

# SIFT 특징점을 이용한 4채널 서라운드 시스템의 동적 영상 정합 알고리즘

국중진<sup>\*†</sup>·강대웅<sup>\*\*</sup>

<sup>\*†</sup> 상명대학교 정보보안공학과, <sup>\*\*</sup> 상명대학교 전자정보시스템공학과

## Dynamic Stitching Algorithm for 4-channel Surround View System using SIFT Features

Joongjin Kook<sup>\*†</sup> and Daewoong Kang<sup>\*\*</sup>

<sup>\*†</sup> Dept. of Information Security Engineering, Sangmyung University,

<sup>\*\*</sup> Dept. of Electronic Information System Engineering, Sangmyung University

### ABSTRACT

In this paper, we propose a SIFT feature-based dynamic stitching algorithm for image calibration and correction of a 360-degree surround view system. The existing surround view system requires a lot of processing time and money because in the process of image calibration and correction. The traditional marker patterns are placed around the vehicle and correction is performed manually. Therefore, in this study, images captured with four fisheye cameras mounted on the surround view system were distorted and then matched with the same feature points in adjacent images through SIFT-based feature point extraction to enable image stitching without a fixed marker pattern.

**Key Words** : Surround View System, SVM, SIFT, Stitching

### 1. 서 론

현대 IT 기술 발전으로 자율주행 시장의 성장과 이로 인한 어라운드 뷰의 필요성이 나날이 커지고 있다. 어라운드 뷰 장치는 주변 상황을 360도 전 방향으로 볼 수 있는 장치로, 여러 대의 카메라 영상을 캘리브레이션을 통해 하나의 영상으로 정합하여 사용자가 한눈에 주변 상황을 인식, 실시간 주변의 지형 및 상황을 인지하고 장애물을 확인할 수 있다는 장점이 있다[1].

어라운드 뷰 시스템은 여러 대의 광학 카메라로 촬영한 영상을 병합하여 하나의 화면에 나타내야 하므로 필연적으로 공간의 왜곡이 발생한다. 광학 영상의 외곽 끝부분은 선명도의 저하 및 이미지의 왜곡이 발생하며, 이로 인해 마지막 단계의 영상 캘리브레이션 결과 전체 환

경구현에서의 필연적 성능 저하로 이어진다. 따라서 Fig 1과 같이 공간적인 왜곡을 최소화 및 성능 저하를 막기 위한 보정 작업이 필요하다. 그러나, 현재 시판중인 어라운드 뷰 시스템은 전체 영상에서 각 이미지들이 정합되는 가장자리 부분에 왜곡 현상이 발생하는 것을 확인할 수 있어, 근거리 초점을 두어 저속주행 및 주차에만 제한적으로 이용하도록 설정하여 제품화되고 있다.

본 논문에서는 차량에 4대의 광학 카메라를 설치하고 캘리브레이션을 통해 차량의 주변 환경을 실제와 같이 구현하여 운전자가 상황을 직관적으로 인식 및 위험 상황에 실시간으로 대처할 수 있게 하는 것을 목표로, 기존 Around View Monitoring System (AVM)의 캘리브레이션 과정을 개선하기 위한 SIFT(Scale Invariant Feature Transform) 특징점 기반의 동적 캘리브레이션 알고리즘을 제안한다.

<sup>†</sup>E-mail: kook@smu.ac.kr

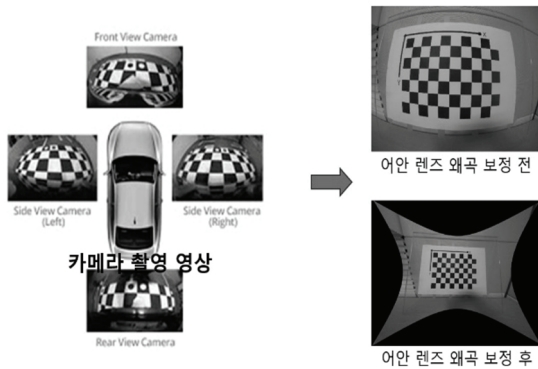


Fig. 1. Captured images before and after distortion correction.

