, 높이 r , 초

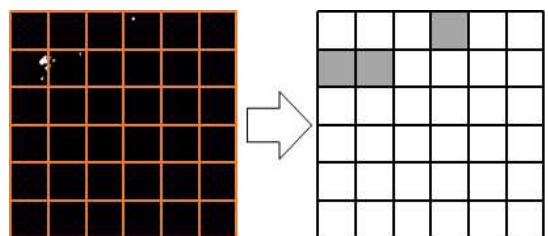


Fig. 10. Predefined ROI mesh and motion selected grids.

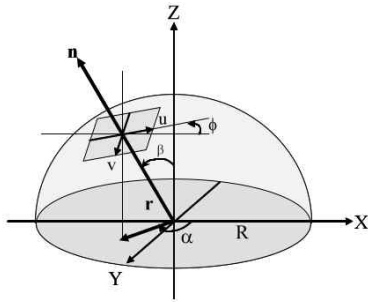


Fig. 11. Equidistant projection model.

점 거리 f 사상하고자 하는 보정좌표 중심과 구 중심의 각도를 β 로 하였을 때, $r = \beta R$ 가 되고, 세계 좌표계 X, Y, Z와 어안렌즈 이미지 평면에서의 왜곡된 좌표 u, v 및 방위각 α , 회전각도 ϕ 로 정리하면 Eq. (2)와 같다.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = R_z(-\alpha)R_x(-\beta)R_z(-\phi) \begin{pmatrix} u \\ v \\ w \end{pmatrix} \quad (2)$$

각 회전행렬을 정리하면 다음과 같다.

$$R_z(-\alpha)R_x(-\beta)R_z(-\phi) = \begin{pmatrix} \cos\alpha - \sin\alpha 0 & 1 & 0 & 0 \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

$$\{R_z(-\alpha)R_x(-\beta)R_z(-\phi)\}^{-1} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} \quad (4)$$

왜곡 보정 방식은 Eq. (4)에서와 같이 보편적으로 많이 활용되는 역방향 사상(inverse mapping)기법을 이용하였다. 이는 보정 영상좌표를 정수형태로 변화시켜 회전행렬의 역변환을 적용하면 왜곡입력 영상좌표의 위치를 알 수 있기 때문에, 단순히 입력된 영상의 해당 픽셀값을 가져오면 된다.

4. 실험 결과 및 고찰

제안한 방식은 CUDA의 활용도를 최대화한 방식에서 구현하기 위해서 미리 만들어진 “drop-in” 가속 GPU 라이브러리나 OpenACC 명령어를 사용하지 않고 직접 CUDA C/C++ 언어 및 Nvidia C/C++ 컴파일러를 사용하여 Linux OS에서 구현하였으며, 성능 평가를 위해 사용된 컴퓨터는 주기억장치 16GB, CPU i7-8700이며, 어안렌즈 카메라 정보는 Table. 1에서와 같다.

Table 1. Camera specification [3]

Model	WV-SFN480
Image sensor	1/2" CMOS
Video Codec	H.264/JPEG
Resolution	9MP(3K×3K)
Frame rate	15 fps
Lens	f=1.38 mm, F1.9

실험에 사용된 ROI 메시 격자는 3×3(격자당 영상 크기는 1,200×1,200)로 설정하였고 카메라는 넓이 6 m², 높이 2.8 5m 방 중앙에 설치하였다. Fig. 12는 원본 영상, 변화 격자 및 격자별 왜곡이 보정된 결과

