

第 35 回 碩士學位論文

指導教授 白 俊 基

파티클 키폰트를 이용한 강건한
특징점 추출 및 영상 분할 기반의
비디오 안정화 방법

Video Stabilization Method

**Based on Robust Feature Detection Using Particle
Keypoints and Motion-Based Optimal Camera Path**

中央大學校 尖端映像大學院
映像學科 映像工學-디지털이미징 專攻
全 世 美
2017年 2月

파티클 키포인트를 이용한 강건한
특징점 추출 및 영상 분할 기반의
비디오 안정화 방법

Video Stabilization Method

**Based on Robust Feature Detection Using Particle
Keypoints and Motion-Based Optimal Camera Path**

이 논문을 석사학위논문으로 제출함

2017 년 2 월

중앙대학교 첨단영상대학원
영상학과 영상공학-디지털이미징 전공
전 세 미

張振範의 碩士學位論文으로 認定함

審査委員長 _____ (印)

審査委員 _____ (印)

審査委員 _____ (印)

中央大學校 尖端映像大學院

2017年 2月

목 차

제 1 장 서 론.....	1
1.1 연구 배경.....	2
1.2 연구 목적 및 내용.....	8
1.3 논문 구성.....	10
제 2 장 배경 이론.....	11
2.1 기하학 기반의 비디오 열화 모델.....	12
제 3 장 제안하는 비디오 안정화 방법.....	19
3.1 움직임 기반 영역 분할.....	22
3.2 파티클 기반 특징점 추출 및 카메라 패스 추정.....	25
3.3 카메라 패스 스무딩.....	31
3.4 영상 워핑을 통한 비디오 안정화.....	33
제 4 장 실험결과.....	35

4.1 파티클 기반 특징점 추출 결과.....	36
4.2 제안하는 알고리즘을 이용한 비디오 안정화 결과.....	39
제 5 장 결 론.....	53
5.1 결론	54
참고 문헌	57
국문 초록	64
Abstract.....	65

그림 목 차

그림 1. CMOS 센서 기반의 휴대용 카메라로부터 얻어진 비디오: (A) 흔들리는 원 카메라 패스 그리고 (B) 안정된 카메라 패스	13
그림 2. 제안하는 비디오 안정화 방법의 블록다이어그램	20
그림 3. 영역 분할 결과: (A) 원본 영상, (B) 추정된 옵티컬 플로우, 그리고 (C) 옵티컬 플로우 기반의 분할된 영상	24
그림 4. 평평한 영역 검출 결과: (A) 입력 영상 그리고 (B) 제안하 는 알고리즘을 이용한 평평한 영역에 대한 맵	27
그림 5. 첫 번째 영역에 대해 추정된 카메라 패스: (A) x축 방 향으로의 원본 카메라 패스 (점선)와 스무딩 된 카메라 패스 (실 선) 그리고 (B) y축 방향으로의 원본 카메라 패스 (점선)와 스 무딩 카메라 패스 (실선)	34
그림 6. 다양한 알고리즘을 이용한 특징점 추출 및 매칭 결과:	

(A) SIFT [32], (B) SURF [33], (C) HARRIS [31]+BRIEF [37], (D)
HARRIS [31] +FREAK [40], (E) FAST [35] +BRIEF [37], (F) FAST
[35] +FREAK [40], (G) BRISK [39], 그리고 (H) 제안하는 파티클
키포인트.....38

그림 7. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A) 65,
66, 67번째 흔들리는 입력 비디오 영상, (B) MATSUSHITA 방법
을 이용한 안정화 영상 [8], (C) LIU 방법을 이용한 안정화 영
상 [23], 그리고 (D) 제안하는 방법을 이용한 안정화 영상...40

그림 8. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A) {(64,
65), (65, 66), (66, 67)}에 대한 입력 비디오 차분 영상, (B)
MATSUSHITA 방법을 이용한 이용한 안정화 차분 영상 [8], (C)
LIU 방법을 이용한 안정화 차분 영상 [23], 그리고 (D) 제안하
는 방법을 이용한 안정화 차분 영상.....42

그림 9. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A) 110,

111, 112번째 흔들리는 입력 비디오 영상, (B) LIU 방법을 이
용한 안정화 영상 [23], (C) LIU 방법을 이용한 안정화 영상
[19], 그리고 (D) 제안하는 방법을 이용한 안정화 영상44

그림 10. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A)

{(109, 110), (110, 111), (111, 112)}에 대한 입력 비디오 차분
영상, (B) LIU 방법을 이용한 안정화 차분 영상 [23], (C) LIU 방
법을 이용한 안정화 차분 영상 [19], 그리고 (D) 제안하는 방
법을 이용한 안정화 차분 영상45

그림 11. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A)

334, 335, 336번째 흔들리는 입력 비디오 영상, (B)
GRUNDMANN 방법을 이용한 안정화 영상 [13], (C) ZHOU 방법
을 이용한 안정화 영상 [20], 그리고 (D) 제안하는 방법을 이
용한 안정화 영상47

그림 12. 다양한 알고리즘을 이용한 비디오 안정화 결과: (A)

{(333, 334), (334, 335), (335, 336)}에 대한 입력 비디오 차분
 영상, (B) GRUNDMANN 방법을 이용한 안정화 차분 영상 [13],
 (C) ZHOU 방법을 이용한 안정화 차분 영상 [20], 그리고 (D)
 제안하는 방법을 이용한 안정화 차분 영상48

제 1장 서 론

본 장에서는 파티클 키포인트와 영상 분할 기반의 비디오 안정화 방법의 개요와 연구 배경, 연구 목적 및 내용, 그리고 본 논문의 구성에 대해 소개한다.

1.1 연구 배경

최근 모바일카메라, 디지털카메라, 캠코더, 드론, 웨어러블 카메라 등의 조 작이 쉽고 소형화 된 카메라의 보급으로 개인용 카메라에 대한 비 전문가 의 수요가 높아지고 있다. 또한 클라우드 서비스의 발전으로 개인이 촬영 한 일상 영상을 공유함으로써 높은 품질의 영상 획득에 대한 요구가 증가 하고 있다. 하지만 비디오 촬영자의 조작 미숙, 손 떨림, 또는 외부 환경 에 따른 의도하지 않은 움직임에 의해 흔들리는 비디오를 얻게 되는 문제 가 있다. 이러한 문제점을 해결하고자 다양한 비디오 안정화 방법이 연구 되고 있다. 일반적인 비디오 안정화 방법은 다음과 같은 3가지로 분류된 다: i) 기계적 안정화 방법, ii) 광학적 안정화 방법, 그리고 iii) 이미지 기반 전자적 안정화 방법이다 [1]. 먼저, 기계적 안정화 방법은 자이로 센서가 부착된 짐볼과 같은 기구를 이용하여 카메라의 흔들림을 방지한다. 다음 으로 광학적 안정화 방법은 카메라 렌즈 내부의 가속도 센서를 이용하여 각 축의 방향으로 움직임을 보정하거나 프리즘의 빛 경로를 변화시켜 움 직임을 보존한다. 마지막으로 이미지 기반의 안정화 방법은 이미지 시퀀 스의 호 분석을 통해 잘못 위치한 화소의 위치를 조절해준다 [1]. 이는 별 도의 하드웨어 추가 비용 없이 주어진 알고리즘을 이용하여 흔들림을 제 거할 수 있으므로, 이미 촬영된 영상의 후 처리 또한 가능하다는 장점이 있다. 따라서 본 논문에서는 이미지 프로세싱 기술을 이용하여 비디오의

흔들림을 제거하는 기술을 제안한다.

