



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

博士學位論文

파노라마 영상 생성을 위한
효과적인 스티칭 방법

濟州大學校 大學院

컴퓨터工學

金 侬 贊

2017 年 02月

파노라마 영상 생성을 위한 효과적인 스티칭 방법

指導教授 李 尙 俊

金 侏 贊

이 論文을 컴퓨터工學 博士學位 論文으로 提出함

2016 年 12 月

金侏贊의 컴퓨터工學 博士學位 論文을 認准함

審査委員長

郭 鎬 榮

委 員

이 삼 준 이삼준

委 員

김 드 현 김드현

委 員

김 한 일 김한일

委 員

박 추 희 박추희

濟州大學校 大學院

2016 年 12 月

An Effective Image Stitching Method to Make Panoramic Images

Hyu-Chan Kim
(Supervised by professor Sang-Joon Lee)

A thesis submitted in partial fulfillment of the requirement for the
degree of Doctor of Computer Engineering

2016. 12.

This thesis has been examined and approved.

Thesis director, Hoyoung Kwak
Thesis director, Sang Joon Lee
Thesis director, Dohyeun Kim
Thesis director, Haril Kim
Thesis director, Chung Hee Park

December 2016

Department of Computer Engineering
GRADUATE SCHOOL
JEJU NATIONAL UNIVERSITY



목 차

그림목차	iii
표 목차	vi
국문초록	viii
영문초록	x
I. 서론	1
1. 연구배경	1
2. 연구목적	3
3. 연구내용	4
4. 논문구성	6
II. 관련연구	7
1. 스티칭 방법	9
2. 이미지 스티칭 종류	14
3. 이미지기반 VR	25
4. 기존 방식의 문제점	34
III. 가변적응성 필터를 적용한 이미지 스티칭	36
1. 이미지 스티칭 과정	36
2. 특징점 추출 및 정합	47

3. 360도 파노라마(VR)영상촬영시스템의 동기화	50
IV. 실험 및 결과	55
1. 실험환경	55
2. 실험데이터	56
3. 기존 알고리즘과의 비교분석 결과	58
4. 동기화 스티칭 결과	65
5. 기존 소프트웨어와의 비교	69
VI. 결론	87
참고문헌	90

그림 목차

Fig. 1. 360VR seoul station	7
Fig. 2. Video acquired with handheld camera	9
Fig. 3. Stitching result of acquired image	10
Fig. 4. Stitching result of acquired image(Images with different exposure values)	10
Fig. 5. Stitching result of acquired image	11
Fig. 6. Stitching result of acquired image(Exposure and focus discrepancies) ..	11
Fig. 7. Stitching result of acquired image	12
Fig. 8. Stitching result of acquired image(After-images(ghost phenomenon)) ..	12
Fig. 9. Two acquired images (time difference occurs with a single camera)	12
Fig. 10. Stitching result of acquired image(overlapping image)	13
Fig. 11. Definition of Homography Relationship	17
Fig. 12. Homography conversion formula	18
Fig. 13. Structure of SIFT algorithm	19
Fig. 14. Structure of SURF algorithm	20
Fig. 15. FAST corner detector	21
Fig. 16. How to make an object VR image	26
Fig. 17. Panorama VR contents	27
Fig. 18. Cylindrical Panorama	28
Fig. 19. Cylinder method	28
Fig. 20. Cylindrical image making method	29
Fig. 21. Cylinder panorama construction principle and Flow Chart	29

Fig. 22. Shooting to overlap the space before and after the left, right, top and bottom	30
Fig. 23. Spherical Panorama image	30
Fig. 24. How to make an image using a general lens	31
Fig. 25. How to make images using fisheye lens	32
Fig. 26. Cubical Panorama image implemented in VR Viewer	32
Fig. 27. Flat panoramic photography	33
Fig. 28. Flat panoramic output image	33
Fig. 29. Image pyramid of SIFT algorithm and SURF algorithm	35
Fig. 30. Panoramic Image Generation Complete Flowchart	37
Fig. 31. Set redundant area of input image	38
Fig. 32. Picture of overlapped area of input image	39
Fig. 33. An example of raw image with lens distortion and its correctin	40
Fig. 34. Overlapping region of input image	44
Fig. 35. Feature Extraction Process with Variable Adaptive Filter	45
Fig. 36. Matching results for extracted minutiae	46
Fig. 37. Matching result	46
Fig. 38. Image stitching result	47
Fig. 39. Area calculation using integral image	48
Fig. 40. Multi camera synchronization system	51
Fig. 41. Simultaneous recording signal Ready (green light) and ready Recording status after the simultaneous recording signal (red)	52
Fig. 42. Rig	53
Fig. 43. 360° VR Rig	53

Fig. 44. Rig for aviation shooting	54
Fig. 45. Comparison image result image	58
Fig. 46. Total run time comparison	60
Fig. 47. Comparing university performance time	61
Fig. 48. Comparison of Practice Time	61
Fig. 49. Kumho World Education Center Lobby Performance Time Comparison	62
Fig. 50. Comparison of feature points of existing algorithms and proposed method	62
Fig. 51. Comparison of average time of total stitching	64
Fig. 52. Comparison of total matching point extraction time	64
Fig. 53. The feature points of the first image and the second image and Distance between two coordinates after stitching using homography H	64
Fig. 54. The feature points of the second image and the third image and distance between two coordinates after stitching using homography H	66
Fig. 55. The feature points of the third image and the fourth image distance between two coordinates after stitching using homography H	66
Fig. 56. Stitching result	67
Fig. 57. Comparison of existing method and proposed method RMSE	70
Fig. 58. Existing method and proposed method Control point comparison result	71
Fig. 59. A process for matching the images captured by the proposed "C- product and the proposed method	72
Fig. 60. Comparison of distance values and control points through feature points extracted from "C product" and proposed method	73
Fig. 61. A process of finding a matching point corresponding to a minutiae of each adjacent image in the photographed image	74
Fig. 62. Matching point check screen	75

Fig. 63. Result of stitching darangshi orum	76
Fig. 64. Initial screen for stitching existing method and proposed method	78
Fig. 65. Stitching screen of university main building	79
Fig. 66. Distance value and control points calculation screen	80
Fig. 67. Stitching result image	81
Fig. 68. Kumho World Education Center lobby stitching initial screen	83
Fig. 69. RMSE calculation screen	84
Fig. 70. Kumho World Education Center lobby stitching result image	85

표 목차

Table 1. 360 Panoramic system configuration	55
Table 2. Depending on the number of camera lenses	56
Table 3. Source data	57
Table 4. Comparison of experimental results	59
Table 5. Feature extraction time and total stitching time comparison	63
Table 6. Stitching error (average)	68
Table 7. Stitching error rate and control points measurement results of existing software and proposed algorithm	70
Table 8. Darangshi orum Stitching Result Data	77
Table 9. Result of stitching image of university image	82
Table 10. Kumho World Education Center lobby image stitching result Data	86

파노라마 영상 생성을 위한 효과적인 스티칭 방법

컴퓨터공학과 김 휴 찬

지도교수 이상준

최근 컴퓨터 비전 기술이 발전함과 동시에 가상현실 및 360VR 콘텐츠에 대한 고해상도 이미지 및 영상물에 대하여 사용자들의 관심이 높아지고 있다. 또한 VR 시장이 활성화되면서 많은 콘텐츠들이 4K, 8K 등 고해상도의 영상을 원하지만 카메라에서 생성된 영상을 스티칭하는 작업에 상당한 시간과 비용이 소요된다.

본 논문에서는 파노라마영상 생성을 위하여 가변 적응형 필터를 적용한 V-SURF알고리즘과 기존 알고리즘(SIFT와 SURF알고리즘)과의 연산 수행속도, 특징점 추출시간 및 총 스티칭시간을 비교 분석하였다.

실험결과에서 볼 수 있듯이 전체적인 스티칭시간이 SIFT와 SURF보다 향상되어 기존의 방법보다 향상되어 제안한 방법이 결과 이미지의 개선과 연산수행 속도에서도 향상된 것을 알 수 있었다.

360도 파노라마 영상을 생성하기 위하여 촬영시스템의 다중 카메라 동기화 시스템을 제안하였다. 여러 대의 카메라가 동일한 프로젝션 센터를 가질 수가 없고 이러한 문제 때문에 카메라간 시차가 발생할 수 있다. 시차는 영상합성을 통해서 해결할 수 없으므로 영상 획득시 최소화 할 있도록 리그를 설계하고 카메라를 리그에 장착하여 카메라를 동기화시켜 촬영되는 카메라의 동기화를 확인하였다.

또한 다중 카메라로 촬영된 360파노라마 영상 Data를 가지고 기존 소프트웨어와 제안한 방법과의 스티칭 오차(Stitching distance)와 특징점 개수(Feature point)를 비교·분석을 하였다.

실험결과를 보면 기존의 방법보다 제안한 방법이 RMSE가 낮음을 알 수 있었

다. RMSE가 낮다고 하는 것은 특징점과 대응점간의 오차(거리)가 적기 때문에 효과적인 스티칭을 할 수 있고, 빠른 연산수행 속도와 고해상도의 이미지를 얻을 수 있음을 확인하였다.

본 연구를 통해 고품질의 카메라를 여러 대 사용하여 영상의 고품질을 유지하면서도, 넓은 화각을 촬영할 수 있어, 전시장 및 대형 광고 시장 등에서 여러 사용자에게 동시에 고해상도의 파노라마 영상을 제공할 수 있을 것으로 기대된다.

주제어 : 파노라마, 360VR, 스티칭, 파노라마 영상, VR콘텐츠

ABSTRACT

An Effective Image Stitching Method to Make Panoramic Images

Kim, Hyu Chan

Department of Computer Engineering

Graduate School

Jeju National University

The consumers are becoming more interested in the recent development of Vision technology and, at the same time, virtual reality and VR 360 content, like high-resolution images and videos. Moreover, as the virtual reality market is becoming more active, a lot of contents require high-resolution images, such as 4k or 8k, however, stitching videos filmed on a camera requires a lot of time and money.

The main goal of this thesis is a comparative analysis of operating speed, feature point extraction time and a total time of stitching of V-SURF algorithm, that uses variable adaptive filters in order to make panoramic videos, and already existing SIFT and SURF algorithms.

As it is shown in the results of experiments, the general time of image stitching is improved compared to the SIFT and SURF algorithms, and since it is better than the already existing methods, this recommended method results also in image enhancement and faster operating speed.

In order to make 360-degree panoramic images, a system of synchronizing multiple cameras was suggested. Multiple cameras, however, cannot have the same center of projection, and because of that there can appear some time differences between the cameras. This problem cannot be solved by video stitching, thus in order to lessen this difference when obtaining the video rigs have to be prepared, cameras have to be installed on those rigs and synchronized. The synchronization of the cameras has to be checked.

Moreover, basing on the data obtained while filming 360-degree panoramic videos on multiple cameras, stitching distance and feature points of existing software and suggested method have been compared and analysed.

Looking at the results of the experiments, it can be said that the RMSE of the recommended method is lower than the RMSE of the existing methods. This means that the distance between the feature point and the reference point is short, and because it is an effective method of stitching, it can improve the operating speed and ensure that the images will be in high-resolution.

Through this study, it is expected that it will be possible to provide panoramic images of high resolution at the same time to various users at the exhibition hall and the large advertisement market because it can shoot a wide angle of view while maintaining the high quality of the image by using a plurality of high quality cameras.

Key word : Panorama, 360VR, Stitching, panorama image, VR Contents

I. 서론

1. 연구배경

우리는 지금 가상현실(VR: Virtual Reality)과 증강현실(AR: Augmented Reality)로 일컬어지는 혼합현실(Mixed Reality)의 시대를 살고 있다. 네트워크 환경이 4G에서 5G 시대로 빠르게 변화하면서 ‘직접 보고 듣고 느끼는 TV’를 만나게 되고, 화려한 픽셀(Painted Pixels) 시대에서 ‘생생한 비트(Tangible Bits)’ 시대로 이동하고 있다. 미래의 실감형-몰입형 장치들이 이제 현실화가 되어가고 있다[1].

VR 콘텐츠는 최근 1~2년 사이에 스마트폰 디바이스에 최적화된 VR 장비들이 등장하면서 빠른 속도로 성장하고 있다. 특히 현실감과 입체감을 요구하는 게임 시장분야에서 인기가 상당히 높다. 또한 유튜브에서 본격적인 VR 기능을 제공하고 있고, 페이스북 역시 360도 영상 지원이 되면서 VR 실감미디어 시대에 커다란 기여를 하고 있다.

실감형 콘텐츠란 단순히 화면에서 재생되는 영상, 즉 시각정보만 받는 것이 아니라 증강, 가상현실, 오감미디어, 홀로그램 등 오감이 느낄 수 있는 콘텐츠다. 영상을 보는 것이 아니라 영상 속에 내가 직접 들어간 것 같은 경험을 할 수 있는 것이 바로 실감형 콘텐츠인 것이다[2]. 실감형 콘텐츠는 영화나 게임, 오락, 군방, 교육, 의료, 자동차 등 미래 전략 분야까지 다양하게 활용이 가능하다[3][4].

현재 우리가 가장 손쉽게 접할 수 있는 VR콘텐츠는 360도 파노라마 영상이다. 현재 많은 콘텐츠들이 360도 카메라로 촬영한 파노라마 동영상으로 제작되고 있으며, 유튜브를 비롯한 다수의 온라인 동영상 서비스들이 실시간 파노라마 영상을 송출하고 있다.

사람의 눈은 최대 1억 2,000만 화소의 해상력을 가지는 반면 현재 모바일 VR의 디스플레이는 스마트폰기기의 해상도를 따르며, HMD의 대명사로 자리를 잡은 오클러스는 1,080×1,200에 불과한 해상도를 갖는다. 현재의 기기들은 어떤 방

식으로든 VR형태로 구현되더라도 몰입감이 떨어지고 멀미를 호소할 수밖에 없는 기술적 한계를 가지고 있다.

전 세계적으로 가상현실이 빠른 발전과 함께 주목을 받고 있다. 이와 함께 상용화가 많이 되고 있는 기술이, 360도 카메라이고, 또한 스마트폰의 panorama 촬영 기능이다. 360도 카메라는 근본적으로 파노라마 촬영과 같은 원리를 갖는다. 360도 카메라는 2개 이상의 렌즈를 이용해 2개 이상의 방향으로 주변을 촬영 한 후, Image Stitching 기술을 통해, 이를 연결하여, 임의의 각도로 보아도 연속된 사진을 보여주게 하는 기술이다[6].

현재의 360도 카메라는 화각(카메라의 시야)을 넓히다보니 화질이 저하되는 현상이 발생한다. 그래서 고화질의 VR 영상을 얻기 위해서는 다른 방법이 필요하다. 한 대의 카메라로 360도를 다 촬영하는 것이 아니라 여러 대의 카메라를 동원하여 촬영을 하는 것이다.

‘스티칭(Stitching)’이란 여러 대의 카메라로 찍은 영상을 이어 붙이는 것을 말한다. 올인원 카메라는 보통 오토스티칭(Auto Stitching) 기능이 있어서 쉽게 이용할 수 있는 반면, 리그를 쓸 경우에는 독립된 카메라로 찍은 영상을 붙여서 360도 영상을 구현해야 한다. VR 영상의 품질은 스티칭에 달려 있다고 할 정도로 중요한 과정이다. 이러한 스티칭 작업을 하기 위해서는 숙련도가 필요한 작업이기도 하다.

고화질 360 파노라마 영상은 넓은 화각을 가진 영상으로 현재의 카메라로는 시야각 정도의 넓은 공간을 시공간적인 왜곡 없이 촬영하기 어렵기 때문에 하나의 이상의 카메라를 동시에 촬영하는 방법이 필요하다. 그러나 이 촬영방법은 다수의 카메라가 마치 하나의 카메라로 찍는 것처럼 촬영해야 하는 문제로 인하여 발생하는 왜곡을 최소화할 수 있는 촬영 기법이 필요하게 된다. 이를 위해 여러 카메라를 하나의 카메라처럼 동일한 특성을 가지도록 멀티카메라 촬영을 통한 스티칭 기술이 필요하다[7].

최근 컴퓨터 비전 기술이 발전함과 동시에 가상현실 및 360VR 콘텐츠에 대한 고해상도 이미지 및 영상물에 대하여 사용자들의 관심이 높아지고 있다. 또한 VR 시장이 활성화되면서 많은 콘텐츠들이 4K, 8K 등 고해상도의 영상을 원하지만 카메라에서 생성된 영상을 스티칭하는 작업에 상당한 시간과 비용이 소요된다.

2. 연구목적

VR 콘텐츠는 게임 같은 엔터테인먼트 분야를 시작으로 교육이나 관광 등 다양한 분야에서 활용 가능성이 큰 디지털 산업으로 떠오르고 있다. VR 콘텐츠 역시 통상의 영상처럼 컴퓨터그래픽과 실사로 나뉘 볼 수 있다. 게임처럼 컴퓨터그래픽만으로 만들어지는 영상과 실사 화면을 촬영, 편집해 제작되는 360도 영상으로 나뉘 생각하면 된다.

360도 파노라마 영상은 최근 들어 가장 뜨거운 영상 콘텐츠 형식 가운데 하나이다. 이용자가 마우스나 손가락으로 화면을 터치하면 360도 전 방향 형태로 시청할 수 있으며, 해외 유명 여행지에 직접 가지 않더라도 VR영상을 통해 집 안에서 생동감 있는 현장을 경험할 수 있다. 음악회 또는 비보이 댄스 팀의 공연도 360도 영상으로 제작되고 있어 VR 기기만 있다면 3D 입체 화면으로 볼 수 있다. VR을 활용한 체험 영상 제작 기술이 활성화되면서 게임, 스포츠, 여행 등의 엔터테인먼트 산업을 포함, 교육, 의료, 비즈니스 등의 분야에서도 크게 활용될 것으로 전망된다[5].

최근 들어 고해상도의 DSLR 카메라가 등장하면서 이미지 화소 수는 기하급수적으로 증가하고 있다. 모니터의 경우도 해상도와 크기가 계속해서 빠른 속도로 증가하고 있다. 텔레비전은 SD(Standard Definition)에서 HD(High Definition)로 그리고 HD가 완전히 보급되기도 전에 기존 풀HD 해상도 대비 약 4배에 해당하는 차세대 초고해상도 화소 수를 가지는 UD(Ultra Definition) 시제품이 선보이고 있다. 수백 장의 사진을 이어 붙여 거대한 이미지를 만드는 실험들과 초고해상도 DSLR 카메라가 등장하고 있으며, 디지털 미디어 기술의 도움으로 입체영상이 주목을 받고 있다[8].

360파노라마 영상은 사용자의 입장에서 해당 공간에 대한 전방향성을 가진다는 특성 때문에 과거에는 가상체험을 제공하는 VR(Virtual Reality)이라는 이름으로 소개되곤 했으나 특정 하드웨어나 소프트웨어에서만 재생이 가능했던 이유로 대중화에는 실패했다. 그러나 최근 오쿨러스 리프트나 기어 VR과 같은 360° 파노라마 영상 관련 하드웨어와 소프트웨어, 그에 따른 개발 환경의 발전에 힘

입어 일반 대중들의 관심이 크게 증가하고 있다.

이처럼 대중적 관심이 빠르게 확대되는 것과는 대조적으로 360파노라마 영상 제작의 발전 단계는 국내외적으로 아직 걸음마 단계라 할 만한 수준이다.

최근 4K나 UHD와 같은 속칭 ‘해상도 싸움’이 한계에 이르렀다는 평가가 나오는 시기에, 360파노라마 영상을 바탕으로 한 인터랙티브 미디어 또는 실감미디어에 대한 기술 개발에 집중될 것이다.

따라서 현실감 있는 파노라마 영상의 생성을 위해서는 콘텐츠의 고해상도와 스티칭 편집의 빠른 처리속도이다. 최근의 VR콘텐츠 생성을 위해서는 최소 4K의 해상도가 필요하다. 이렇게 고해상도의 영상을 획득하기 위해서는 많은 시간과 노력이 필요하다. 본 논문에서는 특징점 추출 영역을 가변 형태로 클러스터화하여 추출하는 방법을 제안하고, 파노라마 영상 생성을 위하여 특징점 기반의 가변 적용형 필터를 적용한 스티칭 방법에 대하여 연구를 수행하였다. 이를 위하여 가변 적용형 필터를 적용하여 특징점을 추출하고, 기존의 알고리즘과 연산 수행 속도를 비교·분석하여 효율성에 대하여 증명하였다.

3. 연구내용

영상처리에서 스티칭 기법은 꾸준히 컴퓨터 비전에서 연구되는 분야로 두 개 이상의 겹치는 영상에서 공통적인 특징점 기술자의 대응점을 찾아 매칭하는 것으로 파노라마 영상과 같은 고해상도의 사진이나 이미지를 얻을 수 있다.[9] 주로 여러 장의 이미지를 연결하여 파노라마 이미지를 생성하는 데 쓰인다[9,10]. 이미지 스티칭 기법은 여러 장의 이미지를 연결하는 것뿐만 아니라 실시간 처리가 가능한 수준에 이르렀다. 최근에는 단순히 여러 장의 이미지를 연결하여 하나의 파노라마 이미지를 생성하는 것보다 이를 적절히 활용하는 것이 중요해지고 있다[11].

기존 파노라마 영상은 여러 대 또는 한 대로 여러 장 사진을 찍어 사진을 이어

불인 후 실린더 형태로 변환하는 방법이다[48]. 360도 파노라마 영상은 사용자가 컴퓨터 사용 시 마우스를 통하거나, 또는 스마트폰 사용 시 손가락 터치나 기기를 움직여 360도로 콘텐츠를 둘러볼 수 있는 신개념 온라인 콘텐츠라고 할 수 있다.

360도 파노라마 영상은 디지털기기가 빠르게 발전함에 따라 영상의 화질도 더욱 선명해지고 해상도도 향상되면서 더 넓은 시야를 한 화면에 기존 평면 영상과 달리 재생 도중 사용자가 상하좌우, 360도 둘러볼 수 있는 현장감을 극대화 할 수 있는 영상이다. 기존의 사용해 왔던 360도 카메라들과의 차별점은 역시 화질에 있다.

최근 3D, UHD 콘텐츠의 등장 이후, 파노라마 영상에 대한 수요가 전시관, 뉴스룸, 옥외 구조물을 통하여 점차 증가하고 있는 추세이다. 3D 콘텐츠의 입체감 제공, UHD 콘텐츠의 사실감 제공과 더불어 고해상도 파노라마 서비스는 기존 HD 영상 미디어에 비하여 넓은 화각을 제공함으로써 사용자에게 극대화된 현장감을 제공한다[12].

본 논문에서는 360도 파노라마 영상을 제작하기 위하여 여러 대의 카메라를 이용하여 영상을 획득하고, 획득된 영상을 통하여 파노라마 영상 생성을 위한 효과적인 스티칭 방법에 대하여 연구하였다.

연구 목적을 달성하기 위하여 다음과 같은 방법을 이용하여 연구를 수행하였다.

- 1) 파노라마 영상 생성을 위하여 효율적인 스티칭 방법을 사용하였다.
 - 파노라마영상 생성을 위하여 본 논문에서 기존의 SURF알고리즘을 개선한 가변 적응형 필터를 적용한 알고리즘(V-SURF : Variable Speed up Robust Feature로 정의 함)과 기존 SIFT와 SURF알고리즘과의 연산 수행속도를 비교·분석함
 - 기존 SIFT와 SURF알고리즘의 연산 수행속도 분석을 통한 제안알고리즘의 효율성 증명
 - 특징점 추출 알고리즘에서 SIFT는 회전, 스케일, 밝기, 변화에 강건하지만 계산 속도가 느리며, SURF는 SIFT만큼 강건하면서도 스케일 공간에서 Fast Hessian Matrix 추출 방법을 사용하고, 수행속도도 더 빠르기 때문에

본 연구에서 SURF알고리즘을 사용하였다. 이는 입력 된 영상에서 적분 (integral) 영상을 구하고, Hessian 행렬에 기반한 검출기를 사용하여 특징을 추출할 때 스케일공간에서 박스 필터(box filter)를 일정한 크기로 증가시켜 계산하기 때문이다.

2) 360도 파노라마 영상 획득 및 다중 카메라 간의 시차를 줄이기 위한 방법을 제시하였다.

- 영상 획득시 시차발생을 최소화하기 위하여 다중카메라를 동시에 촬영할 수 있는 시스템 동기화 방법 제시
- 다중 카메라 동시촬영을 위한 리그 설계 및 제작

3) 가변 적응형 필터를 적용하여 수행된 결과값인 RMSE(Root Mean Square error)와 특징점 개수(Control point) 를 가지고 기존 소프트웨어와의 비교·분석을 하였다.

4. 논문구성

본 논문의 전체 구성은 다음과 같다.

2장은 관련 연구로 스티칭 방법, 기존 특징점 기반의 이미지 스티칭 알고리즘에 대하여 기술하였다.

3장에서는 가변 적응형 필터 이미지 스티칭에 대하여 설명하였다.

4장에서는 실험과정과 기존알고리즘과의 비교분석 하였으며, 기존 소프트웨어와 비교·분석하였다.

끝으로 5장은 본 연구의 결론과 기대효과, 그리고 향후 연구과제에 대하여 기술하였다.

II. 관련연구

360도 촬영 기법은 말 그대로 현실의 어느 시점에서 360도, 즉 사람이 고개를 움직여 볼 수 있을 것이라고 생각되는 모든 장면을 촬영하는 기법이다. 인간의 시야 또는 카메라의 시야의 한계를 일정 거리로 생각할 때, 이 거리를 반지름으로 하는 큰 구 형태의 장면을 촬영하는 기법이라고 말할 수 있다[12].



Fig. 1. 360VR seoul station

그림 1은 실제 360도 영상 촬영한 결과물이다. 최종 형태는 3차원 공간에 ‘구(球)’의 형태로 말아놓은 것과 같은 형태가 나타나게 된다. 이러한 결과물을 만들어내기 위한 촬영을 위해서는 일반 카메라가 아닌 180도 이상의 초 광각렌즈 또는 어안렌즈를 사용해야 한다. 이 어안렌즈(광각렌즈) 카메라를 최소 앞뒤로 2개 이상 조합하여 최소 180도 각도의 영상을 얻고, ‘스티칭’ 편집을 통해 360도 영상을 얻어내는 것이다[12].