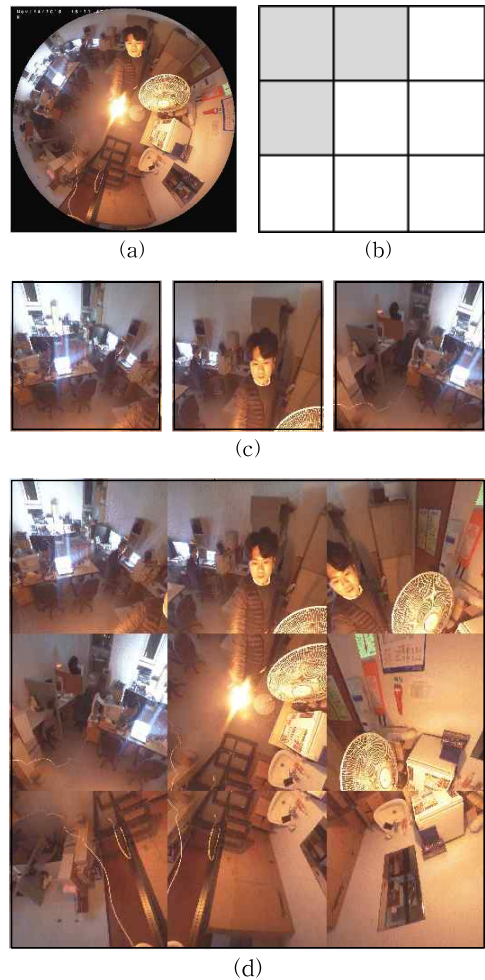


Fig. 12. Experimental results. (a) Input image (b) Motion detected grids (c) Standard view images of each motion grid (d) All grid view

영상을 보여준다.

제안한 격자형 ROI 보정방식의 장점은 전체 영상을 한꺼번에 처리하지 않기 때문에 효율적인 구역연산이 가능하다는 것이다. 왜냐하면, 어안렌즈는 특성상 정방형 이미지 센서에 원형으로 영상이 취득되기 때문에 영상정보가 없는 모서리 부분의 격자는 고려하지 않아도 되며, 움직임이 검출된 격자에서만 보정연산을 처리하면 된다. 그 다음 장점은 각 격자들은 독립적이므로 병렬연산 적용이 수월하다는 것이다. 대부분의 왜곡보정 알고리즘은 최초 계산시 좌표를 계산해두고 참고하는 방법을 사용하지만, 계산시간을 비교하기 위해서 기존 왜곡보정 기법(OpenCV 3.2.x의 fisheye::undistortedImage)과 제안한 기법의 병렬화가 가장 작은 3×3 그리드를 사용하였을 때의 시간을 비교하였다. 제안한 기법은 이미지 피라미드를 구성하고 차 영상을 계산하는 시간을 포함하였다. 실험결과 제안한 방식의 계산시간이 기존방식의 40%로 대략 60%의 계산 이득이 있었다.

기존 보정 방식(cv::fisheye::undistortedImage) : 882ms

제안한 보정 방식(3×3 grid, CUDA C) : 352ms

또한 격자형 방식을 병렬코어에 적용하였을 때 계산감소 효과를 확인하기 위해 보정방식은 Nayer[12] 방법을 채택하여 Fig. 13에 계산시간을 나타내었으며, 병렬연산은 CUDA v10.0/GPU GTX 1050을 사용하였다. 이 결과로 격자형 기법이 왜곡보정에 효과적이라고 할 수 있다. 제안한 방식의 격자를 7×7~3×3으로 변화시켜, 좌표참조방식으로 실험한 결과 1 격자에서만 움직임이 발생된 경우부터 전체격자 모두를 보정한 경우의 계산시간이 14~43 ms로 나타났

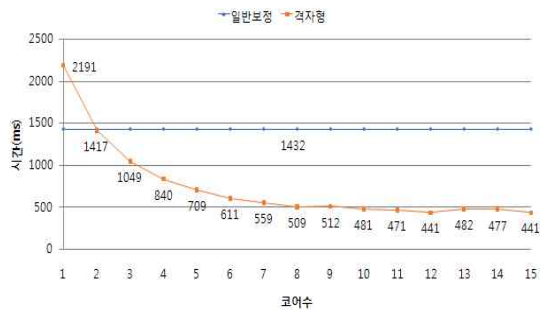


Fig. 13. Computation time.

으며, 영상획득에 병목현상이 없다면 최소 23 fps 이상을 처리할 수 있었다. 끝으로 제안한 방식의 마지막 장점은 보정시 제거되는 사각영역을 최소화할 수 있다는 것이다. 기존의 방식은 앞서 Fig. 4(b)에서 본 것과 같이 오목렌즈로 왜곡보정된 별모양의 방사형 영상을 한 개의 사각형으로 자르기 때문에 제거되는 영역이 모서리 부분에서 크게 나타나지만, 제안한 방식은 Fig. 12에서 본 것과 같이 9~49개의 격자로 나누어 보정되기 때문에 제거되는 영역을 줄일 수 있어 병렬연산뿐만 아니라 고려영역에서도 기존 방법보다 효율적이라고 할 수 있다.