일반적으로 이미지 기반 비디오 안정화 방법은 i) 원 카메라 패스(camera path) 추정, ii) 스무딩 카메라 패스(smoothed camera path) 추정을 통한 흔들리는 움직임 제거, 그리고 iii) 스무딩 카메라 패스를 이용한 결과 영상 렌더링(rendering)으로 분류된다 [2].

기존의 비디오 안정화 알고리즘은 인접한 프레임 간 변환행렬을 추정하여 영상 전체에 대한 카메라 패스를 추정하고, 저역통과필터링(low-pass filtering)을 통해 카메라 패스를 스무딩하여 비디오 안정화를 수행하였다. 초기에는 인접한 프레임의 블록 매칭(block matching) 또는 옵티컬 플로우(optical flow) 기반의 움직임 추정을 이용하여 카메라 패스를 추정하고 흔들린 영상을 안정화하였다. Jang은 영상에 적응적인 블록을 이용하여 인접한 프레임 사이의 최적의 어파인 모델을 추정하여 강건한 카메라 패스를 추정하였다 [3]. Xu는 서클러 블록 매칭(circular block matching)을 이용하여 지역적 패스를 추정하고 least square의 반복을 통해 카메라 패스를 추정하였다 [4]. 이는 간단한 움직임 추정을 이용해 카메라 패스를 추정할 수 있지만, 영상의 밝기, 노이즈, 그리고 가려짐에 민감하고 객체의 움직임이 클 경우 매칭 에러가 발생하는 문제가 있다. Chang은 루카스-카나데(Lucas-Kanade) 알고리즘을 이용한 옵티컬 플로우로 어파인 모델을 추정하여 프레임 간 변환 관계를 정의하였으며, 움직임 보상을 통해 카메라 패스를 안정화 한다 [5]. Xu는 혼-성크(Horn-Schunck) 알고리즘으로 추정된 옵티컬

플로우를 이용하여 프레임 간의 어파인 모델을 추정하고 모델 피팅 필터를 이용하여 카메라 패스를 안정화 한다 [6]. 이는 두 프레임의 변환관계를 간단하고 유연하게 나타내주는 어파인 모델을 추정할 수 있지만, 역동적이고 물체가 많은 장면에서 물체 경계면의 불연속성으로 인하여 왜곡이 발생하는 문제가 있다.

이후 비디오 안정화의 성능을 향상시키기 위해 영상의 회전이나 스케일 변화에 강건한 특징점 매칭을 이용하여 카메라 패스를 추정하는 방법이 제안되었다. Matsushita는 인접한 프레임 간의 기하학적 변환 행렬을 호모그래피로 정의하여 글로벌 카메라 패스를 추정하고 가우시안 커널을 이용하여 스무딩 패스를 계산하였으며, 영상 모자이크(image mosaicing)과 영상 인페인팅(image inpainting)을 결합하여 안정화의 성능을 향상시켰다 [7]. Battiato는 scale invariant feature transform (SIFT)를 이용하여 특정한 점의 움직임을 추적한 후 최소 자승법(least square)을 이용하여 에러를 삭제함으로써 정확한 카메라 패스를 추정하였다 [8]. Lee는 SIFT를 이용한 특징점 매칭을 통해 카메라 패스를 계산하고, 패스 스무딩과 기하학적 왜곡 감소를 위한 에너지 함수를 이용하여 안정화 성능을 향상시켰다 [9]. Xu는 fast accelerated segment test (FAST)를 이용하여 어파인 모델을 추정함으로써 비디오 안정화를 수행하였다 [10]. Pinto는 speed-up robust feature (SURF) 기반의 근사화 된 특징점 추적 방법을 이용하여 글로벌 카메라 패스를 추정하였다 [11]. Grundmann은 특징점을 기반으로 호모그래피를 추정하고 이를

L1 최적화 방법을 통해 안정화 하였으며 [12], Song은 different depth of field (DDOF)와 weighted least square (WLS)를 이용하여 강건한 글로벌 카메라 패스를 추정함으로써 비디오 안정화를 수행하였다 [13]. Kim와 Nejadasl은 카나데-루카스-토마시 트래커 (Kanade-Lucas-Tomasi tracker)를 이용하여 특징점을 추적함으로써 비디오 안정화를 수행하였으며 [14][15], Shene는 SURF 특징점을 이용한 field programmable gate array (FPGA)를 기반의 비디오 안정화 방법을 제안하였다 [16]. Cheng은 SURF와 수정된 random sample consensus (RANSAC)을 이용하여 움직임 추정 방법을 제안했다 [1]. 이와 같은 방법들은 인접한 프레임 간의 선형적 변환에 대한 정보만 추정하기 때문에 계산속도가 빠른 장점이 있다. 하지만 2차원 비디오 안정화 방법은 영상의 거리 정보를 포함하고 있지 않기 때문에 영상에 크게 움직이는 객체가 존재하는 경우 결과 영상에 왜곡이 발생하는 단점이 있다. 또한 영상 전체에 대해 하나의 카메라 패스를 추정하기 때문에 기존의 특징점 추출 알고리즘을 이용한 안정화 방법들은 무늬(texture)가 없는 영역은 모션 정보가 반영되지 못하여 최적의 카메라 패스를 추정할 수 없는 문제점이 있다.

최근에는 비디오 안정화의 성능을 향상시키기 위해 3차원 공간을 재구성함으로써 3차원 상의 카메라 모션을 추정하는 방법이 제안되었다. Liu는 structure from motion을 이용하여 3차원 상의 카메라 모션을 추정하고 시공간적 워핑을 통해 3차원의 구조를 보존함으로써 강건한 안정화를 수

행할 수 있다 [17]. Liu는 2차원 특징점 추적과 3차원 재구성 기반의 렌더링을 이용하여 높은 품질의 안정화 된 영상을 생성한다 [18]. Zhou는 3차원 포인트 클라우드(point cloud)로부터 평면을 재 구성하고 평면의 라벨링을 통해 각각의 호모그래피를 추정함으로써 텍스처가 없던 영역의 모션을 추정하여 기하학적 왜곡을 감소하였다 [19]. Koh는 low rank matrix completion scheme을 적용한 특징점 추적을 통해 강건한 매쉬그리드(mesh grid)기반의 워핑을 수행하였다 [21]. 3차원 비디오 안정화는 장면의 3차원 정보를 이용함으로써 왜곡이 감소한 높은 품질의 결과 영상을 획득할 수 있으나, 2차원 안정화 방법보다 영상 변화에 민감하고 계산량이 많은 단점이 있다.