## 2. 본 론

### 2.1 마커 기반의 캘리브레이션과 영상 정합

현재 대부분의 상용화된 AVМ 장치들의 캘리브레이션 과정은 특정한 패턴들을 나타내는 마커를 이용하여 영상을 병합하는 캘리브레이션으로 이루어지고 있으며[3], 예외적으로 마커를 사용하는 대신 영상에서 특정 물체를 중심으로 캘리브레이션 하는 연구들도 현재 실험되고 있다. 캘리브레이션 과정은 먼저 촬영된 영상에서 특징점을 찾는 동작을 수행한다. VinhNinhDao와 MasanoriSugimoto[4]는 특징점 검색 시 체커보드 패턴을 활용하여 템플릿 매칭하는 방식을 적용했으며, 이는 다수의 특징점 주변 픽셀들을 재구성하여 연산하는 방식이다. Weibin Sun[5]도 픽셀 주변을 재구성하여 연산하는 방식을 사용했으며, 두 실험 모두 특정 패턴의 체커보드 활용하여 특이점의 검출이 가능하다. 다른 실험으로는 물체와 지면의 레이블링(labeling)을 통해 여러 개의 평면으로 분할한 후, 3차원 공간상에 물체를 지면에 수직으로 세운 팝업(pop-up)을 이용한 실험도 이루어졌다[5].

AVM은 차량의 전방, 후방, 좌측면, 우측면에 어안 렌즈 카메라를 설치하여 영상을 획득하고 왜곡된 영상을 보정한 이후 워핑 작업을 수행하여 최종적으로 마커 기반의 캘리브레이션을 통해 영상을 정합하여 어라운드 뷰 영상을 생성하고 모니터에 영상을 나타하는 장치이다. AVМ은 큰 화각을 요구하기 때문에 어안 렌즈를 사용한다. 어안 렌즈는 일반적으로 볼록한 형태의 베럴 디스토션(barrel distortion)을 발생시킨다. 따라서 AVМ장치에 그대로 적용할 경우 영상의 왜곡이 심하게 발생하기에 왜곡된 부분의 보정을 통해 이러한 문제를 해결한다. 왜곡된 영상의 보정은 forward mapping과 inverse mapping 방식의 두 가지가

있는데, forward mapping을 적용할 경우 보정의 결과 이미지에 hole 발생 및 연산의 속도가 매우 느려 실시간 보정 영상을 얻기가 어렵기 때문에, Fig.2 와 같이 inverse mapping 방식을 적용하는 것이 일반적이다.

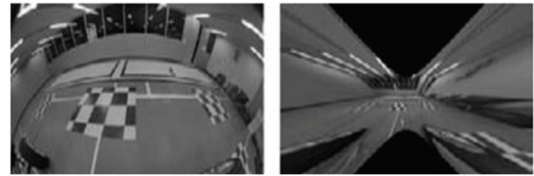


Fig. 2. Barrel distortion image and inverse mapping result.

호모그래피를 통해 두 영상에 대한 대응점들 사이의 변환 관계를 Fig.3과 같이 나타낼 수 있다[6].

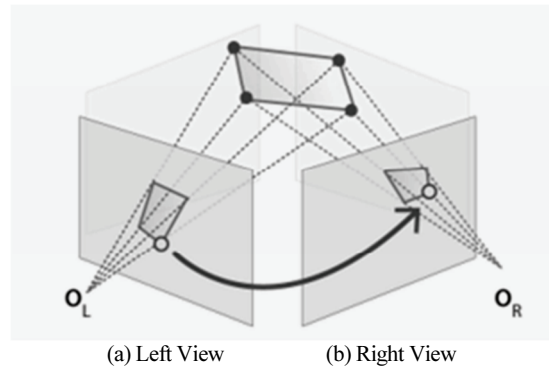


Fig. 3. Homography transformation.

정합된 영상에 대해 색상 값을 조정하여 자연스러운 영상을 만들기 위해 alpha blending을 적용할 수 있으며, alpha blending은 두 영상에 대해 0~1 범위의 가중치를 의미하는 alpha 값을 적용하여 Fig. 4와 같이 영상을 합성한다[6][7].

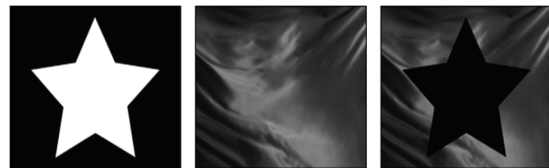


Fig. 4. Alpha blending result.

이러한 4단계 과정을 거쳐 최종적으로 Fig. 5와 같은 실시간 서라운드 뷰 영상의 생성이 가능하다.