5. 결론

본 논문에서는 감시 시스템의 효율성을 높이기 위해 고해상도 어안렌즈 영상을 이용한 격자형 ROI 기반의 표준 화각 관심영역 검출 알고리즘을 제안하였다. 제안한 기법은 감지성능을 높이기 위해 역 피라미드를 이용한 저해상도 영상변환과 움직임이 발생된 영역만을 활용하는 방식을 채택하였다. 또한, 표준 화각의 빠른 관심영역 검출을 위해 미리 설정된 격자형태의 ROI마다 보정된 일대일 사상 좌표를 계산하여 저장해 놓음으로써, 매번 변환 연산을 거치지 않고 보정된 영상을 획득할 수 있게 하였다. 그 결과 ROI 별로 병렬연산을 가능케 하여 멀티코어 프로세서 활용도를 높이는 이점을 발휘할 수 있었고, 실험에서 보정 후 사각지대의 감소효과와 코어 수에 비례하는 효율적인 부하 감소량을 보였다. 향후 움직임 추적을 이용한 객체 검출 성능을 향상시키고 물체의 크기에 따라 적응적으로 격자의 크기를 결정할 수 있는 연구를 수행할 예정이다.

REFERENCE

[1] Suk Yun Hwang, “A study on the methodology of positioning security CCTV cameras in multi-family housings : a study on the methodology of positioning fish-eye lens CCTV cameras,” Master’s Thesis of Sungkyun-kwan University, 2015.

[2] Tang Quang Hieu, Sungho Yeon, and Jaemin Kim, “Deep Learning based Object Detector

for Vehicle Recognition on Images Acquired with Fisheye Lens Cameras,” *Journal of Korea Multimedia Society*, Vol. 22, No. 2, pp. 128–135, 2019.

[3] Panasonic IP Camera, WV-SFN480, <https://security.panasonic.com/products/wv-sfn480>, (accessed March 21, 2020.)

[4] Jina Kang, Byungguk Kim, and Jaemin Park, “The Research for the Wide-Angle Lens Distortion Correction by Photogrammetry Techniques,” *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, Vol. 26, No. 2, pp. 103–110, 2008.

[5] F. Devernay and O. Faugeras, “Straight Lines Have to Be Straight: Automatic Calibration and Removal of Distortion from Scenes of Structured Environments,” *Machine Vision and Applications*, Vol. 13, No. 1, pp. 14–24, 2001.

[6] Z. Zhang, *A Flexible New Technique for Camera Calibration*, Microsoft Research Technical Report, MSR-TR-98-71, 2008.

[7] J. Kannala, “A Generic Camera Model and Calibration Method for Conventional, Wide-angle, and Fish-eye Lenses,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 28, No. 8, pp. 1335–1340, 2006.

[8] G. Piriou, P. Boutheymy, and J.F. Yao, “Recognition of Dynamic Video Contents With Global Probabilistic Models of Visual Motion,” *IEEE Transactions on Image Processing*, Vol. 15, No. 11, pp. 3417–3430, 2006.

[9] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, “Pyramid Methods in Image Processing,” *Radio Cooperation of America Engineer*, Vol. 29, No. 6, pp. 33–41, 1984.

[10] Ju-Hong Shin, “Non-metric Fish-eye Image Correction Using Ellipsoid Geometric Information,” Master’s Thesis of Kyungpook Nat’l University, 2005.

[11] G. Thapa and K. Sharma “Moving Object Detection and Segmentation Using Frame Differencing and Summing,” *International Journal of Computer Applications*, Vol. 102, No. 7, pp. 20–25, 2014.

[12] S.K. Nayar, “Non-metric Calibration of Wide-angle Lenses and Polycameras,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, pp. 1172–1178, 1999.



류 아 침

2017년 금오공과대학교 컴퓨터공학과 공학사
 2019년 금오공과대학교 컴퓨터공학과 공학석사
 관심분야 : 컴퓨터비전, 영상처리, 머신 러닝 임베디드시스템 및 GPU 컴퓨팅 등



한 규 필

1993년 경북대학교 전자공학과 공학사
 1995년 경북대학교 전자공학과 공학석사
 1999년 경북대학교 전자공학과 공학박사

2000년 3월 ~ 현재 금오공과대학교 컴퓨터공학과 교수
 관심분야 : 컴퓨터비전, 영상처리, 컴퓨터그래픽스, 증강 및 가상현실