따라서 비교적 간단하며 영상 내 변화에 강건한 2차원 비디오 안정화 알고리즘의 성능을 향상 시키기 위해, 영상 내에서 지역적 카메라 패스를 추정하는 방법이 제안되었다. Liu는 영상을 블록으로 나누어 특징점을 이용하여 각각 최적의 카메라 패스를 추정함으로써 2차원 안정화의 한계를 극복했다 [22]. 하지만 일정한 블록으로 영역을 나누어 카메라 패스를 추정하기 때문에 객체의 움직임이 큰 경우 불연속적인 영역에서 왜곡이 발생하며, 텍스처가 없는 평평한 영역을 포함한 블록에서는 추정 오류가 발생하는 문제점이 있다. Liu는 영상에서 일관성이 큰 특정 옵티컬 플로우를 이용하여 카메라 패스를 추정함으로써 비디오 안정화의 성능을 향상시켰다 [23].

결과적으로, 다양한 디지털 영상 획득 장치로부터 얻은 흔들리는 비디오로부터 고품질의 안정화 된 비디오를 얻기 위해서는 정확한 카메라 패스 추정이 필수적이며, 최적의 카메라 패스 추정을 통해 기하학적 왜곡이 없는 비디오 안정화를 위한 방법이 필요하다.

1.2 연구 목적 및 내용

기존의 특징점 기반 비디오 안정화 방법은 대체적으로 기존의 특징점 추출 방법을 사용하였기 때문에 바닥 또는 하늘 면과 같은 평평한 영역에서의 특징점 미 검출로 인한 카메라 패스 오추정으로 저품질의 비디오를 얻게 된다. 또한 영상 변화에 강건하며 간단하게 비디오 안정화를 수행할 수 있는 2차원 안정화 방법은 객체의 거리 정보와 같은 3차원 정보를 포함하지 못하기 때문에 배경과 객체의 경계선에서의 불연속적인 움직임으로 인하여 정확한 카메라 패스 추정에 오류가 발생한다.

이러한 문제점을 해결하기 위해, 본 논문에서는 파티클 키포인트 (particle keypoints)를 이용하여 영상의 특징을 재 정의함으로써 분할된 영역에 따라 적응적 카메라 패스 추정을 통한 고품질의 비디오 안정화 방법을 제안한다. 제안하는 알고리즘은 인접한 프레임간의 움직임을 추정하여 영상을 분할하고, 기존의 방법으로 추정된 특징점의 통계적 특성을 이용하여 파티클 키포인트를 생성함으로써 강건한 특징점을 재 정의한다. 이후 분할된 영역에 해당되는 특징점을 이용하여 카메라 패스를 추정하며, 최적화 기반의 스무딩 카메라 패스를 추정한다. 마지막으로 추정된 스무딩 카메라 패스를 적용하여 영상을 워핑함으로써 최종 안정화 영상을 생성한다.

제안하는 방법은 영상 전체에 퍼진 파티클 키포인트 기반의 강건한 특징점을 기반으로 인접 프레임간의 최적의 변환 관계를 추정함으로써, 비디오 안정화를 위한 호모그래피 추정의 정확도를 향상시킬 수 있다. 또한 움직임에 기반하여 영상의 영역을 분할함으로써, 각 분할된 영역에 따른 적응적 카메라 패스의 추정을 통해 불연속적인 움직임에 의한 기하학적 왜곡을 감소할 수 있다. 결과적으로 최적의 카메라 패스를 추정하여 비디오를 안정화함으로써 2차원 안정화 방법의 한계를 보완하고 효과적으로 안정화 된 비디오 결과를 얻을 수 있다.

1.3 논문 구성

본 논문의 구성은 다음과 같다. 2장에서는 비디오 영상을 획득하는 과정에서 발생하는 기하학 기반의 비디오 열화 모델에 대해 설명하고, 3장에서는 제안하는 비디오 안정화 방법에 대해 소개한다. 4장에서는 제안하는 알고리즘의 실험결과를 요약하며, 마지막으로 5장에서 본 논문의 결론을 맺는다.

제 2장 배경 이론

본 장에서는 영상 획득 과정에서 발생하는 비디오의 기하학적 열화 과정 및 왜곡 모델을 설명하고 그에 따른 비디오 안정화 방법에 대해 설명한다.

2.1 기하학 기반의 비디오 열화 모델

다양한 카메라에서 시각적 지각 정보를 보존하기 위해서 고품질의 영상을 획득하는 것은 매우 중요하다. 흔들리는 모션이 포함된 낮은 품질의 비디오 영상은 사용자의 손 떨림이나 바람과 같은 외부 환경으로 인하여 카메라의 위치가 변경되는 동시에 센서의 위치가 변경되면서 획득된 영상의 잘못된 위치로 인해 발생한다.

그림 1은 일반적인 complementary metal-oxide semiconductor (CMOS) 센서로 영상을 획득하는 과정을 간략하게 보인다. 그림 1(a)은 외부 환경으로 인해 흔들림을 포함한 영상을 나타내며, 그림 1(b)는 목표로 하여 보정된 안정화된 영상을 나타낸다. 그림에서 보는 것과 같이, 흔들리는 카메라는 카메라 패스에 평행한 가운데 선을 기준으로 월드 좌표계 기준으로 센서의 각도가 변화한 것을 볼 수 있다. 이는 사영 왜곡(perspective distortion)을 발생시키는 원인이며, 최종적으로 안정화 된 영상은 사영 왜곡이 감소되는 방향으로 수행하게 된다.