스티칭 기법이란 서로 정반대의 방향에서 찍힌 180도 이상의 영상들을 서로 조합하여 최종적으로 구(球) 형태의 360도 영상을 만들어내는 것을 말한다. 반구 형태로 찍힌 서로 다른 두 개의 영상을 겹쳐지는 부분을 정확하게 겹쳐지도록 합쳐 하나의 영상, 온전한 구(球) 형태로 구성하는 기법이다[12]. 180도 이상의 어안 렌즈를 통해 찍힌 두 영상에는 서로 겹치는 부분이 존재할 것이다. 이 겹치는 부분들을 ‘스티칭 포인트’라고 하여 별도의 스티칭 프로그램에서 RGB값이나

빛 등을 고려하여 자동 및 수동으로 설정해 주면 합쳐진 영상이 나타나게 된다. 이러한 과정에서 시뮬레이션 멀미를 유발하는 문제가 발생하게 된다. 첫 번째로는 어안렌즈 자체의 왜곡현상이다. 일반적으로 사람의 시야각과 가장 유사하여 영화 등에서 자주 쓰이는 렌즈는 50mm렌즈이다. 표준렌즈라고도 불리며 이 렌즈를 통해 촬영된 영상을 보며 사람들은 편안하고 자연스럽다고 느끼게 된다. 그러나 어안렌즈의 경우 말 그대로 사람의 눈보다는 물고기의 시야에 가까운 렌즈이다. 게다가 스티칭 포인트를 따기 위해서 정확히 180도의 각도가 아닌 180도보다 조금 더 넓은 각도를 촬영하게 되기 때문에 사람의 시야에 비해 왜곡된 영상을 얻을 수밖에 없으며 스티칭 과정에서, 또는 최종 결과물에서 부자연스러운 느낌을 줄 수밖에 없다. 카메라의 초점이 맞추어져 있는 중앙 부분인 배들이 있는 지점은 비교적 왜곡이 적을 수 있겠지만 시야각의 끝 쪽 부분에는 부자연스러운 왜곡이 느껴질 수 있으며 스티칭 과정에서 이러한 왜곡이 더 심해질 수 있다는 것이다. 따라서 이전에는 다수의 카메라를 붙여서 활용하기도 했지만 비효율적이었으며, 최근에는 삼성의 Gear 360 카메라를 비롯하여 360도 촬영용으로 출시되는 카메라 제품이 등장해서 촬영과 동시에 스티칭까지 자동으로 진행해주고 있기는 하지만 아직까지는 문제가 확실히 해결되었다고는 볼 수 없다. 또한 이러한 스티칭 과정에서의 문제가 해결된다고 하더라도 최종적인 3D 입체감 표현과의 조합에 있어서 문제가 생기게 된다[12].

기존의 이미지 스티칭 방법은 입력영상으로부터 중첩된 영상의 특징점을 추출하고 매칭 시키는 방법을 통해 타깃 영상을 변환하여 스티칭 하는 방법을 사용하였다[13].

1. 스티칭 방법

대화면 스크린의 대중화 및 3D, UHD와 같은 실감미디어 경험을 토대로 보다 사실적이고 현장감을 지원하는 실사 기반 고품질 파노라마 수요가 크게 증가하고 있다. 이러한 파노라마 영상 제작 방법은 여러 장 이미지를 이어 붙여 만들어 한 장 이미지로 찍기 어려운 장면을 생산할 수 있고 적은 비용으로 쉽게 효과를 구현할 수 있다. 또한 HDR(High Dynamic Range)기술을 이용하여 밝은 부분은 좀 더 밝게 어두운 부분은 좀 더 어둡게 표현하여 명암에 의한 세밀함을 살릴 수 있게 하거나 적층초점(Focus stacking)기술을 이용하여 깊은 심도를 확보하여 고품질 파노라마 영상 제작에 활용되기도 한다.

1) Tripod vs handheld from modern cameras[14]



Fig. 2. Video acquired with handheld camera

Fig. 3은 handheld 카메라로 취득한 영상으로 일반인들이 동시에 여러 대의 카메라를 이용하여 파노라마 영상을 취득하기 어려운 경우 한 대의 카메라를 이용하여 영상을 일정 부분 중첩되게(일반적으로 15~30%) 촬영하여 이를 통한 파노라마 영상을 구현한 모습이다. 하지만 동 시간대에 촬영이 불가능함으로 대상 피사체가 움직임이 있을 경우 스티칭 후 잔상이 남거나 스티칭이 잘되지 않는 문제점이 있다.



Fig. 3. Stitching result of acquired image

2) 영상 프레임들간의 노출 유지

Fig. 4는 카메라 노출 값이 달랐을 때 영상의 밝기차이가 발생한다. 이로 인해 영상 스티칭 하였을 때 스티칭 부분에 겹침 현상처럼 하나의 영상으로 인식하기가 힘들다.



Fig. 4. Stitching result of acquired image(Images with different exposure values)

Fig. 5는 노출 값이 다른 영상을 이용해 스티칭 한 결과이다.



Fig. 5. Stitching result of acquired image



Fig. 6. Stitching result of acquired image
(Exposure and focus discrepancies)



Fig. 7. Stitching result of acquired image



Fig. 8. Stitching result of acquired image
(After-image (ghost phenomenon))



Fig. 9. Two acquired images (time difference occurs with a single camera)

Fig. 6 ~ Fig. 10은 파노라마 영상을 만들기 위해 단일 카메라로 영상을 취득할 때 일반적으로 일어나는 문제로, Fig. 6은 포커스 불일치로 블러링이 생겨 스티칭하였을 때 하나의 일관된 영상을 만들기 어렵고, Fig. 7 ~ Fig. 10은 한 대의 카메라를 이용하여 촬영하게 되면 시간차가 발생하고 촬영시간 사이에 피사체의 움직임이 발생하면 스티칭 이후 고스트 현상이 발생하게 된다.



Photo: Russell J. Hewett

Fig. 10. a) Stitching result of acquired image(overlapping image)



Fig. 10. b) Stitching result of acquired image
(overlapping image)

2. 이미지 스티칭 종류

360도 파노라마 영상을 촬영하기 위해서는 카메라 기종은 크게 중요하지 않다. 최근의 LG 360 캠에서 광각렌즈를 달면 화각이 200도까지 지원되기 때문에 두 대만 있어도 360도를 찍을 수 있다. 카메라의 대수가 중요하지 않음에도 카메라를 많이 장착하는 건 영상 촬영 후반부 작업인 스티칭 작업을 수월하게 하기 위함이다.

이미지 스티칭은 고해상도의 영상을 획득하는 영상 기반 렌더링 기법 중의 하나로서, 가상현실, 대화형 멀티 플레이어 게임 등과 같은 다양한 응용분야에서 사용되고 있다[16,17]. 이러한 이미지 스티칭 방법을 사용하여 재구성한 영상을 흔히 파노라마 영상(Panorama Image)이라고 한다. 파노라마 영상이란 사용자 주변의 다양한 방향에 대하여 영상을 촬영한 후 주위의 모든 방향을 한 시점에서 바라볼 수 있도록 하기 위해 단일 영상들을 결합하여 합성한 영상을 의미한다 [18][19][20]. 360°방향의 모든 경치를 촬영하기 위해서는 파노라마 기능을 가지고 있는 고가의 특수 카메라를 사용해야 하지만 실용적이지 못한 단점을 가지고 있다[21]. 반면에 일반 카메라를 사용하면 360°방향의 모든 경치를 촬영할 수 없다. 대신 몇 장으로 나누어 촬영한 뒤 옆으로 길게 이어 붙여 한눈에 볼 수 있도록 영상을 정합하는 이미지 스티칭 기법을 이용하여 한 장의 고해상도 이미지를 생성하여 경치를 감상할 수 있다[22].

스티칭 기법에는 사용자가 수동으로 중첩되는 영역을 찾아 입력하는 수동 스티칭 기법과, 영상처리 알고리즘을 통하여 자동으로 중첩 영역을 찾는 자동 스티칭 기법이 있다[23].

1) 수동 스티칭 기법[23]

수동 스티칭 기법은 두 개 이상의 영상에서 중첩되는 영역을 사용자가 직접 찾아 입력하고, 해당 입력을 기준으로 알파 블렌딩(alpha blending)을 수행하여 영상을 하나로 합성하는 방법이다. 사람이 직접 눈으로 확인하고 중첩 영역을 입

력하기 때문에 비교적 정확도가 높은 편이지만, 사용하기가 번거롭고 시간이 오래 걸리는 단점이 있다.

2) 자동 스티칭 기법[23]

자동 스티칭 기법은 영상처리 알고리즘을 통하여 영상의 중첩 영역을 탐색하고, 탐색된 중첩 영역을 기준으로 두 개 이상의 영상을 하나로 합성하는 방법이다. 알고리즘의 탐색 오류로 인해 부정확한 결과가 도출될 가능성이 있지만, 수동 스티칭과 비교했을 때 수작업이 최소화되어 사용이 편리하고 속도가 빨라 선호도가 높다.

자동 스티칭 기법은 크게 두 단계로 구성된다. 첫 번째, 두 개 이상의 입력 영상에서 중첩 영역을 찾는다. 두 번째, 중첩 영역을 기준으로 입력 영상을 하나로 합성한다. 이 때, 정확한 스티칭을 위하여 중첩 영역을 찾는 알고리즘이 중요하다. 기존에 연구된 중첩 영역을 찾는 방법은 크게 특징 기반 정합 기법(feature-based registration)과 템플릿 기반 정합 기법(template-based registration)의 두 가지로 구분 된다.

(1) 템플릿 기반 기법

영상에서 일정 영역을 지정된 템플릿과 비교하여 가장 유사도가 높은 영역을 판별하는 방법으로, 영상처리의 다양한 분야에서 응용되고 있다[24]. 템플릿 기반 정합의 가장 대표적인 유사도 측정 방법으로는 정규상호상관(NCC, Normalized Cross Correlation)계수가 있다[25]. 정규상호상관계수는 정규화 된 상호 연관성을 찾는 기법으로, 입력 영상과 템플릿 영상 사이에서 밝기 값의 선형적인 차이 및 기하학적인 유사도를 측정한다. 정규상호상관 계수는 두 영상 사이의 밝기 차에 민감하게 반응하지 않기 때문에, 서로 다른 영상의 유사도를 측정하여 합성하는 자동 스티칭 기법에 적합하다[23].

(2) 특징점 기반 방법

특징점 기반의 이미지 스티칭(Feature Based Image Stitching)은 입력된 영상의 기하학적인 특징인, 모서리(Corner), 에지(Edge), 라인(Line)등을 기준 영상에서 추출하여 대응되는 영상과 비교하여 대응점을 찾는 방법이다[22]. 이러한 특징은 밝기 변화에 강인한 특성을 가지고 있다. 그러나 너무 많은 특징점을 추출한 경우 연산량이 많아 수행 속도가 느리며, 잘못된 특징점을 추출하였을 경우에는 정합 과정에서 오류가 발생할 수 있는 단점을 가지고 있다. 따라서 특징점 기반의 이미지 스티칭 방법에서는 특징점 추출이 가장 중요한 요소이다[22].

이미지 스티칭에서 중요한 요소인 특징점 추출을 위해서 사용되는 알고리즘은 SIFT(Scale-Invariant Feature Transform), FAST(Feature from Accelerated Segment Test), SURF (Speeded-Up Robust Features), RANSAC(ransom sample consensus) 등이 많이 사용되고 있다[26][27][28].

특징점(feature point)이라 함은, 이미지의 특징을 나타낼 수 있는 점으로서, 이미지의 스케일, 회전, 왜곡 등의 변화 등에 관계없이 이미지의 특징을 잘 설명할 수 있는 점 또는 점들의 집합을 의미한다. 이러한 특징점은 주어진 이미지의 크기 및 내용 또한 특징점 추출/결정 방법의 종류에 따라 다르지만 사진 한장당 수 백 개 ~ 수천 개의 특징점이 추출될 수 있다. 이러한 특징점은 이미지 프로세싱이나 컴퓨터 비전 분야에서 널리 사용되고 있는데, 특징점을 추출하고 추출된 특징점의 특징 데이터를 이용하여 두 개의 이미지에서 대응되는 부분을 찾아서 물체 인식, 동작 추적, 이미지의 동일성 판별 등의 작업에 다양하게 사용되고 있다.

주어진 이미지에서 특징점 추출 및 결정하는 기술로서는 여러 가지 방법이 제안되어 있는데, LoG(Laplacian of Gaussian) 필터나 DoG(Difference of Gaussians) 필터의 스케일 공간상의 극대/극소값을 이용하는 방법이 알려져 있다. 이외에도, 헤시안 매트릭스(Hessian Matrix)의 디터미넌트(determinant)를 이용하여 특징점을 결정하는 방법 또한 사용되고 있다.

한편, 이미지의 특징점을 추출하는 알고리즘은 본 절에서 설명하는 다양한 방법들이 제안되어 있다. 그러나 지금까지의 특징 데이터 형성 방법은 통상 이미지당 수천 개 가량의 수십 차 특징 벡터를 필요로 하기 때문에 연산 과정이 복잡하고 처리해야 하는 데이터의 양이 많아서 계산 시간이 많이 소요된다는 문제가 있어 대량의 이미지 데이터를 처리해야 하는 경우에는 많은 문제점을 가지고 있다.

기존의 특징점 추출방법은 다양한 환경에서 우수한 성능을 보이고 있다. 그러나 실시간 응용성을 갖기에는 계산 복잡도가 높기 때문에 연산량이 증가하게 되어 특징 추출 시스템의 성능 저하를 유발하게 된다. 이를 극복하기 위해서는 빠른 연산 속도와 우수한 특징 추출을 가능하게 하는 알고리즘 개발이 필요하다.

본 절에서는 기존의 특징점 추출 알고리즘에 대하여 설명하였다.

① Homography

여러 이미지를 Panorama 영상으로 합치기 위해서는, 모든 이미지를 하나의 같은 평면으로 투영(projective)시켜야 한다. 이것이 가능하기 위해서는 영상들이 Homography관계에 있어야 하고, 이의 기하학적 의미는 Fig. 11과 같이, 해당 물체의 투영선에서 임의의 평면으로의 절단면간의 관계와 같다. 수식적으로 보면 Fig. 12와 같이 3x3 행렬로 표현되는 선형변환이다. 다행스럽게도, 카메라 렌즈를 주축으로 회전하면서 찍은 사진은 모두 Homography관계에 있다[6].

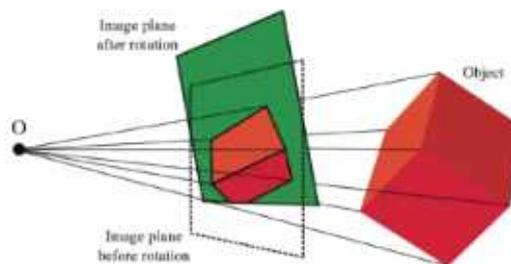


Fig. 11. Definition of Homography Relationship

$$\lambda \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Fig.12 Homography
conversion formula

파노라마 영상을 생성하기 위해서 가장 중요한 요소는 정확한 호모그래피를 추정하는 것이다. 파노라마 영상을 얻을 때 기본 전제 조건은 3차원 실세계에 있는 영상을 카메라의 2차원 이미지 평면에 투영시키는데 두 영상은 동일평면상에 있다고 가정을 한다. 여기서 호모그래피란 3×3 행렬로 기준 영상의 점과 그에 대응되는 대상 영상의 점들 사이의 관계를 나타낸 것이다. 호모그래피를 추정하기 위해서는 기본적으로 기준영상 네 개의 점과 그에 대응되는 대상 영상 네 개의 점이 필요하다[29].

② SIFT(Scale-Invariant Feature Transform)

SIFT 알고리즘은 특징점과 그 점의 최적 Scale의 검출과, 특징점 주변 패턴의 밝기, 기울기 히스토그램에 의한 128차원 벡터에 의한 묘사의 두 가지 처리를 포함한 알고리즘이다. SIFT 특징은 회전, 크기 변화, 밝기 변화뿐만 아니라 일정한 범위의 Affine 변환(시점의 이동)에도 강인하다는 것을 실험을 통해 나타나고 있으며 최근까지도 가장 강력한 기술로 활용되고 있다[30].

SIFT 알고리즘은 D. Lowe에 의해 제안된 회전, 크기변환과 조명변화 및 시점의 변화에 강인한 지역적 특징점 기반 추출 방법 중 하나로, 디지털영상처리 분야에 많이 적용되는 알고리즘이다[31].

SIFT 알고리즘은 Fig. 13에 제시한 바와 같이 다음의 같은 단계로 수행된다 [32].

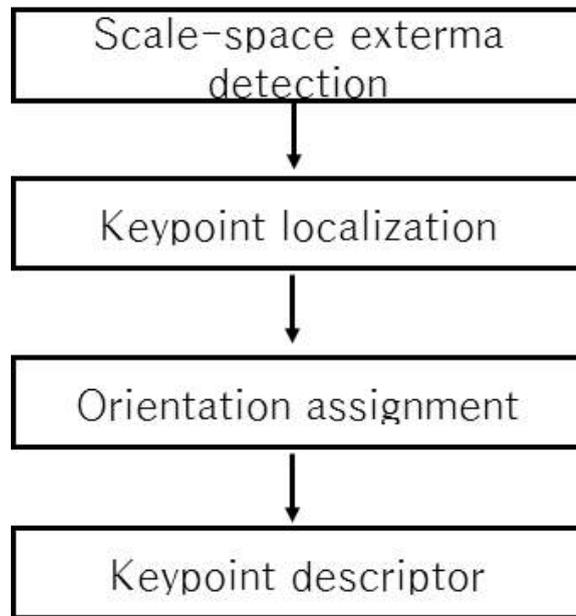


Fig. 13. Structure of SIFT algorithm

SIFT 알고리즘은 크기와 회전 변화에 강인한 알고리즘으로 물체의 크기나 방향에 상관없이 높은 정확도로 해당 물체의 위치를 추적하는 우수한 성능을 보인다[33]. 하지만 연산과정에서 상당히 많은 계산을 필요로 하기 때문에 처리 시간이 오래 걸린다는 단점이 있다. 이러한 이유로 SIFT 알고리즘은 그대로 사용하기보다는 변형을 하거나 다른 방법과 결합하는 형태로 사용되고 있다[33]. 고차원적인 특징 때문에 영상의 크기가 커질수록 계산하는 데이터양이 많아지게 된다. 데이터양이 많아지게 되면 연산량이 많아져 수행 시간이 증가 되는 문제점을 가지게 된다.

얼굴인식, 3차원 객체 인식, 파노라마, 3차원 영상복원 작업의 핵심 알고리즘으로 연구 되고 있다[34].

③ SURF(Speeded Up Robust Feature)

Bay에 의해 제안된 SURF 알고리즘은 여러 개의 영상에서 스케일, 조명, 시점

등의 환경변화를 고려하여 환경변화에 불변하는 특징점을 찾는 알고리즘으로서 [35][36], SIFT가 가지고 있는 추출 및 정합 단계에서 발생하는 연산의 복잡성을 줄여 빠르게 특징점 및 특징 서술자를 추출하고 정합을 수행할 수 있도록 하는 방법이다[26].

SURF 알고리즘은 Fig. 14에 제시한 바와 같이 SURF알고리즘이 속도 향상을 위한 방법은 다음 단계로 수행된다[32].

- ㄱ) 적분 영상(integral image) 이용
- ㄴ) 박스(box) 필터를 이용한 Hessian matrix를 기반으로 특징점 추출
- ㄷ) Non-maximum suppression (NMS)을 이용한 특징점의 위치 판별
- ㄹ) Harr response를 이용한 특징점 방향 계산
- ㅁ) 특징 기술자 생성

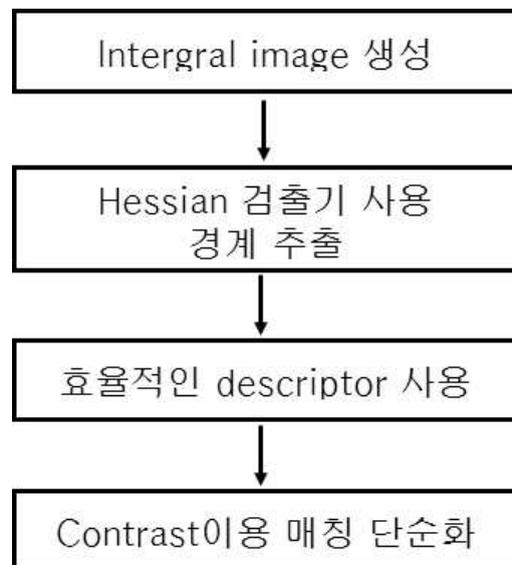


Fig. 14. Structure of SURF algorithm

SURF의 핵심은 적분 영상(Integral Image)을 사용하여 영상에 가해지는 회선(Convolution)작업을 빠르게 수행하고 간편화한 특징점 추출 방법과 특징점 기술자를 사용하여 연산량이 적으면서도 좋은 결과를 얻을 수 있다[27][37]. 또한 명

암 대비(Contrast)를 이용하여 간단한 정합이 가능하다는 장점을 가지고 있다. 이와 같은 특징점 추출 방법들은 이미지 스티칭 분야 뿐만 아니라 지문 인식, 얼굴 인식 및 비디오 스티칭 등의 분야에서 많이 사용되고 있다.

일반적으로 SIFT 방법이 강인한 특징점을 추출하지만 알고리즘의 복잡도가 높고, 연산량이 많아서 특징점 추출 속도가 느린 단점을 가지고 있다. 그러나 SURF는 SIFT 방법과 비교하여 특징점 추출 측면에서는 유사한 성능을 보이지만, 속도 측면에서 크게 향상된 성능을 보인다. 이와 같은 방법으로 특징점을 추출하게 되면 크기 및 회전 등의 왜곡에 강한 특징을 갖게 된다[22].

④ FAST(Features from Accelerated Segment Test)

FAST는 Edward Rosten과 Tom Drummond가 개발한 방법으로 FAST가 속도에 최적화되어 있음에도 불구하고 특징점의 품질(repeatability: 다양한 영상 변화에서도 동일한 특징점이 반복되어 검출되는 정도) 또한 기존의 방법들 (Harris, DoG, ...)을 더 좋다는 점이 있다[38].

FAST 또한 코너점을 찾는 방법 중 하나로서 그 기본 방법은 다음과 같다.

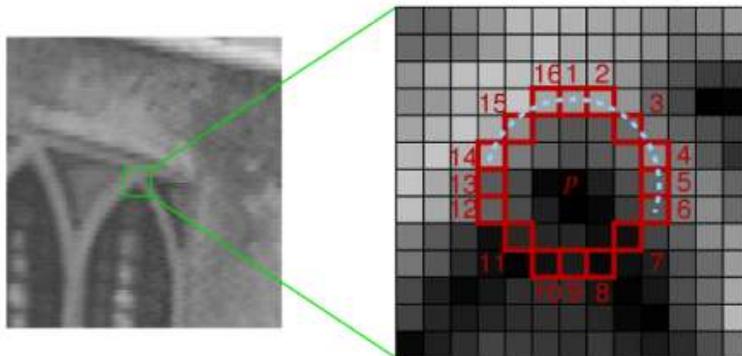


Fig. 15 FAST corner detector

FAST에서는 위 그림과 같이 어떤 점 p 가 코너(corner)인지 여부를 p 를 중심으로 하는 반지름 3인 원 상의 16개 픽셀값을 보고 판단한다. 그래서 p 보다 일정값 이상 밝은($>p+t$) 픽셀들이 n 개 이상 연속되어 있거나 또는 일정값 이상 어두운($<p-t$) 픽셀들이 n 개 이상 연속되어 있으면 p 를 코너점으로 판단한다[39].

FAST 알고리즘은 n 을 어떻게 잡느냐에 따라서 FAST-9, FAST-10, FAST-11, FAST-12,..., FAST-16과 같이 다양한 버전이 가능하다. 즉, FAST-9는 p 보다 일정값 이상 밝거나 어두운 픽셀들이 원을 따라서 연속적으로 9개 이상 존재하는 경우를 코너점으로 검출하는 알고리즘이고, FAST-12는 12개 이상의 점들이 연속적으로 모두 p 보다 충분히 밝거나 모두 p 보다 충분히 어두운 경우를 찾아주는 알고리즘이다[39].

따라서 고속으로 모서리(코너점)를 추출하기 때문에 기존의 방법인 Harris corners, SIFT의 DOG(Difference Of Gaussian), SURF의 Haar wavelet 과 SUSAN(Smallest Uni-Value Segment Assimilating Nucleus Test) 비해 빠른 연산이 가능하며, 변형에 대한 높은 반복성을 지니게 된다. 그러나 노이즈가 많은 환경에서 기존의 특징점 추출 방법에 비해 제대로 된 결과가 나오지 않는 단점을 가지고 있다[40].

⑤ RANSAC(ransom sample consensus)

RANSAC 알고리즘은 Fischler 와 Bolles에 의해 처음 제안되었으며, 측정 노이즈가 심한 원본 데이터로부터 모델 파라미터를 예측하는 방법이다. 에러와 잡음이 섞여있는 초기 데이터들을 최소로 이용하여 일관된 데이터의 집합(Consensus Set)를 확장시켜 나가면서 반복적인 작업을 통해 모델을 예측하여 오차를 최소화하는 알고리즘이다[41].