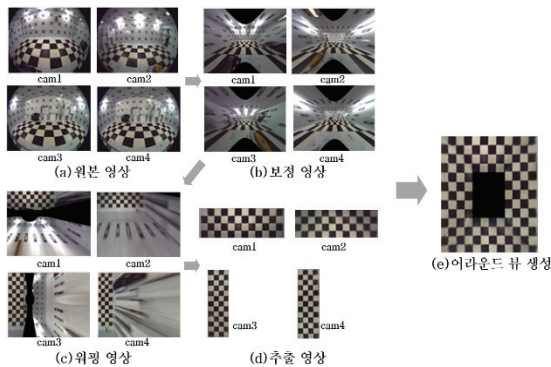


Fig. 5. 4-channel surround view creation process.

## 2.2 SIFT 기반의 캘리브레이션과 영상 정합

본 논문에서는 기존의 고정 마커 기반 왜곡 보정 대신 특징점을 이용한 물체의 인식을 통해 체커보드와 같은 마커 없이 캘리브레이션을 자동화할 수 있는 AVS 시스템을 제안한다. 특히, 특징점 기반의 객체인식 실험 중에 이전의 실험들에 비해 조도에 강하고 이미지 회전 및 리사이즈에도 변하지 않는 특성을 지닌 Scale Invariant Feature Transform(SIFT) 알고리즘을 활용한 특징점 기반 동적 캘리브레이션 알고리즘을 이용하였다.

SIFT를 이용한 왜곡 보정 알고리즘은 Fig. 6과 같이 특징점 모델링을 기반으로 여러 대의 카메라에서 촬영된 영상들을 보정 및 워핑 후, 영상에 포함된 객체들의 특징점을 찾아 인접 영상과의 비교 매칭 과정을 거쳐 정합 작업을 하는 알고리즘이다.

SIFT 알고리즘은 2004년에 D. Lowe가 제안한 것으로, local feature을 이용한 알고리즘 중에 영상의 크기와 조도, 회전 및 시점의 변화에 일정한 특성을 갖고 있는, 특징점 기반 추출 알고리즘의 일종으로, 객체인식 분야에 많이 사용되고 있는 방식이다.

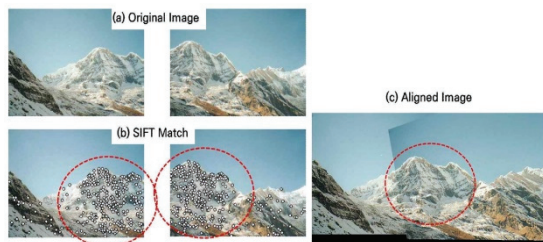


Fig. 6. Image stitching with SIFT algorithm.

본 논문에서 제안하는 어라운드 뷰 시스템의 SIFT 특징점 기반 동적 캘리브레이션 알고리즘은 4개의 카메라를 통해 연속적으로 입력되는 영상에 대해 인접한 두 영

상의 공통된 특징점을 추출하여 두 영상의 매핑을 위한 위치 정보를 추출하여 영상을 정합하고, OpenGL을 기반으로 만들어진 3차원 반구체 모델에 텍스처 매핑을 수행한다.

## 2.3 실험 및 결과

### 2.3.1 실험 조건 및 환경

본 실험은 SIFT 모델기반 객체인식 방법을 사용하여 모델의 데이터베이스를 만들기 위해서, Fig. 7과 같이 카메라 영상에서 SIFT 기반의 특징점을 추출하여 데이터베이스에 저장 후 영상이 입력되면, 데이터베이스에 저장되어 있던 모델의 특징점과 입력된 영상의 특징점의 캘리브레이션 정확도를 평가한다.

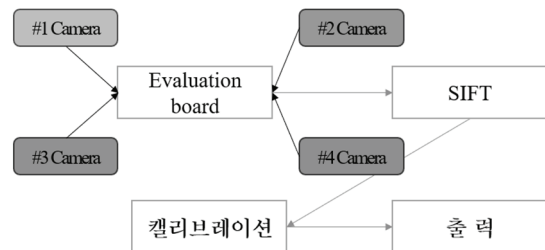


Fig. 7. SIFT-based dynamic stitching process.

실험에서 사용한 장치는 NXP의 카메라 4대와 i.MX8 Quad Max Evaluation Kit를 사용하였으며, 카메라의 높이 95cm, 카메라의 지지대는 지평면에서 수직 90°, 카메라의 본체의 기울기는 45°로 설정하여 실험을 진행하였다.