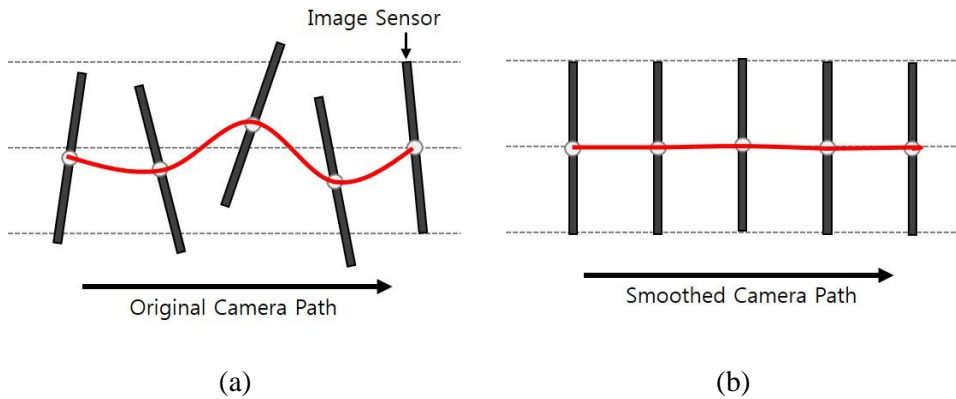


그림 1. CMOS 센서 기반의 휴대용 카메라로부터 얻어진 비디오: (a) 흔들리는 원 카메라 패스 그리고 (b) 안정된 카메라 패스

디지털적 비디오 안정화는 비디오 프레임을 구성하는 화소들의 기하학적 위치 변화에 의한 열화를 개선하고 복원하는 것을 목표로 한다. 이에 따라 화소들의 기하학적 위치 변화에 따른 왜곡은 크게 다음 세가지로: i) 2차원 글로벌 변환 모델, ii) 롤링 셔터 왜곡, 그리고 iii) 3차원 모델이다. 먼저 2차원 글로벌 변환 모델은 가장 단순한 수준의 움직임을 정의할 수 있다. 이는 프레임 전체가 같은 평면에 있다고 가정하여 하나의 2차원 변환 모델 \mathbf{M} 으로 흔들리는 모션을 정의할 수 있으며, 이는 다음과 같다.

$$\mathbf{q} = \mathbf{M}\mathbf{p}, \quad (1)$$

여기서 $\mathbf{q} = [x', y', 1]^T$ 는 호모지니어스 좌표계(homogeneous coordinates)상의 현재 프레임 상의 임의의 화소를 나타내며, $\mathbf{p} = [x, y, 1]^T$ 는 호모지니어스

좌표계 상의 기준이 되는 프레임 상의 임의의 화소를 나타낸다. 여기서 2차원 변환 모델 \mathbf{M} 은 모델링 된 변환에 따라 크게 다음과 같이 나눌 수 있다. 먼저 크기, 평행이동, 그리고 회전에 대해 고려한 유사 변환(similarity transformation)은 다음과 같은 호모지니어스 좌표로 정의할 수 있다.

$$\mathbf{M}_S = \begin{bmatrix} s \cos \theta & -s \sin \theta & t_x \\ s \sin \theta & s \cos \theta & t_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (2)$$

여기서, s 는 크기 요소, θ 는 회전 각도, t_x , t_y 는 각각 x축 y축으로의 평행이동이다. 여기에서 반전, 엇갈림 변환을 추가로 고려하여 어파인 변환(affine transformation)을 정의할 수 있다. 이는 직선, 길이의 비, 평행성을 보존하는 변환이며 다음과 같은 호모지니어스 좌표로 정의할 수 있다.

$$\mathbf{M}_A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix}. \quad (3)$$

마지막으로 같은 평면상의 물체에 대한 2차원적 변환관계에 대한 가장 일반적인 변환 모델은 사영변환(perspective transformation)이며, 이는 호모그래피(homography)라고도 한다. 3차원 공간에서 2차원 이미지로 사영 변환 된 두 이미지 사이의 관계를 모델링 한 것이며, 다음과 같은 호모지니어스 좌표로 정의할 수 있다.

$$\mathbf{M} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (4)$$

다음으로 롤링 셔터에 의한 기하학적 위치 변환이다. 카메라로 이미지를 인식 할 때, 센서에서 모든 화소를 동시에 읽어드리는 글로벌 셔터 방법과 주사선에 기반하여 한 줄씩 차례로 인식하는 롤링 셔터 방법이 있다. 일반적으로 대부분의 CMOS 센서는 비용 줄이기 위하여 롤링 셔터 방법을 사용하고, 주사선이 훑고 지나가는 중에 움직임이 발생하면 왜곡이 생기게 된다. Hong은 롤링셔터 왜곡 모델에 대해 6개의 변수를 갖는 어파인 모델로 정의한다 [24]. 손에 들고 촬영하는 카메라에서는 상, 하, 좌, 우 움직임과 흔들림이 발생하게 되며 회전에 대한 움직임은 평행이동에 비해 상대적으로 작기 때문에 배제하여, 다음과 같은 2차원 변환 모델을 정의하였다.

$$\mathbf{M} = \begin{bmatrix} 1+s_x & -\theta+k & d_x \\ \theta & 1+s_y & d_y \\ 0 & 0 & 1 \end{bmatrix}, \quad (5)$$

여기서 s_x , s_y 는 각각 x, y축으로의 크기 요소, θ 는 회전 각도, t_x , t_y 는 각각 x, y축으로의 평행이동 요소이다.

실 세계에서 카메라의 궤적이 의도된 주요 패스를 따라서 움직이는 과정에서 흔들림이 발생했을 때, 한 장면 내 객체의 거리 정보에 따라 서로 다른 왜곡이 발생한다. 앞서 정의했던 2차원 변환 모델은 프레임 전체

를 한 평면으로 가정하기 때문에 거리 정보를 사용하지 않는다. 3차원 왜곡 모델은 3차원으로 확장된 좌표계에서 카메라의 회전변환 \mathbf{R} 과 평행이동 변환 \mathbf{T} 로 정의될 수 있으며 다음과 같다.

$$\mathbf{q}_c = \mathbf{R}\mathbf{p}_c + \mathbf{T}, \quad (6)$$

여기서 $\mathbf{q}_c = [x', y', z', 1]^T$ 는 호모지니어스 좌표계상의 현재 카메라의 위치를 나타내며, $\mathbf{p}_c = [x, y, z, 1]^T$ 는 호모지니어스 좌표계상의 기준이 되는 카메라의 위치를 나타낸다.

다음 왜곡 모델에 기반한 비디오 안정화 방법은 다음 3단계이다: i) 2차원글로벌 안정화 방법, ii) 3차원 안정화 방법, 그리고 iii) 2차원 지역적 안정화 방법. 먼저 2차원 글로벌 안정화 방법은 2차원 글로벌 변환 모델에 의해 왜곡된 영상을 안정화 하는 방법이다 [3]-[16]. 인접 프레임간의 2차원 움직임을 추정하고 시간에 따른 저역통과필터링을 이용하여 스무딩된 모션을 추정한다. 그 후 전체 프레임의 2차원 위평을 통해 흔들림이 제거된 비디오를 얻어낸다. 이는 실제 3차원 패스를 반영하는 것이 아닌 안정화된 패스를 따라가는 것처럼 만들어 준다. 하지만 프레임 전체를 한 평면으로 보기 때문에 3차원 정보를 반영하지 못하여 시점 변환에 따른 차이에 강건하지 못하다.