RANSAC 방법을 이용한 기본 행렬을 추정의 기본적인 개념은 다음과 같다. 수집한 데이터 집합으로부터 모델을 구성할 수 있는 최소 개수로 이루어진 데이터들을 무작위로 추출한다. 추출한 데이터들을 이용하여 모델의 계수들을 추정하며, 이때 사용되는 추정 방법은 최소 제곱법과 같은 일반적으로 잘 알려진

방법이 될 수 있다. 모델이 추정되면, 모델을 추정하기 위해서 사용한 데이터를 제외한 나머지 데이터들을 모델에 적용한다. 데이터들을 모델에 적용한 결과값과 관측을 통해서 얻어낸 값과의 차이가 RANSAC에 설정한 오차의 임계값(Threshold Value) 보다 작은 값의 개수를 구한다. 위의 과정을 최대 반복횟수까지 반복하여, 모델로 계산된 값과 관측 값의 차이가 오차의 임계값 보다 작을 경우의 개수가 가장 많은 모델을 최종 모델로 선택한다. 최종 모델은 오차를 포함한 데이터(outlier)가 제거된 모델로 판단한다. 이때, 추정된 최종 모델로 데이터들을 계산한 결과 값과 관측을 통해서 얻어낸 값과의 차이가 오차의 임계값 보다 큰 데이터들이 과대 오차를 포함하고 있는 데이터가 된다. RANSAC은 오차의 임계값, 총 반복 횟수 등의 매개 변수를 가지고 있다. 오차의 임계값은 모델을 추정하는 데 적합하지 않은 과대 오차를 포함하는 데이터, 아웃라이어(outlier)와 그렇지 않은 데이터, 인라이어(inlier)의 경계를 나타내는 값이다. 총 반복 횟수는 데이터들이 모델을 추정할 때 최대 몇 번까지 반복할지를 결정하는 값이다[42].

모델 파라미터를 결정하기 위해 전체 원본 데이터 중에서 최소의 데이터를 랜덤하게 샘플링 하면서 반복적으로 해를 계산함으로써 최적의 해를 구하기 때문에 비 결정 알고리즘(Non-Deterministic Algorithm)이라고도 한다. RANSAC 알고리즘을 이용하면 두 영상 사이에 있는 많은 특징점 중에서 정확하게 정합되는 4개의 대응점을 결정할 수 있다[41].

RANSAC 알고리즘의 수행 순서는 다음과 같다[43].

- ㄱ) : 전체 대응점 후보들로부터 N개의 샘플 대응점 획득
- ㄴ) : 획득한 샘플 대응점을 참 값으로 가정하고 Homography Matrix를 예측
- ㄷ) : 예측된 Homography Matrix가 옳은지 판단
- ㄹ) : 참 값이 아닐 경우 ㄱ) - ㄷ)번 과정 반복

위의 과정 ㄱ)에서 N은 파노라마 영상을 구성하기 위해 Homography Matrix를 만드는데 필요한 4개의 대응점 개수다.

⑥ Optimal Seam[44]

영상을 정합 하였을 경우에는 다양한 요인들에 의해 영상들 간에 경계와 기하학적 뒤틀림 오차가 발생하게 된다. Optimal Seam 알고리즘의 경우 기하학적 뒤틀림 오차가 발생 하였을 경우 오차를 최소화 하는 방법으로 정합된 영상의 중첩된 부분에 대해서 화소값이나 칼라 및 텍스처 등의 정합값을 $[0, 1]$ 로 표현하고, 1에 대한 보수 값을 비용으로 간주하여 비용공간으로 표현하면, 유사성이 많은 화소들을 연결하는 경계 라인을 구하게 된다. 경계 라인을 구하는 문제는 비용공간에서의 최적 경로를 찾는 문제로 간주할 수 있다. 이 최적 경로결정에는 동적계획법(Dynamic Programming)[45][46]이 매우 효과적 알고리즘으로 알려져 있으므로 경계 라인의 결정에 활용하기 때문에 최적의 경계를 결정 할 수 있게 된다.

3. 이미지기반 VR

VR(Virtual reality)은 실제 발생할 수 있는 사건이나 상상할 수 있는 판타지를 컴퓨터 시뮬레이션으로 경험할 수 있도록 해주는 기술이다. 대부분의 VR은 비주얼적 재현을 우선으로 하며, 효과를 극대화하기 위하여 사운드 시스템이나 햅틱(haptic)과 같은 터치 시스템을 함께 도입한다. VR 기술은 게임과 같은 단순한 가상적 경험에서부터 의료나 우주개발과 같은 수행과정에 대한 결론을 추론할 수 있는 복잡한 피드백 시스템을 갖춘 솔루션에 이르기까지 여러 분야에서 활용되고 있다. 특히 이러한 가상적 경험에서 시각적 경험과 리얼리티의 진정성을 극대화하기 위하여 특정 공간의 비주얼을 고스란히 전달할 수 있도록 개발된 기술이 VR Photography(사진기반 가상현실)이다.

이미지 기반 VR은 실제 이미지를 기반으로 하기 때문에 사실적이며, 이미지를 입력한 후 모델링 같은 후 처리작업에 많은 시간이 소요되지 않아 제작기간이 짧고 비용이 저렴하다. 이미 만들어진 사진이나 그림을 사용자가 움직이는 대로 신속하게 보여줌으로서 사용자의 입장에서 이를 실제로 느끼게 만드는 기술로서, 오브젝트 파노라마 방식과 파노라마 방식으로 나눌 수 있다.

1) Object VR

Object VR은 단순히 제품의 한 면만을 보여주는 기존의 일반 사진 촬영 기법에서 벗어나 제품을 360도로 회전하며 촬영하여 제품의 다양한 면모를 보여주는 방식이다[47].

자동차, 박물관 및 전시회의 유물 조각 작품, 애완동물, 전자제품, 조형물, 등 사물의 관찰이나 전자상거래 쇼핑몰 제작 등에 많이 사용하고 있다.

파노라마 VR제작방식과 거의 같으나 오브젝트 VR은 턴테이블 위에 물건을 돌려가며 촬영한다. 촬영된 이미지는 Object Modeler를 이용하여 가상현실로 제작된다.

오브젝트 VR은 카메라가 상품/제품/사람 등의 대상물 주위를 같은 거리와 각도를 유지하며 360도 회전하면서 촬영한 후 촬영한 이미지를 연속적으로 보여 줌

으로서 동영상으로 표현하게 된다.

촬영된 이미지를 동영상으로 만드는 방법은 파노라마와는 달리 각각의 이미지들을 그대로 보여주는 것으로, 마우스로 드래그 할 때 빠른 시간에 여러 장을 보여줌으로써 실제로 움직이는 것과 같은 효과를 준다. 마치 필름을 1초에 24컷으로 연속해서 보여주면 하나의 영화가 되는 것과 같은 방법이다[48].

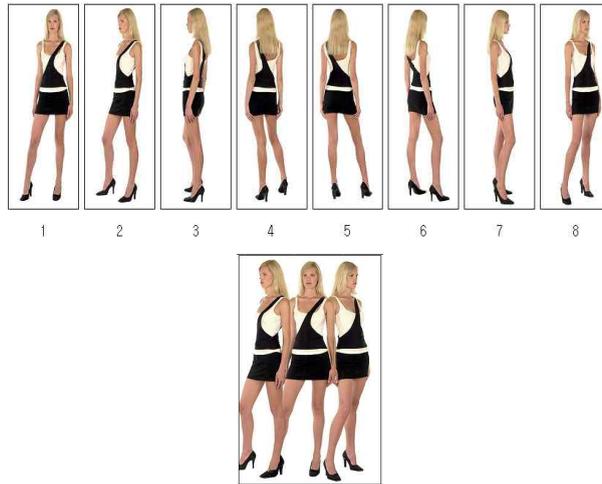


Fig. 16. How to make an object VR image

2) 파노라마 VR

상하좌우에 걸쳐 사진을 촬영하고, 이를 자동으로 합성해 입체적인 사진을 만들어주는 기능이다. VR 파노라마는 일반 파노라마와는 달리 360도 촬영을 하기 때문에 사진이 입체적이다. 반면 일반 파노라마 사진은 180도로 촬영한 사진들을 이어 붙여, 옆으로 길게 퍼진 경치를 보여주기에 평면적이다.

VR 파노라마 기능을 이용하면, 피사체와 가까운 곳에서 촬영을 해도 많은 사람들을 사진 한 장에 모두 담을 수 있다. 또한 기존 파노라마보다 더욱 사실적으로 현장을 기록하고 감상할 수 있다.

파노라마 VR이란 사용자가 공간에서 주위 환경을 둘러보는 효과를 만드는 방법이다. 좌우, 상하 180도~360도까지 볼 수 있어서 특정 장소에 직접가지 않아도

그 곳에 와 있다는 착각이 들 정도의 느낌을 제공한다[49]. 특히 3D로 구현되는 파노라마는 현실감이 좋기 때문에 흥미와 정보를 동시에 제공할 수 있다. 가상여행, 박물관, 유적지, 관광지(산, 바다, 계곡 등), 여행, 펜션, 호텔, 놀이공원, 모델하우스 등 주로 건물의 내외부 그리고 특정장소나 지역을 설명하는데 효과적으로 사용한다.

이미지와 같이 특정 위치에서 삼각대로 카메라를 고정시킨 후 수직 및 수평 360도 모든 방향으로 촬영한 후 하나의 이미지로 겹치는 부분을 스티칭(stitching) 한다. 스티칭된 이미지는 파노라마 VR 뷰어 프로그램을 사용하여 인터넷 상에서 3D로 구현해 낸다[50].

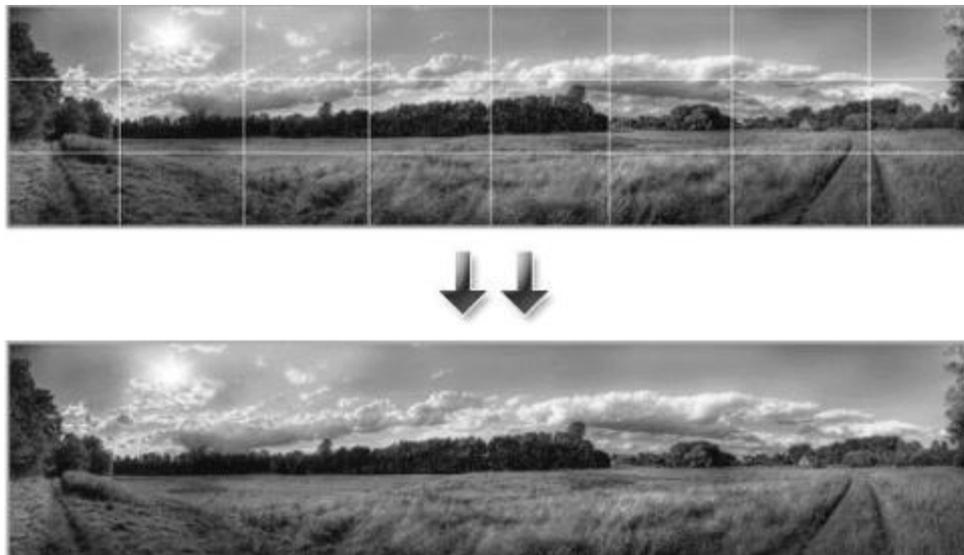


Fig. 17. Panorama VR contents

Panorama VR 제작에는 수평축만을 제작하는 Cylinder 방식과 수평·수직축 전부를 제작하는 Sphere 또는 Cube 방식이 있으며 어떠한 방법으로 제작을 할 것 인지를 사전에 결정을 하여야만 한다. 최종적으로 제작할 방식이 정해 진 후 이 결정에 의해서 촬영 시 사용하는 하드웨어와 합성 시 사용되는 VR 소프트웨어의 선택이 달라진다[48].

(1) Cylinder 파노라마

Cylindrical 파노라마는 원통형으로 영상을 투영하기 때문에 구면 파노라마 영상 만큼은 아니지만 평면 영상과는 구분되는 왜곡이 발생한다. 하지만 원통형 파노라마는 다음과 같은 단점이 있다. 첫째, 상하 시야각이 제한이 있어 카메라의 상단부와 하단부는 촬영되지 않으며, 둘째, 촬영자의 숙련도에 따라 파노라마 영상의 품질이 달라진다는 단점이 있다. Fig.18은 cylindrical 파노라마로, 하단부와 상단부는 촬영되지 않았다[51].



Fig. 18. Cylindrical Panorama

Cylinder 파노라마는 천정과 바닥이 없는 형태로 좌우 360도로 제작하는 방식이며, 수평 360도는 사용하는 렌즈에 따라서 촬영하는 장수가 달라지게 된다[52]. 상하 화각은 사용하는 렌즈의 세로화각에 의해서 결정된다. 주로 실외 파노라마를 제작할 때 사용되며 제작이 간단하며 2D출력물로 사용할 때 적당하다[48].

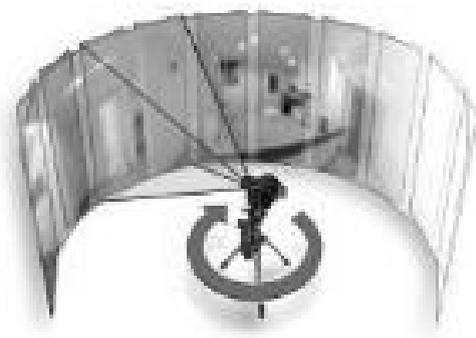


Fig. 19. Cylinder method

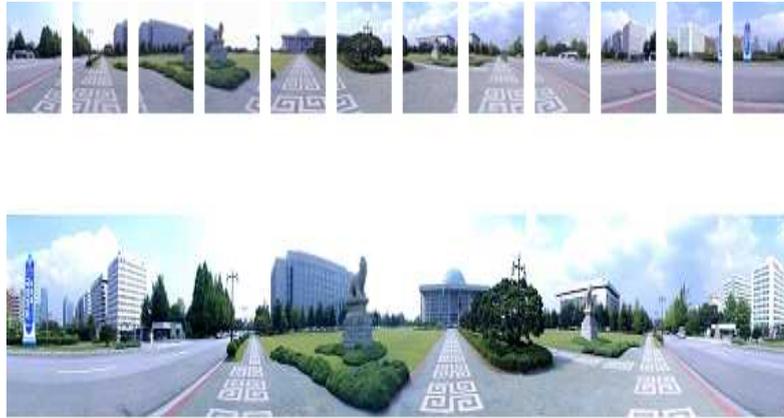


Fig. 20. Cylindrical image making method

실린더 파노라마는 운동시차(motion parallax)가 영상들 사이에 나타나지 않기 위해 카메라를 삼각대에 고정시켜 놓은 다음, 수평(panning)방향으로 회전시켜 얻은 영상을 취득한다. 취득된 영상에서 카메라의 내부 요소인 초점거리를 먼저 추정하고, 영상을 실린더 좌표계로 변환시킨 후, 반복 계산에 의해 영상들 간의 움직임 요소를 복원하여 파라마마 영상을 구성한다[53].

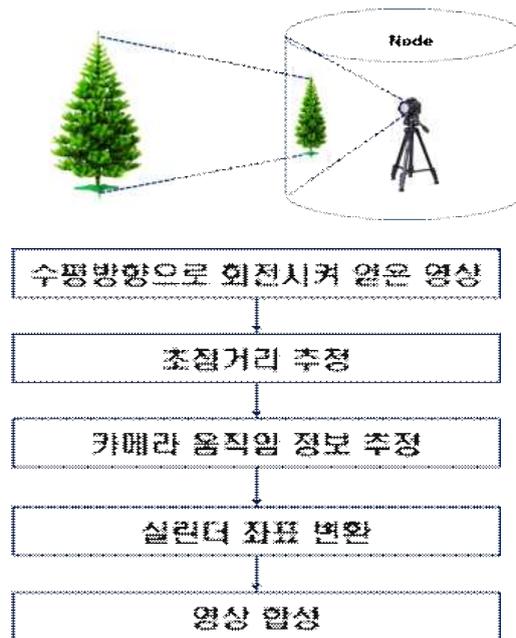


Fig. 21. Cylinder panorama construction principle and Flow Chart

(2) Sphere Panorama

Sphere Panorama 방식은 전후좌우 상하를 빠짐없이 촬영한 뒤 합성하여, 360도를 모두 돌려볼 수 있도록 제작하는 방식이다. Sphere 방식은 360도를 모두 돌려볼 수 있기 때문에 촬영지점 선택이 무엇보다도 중요하다. 어안렌즈나 일반렌즈를 이용하여 전 공간을 촬영하게 된다[54].

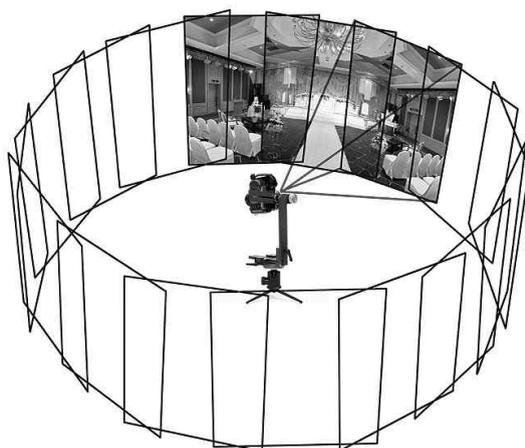


Fig. 22. Shooting to overlap the space before and after the left, right, top and bottom



Fig. 23. Spherical Panorama image

파노라마 영상을 획득하는 방법은 두 가지가 있다. 일반카메라 렌즈를 사용하는 방법과 파노라마용 카메라렌즈(어안렌즈)를 사용하는 방법이 있다[55].

① 일반카메라 렌즈를 이용한 방법

카메라를 파노라마 헤드에 장착할 때 수직으로 장착을 하게 되는데 그 이유는 촬영되는 상하의 각도를 조금이라도 넓게 사용하기 위해서다. 일반렌즈를 사용하면 화각이 좁아서 좌우는 촬영장수를 늘려서 합성해 주면 된다. 카메라를 장착한 파노라마헤드를 수평축으로 360도 회전시키면서 촬영을 하며 이때 파노라마헤드에 장착된 카메라는 움직임이 없어야 한다. 일반렌즈를 이용한 방법은 실시간 상을 얻기 어렵다는 문제가 있다[56].

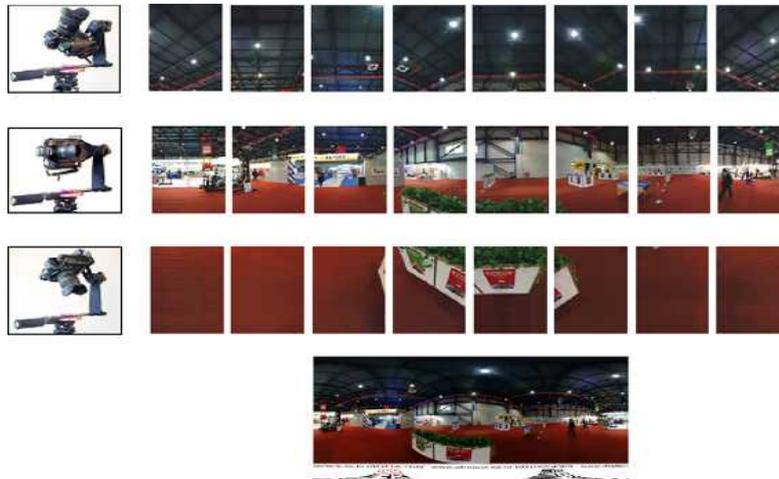


Fig. 24. How to make an image using a general lens

② 파노라마용 카메라렌즈(어안렌즈)를 이용한 방법

파노라마용 카메라렌즈인 어안렌즈를 카메라에 장착하면 좌우상하 180도 이상을 1장에 촬영할 수 있게 된다. 즉 최소 2장을 촬영하여 수평수직 360도 파노라마 VR 제작이 가능해진다[56]. 어안렌즈를 이용하는 방법은 영상의 왜곡을 복원하기가 쉽지 않으며, 매우 복잡한 복원 알고리즘을 필요로 한다[57].



Fig. 25. How to make images using fisheye lens

(3) Cube 파노라마

Cube 파노라마 방식은 Quick 또는 KRPano Viewer로 구현될 때 제작 되는 방식이다. 좌우 360도 상하 180도의 화각을 가지고 있으며, Cubical Panorama로 출력된 이미지는 6장의 평면 이미지이기 때문에 파노라마 수정 및 편집 작업 시 유용하게 사용되는 방식이다[58].



Fig. 26. Cubical Panorama image implemented in VR Viewer

(4) 평면 파노라마(Planar or Flat Panorama)

평면 방식은 좌우180도 이하의 각도를 나누어서 촬영된 이미지를 수평과 수직 상태가 왜곡이 없는 상태로 제작하는 방식이다. 이 방식은 인쇄나 출판을 목적으로 주로 사용된다[58].

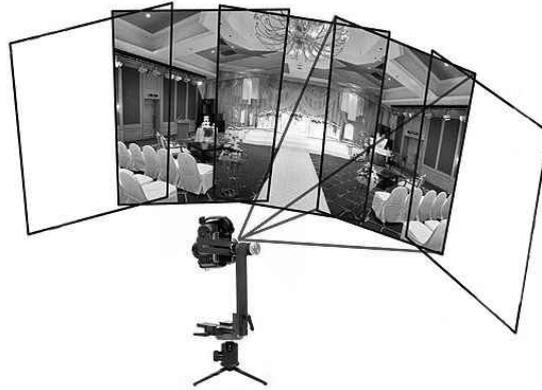


Fig. 27. Flat panoramic photography



Fig. 28. Flat panoramic output image

4. 기존 방식의 문제점

1) SIFT알고리즘

SIFT알고리즘은 아래와 같은 문제점이 있다.

- 일반적으로 SIFT방법이 강인한 특징점을 추출하지만 알고리즘의 복잡도가 높고, 연산과정에서 상당히 많은 계산을 필요로 하기 때문에 처리 속도가 느린 단점이 있다.
- 수행 시간이 많이 소요됨에 따라 실시간 처리를 필요로 하는 시스템에 적용하기 힘들다. 이러한 이유로 SIFT알고리즘을 그대로 사용하기 보다는 변형을 하거나 다른 방법과 결합하는 형태로 많이 사용되고 있다.

2) SURF 알고리즘

기존 SURF 알고리즘은 아래와 같은 문제점을 가지고 있다.

- 이미지 피라미드를 사용하여 크기 변화 및 회전에 강한 특징점을 보장하기 때문에 많은 연산량을 필요로 한다.
- Non-maximum suppression (NMS)을 이용한 특징점의 위치 판별을 하기 때문에 찾고자 하는 대상이 위치한 영역이 아닌 다른 곳에 배경이 위치한 영역에서도 연산을 수행하여 효율성이 떨어진다.

기존의 SURF 알고리즘은 크기 변화와 회전에 강한 특징점을 추출하기 위해 이미지 피라미드를 구성한다. 서로 다른 크기의 이미지를 적용하는 방식을 사용하는 SIFT 알고리즘과는 달리, SURF 알고리즘은 Fig.29와 같은 사각 필터의 크기를 변화하는 방식을 사용하여 이미지 피라미드를 구성함으로써 연산의 효율성을 높였다. 하지만, 이미지 피라미드의 구축은 그 자체로 많은 연산을 필요로 하기 때문에 효율성이 저하되는 단점이 있다.

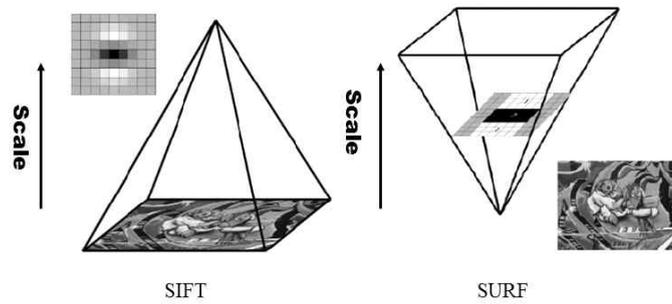


Fig. 29. Image pyramid of SIFT algorithm and SURF algorithm

3) 360도 파노라마 영상 촬영시 문제점

고화질 360VR 영상은 대 화각을 가진 영상으로 현재의 카메라로는 시야각 정도의 넓은 공간을 시공간적인 왜곡 없이 촬영하기 어렵기 때문에 시야각 확보를 위해 큰 왜곡이 생긴다. 따라서 다수의 카메라로 영상을 획득할 경우, 각각의 카메라에서 촬영된 영상간의 공간적 끊김이 없도록 영상의 일부분이 겹치도록 카메라를 배치하여야 한다. 이 때, 카메라들은 동일한 프로젝션 센터를 가질 수 없기 때문에 시차(parallax)가 발생하게 된다. 시차는 영상합성을 통하여 해결할 수 없으므로 영상 획득시 이를 최소화 할 수 있도록 카메라 간의 간격을 줄여 시차를 최소화하여야 한다.

Ⅲ. 가변 적응형 필터를 적용한 이미지 스티칭

1. 이미지 스티칭 과정

기존의 스티칭 기법의 대표적인 상용소프트웨어는 Kolor사의 Autopano와 Videostitch사의 Videostitch, PTGui, Adobe Premiere 등이 있다.

Autopano의 경우 프로그램을 실행시킨 후 하나로 합칠 동영상을 선택한 후 시작점을 확인하고 스티칭 버튼을 누르면 하나의 동영상으로 만들 수 있다. 하지만 스티칭 과정은 동영상을 프레임 단위로 분석해서 공통된 부분을 찾아내 자동으로 연결을 해주는 과정을 반복해야 한다. 특히 촬영 과정에서 왜곡이 심하거나 잘못 촬영된 경우에는 스티칭의 결과 역시 좋지 못하거나 작업이 완료되지 않을 수 있다. 그렇기 때문에 좋은 스티칭 결과물을 만들기 위해서는 촬영 단계부터 스티칭 과정을 고려해서 촬영을 진행해야 한다. 기존의 소프트웨어는 광학적 초점거리와 왜곡을 정보를 가지고 스티칭을 하게 된다.

본 논문에서 수행한 스티칭 과정은 먼저 대상이 되는 자연경관이나 건물 등을 여러 장으로 분할하여 영상을 생성하고, 입력받은 영상에서 기본적인 전처리과정을 수행한다. 입력된 영상의 렌즈 왜곡이 필요한 경우에는 입력된 영상에 대하여 preprocessing 과정을 거친다. 이러한 preprocessing과정이 필요하지 않는 경우에는 영역설정으로 진행한다. 영역설정에서는 획득되는 렌즈특성(일반, 광각, 어안)을 고려하여 영역 설정을 한다.