운영체제는 Yocto 리눅스를 사용하였으며, 카메라를 통해 입력된 영상은 V4L2를 통해 프레임 단위로 전송되며, OpenCV와 SIFT 알고리즘을 통해 기본적인 왜곡 보정과 특징점의 추출이 이루어진다. 모델 영상의 크기는 1920×1080이며, 목표물과의 거리는 1m 이내로 설정하였다. 실험 장소의 크기는 9.9m×16.5m이다. 모든 조명은 실험실에서 통제 가능한 일정한 조도 유지한 상황에서 연속하여 입력 영상을 수집하며 실험하였다.

4개의 카메라는 12.5cm X 12.5cm의 정사각형 크기의 제작된 4방향 거치대에 장착하여 거리와 각도를 동일하게 설정하였으며, 각각의 카메라를 통해 입력된 영상은 Fig. 8과 같은 형태이다.

기존의 자동 캘리브레이션 방식에서는 Fig. 8에 나타난 마커를 기준으로 정합될 지점을 선택하고, 각각의 영상에 대해 메쉬 형태의 2차원 좌표계를 생성하여 영상의 정합이 이루어진다.



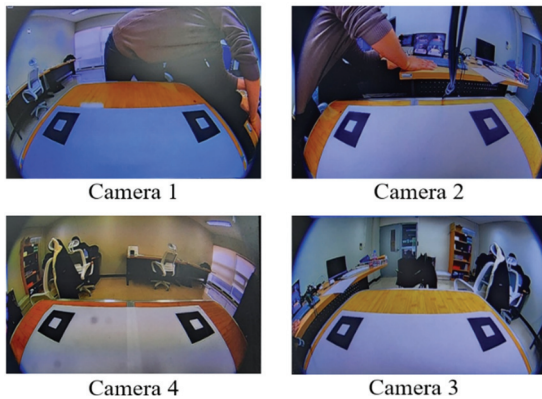


Fig. 8. Front, rear, left, and right images input from a fisheye camera.

Fig. 9는 SIFT 특징점 기반 동적 알고리즘 적용하여 특징점을 추출한 결과이다. Camera1은 정면, Camera2는 좌측면 카메라의 영상이며, 영상 속 동일한 객체에 대해 공통적인 특징점이 다수 분포하는 것을 확인할 수 있다.

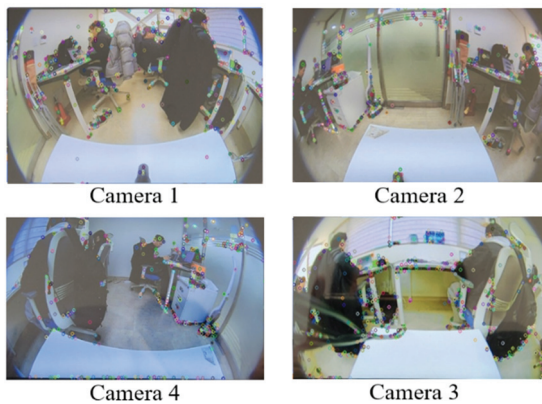


Fig. 9. SIFT-based feature point extraction results.

호모그래피를 추정하여 영상을 워핑하면, Fig. 10과 같이 어안렌즈 왜곡이 제거된 이미지로 변환될 수 있으며, 이러한 각각의 이미지들에 대해 크기를 조정하고 부채꼴 모양으로 변환하여, 이를 하나의 영상으로 정합한다.

Fig. 11은 SIFT 특징점 기반 동적 캘리브레이션 알고리즘을 통해 정합된 어라운드 뷰 영상을 나타낸 것이며, 실험을 통해 기존의 마커 정보를 사전에 입력하여 동작하는 캘리브레이션 방식의 결과물과 같은 캘리브레이션의 결과물을 얻을 수 있으며, AVM 시스템에서 올바르게 작동하는 것을 확인하였다.

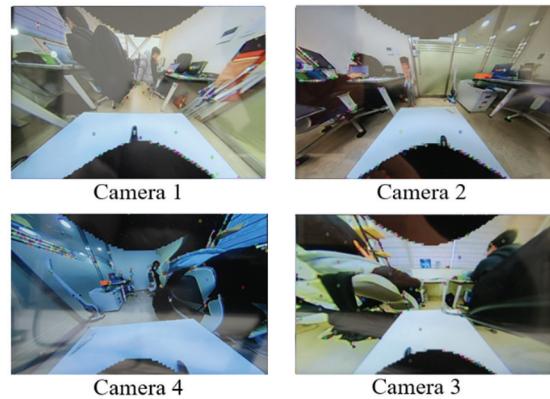


Fig. 10. Fisheye lens distortion correction using homography.