3차원 정보를 보완하기 위한 3차원 안정화 방법은 3차원 재구성 알고리즘을 이용하여 3차원 상의 카메라 위치를 재구성한다 [17]-[21]. 따라서

이는 앞서 살펴보았던 3차원 왜곡 모델에 의해 열화 된 비디오를 안정화할 수 있다. 이는 structure-from-motion 또는 3차원 포인트 클라우드 등을 이용하여 3차원 카메라 궤적을 추정하고 최적화 방법을 통해 최적의 회전과 평행이동 행렬을 찾음으로써 패스를 스무딩한다. 마지막으로 이미지 기반의 렌더링을 통해 흔들림이 제거된 비디오를 생성한다. 이는 3차원 정보에 대한 시차에는 강건하지만 특징이 잘못 검출된 것에 대해 2차원 방법보다 민감하며, 계산량이 많다는 단점이 있다.

이에 따라 마지막으로 2차원 지역적 안정화 방법은 2차원의 간단한 변환 모델에 기반하여 3차원의 정보를 보완 하기 위해 제안되었다 [22]-[23]. 또한 이는 롤링 셔터 왜곡과 같은 비강체 효과(non-rigid effects)에 의해 왜곡된 비디오를 지역적으로 안정화함으로써 안정화의 성능을 높인다. 먼저 영상을 특정 영역으로 나누어 각각의 2차원 카메라 모션을 추정하고 지역통과필터링을 이용하여 스무딩한다. 마지막으로 영역별로 추정된 움직임으로 워핑함으로써 흔들림이 제거된 비디오를 얻는다.

제안하는 알고리즘은 다음과 같은 2차원 지역적 비디오 안정화 방법을 이용하여 간단한 2차원 모델을 이용하여 3차원의 정보를 보완한다. 먼저 같은 움직임을 가진 것들은 같은 평면에 존재한다고 가정하고 움직임을 기반으로 영상을 분할한다. 따라서 분할된 영상에 대해 각각 2차원 움직임을 추정하며, 식 (10)와 같이 흔들리는 움직임을 호모그래피로 정의하여 안정화를 수행한다. 따라서 본 논문에서는 영상에서의 강건한 특징점

을 이용하여 최적의 호모그래피를 추정함으로써 앞서 정의한 2차원 변환 모델에 대해 안정화 된 비디오를 얻을 수 있으며, 분할된 영역에 적응적인 카메라 패스를 이용하기 때문에 3차원 모델에 대한 보완이 가능하다.

제 3장 제안하는 비디오 안정화 방법

본 장에서는 제안하는 파티클 키포인트를 이용한 강건한 특징점 추출 방법과 영상 분할 기반의 적응적 카메라 패스 추정을 통한 비디오 안정화 방법에 대해 설명한다.

본 장에서는 제안된 파티클 키포인트 추정 및 영상 분할 기반 비디오 안정화 알고리즘을 자세하게 설명한다. 제안된 알고리즘은 다음과 같은 4단계로: i) 움직임 기반 영역 분할, ii) 파티클 기반의 특징점 추출, iii) 카메라 패스 스무딩, 그리고 iv) 영상 워핑을 통한 안정화이다. 그림 2에서는 제안된 방법의 블록다이어그램을 보인다.

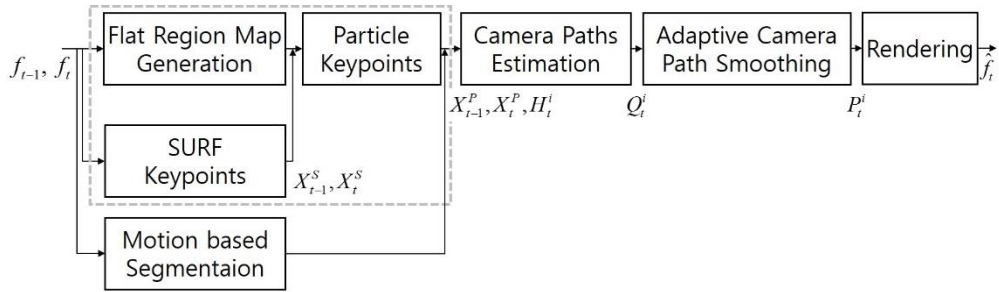


그림 2. 제안하는 비디오 안정화 방법의 블록다이어그램

제안하는 알고리즘은 영상에서 인접한 프레임 f_{t-1} 와 f_t 사이의 움직임을 구하고 움직임에 따라 영역을 분할한다. 다음으로 SURF 알고리즘을 이용하여 f_{t-1} 와 f_t 사이의 대응점 X_{t-1}^S , X_t^S 을 추출한다. 추출된 특징점의 통계적 특성을 이용하여 평평한 영역에 대한 파티클 키포인트 X_{t-1}^P ,

X_t^p 를 생성한다. 이를 이용하여 영상의 특징점을 재 정의하고 인접 프레임의 매칭점을 추정한다. 이후 추정된 매칭점을 이용하여 분할된 영역에 따라 각각 호모그래피 H_t^i 를 추정하여 움직임에 적응적인 카메라 패스 Q_t^i 를 갱신한다. 결과적으로 추정된 카메라 패스를 최적화 방법을 이용하여 스무딩함으로써, 각 영역에 해당되는 스무딩 된 카메라 패스 P_t^i 로 영상을 재 구성하여 안정화 된 비디오 프레임 \hat{f}_t 을 생성한다.

3.1 움직임 기반 영역 분할

일반적으로, 기존 비디오 안정화 방법은 단일의 글로벌 카메라 패스를 추정하여 스무딩함으로써 흔들리는 비디오의 성능을 향상시킨다. 기존 안정화 방법은 영상의 변화에 강건한 특징점의 매칭을 통하여 사영변환을 추정함으로써 글로벌 카메라 패스를 추정한다. 하지만, 이와 같은 방법은 영상 내의 거리 정보가 상이한 경우, 3차원적인 요소를 고려하지 않기 때문에 안정화 수행 시 기하학적 왜곡이 발생하는 문제가 있다. 또한 임의로 움직이는 물체 영역에서의 변환 행렬의 오 추정으로 인하여 안정화 된 카메라 패스를 추출하지 못하는 문제점이 있다. 이를 해결하기 위하여, 제안하는 알고리즘은 움직임 정보에 따라 영상을 분할하고, 각 영역에 맞는 호모그래피를 추정한다. 호모그래피는 같은 평면에 있는 점들의 변환 관계를 나타낼 수 있기 때문에, 제안하는 알고리즘은 주변 영역과 움직임이 비슷한 영역은 동일 평면에 존재하고, 움직임이 상이한 영역은 다른 평면에 존재한다고 가정하여 3차원적인 영상 정보를 보완한다. 이 때, 각 영역의 모양에 대해 추정하는 것이 아닌 하나의 평면을 추정하는 것이기 때문에 호모그래피를 추정하였으며, 분할된 영역의 모양은 3.4절에서 영상 워핑 시 사용된다.