다음 단계로 특징점 추출을 하고 여기서 가변 적응형 필터를 적용한다. 가변 적응형 필터는 파노라마영상 생성속도를 향상하기 위하여 입력된 이웃하는 두 영상의 중첩되는 영역을 중심으로 각각 추출 영역을 설정하고, 그 영역 안에서 의미 있는 특징점만을 추출하는 방법이다. 파노라마 영상 생성과정에서 실제 대응점들을 추출하기 때문에 불필요한 영역의 특징점들은 연산속도를 떨어지게 하는 요인이 된다. 따라서 효과적인 영역 설정의 크기에 따라 추출할 특징점의 개수는 감소되고, 수행속도는 달라진다. 마지막으로 호모그래피를 추정하여 스티칭하고 최종 영상을 생성한다. 본 논문에서는 효과적인 특징점을 추출하기 위하여

가변 적응형 필터를 적용한 V-SURF(Variable Speeded up Robust Feature)를 제안하였다.

Fig. 30은 본 논문에서 파노라마 영상생성 전체 흐름도이다.

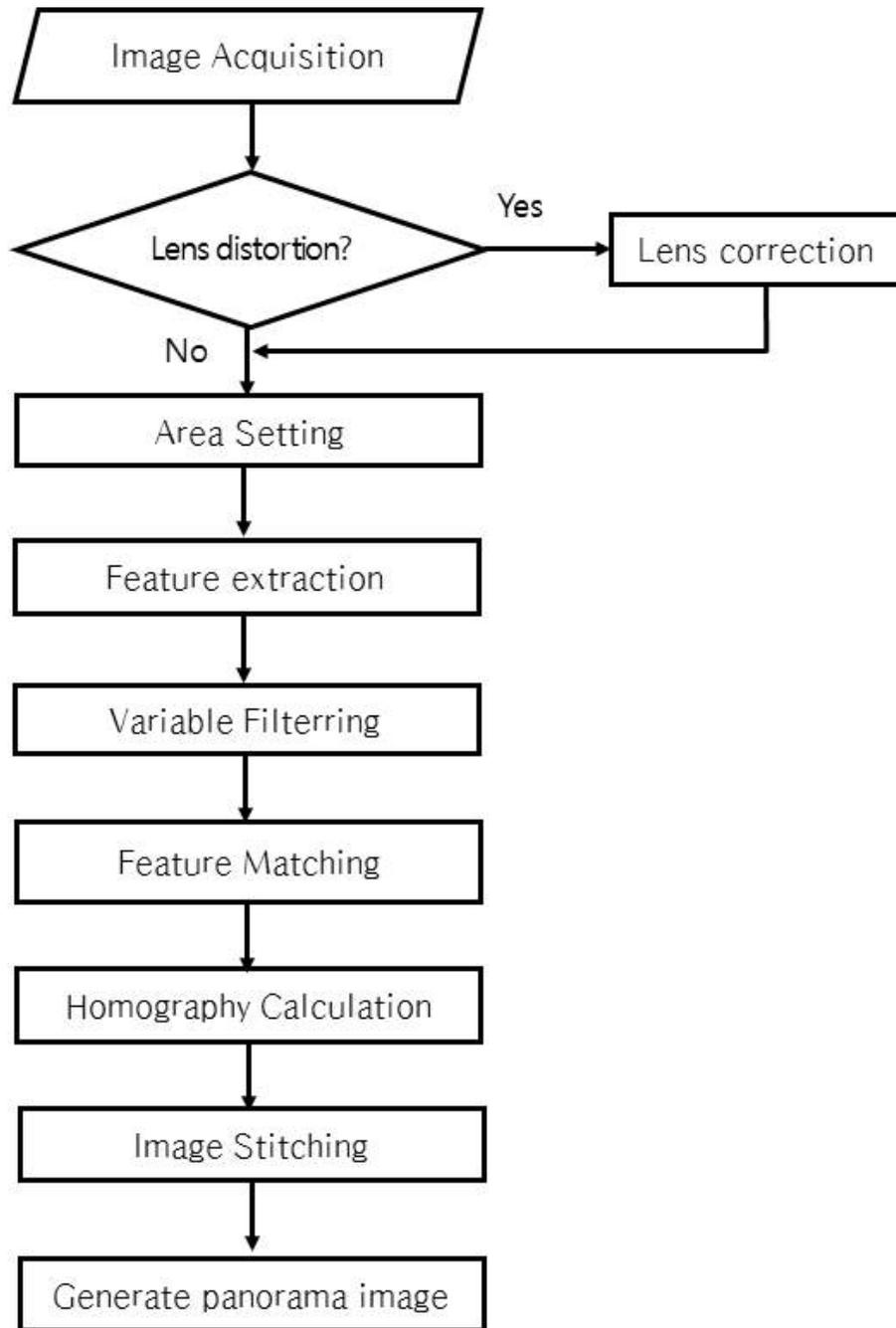


Fig. 30. Panoramic Image Generation

파노라마 영상생성을 위한 절차는 다음과 같다.

1) 렌즈특성을 고려한 영역 설정

- 카메라 렌즈의 특성 즉, 일반렌즈 및 광각렌즈, 어안렌즈의 특성을 고려하여 영역 설정 필요하다.
- 어안렌즈로 인해 생기는 왜곡을 보정한다.
- 중복되는 영상에서 가변적응형 필터를 적용하여 핵심 영역에서만 연산을 수행하도록 하여 연산량을 줄임으로써 연산 수행시간을 줄일 수 있다(효율성을 높였다).

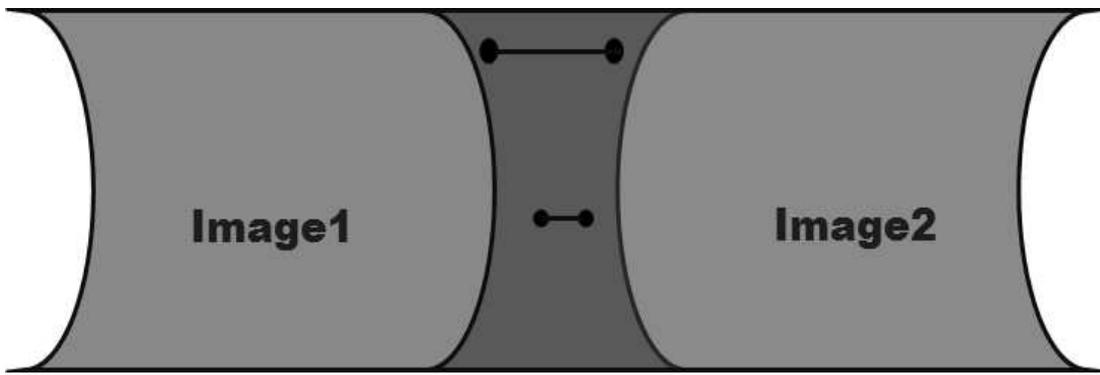


Fig. 31. Set redundant area of input image

Fig. 32에서와 같이 360도 파노라마 영상을 제작하기 위하여 화각이 넓은 어안 렌즈를 이용하게 되면 화면 중심에서 바깥으로 갈수록 왜곡이 발생하게 된다. 즉 중심점에서 멀어질수록 추출된 특징점들의 유클리디언 거리 차이가 많이 발생하게 된다. 추출된 특징점들의 매칭을 하기 위하여 search window mask 사이즈를 크게 할 필요성이 있다.

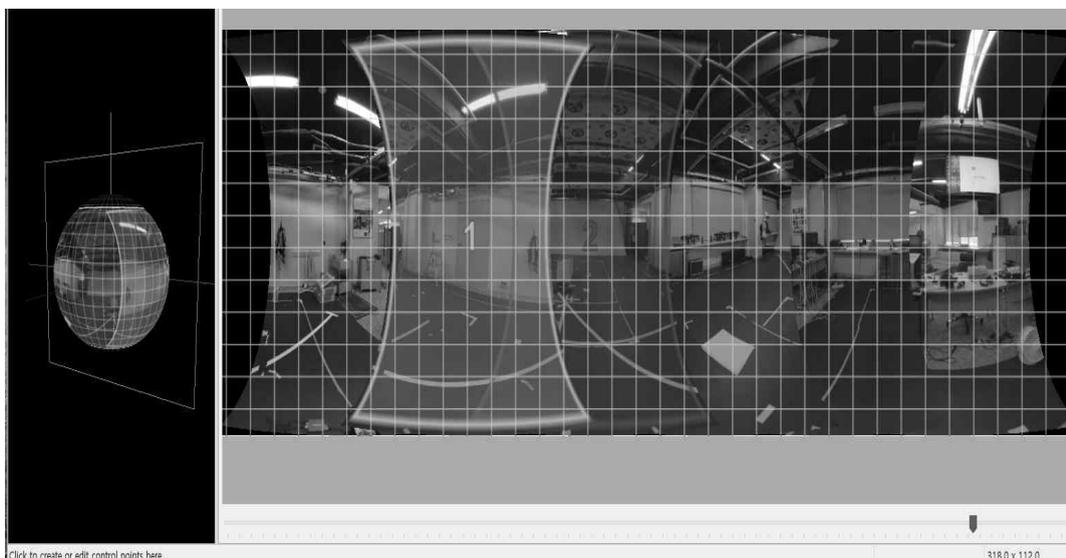


Fig. 32. Picture of overlapped area of input image

Fig. 32는 실제 입력된 영상에서 영상 1과 영상2 사이의 중복된 영역을 보여주고 있으며 중심에서 바깥으로 갈수록 왜곡이 크게 발생하는 것을 확인할 수 있다.

2) 왜곡을 보정

영상을 정합하는 과정에서 조명이나 시점이 달라 생기는 경계선, 카메라 렌즈의 특성 등에 따른 왜곡이 생겼을 때에는 왜곡을 자연스럽게 보정하기 위해서 왜곡을 보정식 함수를 이용해 파노라마 영상을 생성할 수 있다.

일반적으로 촬영된 영상들은 렌즈의 왜곡으로 취득된 영상에 특정 왜곡을 가지고 있다. 카메라에 시야각(FOV)이 넓은 광각렌즈나 초광각 렌즈를 사용하면 넓은 범위를 볼 수 있지만 이로 인해 아래 그림과 같이 상대적으로 영상왜곡이 심해지는 문제가 있다.

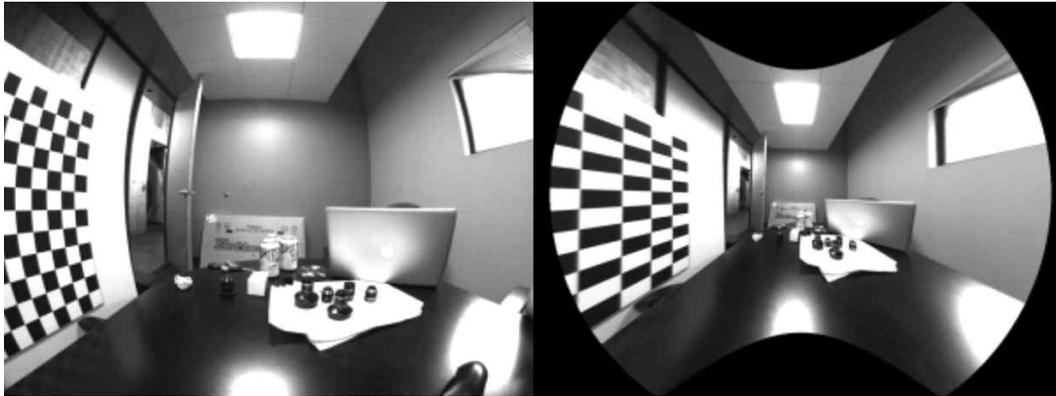


Fig. 33. An example of raw image with lens distortion and its correction

Fig.33과 같은 영상 왜곡은 시각적인 문제 외에도 영상분석을 통해 정확한 수치 계산이 필요할 때 문제가 된다. 예를 들면, 영상에서 추출한 물체의 실제 위치를 알기 위해 영상좌표를 물리적인 좌표로 변환한다면 영상왜곡 정도에 따라 오차가 발생할 것이다[63].

(1) 렌즈왜곡의 종류

렌즈 왜곡에는 크게 방사왜곡(radial distortion)과 접선왜곡(tangential distortion)이 있다.

방사왜곡은 볼록렌즈의 굴절률에 의한 것으로서 영상의 왜곡 정도가 중심에서의 거리에 의해 결정되는 왜곡이다. 접선왜곡(tangential distortion)은 카메라 제조과정에서 카메라 렌즈와 이미지 센서(CCD, CMOS)의 수평이 맞지 않거나 또는 렌즈 자체의 centering이 맞지 않아서 발생하는 왜곡이다.

(2) 렌즈계 왜곡의 수학적 모델

일반적으로 렌즈 왜곡의 수학적 모델은 카메라 내부 파라미터의 영향이 제거된 normalized image plane에서 정의된다.

렌즈계의 왜곡이 없다고 할 경우, 3차원 공간상의 한 점 (X_c, Y_c, Z_c)는 central projection (pinhole projection)에 의해 normalized image plane상의 한 점 (x_{n_u}, y_{n_u})로 투영된다(첨자 n: normalized, u: undistorted). :

$$\begin{bmatrix} x_{n_u} \\ y_{n_u} \end{bmatrix} = \begin{bmatrix} X_c/Z_c \\ Y_c/Z_c \end{bmatrix} \quad 1)$$

그러나 실제로는 (x_{n_u}, y_{n_u})는 렌즈계의 비선형성에 의해 왜곡(주로 radial distortion)이 된다. (x_{n_d}, y_{n_d})를 렌즈계의 왜곡이 반영된 normalized 좌표라면, 렌즈계 왜곡 모델은 다음과 같다(첨자 d: distorted). :

$$\begin{bmatrix} x_{n_d} \\ y_{n_d} \end{bmatrix} = (1 + k_1 r_u^2 + k_2 r_u^4 + k_3 r_u^6) \begin{bmatrix} x_{n_u} \\ y_{n_u} \end{bmatrix} + \begin{bmatrix} 2p_1 x_{n_u} y_{n_u} + p_2 (r_u^2 + 2x_{n_u}^2) \\ p_1 (r_u^2 + 2y_{n_u}^2) + 2p_2 x_{n_u} y_{n_u} \end{bmatrix} \quad 2)$$

$$\text{단, } r_u^2 = x_{n_u}^2 + y_{n_u}^2 \quad 3)$$

위 수식에서, 우변의 첫 번째 항은 radial distortion, 두 번째 항은 tangential distortion을 나타낸다(k_1, k_2, k_3 는 방사왜곡계수:radial distortion coefficient, p_1, p_2 는 접선왜곡 계수:tangential distortion coefficient). r_u 는 왜곡이 없을 때의 중심(principal point)까지의 거리(반지름)이다.

이 때, (x_{n_d}, y_{n_d})는 normalized image plane에서의 좌표이며 실제 영상 픽셀 좌표 (x_{p_d}, y_{p_d})는 카메라 내부 파라미터를 반영하여 다음과 같이 구해진다.

$$\begin{bmatrix} x_{p-d} \\ y_{p-d} \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & skew_c & f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{n-d} \\ y_{n-d} \\ 1 \end{bmatrix} \quad 4)$$

$$\begin{aligned} \text{즉, } x_{p-d} &= f_x(x_{n-d} + skew_c y_{n-d}) + c_x \\ y_{p-d} &= f_y y_{n-d} + c_y \end{aligned} \quad 5)$$

여기서, f_x , f_y 는 초점거리, c_x , c_y 는 렌즈 중심 영상좌표(principal point), $skew_c$ 는 비대칭계수를 나타내는 카메라 내부 파라미터이다.

(3) 영상 왜곡 보정

왜곡된 영상으로부터 보정된 영상을 생성하는 것은 다음과 같은 식에 의해 계산되어진다.

왜곡된 영상을 I_d , 보정된 I_u 라 한다. 기본적으로 I_u 의 각 픽셀값을 해당 픽셀 좌표를 왜곡시켰을 때의 I_d 의 대응되는 픽셀값으로 채우는 것이다.

영상 I_u 의 한 점을 (x_{p-u}, y_{p-u}) 라 하면, 일단 이것을 카메라 파라미터를 역으로 적용하여 normalized 좌표 (x_{n-u}, y_{n-u}) 로 변환한다.

$$\begin{bmatrix} x_{n-u} \\ y_{n-u} \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & skew_c & f_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{p-u} \\ y_{p-u} \\ 1 \end{bmatrix} \quad 6)$$

즉,

$$\begin{aligned} x_{n-u} &= (x_{p-u} - c_x) / f_x - skew_c y_{n-u} \\ y_{n-u} &= (y_{p-u} - c_y) / f_y \end{aligned} \quad 7)$$

다음으로, 중심까지의 거리 r_u ($r_u^2 = x_{n_u}^2 + y_{n_u}^2$)를 구하고 식(2) 왜곡모델을 적용하여 왜곡된 좌표 (x_{n_d}, y_{n_d})를 구한다.

$$\begin{bmatrix} x_{p_d} \\ y_{p_d} \end{bmatrix} = (1 + k_1 r_u^2 + k_2 r_u^4 + k_3 r_u^6) \begin{bmatrix} x_{n_u} \\ y_{n_u} \end{bmatrix} + \begin{bmatrix} 2p_1 x_{n_u} y_{n_u} + p_2 (r_u^2 + 2x_{n_u}^2) \\ p_1 (r_u^2 + 2y_{n_u}^2) + 2p_2 x_{n_u} y_{n_u} \end{bmatrix} \quad 8)$$

마지막으로 (x_{n_d}, y_{n_d})를 다시 픽셀 좌표계로 변환하면 (x_{p_u}, y_{p_u})의 왜곡된 영상에서의 좌표 (x_{p_d}, y_{p_d})를 구할 수 있다.

$$\begin{bmatrix} x_{p_d} \\ y_{p_d} \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & skew_cf_x & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{n_d} \\ y_{n_d} \\ 1 \end{bmatrix} \quad 9)$$

(4) 영상좌표 왜곡 보정

영상처리에서는 보정된 영상을 구하는 것보다는 보정된 영상좌표를 구하는 것이 훨씬 중요하다. 왜냐하면 영상전체를 보정한 후에 영상처리 알고리즘을 적용하는 것보다는 원래 왜곡된 영상에서 처리를 한 후에 그 결과만 보정(undistort)을 하는 것이 효과적일 수 있기 때문이다.

왜곡된 영상좌표 p_d 로부터 왜곡 보정된 좌표 p_u 를 구하는 한 방법은 $p_u = p_d$ 로 가정하고 식 (2)를 이용해서 p_u 를 왜곡시켜 보는 것이다. 이렇게 구한 점이 실제 p_d 와 얼마나 오차가 나는지를 계산해서 그 오차를 p_u 에 역으로 반영시킨다. 예를 들어, p_u 를 왜곡시킨 점이 p_d 보다 δ 만큼 크다면 p_u 를 δ 만큼 감소시킨다. 이러한 과정을 오차가 임계값 이하가 될 때까지 반복한다.

카메라 왜곡을 보정하기 위하여 렌즈왜곡 계수를 구한 후 카메라 보정을 수행한다.

3) 가변 적응형 필터를 적용한 특징점 추출 및 매칭

180도 이상의 어안렌즈를 통해 촬영된 두 영상에는 서로 겹치는 부분이 존재할 것이다. 이 겹치는 부분들을 ‘스티칭 포인트’ 라고 하여 스티칭이 필요하다.

어안렌즈를 사용해서 촬영을 하게 되면 사람의 시야에 비해 왜곡된 영상을 얻을 수밖에 없으며 스티칭 과정에서 부자연스러운 느낌을 줄 수밖에 없다. 카메라의 초점이 맞추어져 있는 중앙 부분의 지점에서는 비교적 왜곡이 적을 수 있겠지만 시야각의 끝쪽 가장자리 부분에서는 부자연스러운 왜곡이 느껴질 수 있어 스티칭 과정에서 이러한 왜곡이 더 심해질 수 있다는 것이다.

이러한 부분을 해결하기 위하여 획득된 영상 특성을 활용하여 중심점을 기준으로 바깥(영상에서 위-아래) 방향으로 영역을 구분하여 각 영상으로부터 특징점을 추출 및 클러스터링을 하고 영역별 가변적 search window mask를 적용하여 매칭을 하게 된다. Fig. 34와 같이 입력된 영상 1과 2의 중첩된 영역에서 중심에 해당되는 영역 1(카메라 중심에서 가까운 거리, Region1 : R1)에서는 search window mask 사이즈를 5x5, 영역 2(Region2 : R2)에서는 7x7, 영역 3(Region3 : R3)에서는 15x15를 설정하여 가변적인 search window mask를 적용함으로써 모든 영역에서 연산을 수행하지 않고 cluster화함으로써 영역 안에서만 효과적인 특징점만을 추출하는 방법이기 때문에 연산 비용을 줄이고 특징점 추출 및 매칭 결과와 수행속도가 개선됨을 확인할 수 있다.

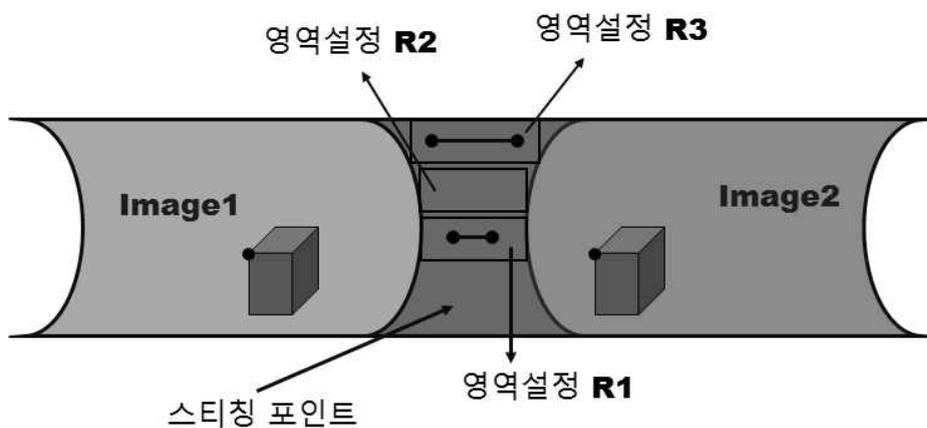


Fig. 34. Overlapping region of input image

Fig. 35는 가변 적응형 필터를 적용한 특징점 매칭 추출과정을 보여준다.

입력된 영상으로부터 렌즈왜곡이 있는 경우 렌즈 보정을 하고, 영역설정 후 각 영역에 해당되는 가변필터를 적용하여 특징점을 추출하고 특징점을 매칭하게 된다.

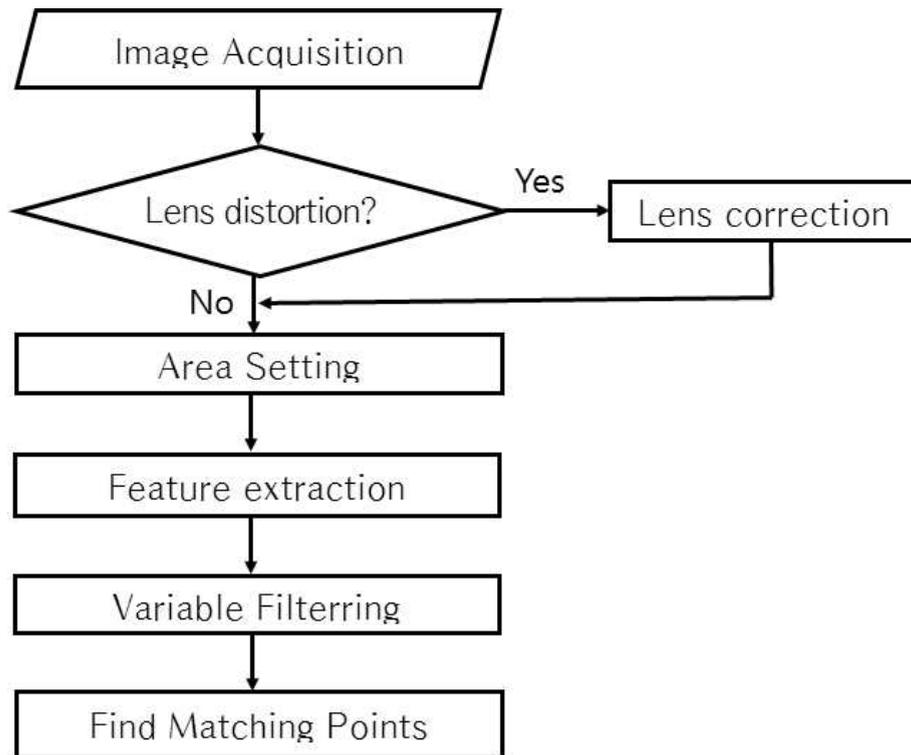


Fig. 35. Feature Extraction Process with Variable Adaptive Filter

Fig. 36은 제안된 알고리즘을 통해서 각 영역별로 추출된 특징점에 대한 매칭 결과를 보여주고 있다. 영역 1에서는 총 12개의 매칭 결과, 영역 2에서는 25개, 영역 3에서는 40개의 매칭 결과를 확인할 수 있다. 결과적으로 입력된 영상에서 영상 1과 영상 2의 총 매칭점의 수는 77개이며 매칭 에러(RMSE)는 2.4의 결과 값을 갖는다.



Fig. 36. Matching results for extracted minutiae

Fig. 37은 특징점 및 매칭점 추출 후 이미지 6장이 정합되는 과정을 보여준다.



Fig. 37. Matching result

Fig. 38은 실습실 영상의 정합된 과정을 통하여 최종 스티칭 한 결과를 보여주고 있다.



Fig. 38. Image stitching result

2. 특징점 추출 및 정합

수평축에 6대의 카메라 수직축에 1대의 카메라를 리그에 연결한 촬영시스템을 통하여 영상을 획득한다. 영상마다 카메라와 카메라간의 거리와 초점거리가 차이가 있기 때문에 파노라마 영상을 만들었을 때 영상이 늘어지는 왜곡 현상이 가끔씩 발생할 수 있다.

영역 선정을 위하여 먼저 입력된 영상을 통해서 입력된 영상의 렌즈 왜곡 (Lens Distortion) 이 필요한 경우에는 입력된 영상에 대하여 preprocessing 과정을 거친다. preprocessing 과정이 필요하지 않는 경우에는 바로 특징점 추출 단계로 넘어간다.

파노라마 영상 생성을 위해 선행되어야 할 과정은 호모그래피 행렬을 구하는 것이다. 호모그래피 행렬은 임의의 네 개 이상의 점이 필요하고, 이 네 개 이상의 매칭쌍을 구하기 위해서 특징점을 찾는 과정이 필요하다. 이미지 스티칭에서 입력 영상의 크기 및 회전 등의 왜곡에 강인한 특징점 추출은 중요한 요소이다. 기존의 방법들은 특징점 추출에 있어 높은 정확도를 보이지만 연산에 대한 복잡도가 높아 수행 속도가 느린 단점을 가지고 있다.