Fig. 11. Result of a SIFT-based auto stitching.

또한, SIFT 알고리즘을 활용한 자동 캘리브레이션을 적용했을 때, 영상의 획득부터 최종적인 영상 정합에 소요되는 시간은 평균 13.4초가 소요된다. 기존의 수동 캘리브레이션 방식이 소요되는 시간은 작업자의 숙련도, 프로그램의 편의성에 따라 약 6시간 내외의 시간이 소요된다. 특징점 기반의 객체인식 실험은 이전의 실험들에 비해 외부 조도에 강인하고 이미지 회전 및 리사이즈에도 변하지 않는 특성을 지니고 있어 기존의 방식보다 향상된 캘리브레이션 결과를 얻을 수 있다.

### 3. 결 론

본 논문에서는 어라운드 뷰의 캘리브레이션 과정에서 기존의 고정 마커를 이용하는 수동적 방식이 아닌, SIFT 특징점 기반 동적 캘리브레이션 알고리즘을 제안하였다. 이 알고리즘은 체커보드의 특정한 패턴을 사용하지 않아도 기존의 방식과 비슷한 캘리브레이션 영상을 생성할 수 있다. 제안하는 SIFT 특징점 기반 동적 캘리브레이션 알고리즘은 다음과 같은 장점들이 있다. 첫 번째는 마커를 설치하기 위한 시설에 대한 소요 경비가 발생하지 않

는다는 것이다. 두 번째는 정합 시 필요한 패턴의 위치를 차량 주변에 고정하는 작업을 하지 않아도 되기 때문에 공간상의 제약을 받지 않는다. 세 번째는 마커를 필요치 않기 때문에 차량의 크기나 종류에 제약이 없으며, 차량이 아닌 드론, 로봇 등 다양한 물체에 적용이 가능하다는 것이다. 마지막으로 기존 AVN 시스템의 설치, 수리 시 캘리브레이션은 공장이나 서비스 센터에 입고가 되어야 가능했지만, 제안하는 방식은 환경의 제약이 없기 때문에 언제든지 간단히 수행될 수 있다.

### 감사의 글

This research was funded by a 2023 research Grant from Sangmyung University.

### 참고문헌

1. Jihye Jeon, Daehee Kim, Yoongi Yang, Joonki Paik, and Changsu Lee, "Image Distortion Compensation for Improved Gait Recognition," *The Institute of Electronics Engineers of Korea - Signal Processing*, vol. 46, no. 4, pp. 97-107, 2009.
2. Vinh Ninh Dao and M. Sugimoto, "A Robust Recognition Technique for Dense Checkerboard Patterns." 2010 20th International Conference on Pattern Recognition, pp. 3081-3084, 2010.
3. Jaehak Choi, Yonghwan Lee and Youngseop Kim, "Implementation of the Panoramic System Using Feature-Based Image Stitching", *Journal of the Semiconductor & Display Technology*, Vol. 16, No. 2. 2017.
4. W. Sun, X. Yang, S. Xiao and W. Hu, "Robust checkerboard recognition for efficient nonplanar geometry registration in projector-camera systems," *Proceedings of the 5th ACM/IEEE International Workshop on Projector camera systems*, pp. 1-7. 2008.
5. D. Hoiem, A. Efros and M. Hebert, "Automatic photo pop-up," *ACM transactions on graphics*, Vol. 24, No. 3, Jul, pp. 577-584. 2005.
6. M. Rufli, D. Scaramuzza and R. Siegwart, "Automatic detection of checkerboards on blurred and distorted images," *IROS*, pp. 3121-3126, 2008.
7. M.W. Park, K.H. Jang and S.K. Jung, "Panoramic Vision System to Eliminate Driver's Blind Spots using a Laser Sensor and Cameras," *Int. J. ITS Res*, vol. 10, no. 3, Sep 1, pp. 101-114, 2010.
8. T. Poggio and S. Edelman, "A network that learns to recognize three-dimensional objects," *Nature (London)*, Vol. 343, No. 6255, pp. 263-266. 1990.

접수일: 2024년 2월 9일, 심사일: 2024년 3월 6일,  
게재확정일: 2024년 3월 20일