제안하는 알고리즘은 영상 분할을 위해 Liu의 옵티컬 플로우 방법을

이용하여 인접 프레임 f_{t-1} 와 f_t 사이의 움직임을 추정한다 [25]. 이는 밝기 항상성과 기울기 항상성을 가정하여 데이터항을 정의하고, 불연속적인 부분을 보완하는 스무딩항을 이용한 에너지 함수를 정의함으로써 강건한 옵티컬 플로우를 추정한다. 다음으로는 Hoiem의 슈퍼 픽셀(super-pixels) 방법으로부터 움직임에 따른 영역 분할을 수행한다 [26][27]. Hoiem는 영상을 분할하기 위하여 완전한 기하학적 클래스 집합을 제안하였지만, 본 논문에서는 움직임 기반의 분할에 대한 목적을 위하여 간단화하여 사용하였다. 먼저, 영상 분할을 위하여 입력 비디오 프레임에 대해 작고 균일한 영역에 대한 슈퍼 픽셀을 정의한다. 작은 단일 화소로부터 시작하여, 경계를 가로지르는 움직임 차의 최솟값이 영역 안의 움직임 차의 최댓값보다 클 때 인접한 영역을 합하여 슈퍼 픽셀을 생성한다. 다음으로, 밝기 또는 색상과 같은 영역의 간단한 특징을 이용하여 영역의 라벨링(labeling)을 수행한다. 제안하는 분할 방법은 신뢰 할 수 있는 확률 값을 기반으로 추정되었기 때문에 영역마다 가장 적합한 클래스로 분류 가능하다. 본 논문에서는, 영상 분할을 위하여 반복적으로 인접한 영역을 합하여 라벨링 된 영상을 생성하였으며, 실험적으로 10번의 반복을 수행하였다.

제안된 알고리즘은 라벨링 된 이미지를 이용하여 각 영역의 최적의 카메라 패스를 추정한다. 그림 3은 제안하는 알고리즘을 이용한 영역 분할 결과를 보인다. 그림 3(a)의 t번째 입력 영상에 대해, 그림 3(b)와 같이 옵티컬 플로우를 이용하여 움직임을 추정한다. 이후 그림 3(c)와 같이 제

안하는 방법을 이용하여 분할된 영상을 생성할 수 있다.



(a)

(b)

(c)

그림 3. 영역 분할 결과: (a) 원본 영상, (b) 추정된 옵티컬 플로우, 그리고 (c) 옵티컬 플로우 기반의 분할된 영상

3.2 파티클 기반 특징점 추출 및 카메라 패스 추정

영상의 특징점을 추출하고 기술하는 것은 인접한 프레임 사이의 특징점 매칭을 통해 두 프레임 사이의 기하학적 관계를 분석함에 있어서 매우 중요하다 [28][29]. 따라서 영상 처리에서 특징점은 서로 다른 영상에서 공통 영역을 찾는 방법 중의 하나로 다양한 응용 프로그램에 활용 되고 있으며, 기존에 연구가 활발히 진행되어 왔다.

먼저 Harris는 코너 검출기의 가장 초기 모델을 제안하였다 [30]. 입력 영상 내 작은 윈도우를 움직여봄으로써, 윈도우 내에서 가장 큰 변화량을 갖는 점을 주요 특징점으로 찾아낸다. Lowe는 가우시안의 차(difference of Gaussian; DoG)를 이용하여 이미지 피라미드를 생성하고, 생성된 피라미드 내에서 국부 최댓값(local maxima)을 특징점으로 검출하는 알고리즘인 SIFT를 제안하였다 [31]. 이는 크기, 회전, 조명 변화에 강건하지만 시간복잡도가 높은 문제가 있다. 이를 해결하기 위해, Bay는 근사화된 필터와 적분 영상(integral image)을 적용시킴으로써 향상된 속도로 스케일, 회전에 강건한 특징점을 추출한다. [32]. Tola는 SIFT의 알고리즘에서 출발하여 새로운 기술자 패턴을 사용함으로써 효과적으로 영상에서의 밀집한 특징점을 검출하였다 [33]. 최근에는 특징점 추출 속도의 향상을 위하여 영상의 밝기 비교를 통해 특징점을 추출하고 기술한다. Rosten은 가속화된 분할

테스트(accelerated segment test)를 이용하여 빠른 속도로 영상의 코너를 추출한다 [34][35]. Calonder는 단순한 영상의 밝기 비교를 통해 이진 기술자(binary descriptor)를 생성하는 방법을 제안했다 [36]. 이후 이를 기반으로 샘플링 패턴을 이용하여 방향성을 보완한 다양한 이진 기술자들이 제안되었다 [37][38][39].

기존의 비디오 안정화 알고리즘은 영상의 변화에 강건한 기존의 특징점 추출 알고리즘을 통해 인접 프레임 사이의 움직임을 추정하여 카메라 패스를 구한다. 하지만 이는 입력 영상이 평평한 배경 영역을 많이 포함하고 있거나, 객체가 작거나 존재하지 않는 영상에서 특징점이 잘 추출되지 않는 문제점이 있다. 이에 따라 영상의 특성을 반영하지 못한 호모그래피가 추출됨으로써, 안정화 수행 시 기하학적 왜곡이 발생하는 문제가 있다. 따라서 이러한 문제를 해결하기 위해, 제안하는 알고리즘은 평평한 영역을 검출하고, 그 영역 내의 특징점을 재 정의함으로써 최적의 카메라 패스를 추정한다.

제안하는 알고리즘은 영상의 평평한 영역을 추정하기 위하여, 평평한 영역에 대한 맵을 생성한다. 먼저 입력 영상 f_{t-1} 와 f_t 에 3×3 크기의 가우시안 필터를 적용시켜 스무딩 된 영상을 얻는다. 그 후 원본 영상과 스무딩 된 영상의 차를 이용하여 에지 영역과 평평한 영역으로 나눈다. 마지막으로 확장 연산을 통해 에지 영역을 확장시킴으로써 최종적인 평평한 영역에 대한 맵을 생성한다. 결과적으로, 생성된 평평한 영역에 대한 맵은

강건한 특징점 재 정의를 위해서 사용된다. 그림 4는 t 번째 입력 영상에 대응하는 평평한 영역에 대한 맵을 나타낸다.



그림 4. 평평한 영역 검출 결과: (a) 입력 영상 그리고 (b) 제안하는 알고리즘을 이용한 평평한 영역에 대한 맵

제안하는 안정화 방법은 초기의 랜덤한 파티클 특징점의 분포를 결정하기 위하여 SURF를 사용하여 정확하고 빠른 특징점을 추출한다 [32]. SURF는 2차 가우시안 미분 형태인 헤시안 행렬(hessian matrix) 기반으로 영상의 특징점을 검출한다. 근사화 된 헤시안 행렬 H_{approx} 의 행렬식을 이용하여 특징점의 후보군을 결정하게 되며, 이는 식 (7)로 표현할 수 있다.