SURF는 특징점 추출을 위해서 고속 헤이시안 검출기(Fast Hessian Detector)를 사용하여 추출한다. 적분영상은 원본 영상의 원점으로부터 각 위치까지의 픽

셀값의 합을 저장한 영상을 말하는데 적분영상을 이용하여 바로 부분 영상을 표현할 수 있고, 부분 영역에 대하여 합을 바로 구해낼 수 있다. 적분 영상을 생성하여 이용하면 필터의 크기와 상관없이 임의의 크기의 영상 영역을 선택하더라도 4번의 연산을 수행하면 선택된 영역의 모든 화소의 합을 구할 수 있다.

적분 영상을 이용하여 영역 계산 방법은 다음과 같다.

적분영상은 원본 이미지의 픽셀값의 합을 계산하는 방법이므로 네 개의 점 (a,b,c,d)의 위치만 알면 사각형 영역내의 면적 값을 쉽게 계산할 수 있기 때문에 빠른 시간내에 연산이 가능하고, 연산 수행 속도를 효율적으로 향상시킬 수 있다.

먼저 입력 영상의 사각형 영역의 특정위치에서 모든 화소값을 더하면 적분 영상을 생성할 수 있다.

d영역의 픽셀 값을 얻기 위해서 점 d까지의 넓이에서 점 b까지의 넓이와 점 c까지의 넓이를 뺀 후 두 번 빼진 점 a 까지의 넓이를 한 번 더해줌으로써 점 d 영역의 넓이를 구할 수 있는 것이다.

$$([4] = [1+2+3+4] - [1+3] - [1+2] + [1])$$

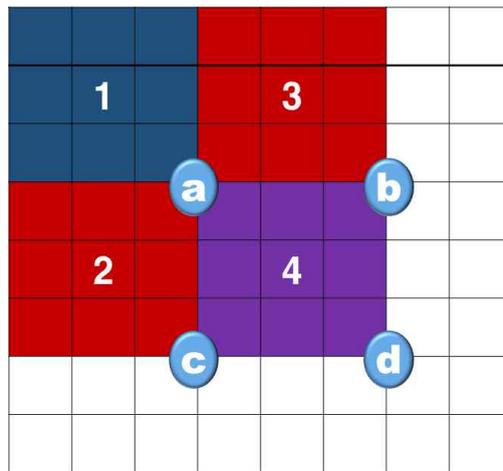


Fig. 39. Area calculation using integral image

Hessian Detector는 Hessian Matrix를 기반을 둔 특징점 추출 알고리즘으로 속도와 정확도에서 좋은 성능을 보인다. 또한 적분 영상을 이용하여 연산량을 효

율으로 감소시킨다. Hessian Detector는 박스 필터를 사용하여 원본 영상의 스케일은 변화하지 않고 박스 필터의 크기만 증가시켜 다양한 스케일의 변화에 대한 특징점을 추출한다.

획득된 영상에서 많은 특징점들이 추출되게 되면 정합 과정에서 연산량이 많게 되어 처리 속도가 떨어지게 된다. 또한 특징점들의 수가 너무 적게 되면 정보가 부족하여 이미지 스티칭이 정확하게 이루어질 수 없게 된다.

이러한 문제점을 해결하기 위해서 입력된 이웃하는 두 영상의 중첩되는 영역에서 중심에 해당되는 영역 1과 영역 2, 영역 3에 가변적인 search window mask를 적용함으로써 모든 영역에서 연산을 수행하지 않고 영역 안에서만 의미 있는 특징점만을 추출하는 방법을 제안하였다. 가변 적응형 필터를 적용하여 영역내의 존재하는 특징점만을 선별하여 추출하는 방법을 적용한다. 영상에서 의미 있는 특징점들은 대부분 모여 있는 특성을 가지고 있다. 반대로 중심 지역에서 벗어난 지역의 특징점들은 잡음 및 오류 등으로 분류되어 의미 없는 특징점들이 될 수 있다. 본 논문에서는 추출된 특징점들 중에서 영역 내의 효과적인 특징점들을 선별하여 추출할 수 있도록 하였다. 가변 적응력 필터를 통한 특징점 추출은 헤이시안 행렬식으로 추출된 특징점들 중에서 가변 필터내의 특징점 개수를 계산한다.

입력 영상과 이에 대응되는 영상에서 유사한 부분이 있다면 각각의 영상이 가지고 있는 특징점들 또한 유사하게 된다. 즉, 입력 영상과 유사한 부분을 찾기 위해서 대응되는 영상에서 유사한 특징점을 찾아 대응되는 위치를 계산할 수 있다. 이렇게 서로 대응되는 유사한 특징점을 찾기 위해서는 특징점 정합(Feature Matching) 과정이 필요하다. 특징점 정합과정은 두 영상의 특징점들을 서로 비교하여 가장 유사한 특징점과 정합시키는 과정이다[22].

이미지 스티칭은 입력 영상에서 추출된 특징점과 대응되는 점들을 이용하여 수행 된다. 따라서 추출된 특징점 데이터가 부정확하면 부정확한 이미지 스티칭을 수행하기 때문에 추출된 특징점의 정확도를 높이는 과정이 중요하다. 그러나 특징점 추출 과정 중에 불필요한 특징점으로 인하여 잘못된 오차가 추출될 가능성은 항상 존재하며, 이러한 오차가 많을수록 심각한 오류 결과를 가져올 수 있다.

1) RANSAC을 이용한 호모그래피 계산

이웃하는 두 영상을 정합하기 위해서는 하나의 영상에서 추출된 특징점 중에서 다른 영상에서 대응되는 특징점을 찾아 대응 시키는 과정이 필요하다[60]. 특징점 기반의 정합 방법은 호모그래피(Homography)를 통하여 이루어진다. 호모그래피 행렬은 정합된 네 점의 좌표를 가지고 구할 수 있다. 행렬을 구할 네 점의 좌표의 특징점들은 선택하는 과정에서 이상점(outlier)을 스티칭을 하기 위한 특징점 좌표로 이용될 경우 잘못된 호모그래피 행렬을 구할 수 있다. 이러한 이상점들을 제거하기 위하여 RANSAC기법을 이용한다. RANSAC기법은 에러와 잡음이 함께 있는 데이터들로부터 적절한 모델을 예측하여 오차를 최소화하는 방법이다[30]. 각 영상에서 가장 정확하게 일치된 네 점의 좌표를 찾고 영상간의 호모그래피를 계산한다. 계산된 호모그래피를 이용하여 한 영상을 좌표 변환한 뒤 정합하게 되면 영상이 생성된다. 호모그래피를 구하는 방법은 호모그래피 행렬식에 의하여 구한다.

3. 360도 파노라마 영상 촬영시스템의 동기화

현실감을 증대시키고 넓은 영상 시야각을 확보하기 위해서는 고해상도의 360도 파노라마 영상을 획득할 수 있는 방법은 한 대의 카메라를 축(axis)을 중심으로 회전시켜 영상을 획득하거나 다수의 카메라를 활용하여 동시에 다수의 영상을 획득하는 방법이 있다.

다수의 카메라로 영상을 획득할 경우, 각각의 카메라에서 촬영된 영상간의 공간적 끊김이 없고 영상의 일부분이 겹치도록 카메라를 배치하여야 한다. 이 때, 카메라들은 동일한 프로젝션 센터를 가질 수 없기 때문에 시차(parallax)가 발생하게 된다. 시차는 영상합성을 통하여 해결할 수 없으므로 영상획득 시 이를 최소화 할 수 있도록 카메라 간의 간격을 줄여 시차를 최소화 하여야 한다.

앞장에서도 설명을 했지만 영상 획득시 시차발생을 최소화하기 위하여 카메라를 동시 촬영하는 방법이 필요하다. 본 논문에서는 여러 대의 카메라를 리그에 장착하여 카메라를 동기화를 시키기 위하여 다중카메라 동기화시스템을 다음과 같이 제안 하였다.

1) 시스템 동기화

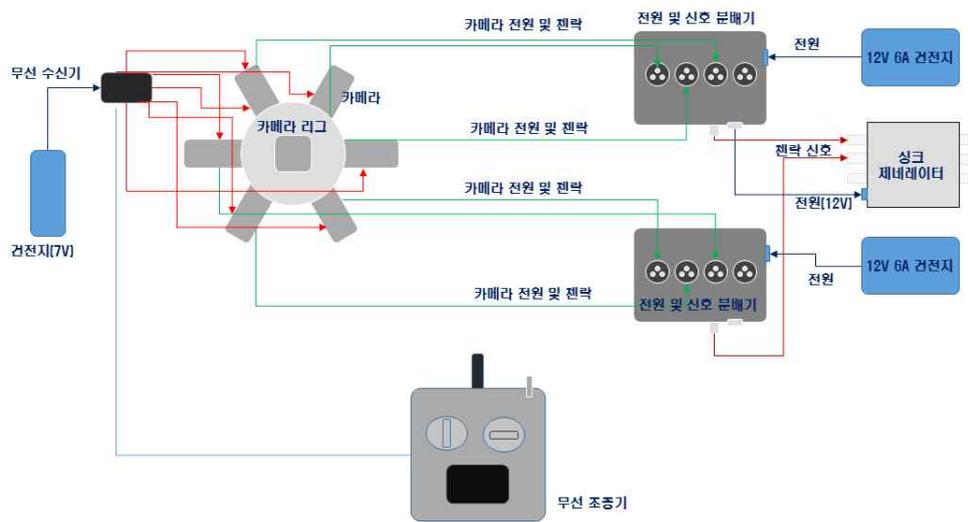


Fig. 40. Multi camera synchronization system

시스템 동기화 절차는 다음과 같다.

- ㄱ) Sync generator로부터 나오는 신호를 분배기를 통해 오실로스코프에 연결한다.
- ㄴ) 분배기를 통해 나오는 신호 파형을 측정된 모습을 확인한다. - 동일한 파형이 서로 겹치는지 확인한다.
- ㄷ) 신호 파형 스케일 확대하여 파형이 중복되어 있는지 확인한다.
- ㄹ) 신호 파형 간 지연 시간을 측정한다.
- ㅁ) 오실로스코프 파형 분석을 통해 동기화가 이루어짐과 동시에 동시 신호를 받은 다중 카메라의 동시 촬영이 이루어지는지(녹색--> 빨간색) 확인한다.

또한 시스템 동기화가 이루어진 후에는 동시촬영을 위한 기본적인 조건을 다음과 같이 제시하였다.

- ㄱ) 동일한 영상촬영을 위해 색온도 /셔터 스피드 /노출 등 수동 셋팅을 하고 촬영한다.
- ㄴ) 동일한 화각을 유지하기 위해서 동일한 촬영거리(5m 이상, 50%이상 중복 ; BlackMagic Micro Studio Camera 4K를 사용하여 인접한 카메라와의 중복 촬영을 50%이상 겹치게 촬영)를 유지한다.
- ㄷ) 카메라를 동시에 동기화를 시킨다 : 카메라 7대를 같은 조건으로 셋팅하여 동일한 조건에서 촬영한다.
 - 영상획득을 위해 사용되는 카메라 : BlackMagic Micro Studio Camera 4K, 150도 어안렌즈 사용)

2) 동시 촬영 여부 확인

Fig. 41에서와 같이 동시 촬영 신호가 들어가기 전에는 카메라가 대기상태(준비 상태)로 초록색 불이 들어와 있으며 동기 신호를 통해 동시 촬영 신호를 주게 되면 카메라는 붉은색으로 촬영 단계에 들어간다. 총 6대의 카메라가 동시에 촬영되고 준비 단계로 바뀌는 것을 확인할 수 있으며 동시 촬영 100% 달성 하였다.



Fig. 41. Simultaneous recording signal Ready (greent) and ready Recording status after the simultaneous recording signal (red)

고해상도의 360도 파노라마 영상을 촬영하기 위해서 여러 대의 카메라를 이용하여 전 방향 영상을 획득하게 되는데 이를 위해 여러 대의 카메라를 지지할 수 있는 리그를 설계하였다. Fig. 42는 파노라마 영상 촬영을 위해 제작된 리그이다.



Fig. 42. Rig

Fig.43과 같이 카메라를 배치하게 되면 무인항공 촬영시 기체가 화면에 잡혀 360도 파노라마 영상을 제작하기가 어려워 항공 촬영에는 적합하지가 않다.

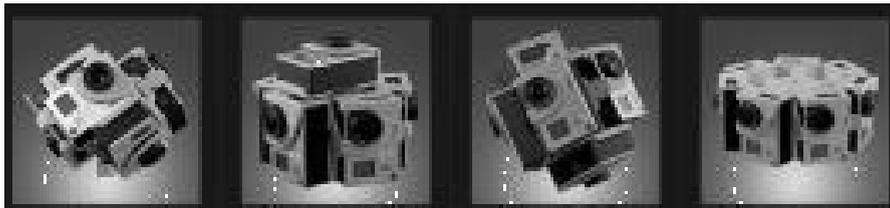


Fig. 43. 360° VR Rig

따라서 Fig.44와 같이 긴 지지대의 상단부에 1대의 카메라를 장착하고 하단에 2대의 카메라를 이용한 무인 촬영용 360도 파노라마 VR 리그를 제작하였다. 각 카메라에는 220도 화각을 가지는 어안렌즈를 장착하여 360도 전방향 촬영이 가능하도록 하였다. 최소 3대의 카메라를 이용함으로써 인해 무게를 줄이고 후반 작업에서 복잡한 스티칭 문제를 쉽게 해결할 수 있는 장점이 있다.



Fig. 44. Rig for aviation shooting

리그(Rig)는 여러 대의 소형카메라 또는 렌즈를 하나의 어레이에 균일하게 장착하여 각 카메라에서 촬영된 영상을 이어 붙여 하나의 360도 파노라마 영상을 완성할 수 있도록 해 주는 것이다. 이 리그는 VR 카메라 5대 이상이 장착되어 고화질의 영상뿐만 아니라 위아래 왜곡 없는 자연스러운 1인칭 촬영이 가능하도록 설계하였다.

IV. 실험 및 결과

1. 실험 환경

본 논문에서는 Table 1과 같은 성능과 환경에서 구현하였다. 실험환경은 Microsoft Visual Studio C++ 2015 와 OpenCV 2.4.8 라이브러리를 이용하여 구현하고, 3.40GHz 의 인텔 i5 쿼드코어 프로세서를 이용하였다. 사용 이미지 Size 는 1920*1080이다.

시스템의 구성은 다음과 같다.

Table 1. 360 Panoramic system configuration

구분	세부사항
▪ 카메라	BlackMagic Micro Studio Camera 7대(Micro Studio Camera 4K Expansion Cable 7개 등)
▪ 어안렌즈(FishLens)	화각 180도 7대(F3.5/7.5mm)
▪ 모니터	시스템 체크용
▪ 리그	자체제작
▪ PC	Pentium i5/8GB PC
▪ 운영체제	Original software for Windows10
▪ Panorama Editor	Video Stitcher

2. 실험 데이터

본 논문에서 사용된 모든 실험 데이터는 360도 파노라마 영상을 생성하기 위해 3840x2160 해상도의 BlackMagic 사의 BlackMagic Micro Studio Camera로 직접 촬영한 영상들을 실험 영상으로 사용하였다. 영상을 촬영할 때 렌즈의 왜곡, 컬러 보정 등의 부가기능을 제외하고 촬영하였으며, 실험영상은 다량쉬 오름, 교내의 건물과 산굼부리, 성읍 민속마을 등이 포함된 영상이다.

Table 2. Depending on the number of camera lenses

구분	장점	단점	사용환경
렌즈 2개	<ul style="list-style-type: none"> • 카메라 작고 가벼워 기동성 좋음 • 렌즈개수 작아 스티치 작업 간편 	<ul style="list-style-type: none"> • 렌즈 개수가 적어 화질 떨어짐 • 렌즈이동을 통한 스티치 수동 불가 	<ul style="list-style-type: none"> • 가벼운 촬영
렌즈 6개	<ul style="list-style-type: none"> • 고프로 등에 하드케이스 설치 후 장착 가능 • 방수 및 방진에 용이, 거친 촬영 적합 	<ul style="list-style-type: none"> • 카메라간의 거리 존재 • 6개의 카메라만 이용하여 다소 왜곡 있음 	<ul style="list-style-type: none"> • 힘든 환경
렌즈 7개	<ul style="list-style-type: none"> • 마운트 재질 따라 가벼움 • 7개의 카메라를 사용하는 방식, 왜곡 작음 • 수평 촬영을 통한 다양한 분야 적용가능 	<ul style="list-style-type: none"> • 나사 형식이라 탈부착 불편 • 물 먼지에 취약 	<ul style="list-style-type: none"> • 인물 및 공간

실험을 위해 사용한 영상데이터는 직접 촬영한 것으로서, 표 4와 같이 총 14개의 원천데이터를 확보하였다. 각 영상데이터는 제주의 오름, 공원과 유명 관광지(풍경), 해수욕장(바다), 학교 및 실습실 등에서 촬영하였다.

본 논문에서는 실험 원천데이터의 평균 특징점 개수의 분포에서 상위 10%와 하위 10%를 임계치로 하여 설정하였다.

Table 3. Source data

구분	영상 source	촬영영상수	영상크기	비고
1	다랑쉬오름(실외)	7	360*180	
2	용눈이오름(실외)	7	360*180	
3	아부오름(실외)	7	360*180	
4	금호세계교육관 컨퍼런스홀(실내)	7	360*180	
5	수목원공원(실외)	7	360*180	
6	대학 공용주차장 1(실외)	7	360*180	
7	대학 공용주차장 2	7	360*180	
8	IT실습실(실내)	7	360*180	
9	탐동공연장(실외)	7	360*180	
10	이호해수욕장(실외)	7	360*180	
11	함덕 다려도(실외)	3	360*180	
12	산굼부리(실외)	6	360*180	
13	성읍민속마을(실외)	6	360*180	
14	더마클럽(실외)	2	360*180	

본 논문에서는 다음과 같은 몇 가지 사항을 고려한다.

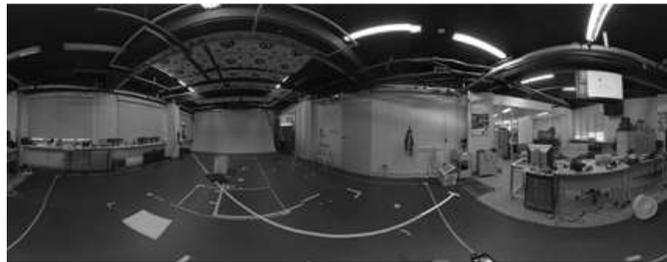
- ① 촬영된 모든 영상들은 25~50% 겹치는 영상들이다.
- ② 촬영한 영상의 size는 같다.
- ③ 촬영방향은 동일하며, 카메라 회전은 없다.
- ④ 스티칭 한 후 영상에 대한 후보정은 하지 않는다.
- ⑤ 촬영대상과의 촬영거리는 5m이상으로 한다.

3. 기존 알고리즘과의 비교분석 결과

실험은 최근 특징점 추출에 많이 사용되는 SIFT 및 SURF 알고리즘과 제안한 알고리즘을 비교하였다. Test이미지는 직접 촬영한 대학전경, 실습실전경, 금호세계교육관 로비 영상을 사용하였다.



a) 대학전경



b) 실습실 전경



c) 금호세계교육관 로비

Fig. 45. Comparison image result image

실험은 동일한 영상을 SIFT, SURF 알고리즘과 제안한 알고리즘에 적용하여 수행시간 및 특징점 개수를 비교하는 방식으로 수행되었다. 실험결과는 아래 Table 4에 나타난 것과 같다. 제안한 알고리즘은 수행시간에서 SIFT와 SURF보다 우수한 성능을 나타내었다.

Table 4. Comparison of experimental results

이미지	알고리즘	평균 특징점 추출수행시간(s)	특징점수 (개)
대학전경 (image1)	SIFT 알고리즘	3.18	1,566
	SURF 알고리즘	3.07	1,274
	제안 알고리즘	2.62	190
실습실 전경 (image2)	SIFT 알고리즘	2.66	1,430
	SURF 알고리즘	2.56	1,325
	제안 알고리즘	2.04	198
금호세계교육관 로비 (image3)	SIFT 알고리즘	1.66	783
	SURF 알고리즘	1.57	679
	제안 알고리즘	1.26	143

Table 4에서 보듯이 기존의 알고리즘과 비교하여 제안한 방법이 향상된 성능을 보여주고 있다. 다량취 오름의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 3.18초, 3.07초 및 2.62초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 18.7%, 14.7% 감소되었다. 총 특징점 개수는 SIFT, SURF 및 제안 방법에서 각각 1,566개, 1,274개 및 190개로 나타난다.

실습실 전경의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 2.66초, 2.56초 및 2.04초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 23.3%, 20.3% 감소되었다. 총 특징점 개수는 SIFT, SURF 및 제안 방법에서 각각 1,430개, 1,325개 및 198개로 나타난다.

금호세계교육관 로비의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 1.66초, 1.57초 및 1.26초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 19.7%, 18.5% 감소되었다. 총 특징점 개수는 SIFT, SURF 및 제안 방법에서 각각 783개, 679개 및 143개로 나타난다.

실험결과에서 전체적인 특징점 추출 시간이 SIFT와 SURF보다 제안한 방법이 크게 향상된 것을 알 수 있었다.

제안된 알고리즘은 가변 적응형 필터 특징점 추출 영역을 설정하여 입력된 두 영상의 중첩되는 영역을 중심으로 각각 추출 영역을 설정하고, 그 안에서만 특징점을 추출하는 방법이다. 파노라마 이미지 생성과정에 있어서 실제 대응하는 점들을 크게 다루기 때문에 불필요한 영역의 특징들은 오히려 연산속도에 저해요인이 될 수 있다. 영역설정의 크기에 따라 추출할 특징점의 개수가 감소되고, 알고리즘의 연산 수행 속도가 달라진다. 따라서 특징점 추출속도 뿐만 아니라 대응점을 찾기 위한 매칭 횟수가 감소되고 속도는 향상된다.

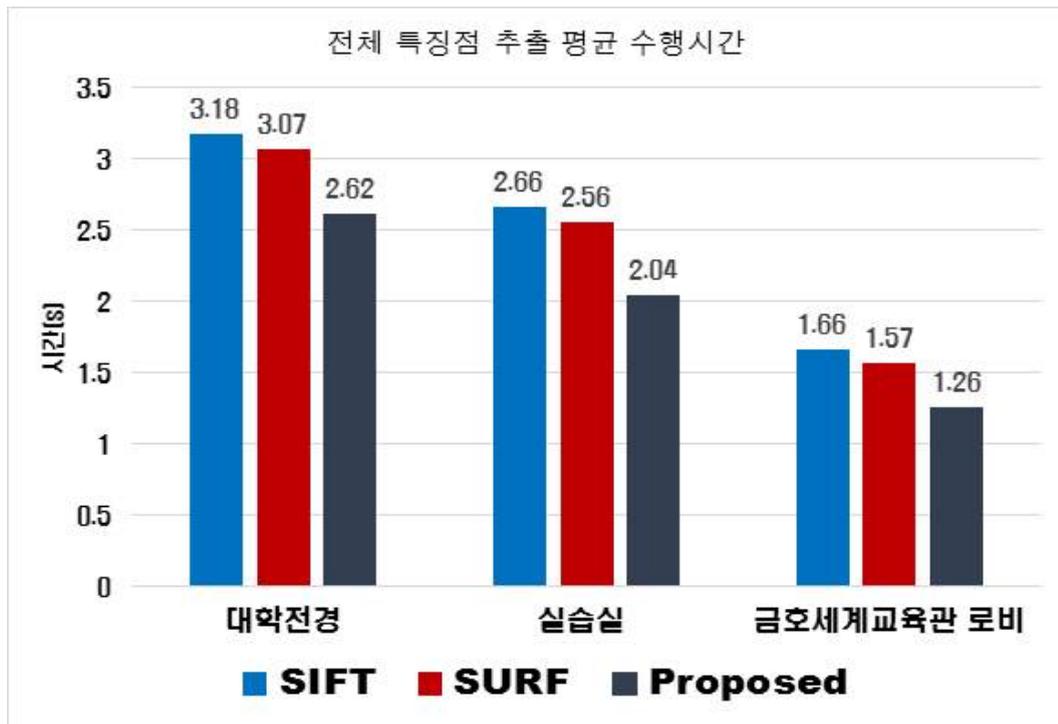


Fig. 46. Total run time comparison



Fig. 47. Comparing university performance time

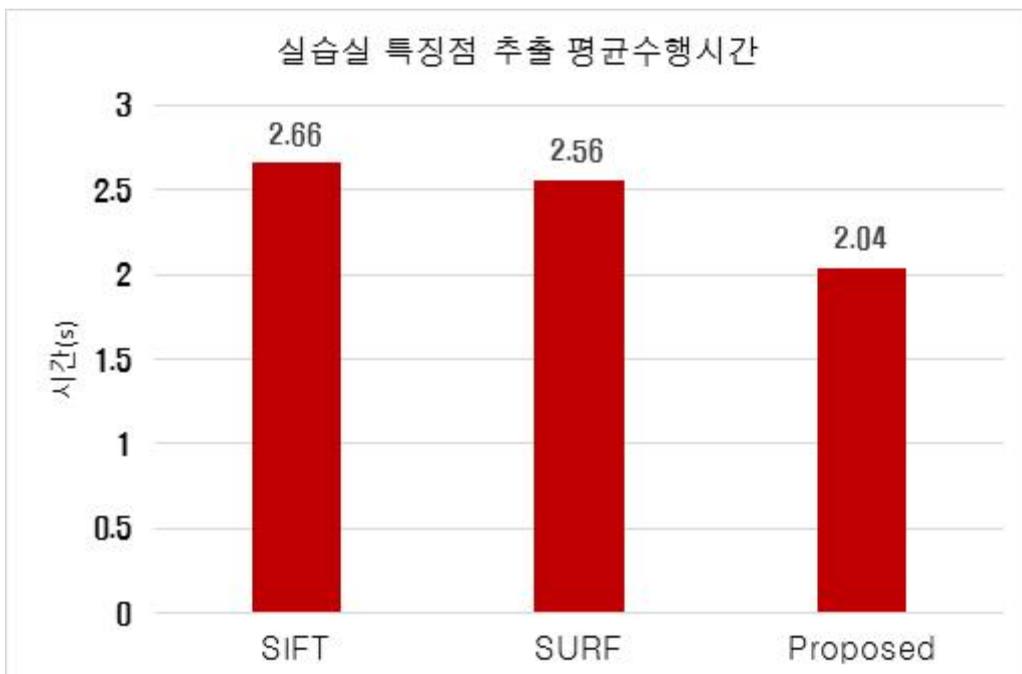


Fig. 48. Comparison of Practice Time

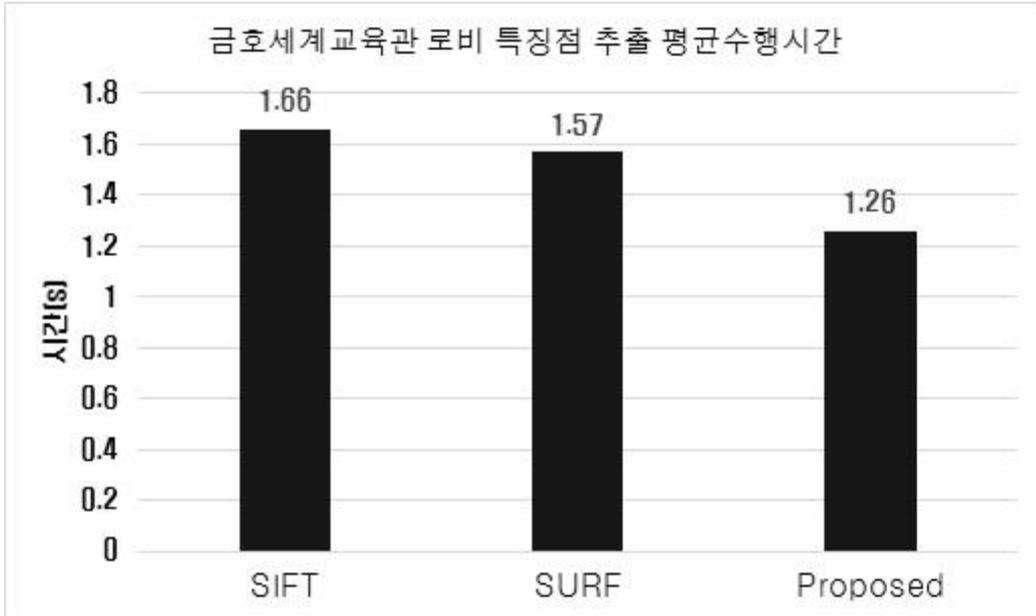


Fig. 49. Kumho World Education Center Lobby Performance Time Comparison

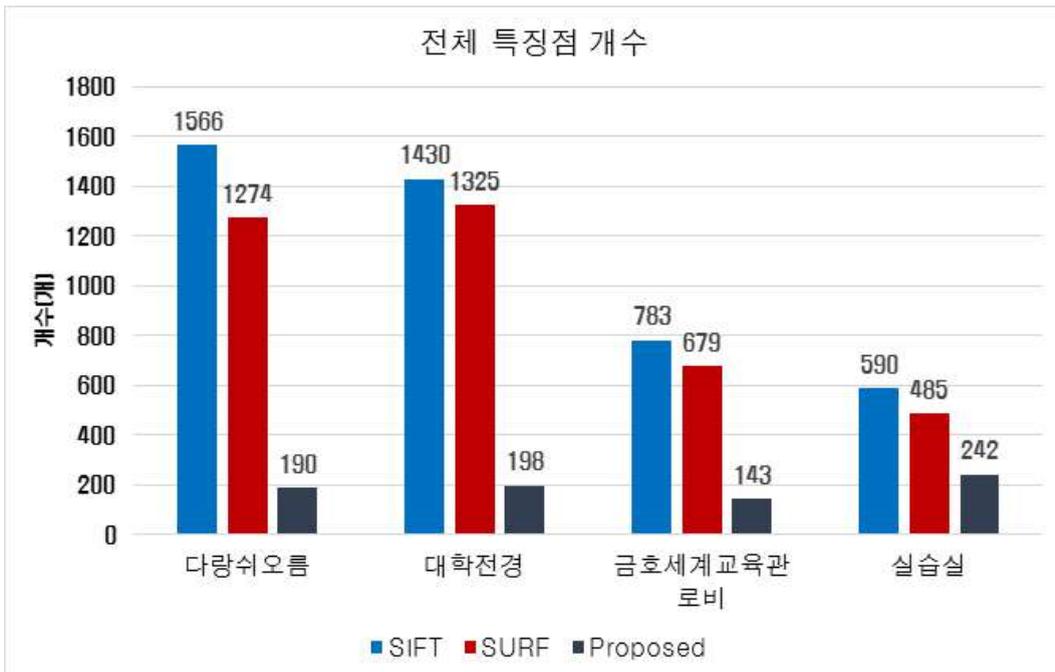


Fig. 50. Comparison of feature points of existing algorithms and proposed method

Table 5에서는 기존의 SIFT, SURF알고리즘과 제안 방법 중에서 속도향상 및 성능에 대한 부분을 확인하기 위하여 대학전경, 실습실전경 금호세계교육관 로비에 대한 매칭점 추출시간과 총 스티칭 시간을 측정하였다.

Table 5. Feature extraction time and total stitching time comparison(sec)

구분		매칭점 평균 추출시간	총 스티칭 평균시간
대학전경	SIFT	4.59	8.07
	SURF	3.35	7.28
	제안	3.24	6.44
실습실전경	SIFT	4.03	6.95
	SURF	4.12	6.85
	제안	3.36	5.58
금호세계교육관 로비	SIFT	2.37	4.17
	SURF	2.16	3.93
	제안	1.98	3.53

Table 5에서 알 수 있듯이 기존의 알고리즘과 비교하여 제안한 방법이 향상된 성능을 보여주고 있다. 대학전경의 평균 매칭점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 4.59초, 4.21초 및 3.82초로 나타난다. 총 스티칭 평균시간은 SIFT, SURF 및 제안 방법에서는 각각 8.07초, 7.28초 및 6.44초로 나타난다. 따라서 이미지 스티칭 총 시간은 SIFT보다 20.2%, SURF 보다는 11.5% 감소되어 전체적으로 향상된 것을 알 수 있다.

실습실 전경의 평균 매칭점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 4.03초, 4.12초 및 3.36초로 나타난다. 총 스티칭 평균시간은 SIFT, SURF 및 제안 방법에서는 각각 6.95초, 6.85초 및 5.58초로 나타난다. 이미지 스티칭 총 시간은 SIFT보다 19.3%, SURF 보다는 18.5% 감소되었다.

금호세계교육관 로비의 평균 매칭점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 2.37초, 2.16초 및 1.98초로 나타난다. 총 스티칭 평균시간은 SIFT,

SURF 및 제안 방법에서는 각각 4.17초, 3.93초 및 3.53초로 나타난다. 이미지 스티칭 총 시간은 SIFT보다 15.4%, SURF 보다는 10.2% 감소되었다.

실험결과에서 볼 수 있듯이 전체적인 스티칭 시간이 SIFT와 SURF보다 제안한 방법이 크게 향상되어 결과적으로 연산 수행 속도가 향상된 것을 알 수 있었다.

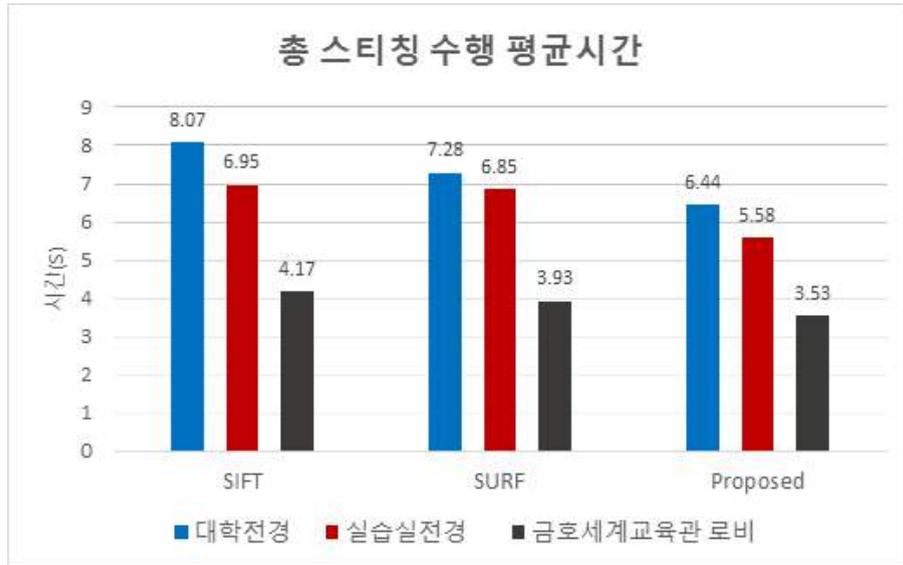


Fig. 51. Comparison of average time of total stitching

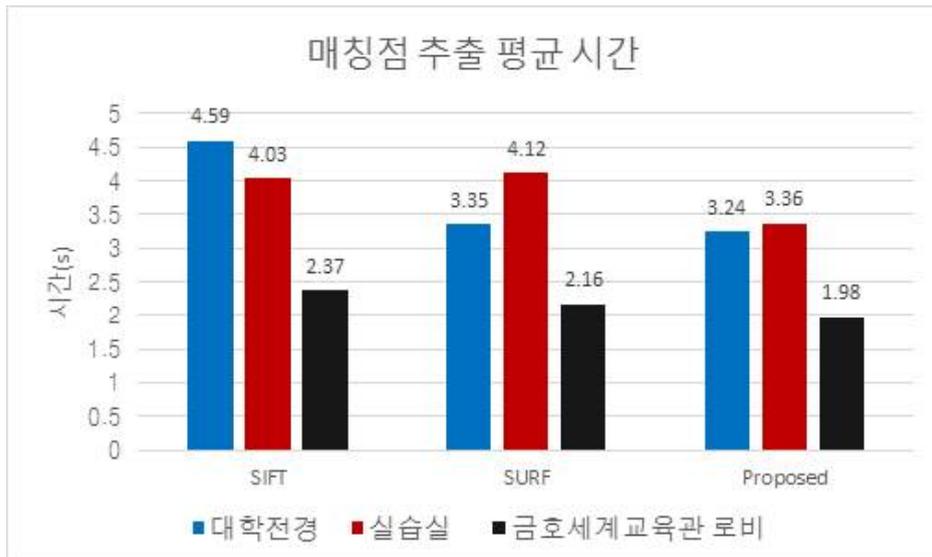


Fig. 52. Comparison of total matching point extraction time

4. 동기화 스티칭 결과

360도 파노라마 영상 획득 및 다중 카메라 간의 시차를 줄이기 위하여 다중카메라 촬영시스템을 제작하였다. 이 시스템을 통하여 시스템 동기화 및 동시에 촬영한 영상을 가지고 스티칭을 한 결과 목표값보다 적은 결과값을 확인할 수 있었다.

① 스티칭 방법

첫 번째 영상과 두 번째 영상의 중첩되는 영역에서 특징점을 추출하고 추출된 특징점을 매칭하여 호모그래피를 구한 후 호모그래피를 이용한 두 번째 영상을 역변환하여 첫 번째 영상에 중첩시킨다.

호모그래피를 구하는 방법은 호모그래피 행렬식에 의하여 구한다.

Fig. 53은 첫 번째 영상과 두 번째 영상의 특징점을 추출하여 추출된 좌표를 나타내고 있으며, 추출된 좌표를 통해 호모그래피를 구하고 이를 통한 스티칭 후 두 좌표간의 거리값(스티칭 오차)을 계산한 결과를 보여주고 있다.



Fig. 53. The feature points of the first image and the second image and Distance between two coordinates after stitching using homography H (Stitching error)

예) 왼쪽 영상의 바닥에 있는 5번 좌표(x:1692.90, y:237.33) 는 오른쪽 영상의 바닥 5번 좌표(x:1685.73, y:946.85)와 매칭된 결과이며 이를 통해 스티칭 후 두 좌표간의 거리값은 1.38 이다. 이 거리 값이 곧 스티칭 오차이다.

Fig. 54~55은 각각 두 번째와 세 번째, 세 번째와 네 번째 영상에 대한 추출된 대응점과 거리값을 나타내고 있다.

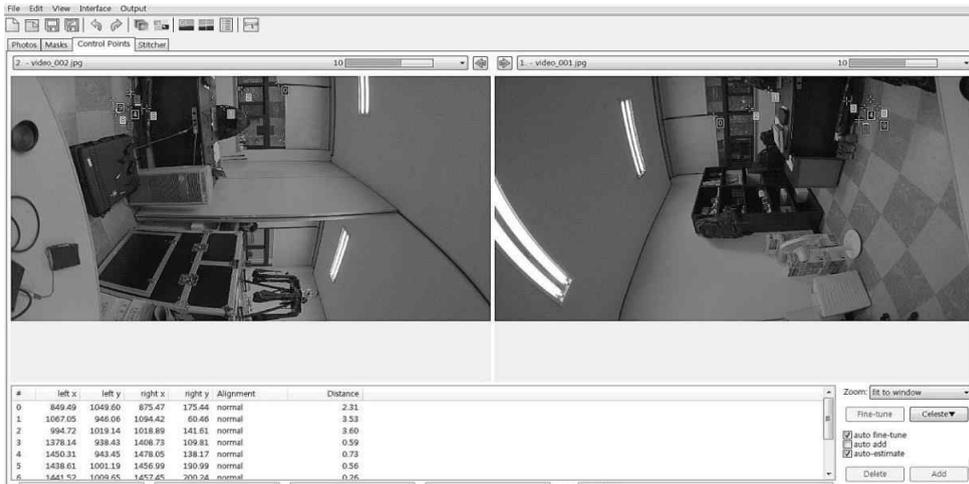


Fig. 54. The feature points of the second image and the third image and distance between two coordinates after stitching using homography H



Fig. 55. The feature points of the third image and the fourth image and Distance between two coordinates after stitching using homography H

Fig. 56은 최종 스티칭 된 결과를 보여주고 있다.

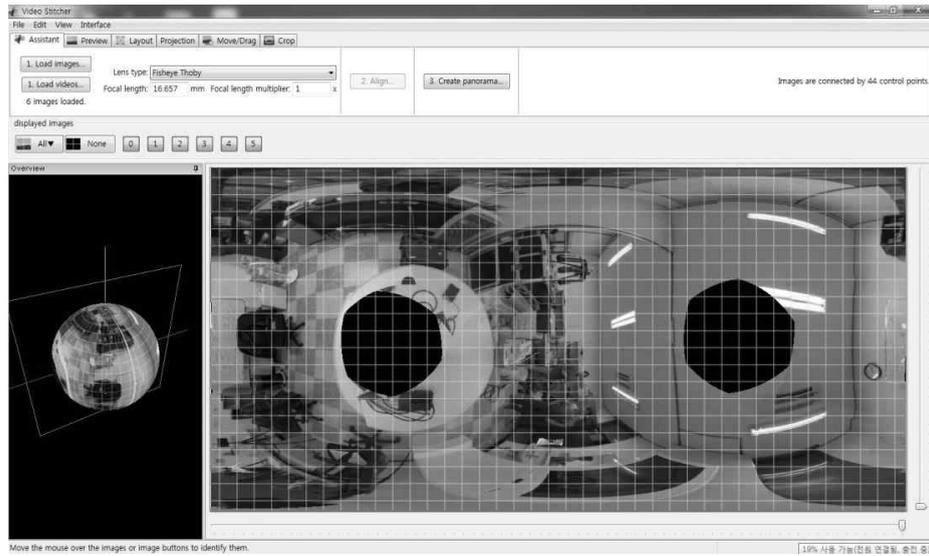


Fig. 56. Stitching result

다음의 표는 총 6대 카메라로 촬영된 영상에 대하여 추출된 특징점을 통한 거리값을 나타내고 있다. 총 6대 카메라로부터 이웃한 영상과의 총 특징점 개수는 42개이며 이를 통한 거리값은 평균 1.78로 목표치 2픽셀 이내를 확인할 수 있다.

Table 6. Stitching error (average)

Index	Distance	Index	Distance	Index	Distance
1	3.91	15	2.96	29	2.53
2	3.43	16	0.96	30	1.83
3	0.3	17	1.51	31	3.04
4	1.33	18	1.9	32	1.34
5	2.16	19	2.94	33	0.36
6	2.88	20	0.95	34	1.46
7	1.5	21	2.62	35	1.89
8	2.31	22	1.26	36	2.69
9	3.53	23	1.39	37	0.46
10	3.6	24	0.71	38	2
11	0.59	25	2.22	39	0.34
12	0.73	26	1.1	40	2.02
13	0.56	27	1.87	41	1.8
14	0.26	28	2.26	42	1.38
Average 1.783					

360도 파노라마 영상은 하나 이상의 카메라로부터 획득된 여러 장의 영상을 합성하여 넓은 시야각을 가지는 하나의 영상을 만드는 기술로, 좌우 최대 360°, 상하 최대 180°등 사용자에게 넓은 시야각을 제공함으로써 특정 장소에 가지 않더라도 마치 현장에 와 있는 듯한 현실감을 제공할 수 있다. 이러한 파노라마 영상은 주요 관광지, 유적지, 관공서, 전시장 등 특정장소나 지역 등을 표현하는데 효과적으로 쓰일 수 있다.

5. 기존 소프트웨어와의 비교

본 실험에서는 동일한 영상데이터를 가지고 각각의 프로그램에서 실행하여 스티칭 오차(RMSE, 거리값)와 특징점 수를 계산하였다. 계산된 값을 가지고 기존 스티칭 프로그램으로 많이 사용되고 있는 “C제품”과의 비교를 하였다.

각 영상에서 중첩되는 1번~7번 영상이미지 영역에서 특징점을 추출하고 추출된 특징점을 매칭하여 호모그래피를 구한 후 호모그래피를 이용한 두 번째 영상을 역변환하여 첫 번째 영상에 중첩시킨다. 이러한 방법으로 7번째 영상이미지까지 반복하여 거리값과 특징점 개수를 계산한다. 호모그래피를 구하는 방법은 호모그래피 행렬식에 의하여 구하고 이를 통한 스티칭 후 두 좌표간의 거리값을 계산한다.

전체적으로 수행한 결과 값은 기존소프트웨어와 제안 방법과 비교했을 때 7대의 카메라로부터 이웃한 영상과의 총 특징점 개수는 다량쉬 오름 1,235개, 대학전경 1,272, 금호세계교육관 로비 558개이며 이를 통한 거리값은 2.78, 2.97, 3.42이다. 제안한 알고리즘에서는 특징점 개수는 다량쉬 오름 172개, 대학전경 183, 금호세계교육관 로비 103개이며 이를 통한 거리값은 2.50, 1.02, 3.10으로 기존 소프트웨어의 결과값 보다 낮음을 알 수 있다. 이는 이웃한 영상간의 거리값이 줄어들어 스티칭 정확도가 높음을 알 수 있었다.

동일한 영상에 대하여 자동 스티칭의 결과로서 제안기법과 기존소프트웨어는 올바른 매칭점을 찾아서 합성에 성공하였으나, 일부에서는 매칭점 판단 오류로 인하여 잘못된 스티칭오류의 결과를 보여주기도 하였다.

자동 스티칭의 매칭 포인트와 수동 스티칭의 매칭 포인트와의 평균 거리오차가 5픽셀 이하일 경우 스티칭에 성공했다고 판단하였다.

Table 6에서 보는 바와 같이 평균 오차 거리(오차율)는 제안한 방법이 기존의 소프트웨어인 “C제품”보다 좋은 것을 확인하였다. 또한 특징점 개수를 찾는 것도 “C제품”과 비교해도 현저하게 적음을 알 수 있다.

Table 7은 본 실험에서 테스트 한 3개의 영상에 대하여 기존 방법과 제안방법의 RMSE와 특징점 개수를 계산한 결과이다.

Table 7. Stitching error rate and control points measurement results of existing software and proposed algorithm

Test 영상	previous “C제품”		Proposed	
	Control Points	RMSE	Control Points	RMSE
다랑쉬 오름	1,235	2.78	172	2.10
대학 전경	1,272	2.97	183	1.02
금호세계교육관 로비	558	3.42	103	3.10

Fig. 57은 기존소프트웨어와 제안방법의 특징점 개수를 비교한 결과이다



Fig. 57. Comparison of existing method and proposed method RMSE

Fig. 58은 기존소프트웨어와 제안방법의 평균 거리값(RMSE)를 비교한 결과이다

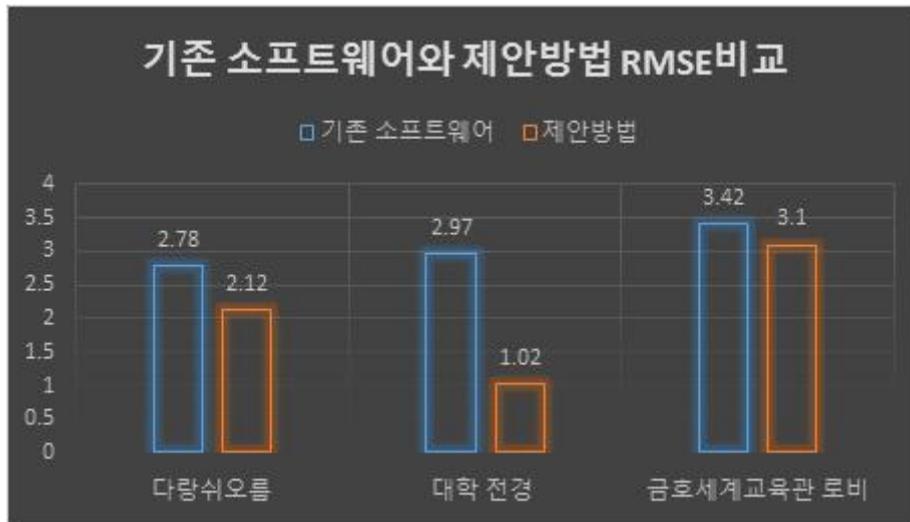
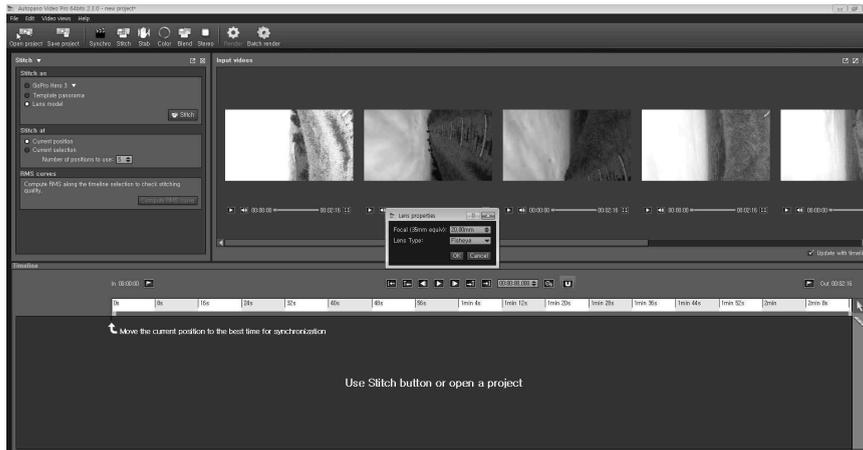


Fig. 58. Existing method and proposed method Control point comparison result

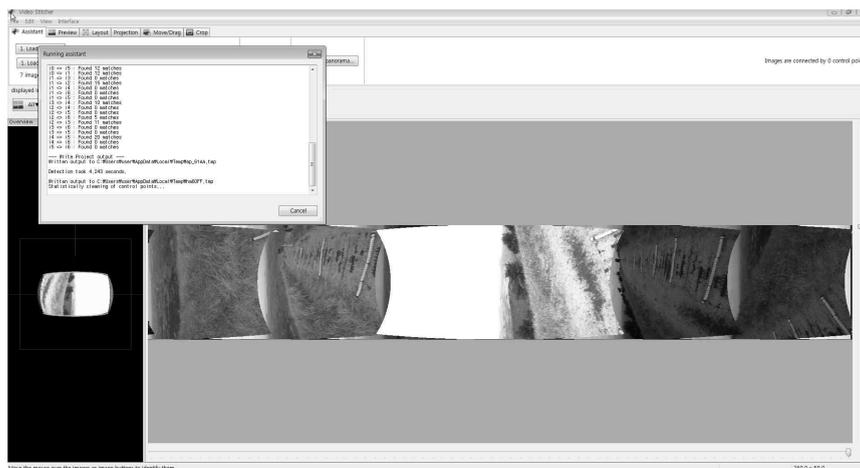
1) 다량쉬 오름

다량쉬오름 정상에서의 영상으로 총 7장으로 구성되어 있다. 기존방법과 제안 방법의 스티칭 결과를 비교하였다.

Fig. 59는 기존 소프트웨어인 “C제품”과 제안 방법에서 7개의 카메라에서 촬영된 영상을 정합하기 위한 과정이다.