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2, \quad (7)$$

여기서, D_{xx} , D_{yy} , 그리고 D_{xy} 는 각각 근사화 된 x축, y축으로의 2차 미분을 나타낸다. SURF는 근사화 된 필터와 적분 영상(integral image)를 이용함으로써, 연산을 단순화 시켜 속도를 향상시킨다. 또한 근사화 된 필터의 사이즈를 변경함으로써 크기 변화에 강건한 특징점을 검출할 수 있다. 이후 하르 웨이블릿 반응(haar wavelet response)을 이용하여 검출된 특징점의 방향 성분에 대해 64차원의 벡터로 기술한다. 따라서 영상의 회전 변화에 강건한 특징점을 검출할 수 있다.

본 논문에서는 평평한 영역에 대한 맵을 바탕으로 인접한 프레임간의 강건한 특징점을 추출한다. 먼저 각 프레임에 적합하며 신뢰성 있는 파티클 키포인트의 분포를 결정하기 위해 SURF를 사용하여 인접한 프레임 f_{t-1} 와 f_t 사이의 특징점을 추출한다. 이를 통해 각 영상의 M 개의 특징점 $X_{t-1}^S = \{(x_{t-1}^1, y_{t-1}^1), \dots, (x_{t-1}^M, y_{t-1}^M)\}$ 와 $X_t^S = \{(x_t^1, y_t^1), \dots, (x_t^M, y_t^M)\}$ 을 추출할 수 있다. 다음으로 강건한 특징점 추출을 위하여 flat한 배경 영역에서의 particle keypoints를 생성한다. 영상 f_{t-1} 와 f_t 의 flat한 영역 내에서 $G(\bar{X}_{t-1}^S, \Sigma_{t-1})$, $G(\bar{X}_t^S, \Sigma_t)$ 의 가우시안 분포를 갖는 개의 N 개의 랜덤한 particle keypoints $X_{t-1}^P = \{(x_{t-1}^1, y_{t-1}^1), \dots, (x_{t-1}^N, y_{t-1}^N)\}$ 와 $X_t^P = \{(x_t^1, y_t^1), \dots, (x_t^N, y_t^N)\}$ 을 생성한다. 여기서 $G(\cdot)$ 는 평균 μ 와 표준편차 σ 을 갖는 가우시안 분포를 나타내며, 식 (8)로 표현할 수 있다.

$$G(\mu, \sigma^2) = \frac{e^{\frac{-(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}. \quad (8)$$

생성된 파티클 키포인트와 SURF 기반의 특징점 사이의 거리를 이용하여 매칭을 위한 특징점 간의 기술자를 재 정의하며, 이는 다음 식 (9)와 같다.

$$D_t = X_t^P - X_t^{FB} \quad (9)$$

여기서 D_t 는 t 번째 프레임에서 파티클 키포인트를 이용하여 생성된 기술자를 나타낸다. 인접한 프레임에서 각각 생성된 기술자를 이용하여 파티클 키포인트를 매칭함으로써 평평한 영역에서도 강건한 특징점을 재 정의할 수 있다. 마지막으로 RANSAC 알고리즘을 이용하여 특징점 사이의 반복적인 랜덤 샘플링을 통해 오 추정된 특징점을 제거한다 [40]. 특징점을 분할된 영역에 맞게 라벨링하고 이를 이용하여 f_{t-1} 와 f_t 간의 i 번째 영역의 호모그래피 H_t^i 를 추정한다. 본 논문에서는 비디오의 카메라 패스 Q_t^i 를 프레임 사이의 관계를 나타내는 호모그래피의 누적 곱으로 나타내며, 이는 식 (10)로 표현할 수 있다.

$$Q_t^i = H_1^i \cdots H_{t-1}^i H_t^i, \quad (10)$$

여기서

$$\mathbf{H}_t^i = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}. \quad (11)$$

제안하는 알고리즘은 f_{t-1} 와 f_t 사이의 독립적인 움직임만을 스무딩 하는 것이 아니라, 영상 전체에 대해 불연속적으로 흔들리는 카메라 패스를 안정화 해야 하기 때문에 누적 곱을 통해 영상 전체에 대한 카메라 패스를 추정한다. 하지만 카메라 패스를 스무딩 하는 단계를 단순화하기 위하여 특정 개수의 프레임을 이용할 수 있다. 본 논문에서는 제안하는 알고리즘의 단순화와 호모그래피 누적 곱에 대한 에러를 보상하기 위하여, 100프레임마다 호모그래피를 초기화 하여 영상을 안정화 한다.

3.3 카메라 패스 스무딩

기존의 비디오 안정화 알고리즘은 칼만 필터와 가우시안 커널기반의 이동 평균(moving average)를 이용하여 카메라 패스를 스무딩하였다 [41][7]. 칼만 필터는 긴 기간의 움직임은 보존하면서 짧은 기간의 흔들림을 제거할 수 있다. 이동 평균은 시간적으로 인접 프레임간의 패스의 평균을 구함으로써 추정된 카메라 패스를 스무딩 할 수 있다. 또한 가우시안 커널을 글로벌 변환에 적용함으로써 움직임을 보정하여 스무딩 패스를 추정할 수 있다. 하지만 기존의 방법은 스무딩 할 때 시간 지연 현상이 발생하고 샘플 간의 급격한 움직임을 따라가지 못한다. 이에 따라 스무딩 패스가 원 패스와 비슷하지 않기 때문에 렌더링 결과에서 크롭(cropped) 된 영역과 기하학적 왜곡이 증가하여 낮은 품질의 결과 영상을 얻는 문제점이 있다. 또한 비디오 안정화 결과는 하나의 장면에 대해 얻게 되므로 분할된 영역들이 서로 독립적이지 않도록 카메라 패스를 스무딩 하는 것이 필요하다.