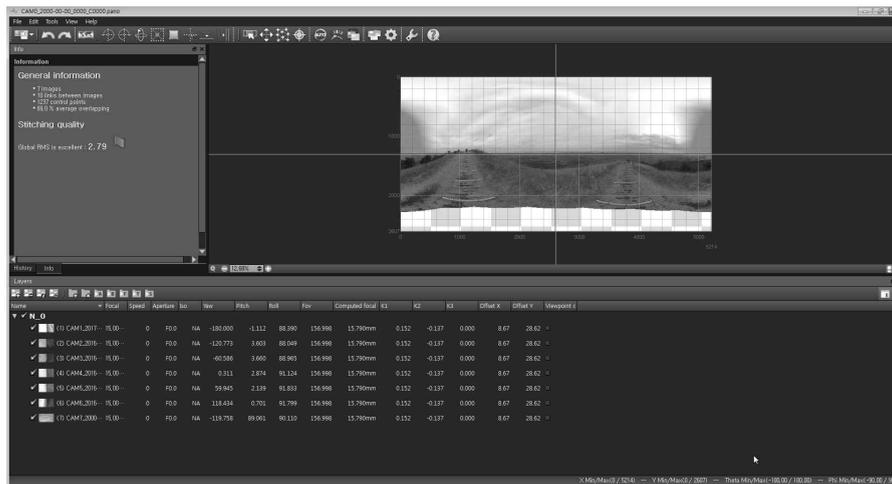
(a) Previous method



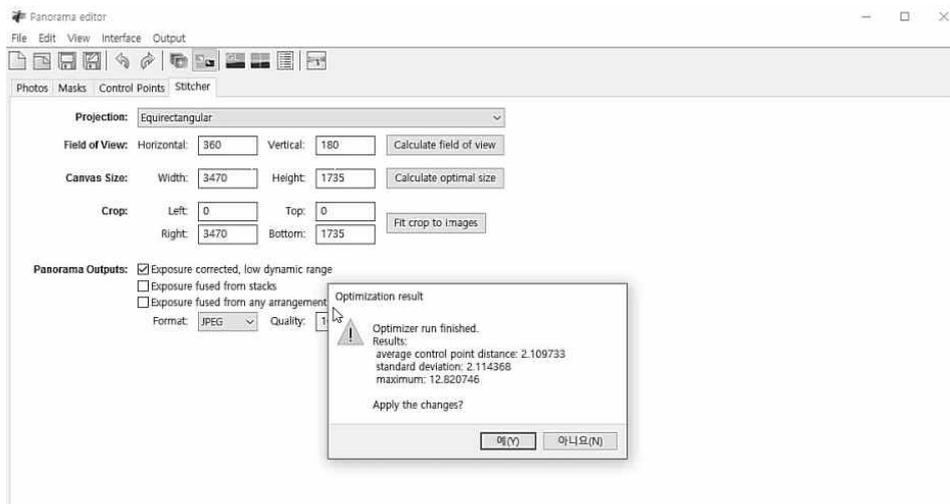
b) Proposed method

Fig. 59. A process for matching the images captured by the proposed c-product and the proposed method

Fig. 60은 기존 소프트웨어인 “C제품”과 제안방법에서 추출된 특징점을 통한 거리 값과 특징점 개수를 나타내고 있다. 총 7대의 카메라로부터 이웃한 영상과의 특징점의 총 개수는 각각 1,235와 185개 이며, 이를 통한 거리 값은 2.79와 2.1임을 보여주고 있다.



(a) Previous method



(b) Proposed method

Fig. 60. Comparison of distance values and control points through feature points extracted from "C product" and proposed method

Fig. 61 a)와 b)는 1-7번까지의 카메라에서 촬영된 인접한 각 영상의 특징점과 대응되는 정합점을 찾는 과정이다.

1번 카메라의 경우 인접한 2, 6, 2번 카메라의 경우 1,3,7카메라, 3번 카메라는 2,4,7, 4번 카메라는 3,5,7, 5번 카메라는 4,6,7, 6번 카메라는 1,5,7번 카메라는 2,3,4,5카메라와의 대응점을 찾고 있다.

Link	Points	RMS	Visual similarity
1 - 2	186	2.20	100%
1 - 6	128	2.08	100%
2 - 3	159	2.31	100%
2 - 7	118	3.13	100%
3 - 4	115	2.87	100%
3 - 7	55	3.34	100%
4 - 5	145	3.44	100%
4 - 7	61	3.56	100%
5 - 6	194	2.72	100%
5 - 7	76	2.94	100%

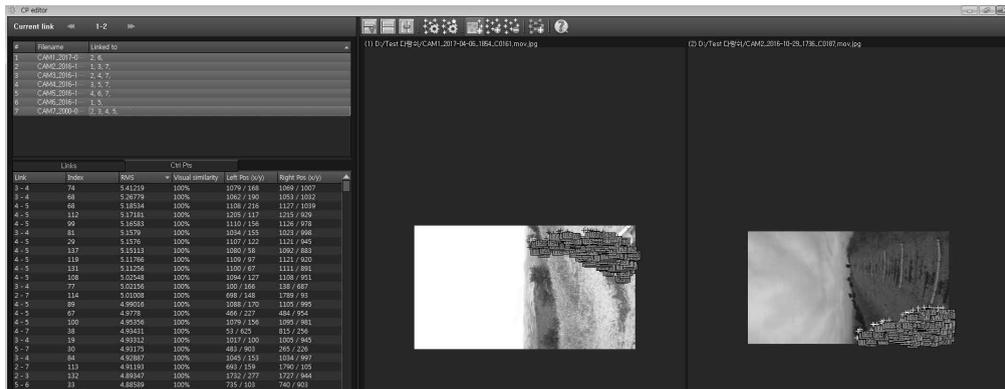
(a) Previous method

G CP#	Left img	Right L.	P CP#	Alignment	Distance
0	0	1	0	normal	2.52
1	0	1	1	normal	3.18
2	0	1	2	normal	2.59
3	0	1	3	normal	0.82
4	0	1	4	normal	1.28
5	0	1	5	normal	1.56
6	0	1	6	normal	0.58
7	0	1	7	normal	2.14
8	0	1	8	normal	2.28
9	0	1	9	normal	1.65
10	0	5	0	normal	3.30
11	0	5	1	normal	1.05
12	0	5	2	normal	1.87
13	0	5	3	normal	3.07
14	0	5	4	normal	1.16
15	0	5	5	normal	0.44
16	0	5	6	normal	1.01
17	0	5	7	normal	0.82
18	0	5	8	normal	1.30
19	0	5	9	normal	0.22
20	0	5	10	normal	0.35
21	0	5	11	normal	1.33
22	1	2	0	normal	1.71
23	1	2	1	normal	4.94
24	1	2	2	normal	1.99
25	1	2	3	normal	1.17
26	1	2	4	normal	0.78
27	1	2	5	normal	0.83
28	1	2	6	normal	0.53
29	1	2	7	normal	1.00
30	1	2	8	normal	0.48
31	1	2	9	normal	0.97
32	1	2	10	normal	0.67

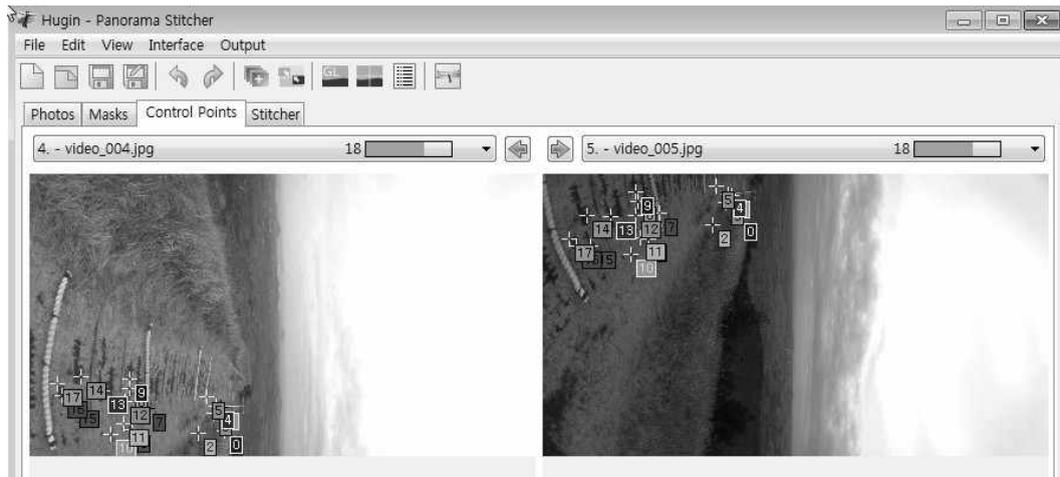
b) Proposed method

Fig. 61. A process of finding a matching point corresponding to a minutiae of each adjacent image in the photographed image

Fig. 62는 다량쉬 오름에서의 기존 소프트웨어인 “C제품”과 제안 방법에서의 대칭점 찾는 과정의 모습이다. 기존 소프트웨어에서는 특징점 추출이 상당히 많음을 확인할 수 있으며, 제안한 방법에서는 의미있는 특징점 추출이 이루어져 특징점 개수가 줄어들음을 알 수 있다. 특징점 개수가 적게 추출됨으로서 연산수행 속도에도 영향을 미치는 것을 확인할 수 있었다.



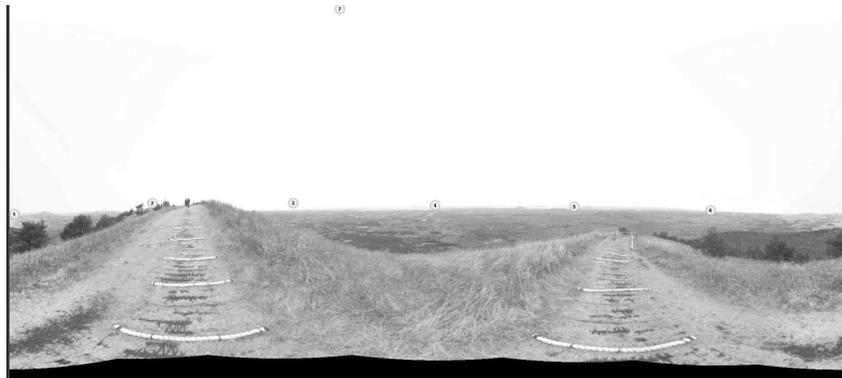
(a) Previous method



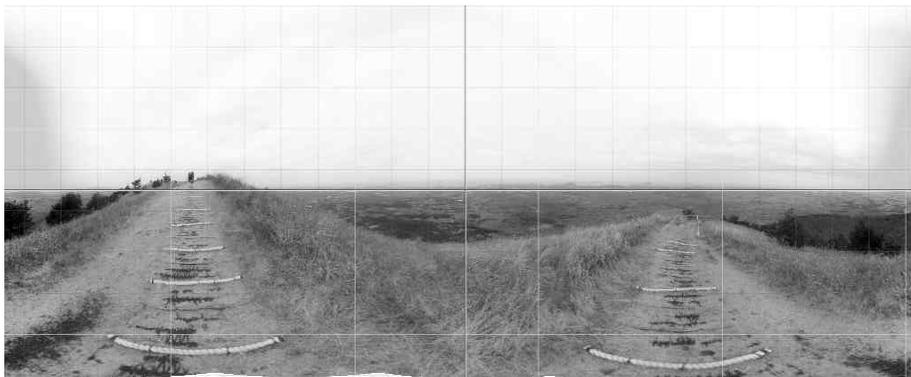
b) Proposed method

Fig. 62. Matching point check screen

Fig. 63 a)와 b)는 기존 소프트웨어인 “C제품”과 제안 방법에서의 다랑쉬오름에서 촬영한 최종 스티칭 결과 이미지이다. 기존과 제안한 방법의 이미지에는 큰 차이는 없었으나 제안한 방법에서 좌우 가장자리 부분에 일부 회색계열의 색상이 나타났다. 이러한 원인은 촬영당시 날씨도 영향이 있는 것으로 보고 있다(촬영당시 날씨는 흐리고 구름이 검은 조금 있는 상태였음). 7개의 카메라에서 촬영한 서로 이웃한 이미지를 정합하는 과정에서 수평이미지와 수직(하늘방향) 이미지를 정합하는 과정에서 약간의 오차가 발생한 것으로 판단된다.



(a) Previous method



b) Proposed method

Fig. 63. Result of Stitching darangshi orum

(1) 다랑쉬 오름 스티칭 결과

Table 8은 제주동부에 위치한 다랑쉬 오름 정상에서 날씨가 흐린날 촬영한 영상이다. 기존방법과 제안방법의 각 영상의 특징점 개수와 RMSE를 보여주고 있다. 기존의 방법과 제안한 방법을 수행했을 때 각 카메라에 인접된 카메라와 연결된 특징점의 개수와 RMSE를 수행한 결과를 나타내고 있다.

Table 8에서 1번부터 7번 카메라의 영상에서 특징점 개수를 보면 기존의 방법에서는 총 1,235개, 제안한 방법에서는 172개로 약 86%이상 줄어들었다. 특징점 개수가 줄어든 것은 의미 있는 특징점만을 가지고 정합 과정을 수행하기 때문에 총 특징점 개수가 줄어들었다고 볼 수 있다.

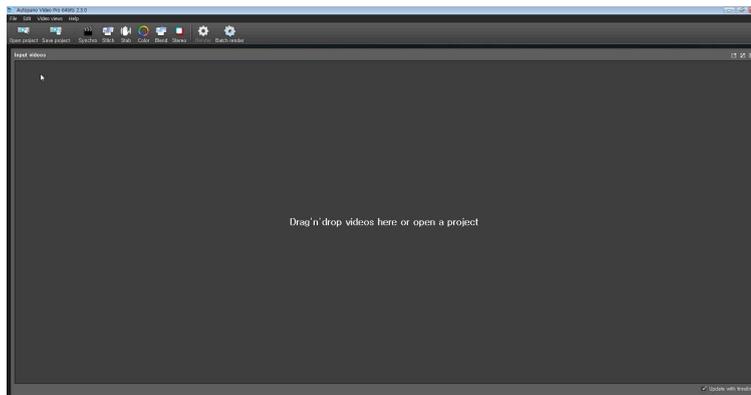
또한 각각의 이웃한 영상에 대한 추출된 대응점과 거리값은 기존의 방법은 평균 2.78, 제안한 방법은 2.10로 제안한 방법이 우수함을 보여주고 있다. 이는 전체적인 스티칭 성능에 영향을 줄 수 있을 것으로 보여진다.

Table 8. Darangshi orum Stitching Result Data

Test 영상	previous					proposed					
	#	Linked to	Links	C.P 개수	RMSE	#	Linked to	Links	C.P 개수	RMSE	
다랑쉬 오름	1	2,6	1-2	186	2.21	1	2,6	1-2	22	2.08	
	2	1,3,7	1-6	128	2.08	2	1,3,7	1-6	36	1.47	
	3	2,4,7	2-3	159	2.30	3	2,4,7	2-3	30	2.05	
	4	3,5,7	2-7	118	3.11	4	3,5,7	2-7			
	5	4,6,7	3-4	114	2.85	5	5,6,7	3-4	46	4.07	
	6	1,5	3-7	55	3.33	6	1,5	3-7			
	7	2,3,4,5	4-5	144	3.43	7	2,3,4,5	4-5			
			4-7	61	3.54			4-7			
			5-6	194	2.72			5-6	38	2.55	
			5-7	76	2.93			5-7			
합계				1,235	2.78	합계				172	2.10

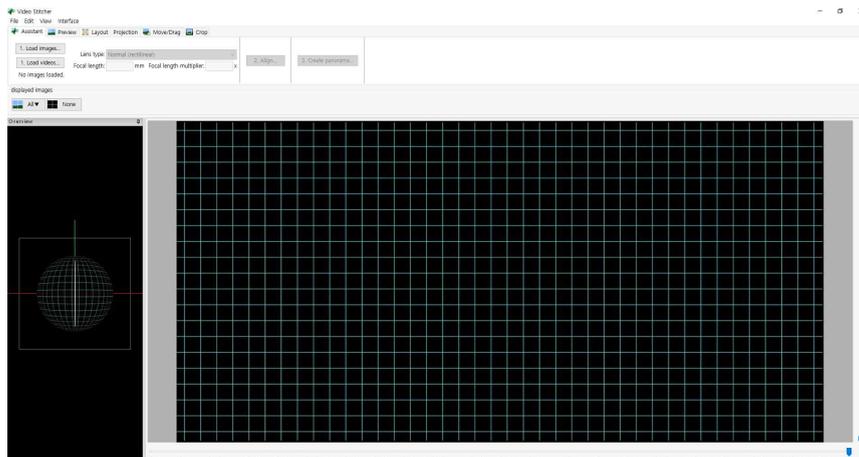
2) 대학 전경

기존 소프트웨어인 “C제품”을 실행하면 아래 Fig. 64 a)와 같이 실행되면서 스티칭을 위한 준비상태의 화면이 실행된다. 스티칭 실행방법이 간단한 화면으로 구성되어 있다



(a) Previous method

(b)는 Video Stitch라고 이름을 붙인 제안한 화면으로 이웃한 영상을 불러오기 위한 초기화면이다.



b) Proposed method

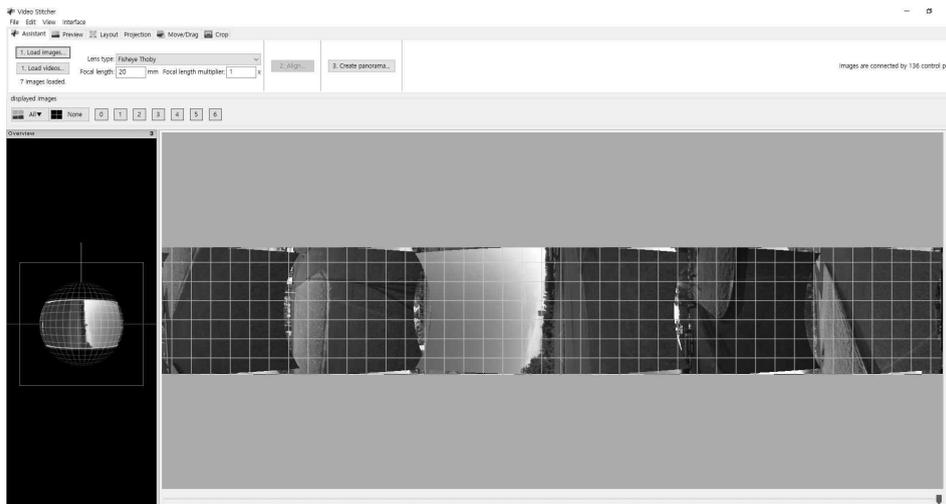
Fig. 64. Initial screen for stitching existing method and proposed method

대학 내의 전경을 촬영한 영상으로 총 7장으로 구성되어 있다.
 Fig. 65 a)는 기존 소프트웨어인 “C제품”에서 7개의 카메라에서 촬영된 영상을 불러오기 하여 정합하는 과정이다.



(a) Previous method

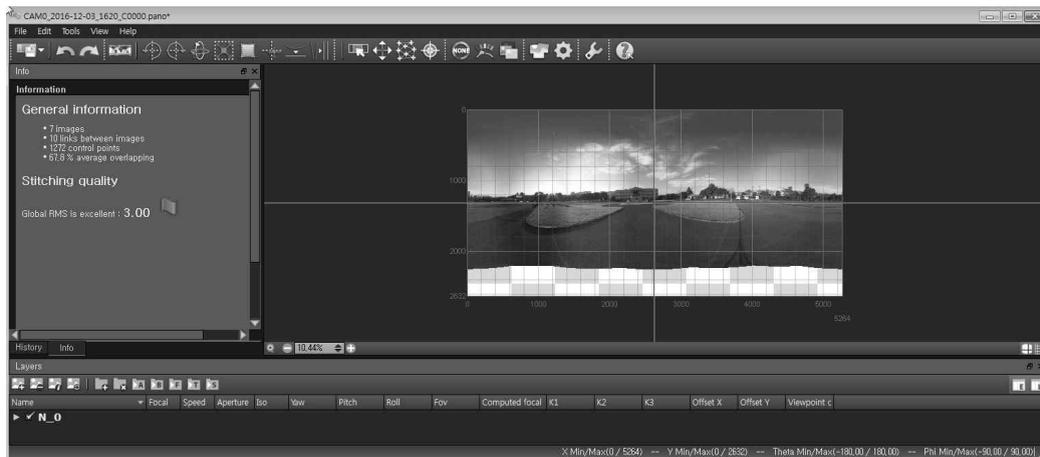
Fig. 65 b)는 제안한 방법에서 7개의 카메라에서 촬영된 영상을 정합하기 위한 과정이다.



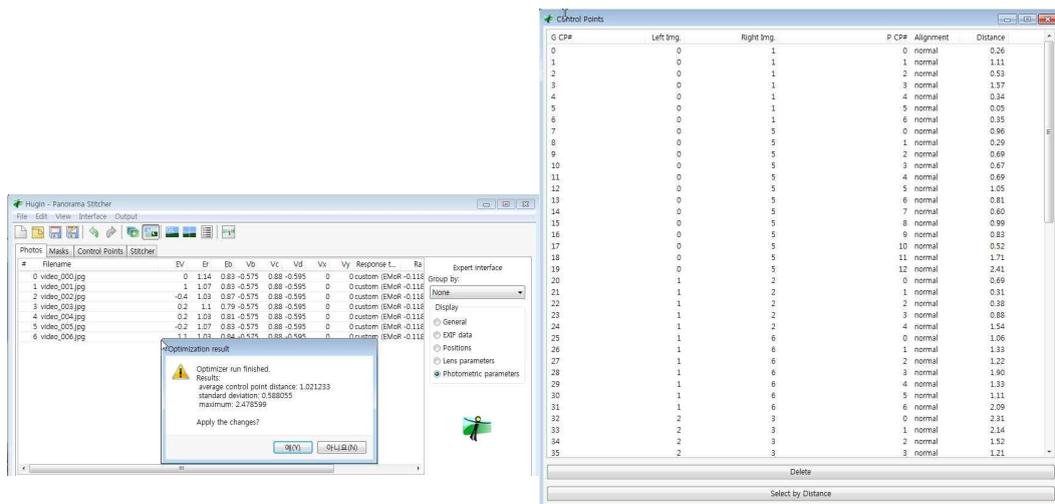
b) Proposed method

Fig. 65. Stitching screen of university main building

Fig. 66 a)와 b)는 기존 방법과 제안 방법에서 추출된 특징점을 통한 거리값 오차를 나타내고 있다. 총 7대의 카메라로부터 이웃한 영상과의 특징점의 총 개수는 기존방법은 1,272개이며, 제안방법은 183개 이다. 이를 통한 거리 값은 기존 방법은 3.00, 제안 방법은 1.02임을 보여주고 있다.



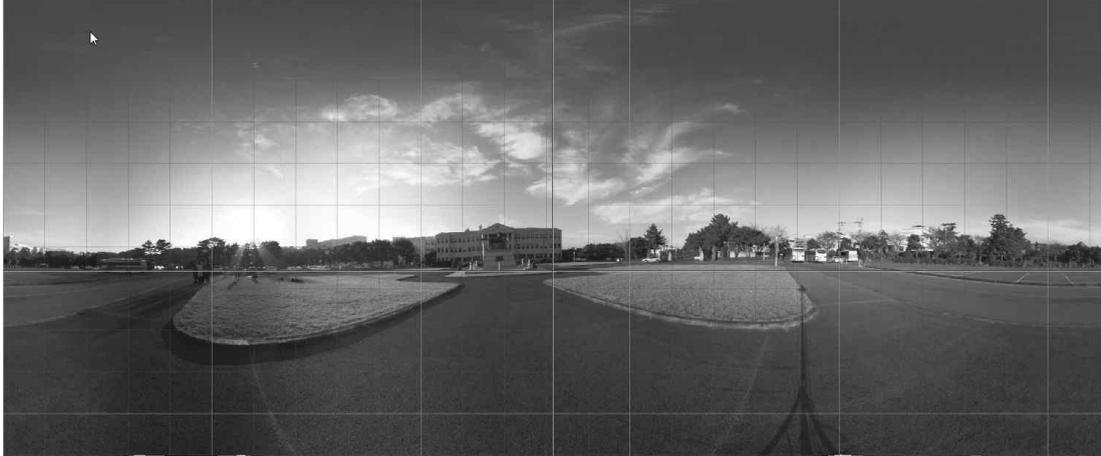
(a) Previous method



b) Proposed method

Fig. 66. Distance value and control points calculation screen

Fig. 67 a)와 b)는 대학전경의 스티칭 최종 화면 이미지로서, 파노라마 영상을 생성할 때, 조명이나 시점의 변화에 따라 정합 영역의 경계선이 생기거나 색상이 불연속인 경우가 많이 발생한다. 이러한 경우 필요에 따라 색상보정 작업이 필요한 경우이다.



a) Previous method



b) Proposed method

Fig. 67. Stitching result image

(1) 대학 전경 이미지 스티칭 결과

Table 9는 대학전경 영상에서 각 영상의 특징점 개수와 RMSE를 보여주고 있다. 기존의 방법과 제안한 방법을 수행했을 때 각 카메라에 인접된 카메라와 연결된 특징점의 개수와 RMSE를 수행한 결과를 나타내고 있다.

Table 9에서 1번부터 7번 카메라의 영상에서 특징점 개수를 보면 기존의 방법에서는 총 1,272개, 제안한 방법에서는 183개로 약 86%이상 줄어든 것을 볼 수가 있다. 특징점 개수가 줄어든 것은 의미 있는 특징점만을 가지고 정합 과정을 수행하기 때문에 총 특징점 개수가 줄어들었다고 볼 수 있다.

또한 각각의 이웃한 영상에 대한 추출된 대응점과 거리값은 기존의 방법은 평균 2.97, 제안한 방법은 1.02로 제안한 방법이 우수함을 보여주고 있다. 이는 전체적인 스티칭 성능에 영향을 줄 수 있을 것으로 보여진다.

자동스티칭의 매칭 포인트와 수동 스티칭의 매칭 포인트와의 평균 거리오차가 5 픽셀 이하일 경우 스티칭에 성공했다고 판단하였다. 표 9에서 처럼 평균오차거리(오차율)는 제안한 방법이 기존 방법보다 좋은 것을 볼 수가 있다.

Table 9. Result of stitching image of university image

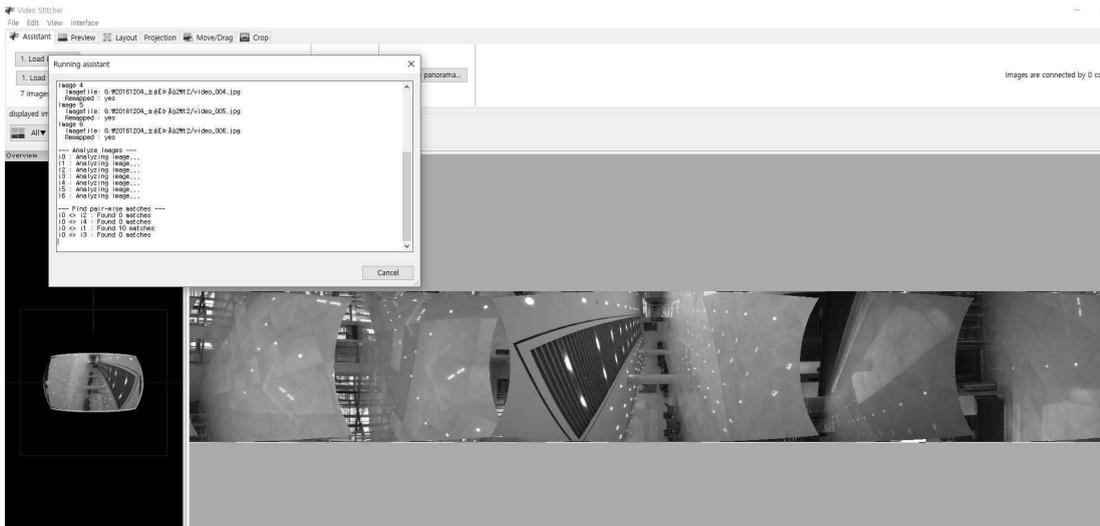
Test 영상	previous					proposed					
	#	Linked to	Links	C.P 개수	RMSE	#	Linked to	Links	C.P 개수	RMSE	
대학 전경	1	2,6	1-2	114	2.46	1	2,6	1-2	12	0.59	
	2	1,3,7	1-6	131	2.2	2	1,3,7	1-6	23	0.92	
	3	2,4,7	2-3	80	2.6	3	2,4,7	2-3	12	0.75	
	4	3,5,7	2-7	174	3.53	4	3,5,7	2-7	18	1.43	
	5	4,6,7	3-4	141	3.17	5	5,6,7	3-4	35	1.36	
	6	1,5	3-7	80	3.45	6	1,5	3-7			
	7	2,3,4,5	4-5	153	2.87	7	2,3,4,5	4-5	22	0.64	
			4-7	106	3.07			4-7	16	1.28	
			5-6	123	3.05			5-6	19	0.8	
			5-7	170	3.26			5-7	26	1.12	
합계				1,272	2.97	합계				183	1.02

3) 금호세계교육관 로비

Fig. 68 a)와 b)는 기존 방법과 제안방법에서 7개의 카메라에서 촬영된 영상을 정합하기 위한 과정을 보여주고 있다.



(a) Previous method



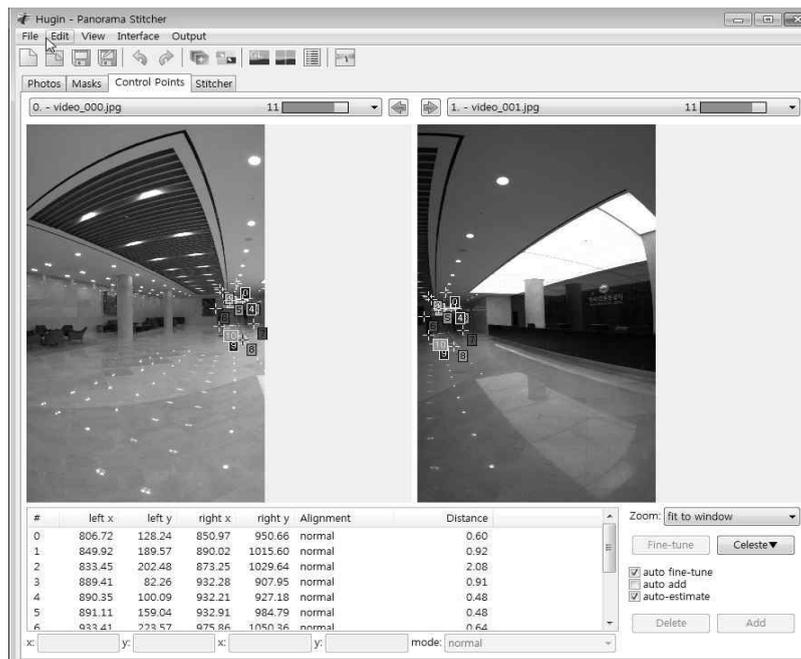
b) Proposed method

Fig. 68. Kumho World Education Center lobby stitching initial screen

Fig. 69 a)와 b)는 기존 소프트웨어인 “C제품”에서 추출된 특징점을 통한 RMSE를 나타내고 있다. 총 7대의 카메라로부터 이웃한 영상과의 특징점의 총 개수는 기존방법은 558개이며, 제안방법은 103개 이며, 이를 통한 거리 값은(스티칭오차)는 기존 방법은 3.42, 제안 방법은 3.10임을 보여주고 있다.



(a) Previous method



b) Proposed method

Fig. 69. RMSE calculation screen



a) Previous method



b) Proposed method

Fig. 70. Kumho World Education Center lobby stitching result image

금호세계교육관 로비의 경우 실내에서 촬영을 한 경우로, 일부 천장부분에서 스티칭이 일부 오류가 있었다. 그 이유는 촬영하는 바로 그 장면 이외에는 조명, 연출용 기구, 감독을 비롯한 스태프 등이 있어도 영상에 영향을 미치지 않았던 기존의 영상 촬영과는 달리 360도 영상에서는 카메라가 보이는 모든 지점에는 피사체를 제외한 관계자(스태프) 등이 보여서는 안 되고, 조명 등의 연출이 어려워지기 때문에 원하는 영상을 쉽게 얻기가 어렵다는 단점도 있다. 여기에서도 스티칭이 일부 오류가 있었던 것도 천장에 있는 불빛이 과장되면서 다른 곳에 영향을 미쳤을 것으로 보여진다.

(1) 금호세계교육관 로비 이미지 스티칭 결과

Table 10은 금호세계교육관로비 영상에서 각각 영상의 특징점 개수와 RMSE를 보여주고 있다.

Table 10에서 1번부터 7번 카메라의 영상에서 특징점 개수를 보면 기존의 방법에서는 총 558개, 제안한 방법에서는 103개로 약 82%이상 줄어든 것을 볼 수가 있다.

또한 각각의 이웃한 영상에 대한 추출된 대응점과 거리 값(RMSE)은 기존의 방법은 평균 3.42, 제안한 방법은 3.10로 유사한 결과를 보여주고 있는데 실제 영상 스티칭 과정에서 영상정합의 잘 되지 않아 오류가 많이 발생하였다. 특징점 개수가 많다고 좋은 것이 아니라 의미 있는 특징점을 선별하는 것이 중요하다

Table 10. Kumho World Education Center lobby image stitching result Data

Test 영상	previous					proposed					
	#	Linked to	Links	C.P 개수	RMSE	#	Linked to	Links	C.P 개수	RMSE	
금호 세계 교육 관	1	2,6,7	1-2	92	2.63	1	2,6	1-2	23	1.02	
	2	1,3	1-6	81	3.18	2	1,3,7	1-6	18	0.67	
	3	2,4,7	1-7	12	3.76	3	2,4,7	2-3	7	4.3	
	4	3,5,7	2-3	54	3.69	4	3,5,7	3-4	13	3.0	
	5	4,6,7	3-4	62	2.99	5	5,6,7	3-7	15	8.49	
	6	1,5,7	3-7	31	4.29	6	1,5	4-5	9	3.22	
	7	1,3,4,5,6	4-5	53	2.67	7	2,3,4,5	5-6	13	0.93	
			4-7	14	3.96			6-7	5	3.15	
			5-6	99	2.86						
			5-7	20	4.07						
		6-7	40	3.55							
합계				558	3.42	합계				103	3.10

Table 8~10까지의 결과 데이터를 종합해 보면 각 RMSE(거리 오차)와 특징점의 개수가 기존의 방법보다 제안한 방법이 줄어드는 결과를 볼 수 있다. 의미 있는 특징점을 추출한다면 특징점 추출 후 특징점간의 거리를 줄여 영상의 화질이 나 처리속도에 영향을 미치는 것을 알 수 있었다.

VI. 결론

최근 컴퓨터 비전 기술이 발전함과 동시에 가상현실 및 360VR 콘텐츠에 대한 고해상도 이미지 및 영상물에 대하여 사용자들의 관심이 높아지고 있다. 또한 VR 시장이 활성화되면서 많은 콘텐츠들이 4K, 8K 등 고해상도의 영상을 원하지만 카메라에서 생성된 영상을 스티칭하는 작업에 상당한 시간과 비용이 소요된다.

고해상도의 360도 영상에 대한 Image Stitching을 하려다 보면 점점 더 영상 생성 속도 및 영상의 해상도 향상이 필요로 하게 될 것이다[6].

본 논문에서는 파노라마 영상 생성을 위하여 효과적인 스티칭 방법을 제안하였다. 가변 적응형 필터를 적용한 V-SURF 알고리즘을 제안하여 기존 알고리즘인 SIFT와 SURF 알고리즘과의 연산 수행속도, 특징점 추출시간 및 총스티칭 시간을 비교 분석을 통하여 연산수행 속도가 향상됨을 확인하였다. 또한 다중 카메라로 촬영된 360파노라마 영상 Data를 통하여 기존 소프트웨어인 “C제품”과 제안한 방법과의 스티칭 오차와 특징점 개수를 비교·분석을 하였다. 영상의 사이즈가 클수록 즉 픽셀이 많을수록 특징점 추출시간과 대응점 매칭시간이 증가하게 되므로 그만큼 스티칭 속도는 느려지게 된다. 기존 소프트웨어와의 비교를 통하여 제안한 알고리즘이 영상 정합시에 효과적인 특징점만을 추출하고 스티칭 하도록 하여 RMSE 및 특징점 개수를 줄임으로써 스티칭 처리속도가 향상됨을 확인하였다.

360도 파노라마 영상을 생성하기 위하여 촬영시스템의 다중 카메라 동기화 시스템을 제안하였다. 여러 대의 카메라가 동일한 프로젝션 센터를 가질 수가 없고 이러한 문제 때문에 카메라간 시차가 발생할 수 있다. 시차는 영상합성을 통해서 해결할 수 없으므로 영상 획득시 최소화 할 있도록 리그를 설계하고 카메라를 리그에 장착하여 카메라를 동기화시켜 촬영되는 카메라의 동기화를 확인하였다.

본 논문의 실험을 위하여 자연, 풍경, 건물, 바다 등 다양한 영상을 포함한 영상 데이터 15개를 사용하였다.

Table 5의 결과에서도 알 수 있듯이 기존의 알고리즘과 비교하여 제안한 방법

이 향상된 성능을 보여주고 있다. 다량쉬오름의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 3.18초, 3.07초 및 2.62초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 18.7%, 14.7% 감소되었다. 총 스티칭 평균시간은 SIFT, SURF 및 제안 방법에서는 각각 8.07초, 7.28초 및 6.44초로 나타난다. 이미지 스티칭 총 시간은 SIFT보다 20.2%, SURF 보다는 11.5% 감소되었다.

실습실 전경의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 2.66초, 2.56초 및 2.04초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 23.3%, 20.3% 감소되었다. 총 스티칭 평균시간은 SIFT, SURF 및 제안 방법에서는 각각 6.95초, 6.85초 및 5.58초로 나타난다. 이미지 스티칭 총 시간은 SIFT보다 24.1%, SURF 보다는 19.7% 감소되었다.

금호세계교육관 로비의 평균 특징점 추출시간은 SIFT, SURF 및 제안 방법에서는 각각 1.66초, 1.57초 및 1.26초로 나타난다. 따라서 특징점 추출시간은 SIFT와 SURF보다 19.7%, 18.5% 감소되었다. 총 스티칭 평균시간은 SIFT, SURF 및 제안 방법에서는 각각 4.17초, 3.93초 및 3.53초로 나타난다. 이미지 스티칭 총 시간은 SIFT보다 15.3%, SURF 보다는 10.2% 감소되었다.

실험결과에서 볼 수 있듯이 전체적인 스티칭 시간에서 제안한 방법이 SIFT와 SURF보다 향상된 결과를 보였으며, 가변 적응형 필터를 적용하여 윈도우 마스크의 크기를 조절함으로써 핵심영역에서만 특징점 추출 및 정합이 이루어지도록 하여 연산량 및 연산시간을 줄일 수 있었다. 스티칭 한 결과 총스티칭 평균시간 및 특징점 추출시간이 단축되는 성능을 보였다. 따라서 제안한 방법이 전체적으로 효율성 있는 스티칭을 수행하는 것을 보여 주었다.

본 연구에서 기존 스티칭 편집소프트웨어인 “C제품”을 사용하여 제안알고리즘과 스티칭 오차와 특징점 개수에 대하여 비교하였다.

다량쉬오름의 RMSE는 기존방법이 2.78, 제안한 방법은 2.1로 약 24.5% 낮은 결과를 보였다. 또한 특징점 개수도 기존방법은 1, 235개 제안한 방법에서는 185개의 결과를 얻었다. 대학 전경의 RMSE는 기존방법이 2.97 제안한 방법은 1.02로 약 65.7% 낮은 결과를 보였다. 또한 특징점 개수는 기존방법은 1, 272개 제안한 방법에서는 175개의 낮은 결과를 얻었다. 금호세계교육관의 경우 RMSE는 기존방법이 3.42, 제안한 방법은 3.10로 약 10% 낮은 결과를 보였다. 또한 특징점

개수가 기존 방법은 558개 제안한 방법에서는 103개의 결과를 얻었다.

실험결과를 보면 기존의 방법보다 제안한 방법이 RMSE가 낮음을 알 수 있다. RMSE가 낮다고 하는 것은 특징점과 대응점간의 오차(거리)가 적기 때문에 효과적인 스티칭을 할 수 있고, 빠른 연산수행 속도와 고해상도의 이미지를 얻을 수 있음을 확인하였다.

본 연구를 통해서 고품질의 카메라를 여러 대 사용하여 동영상의 고품질을 유지하면서도, 넓은 화각을 촬영할 수 있어, 전시장 및 대형 광고 시장 등에서 여러 사용자에게 동시에 고해상도의 파노라마 영상을 제공할 수 있을 것으로 기대된다. 전방향성 파노라마 영상 제작 기술은 다양한 형식의 문화콘텐츠 제작에 활용 가능하며, 방송 콘텐츠와 영화 특수효과 등과 같은 크로마키 촬영 기법이 사용되는 분야에서 활용도가 높을 것으로 기대된다. 특히 영상기반 모델링 및 텍스처링(Texturing) 기술과 접목하게 되면 파노라마 영상만으로도 촬영된 공간을 3차원 폴리곤 데이터로 전환하는 것이 가능해지는데, 이러한 기술은 사전시각화 장비와 결합했을 때 효율적인 컴퓨터그래픽 영상 제작 공정으로 발전할 수 있을 것이다.

그리고 영상기반 모델링 및 텍스처링 기술과 접목하게 되면 파노라마 영상만으로도 촬영된 공간을 3차원 폴리곤 데이터로 전환하는 것이 가능해지는데, 이러한 기술은 사전시각화 장비와 결합했을 때 효율적인 컴퓨터그래픽 영상 제작에 활용할 수 있으며, 직접 현장을 방문한 것처럼 관광지의 상황 및 풍경을 홍보하기 위한 360도 홍보용 영상물 제작할 수 있어 인터넷을 활용한 손쉬운 정보를 얻을 수 있는 편리함을 제공 할 수 있을 것이다.

본 연구를 통하여 이미지나 영상의 품질을 높이기 위해서는 스티칭의 중요성이 다시 한번 대두되었고, 제안된 RMSE 및 특징점 개수 추출방법을 다른 방법과 조합하면 보다 향상된 결과를 얻을 수 있을 것으로 보여진다. 향후 연구에서는 고해상도의 파노라마 VR영상콘텐츠를 제작하기 위하여 다양한 화각을 가진 카메라 및 다양한 촬영조건을 고려하여 스티칭 기법과 모바일환경에서 실시간으로 고화질 편집이 가능한 방법의 연구가 필요하다.

참고문헌

- [1] 손현진, 2015년 6월, 5G시대에 주목 받는 실감형 콘텐츠, KT경제경영연구소 디지에코 보고서
- [2] <http://www.ddaily.co.kr>, 2015년 6월5일 기사
- [3] 조용상, 실감형콘텐츠, ICT EXPERT INTERVIEW _Special Theme, TTA Journal Vol.162, pp.9~10
- [4] http://www.ddaily.co.kr/m/m_article.html?no=131157, 2015.06.05.
- [5] <http://www.sindohblog.com/947>
- [6] Jung Won Bang, 2016.07, Image Stitching for generating panorama image, Proceedings of the Korean Society of Computer Information Conference 24(2), pp.287-288.
- [7] Seok Joo Myoung, Lim Seong Yong, Cha Jihun, 2014, A Development on the Multi-camera Monitoring Systems using User Selected features based preview stitching, KICS, vol.57, pp.862-863.
- [8] Kwon, Jong Soo, Kim, Hyoyoung, Park, Jin Wan, 2013.1, “A Study on High Resolution Lenticular Photographs Using Stitching Technique : Mainly with ‘Feather’ Series”, Journal of Digital Design 13(1), pp.183-192.
- [9] R. Szeliski, Jan 2006, “Image alignment and stitching: A tutorial,” Foundations and Trends® in Computer Graphics and Vision, vol. 2, pp. 1-104.
- [10] D. Milgram, 1975. “Computer methods for creating photomosaics,” IEEE Transactions on Computers, vol. C-24, Nov 1975, pp. 1113-111.
- [11] Bang ga-hee, 2016, “Removal of moving objects using Median filter in Panorama image”, Dept.of Computer & Software Engineering Graduate School of HanYang University, pp1-5.
- [12] <http://bolg.naver.com/grectus/22081263012>

- [13] Sang-Bong Byun, Yong-Hwan Lee, “Study of Image Stitching Algorithm for Hardware Implementation”, Proceedings of KIIT, Summer Conference, pp.217-219.
- [14] Computational Photography Derek Hoiem, 2010, University of Illinois.
- [15] Sang-Bong Byun, Young-Hwan Lee, 2014.5, Hardware Algorithm for Panorama Image Proceedings of KIIT Summer Conference, pp.76~78.
- [16] Y. Xiong, K. Pulli, 2010, Fast Image Stitching and Editing for Panorama Painting on Mobile Phones, IEEE International Workshop on Mobile Vision in Conjunction with CVPR '2010(IWMV2010), San Francisco, USA, pp.298-306, June 18.
- [17] D. Chen, S. S. Tsai, V. Chandrasekhar, G. Takacs, J. Singh, and B. Girod, 2009, Robust image retrieval using multiview scalable vocabulary trees, in Visual Communications and Image Processing, Majid Rabbani and Robert L. Stevenson, Eds., vol. 7257, pp.72570V.
- [18] Jeongho Kim and Daewon Kim, April 2014, Matching Points Filtering Applied Panorama Image Processing Using SURF and RANSAC Algorithm, Journal of The Institute of Electronics and Information Engineers Vol. 51, NO. 4.
- [19] Daehyun Kim, Jongsoo Choi, Nov.2003, View Interpolation Algorithm for Continuously Changing Viewpoints in the Multi-panorama Based Navigation, IEIE Journal(SP), vol. 40, no. 6, pp.141-148.
- [20] Sehwan Kim, Kiyoun Kim, Woontack Woo, Mar. 2004, Multiple Camera Calibration for Panoramic 3D Virtual Environment, IEIE Journal(CI), vol. 41,no. 2, pp.137-148.
- [21] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, 2004, Interactive digital photomontage, ACM Trans. Graph, vol. 23, pp.294 - 302.
- [22] Ahn hyo Chang, 2011, “A Study on Efficient Image Stitching Using Feature Extraction on Mobile Phones”, Dankuk university

- [23] Hyunji Cho, Heewon Kye, Jeongjin Lee, June 2015, Rapid Stitching Method of Digital X-ray Images Using Template-based Registration, Journal of Korea Multimedia Society Vol. 18, No. 6, pp.701-709.
- [24] T. Mahalakshmi, R. Muthaiah, and P. Swaminathan, 2012, Review Article: An Overview of Template Matching Technique in Image Processing, Research Journal of Applied Sciences, Engineering and Technology, Vol.4, No. 24, pp.5469-5473.
- [25] L.D. Stefano, S. Mattocchia, and F. Tombari, 2004, An Algorithm for Efficient and Exhaustive Template Matching, Lecture Notes in Computer Science, Vol. 3211, No. 1, pp.408-415.
- [26] L. Juan and O. Gwun, 2009, A Comparison of SIFT, PCA-SIFT and SURF, IJIP, Vol.3, Issue 4.
- [27] H. Ahn, S. Rhee, April. 2012, Fast Image Stitching Based on Improved SURF Algorithm Using Meaningful Features, The KIPS Transaction:Part B, Vol.19B, No.2, pp.93-98.
- [28] Dong-Kyu Choi, Gun-Bum Kim, So-Yun Min, Kyung-Chun Park, Teng Yu, Hyunchul Shin, Improved Techniques of Stitching and Blending for Panoramic Images, THE INSTITUTE OF ELECTRONICS ENGINEERS OF KOREA, Vol.34, No.1, pp.754~755.
- [29] Sangwon Seo, Soowoong Joeng, Yunsang Han, Jongsoo Choi, and Sangkeun Lee, August 2013, Efficient Homography Estimation for Panoramic Image Generation, Journal of The Institute of Electronics Engineers of Korea Vol. 50, NO. 8.
- [30] 오원근, “실감형 영상 콘텐츠 검색 기술의 오늘과 내일”, TTA Journal.162, pp38~39.
- [31] Seung Heon Kang, 2014, Parallelization of Image Processing Algorithms on Mobile GPU using OpenCL, IInHa University, p27.
- [32] Seung Heon Kang, Seung-Jae Lee, Man Hee Lee, and In Kyu Parka, 2014.5, Pallelization of Feature Detection and Panorama Image Generation

using OpenCL and Embedded GPU, JOURNAL OF BROADCAST ENGINEERING 19(3), pp.316-323.

- [32] Myung-ju Moon, 2011, A Modified SURF Algorithm with Effective Feature Extraction Scheme Based on Edge, Dept.of Computer Engineering, Graduate School, Kumoh National Institute of Technology, p17.
- [35] Chan-Il Park, Su-Hyun Lee, and Yong-Jin Jeong, 2009, A Hardware Design of Feature Detector for Realtime Processing of SIFT(Scale Invariant Feature Transform) Algorithm in Embedded Systems, pp86-95.
- [34] Oh Hyang Ian, 2014, A Study on Fast Stitching Method for Avionic Images with Positioning Information, Dept.of Internet & Multimedia Engineering Graduate School of KonKuk Univ., pp.3~4.
- [36] <http://www.haenki.com/81>, <http://perpet.tistory.com/51>
- [37] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, vol. 60, no. 2, pp.91 - 110.
- [38] <http://m.blog.naver.com/samsjang/220655420471>
- [39] <http://darkpgmr.tistory.com/131>.
- [40] Siyoung Park, Jongho Kim, Jisang Yoo, 2015, Fast stitching algorithm by using feature tracking, The Korean Society Of Broad Engineers, pp.177-180
- [41] Kim, Jongho, Park, Siyoung, Yoo, Jisang, 2014, Panorama image generation using SURF and cylindrical projection, The Korean Society Of Broad Engineers, pp.242-244.
- [42] Kim Sanghoon, 2008, A Study on a Fundamental Matrix Estimation and a 3D Reconstruction using Projection Matrix Alignment under Circular Motion, JoongAng University, pp.36~37.
- [43] Jeongho Kim and Daewon Kim, April. 2015, Matching Points Filtering Applied Panorama Image Processing Using SURF and RANSAC Algorithm, Journal of The Institute of Electronics and Information

Engineers Vol. 51, NO. 4.

- [44] Park JunHoon, 2012, A Study on the Generation of Panoramic Image using Effective Boundary Correction Method, JoongAng University, p16.
- [45] C. Sun, Dec. 1997, A fast stereo matching method, in Digital Image Computing: Techniques and Applications, pp.95-100.
- [46] K.T.Kim, S.B.Park and Y.Huh, 2002, A fast stereo matching method for the effective removal of the occlusions using Dynamic Programming, International Conference on Electrical Engineering.
- [47] <http://www.spacevr.com/vr1/05.php>
- [48] Kim Tae Hyeon , 2003, A Comparative Study on the Image Based Virtual Reality and the Modeling Based Virtual Reality, Vol.13, pp.15-26
- [49] 한형만 저, 2003, “디지털카메라 & VR웹사이트”, 정보문화사.
- [50] <http://www.spacevr.com/vr1/02.php>
- [51] Bosung Kim, 2016, An Exemplar-based Image Inpainting Technique for Spherical Panorama using Geographic Coordinate System, InChun University, p5.
- [52] <http://blog.naver.com/cubicpan/220619355043>
- [53] Kang-Ho Lee, 2006.5, Simplified projective transform for reconstruction of cylindrical panorama, Journal of The Korea Society of Computer and Information 11(2), pp.169-175.
- [54] <http://m.blog.naver.com/cubicpan/220619355043>
- [55] Jun-Seong Joe, Bum-Jong Lee, Jong-Seung Park, 2007, Creating Full View Panorama Image from Multiple Images, KOREA INFORMATION SCIENCE SOCIETY Vol.34.No.2(B), pp162-165.
- [56] Soon-Cheol Kim, Soo-Yeong Yi, 2015, Wide FOV Panorama Image Acquisition Method, Journal of the Korea Academia-Industrial cooperation Society Vol. 16, No. 3 pp.2117-2122.
- [57] Kim and Y. Choi, 2010, Image-processing Based Panoramic Camera Employing Single Fisheye Lens, Jour. of Optical Society of Korea, Vol. 14, no3, pp.245-259.

- [58] 김영상 공저, 2006, 포토그래픽 VR응용실습, 제주한라대학, p44.
- [59] 이인재,조용주,기명석,임성용,이희경,차지훈, 2011.5, 고품질 휴먼융합형 파노라마 서비스, 한국통신학회지(정보와통신) 제28권 제6호, pp.11-20.
- [60] Seok Joo Myoung, Lim Seong Yong, Cha Jihun, 2014, A Development on the Multi-camera Monitoring Systems using User Selected features based preview stitching, ETRI.
- [61] <http://www.ivs.auckland.ac.nz/web/calibration.php>
- [62] 한국정보통신기술협회, 2014, 고품질 파노라마영상 기술보고서, pp.1~76.
- [63] <http://darkpgmr.tistory.com/31>