본 논문에서는 이를 보완하기 위하여 카메라 패스 안정화를 위한 두 개의 제약 조건을 포함하고 있는 에너지 함수를 최소화시킴으로써 비디오를 안정화한다. 제안하는 스무딩 카메라 패스 생성 방법은 영상을 블록으로 나누어 처리할 때 블록간의 상관관계를 유지시켜 영상의 형태를 보존해주는 에너지 함수를 사용하는 Liu의 방법의 발전된 형태이다 [22]. 제안

하는 영역에 따라 스무딩 된 카메라 패스 P_t^i 는 다음 식의 에너지 함수의 최적화에 의해 결정된다.

$$E(P_t^i) = \sum_i (\|P_t^i - Q_t^i\|^2 + \lambda_1 \sum_n w \|P_t^i - P_n^i\|^2 + \lambda_2 \sum_j \|P_t^i - P_t^j\|^2), \quad (12)$$

여기서 Q_t^i 는 원 카메라 패스, P_t^i 는 t 번째 프레임의 i 번째 영역에서 스무딩 된 카메라 패스, P_n^i 는 n 번째 프레임의 i 번째 영역에서 스무딩 된 카메라 패스, P_t^j 는 t 번째 프레임의 j 번째 영역에서 스무딩 된 카메라 패스이다. λ_1 과 λ_2 는 각 항의 균형을 맞춰주는 가중치 상수이며, 실험적으로 $\lambda_1 = 20$ 그리고 $\lambda_2 = 1$ 으로 결정하였다. w 는 시간적으로 인접한 카메라 패스에 대한 가중치를 의미하며, w 는 가우시안 함수를 이용하여 시간적으로 인접한 카메라 패스일수록 가중치를 높게 설정하였다.

다음으로 각 항이 의미하는 것은 아래와 같다. 첫 번째 항은 스무딩 된 카메라 패스가 원 카메라 패스를 유지하도록 하는 역할을 하며, 두 번째 항은 카메라 패스를 스무딩 하는 역할을 한다. 세 번째 항은 분할된 영역의 카메라 패스들이 상이하지 않도록 보완해주는 역할을 한다. 이때, 식 (12)와 같은 에너지 함수가 볼록한 함수이므로, 선형 풀이 방법을 이용하여 함수를 풀 수 있다.

3.4 영상 워핑을 통한 비디오 안정화

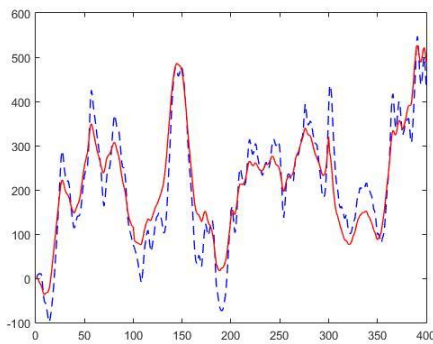
마지막으로 스무딩 카메라 패스를 적용하여 영상을 워핑함으로써 안정화된 영상을 생성한다. 우리는 움직임에 기반하여 분할된 영역에 따른 스무딩 카메라 패스를 적용하기 위하여, $M \times N$ 개수의 작은 매쉬 그리드로 영상을 나눈다. 이때, t 번째 프레임에서 나뉜 영상의 집합은 $V_t = \{V_t^1, \dots, V_t^{MN}\}$ 로 정의한다. 본 논문에서는 복잡한 렌더링 과정을 거치지 않고 앞서 나누었던 각 영역의 모양을 보존하기 위하여 각 매쉬 그리드의 4개의 꼭지점을 이용한다. 우리는 각 그리드에 해당하는 영역의 호모그래피를 이용하여 영사아을 워핑함으로써 안정화 된 영상의 매쉬 그리드 집합 $\hat{V}_t = \{\hat{V}_t^1, \dots, \hat{V}_t^{MN}\}$ 을 추정하며, 이는 다음과 같이 정의할 수 있다.

$$\hat{V}_t^i = \tilde{H}_t^j V_t^i, \quad (13)$$

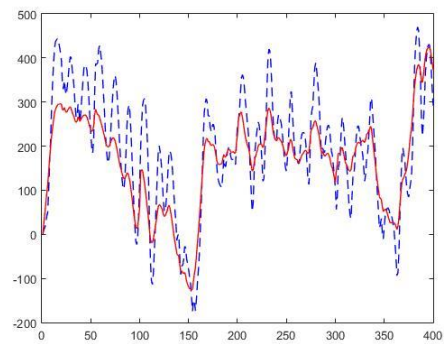
여기서 $i=1, \dots, MN$ 일 때 $V_t^i = \{v_i^1(x, y), v_i^2(x, y), v_i^3(x, y), v_i^4(x, y)\}$ 는 t 번째 프레임의 i 번째 그리드의 4개의 꼭지점을 나타내며, $j=1, \dots, K$ 일 때 $\tilde{H}_t^j = (P_{t-1}^j)^{-1} P_t^j$ 는 t 번째 프레임의 j 번째 영역의 스무딩 된 호모그래피를 의미한다. 이때, 만약 하나의 그리드에서 여러 평면이 만나게 된다면, 그 중 더 많은 영역을 차지하는 평면을 현재 그리드가 해당되는 평면으로 지정한다. 결과적으로 움직임에 따른 영역 분할에 기반하여 기하학적 왜곡

이 없는 안정화 비디오 프레임 \hat{f}_t 를 생성한다.

그림 5은 제안하는 방법으로 추정한 카메라 패스를 나타낸다. 그림 5(a)와 그림 5(b)에서 보듯, 원 카메라 패스와 비슷하며 흔들리는 모션이 감소한 스무딩 패스가 생성된 것을 확인 할 수 있다.



(a)



(b)

그림 5. 첫 번째 영역에 대해 추정된 카메라 패스: (a) x축 방향으로의 원본 카메라 패스 (점선)와 스무딩 된 카메라 패스 (실선) 그리고 (b) y축 방향으로의 원본 카메라 패스 (점선)와 스무딩 카메라 패스 (실선)

제 4장 실험결과

본 장에서는 3장에서 제안된 파티클 기반의 특징점 추출 결과와 제안하는 알고리즘을 이용한 비디오 안정화 방법의 실험 결과를 확인한다.

4.1 파티클 기반 특징점 추출 결과

본 장에서는 제안된 파티클 키폰트 방법과 기존 특징점 추출 방법들의 비교를 통해 제안하는 방법의 우수성을 증명한다. 제안하는 방법은 파티클 키폰트를 이용하여 최적의 호모그래피를 추정함으로써 안정화를 수행한다. 따라서 인접 프레임 f_{t-1} 와 f_t 간의 추정된 호모그래피 H_t 의 정확도를 측정하기 위하여, 두 영상에서 매칭된 특징점 X_{t-1} 와 X_t 을 이용하여 에러를 추정하며 이는 다음과 같이 정의된다.

$$Error = \frac{1}{n} \sum_n \|\tilde{X}_t - X_t\|, \quad (14)$$

여기서 $\tilde{X}_t = H_t X_{t-1}$ 는 이전 프레임의 특징점 X_{t-1} 가 추정된 호모그래피 H_t 로부터 변환된 점을 나타낸다. 본 논문에서는 기존 방법 중 Harris [30], SIFT [31], SURF [32], FAST [34], BRIEF [36], BRISK [38], FREAK [39] 방법과 제안하는 방법을 이용하여 각각 호모그래피 H_t 를 추정하여 에러 값을 추정한다. 공정한 비교를 위하여, 알고리즘의 비교에 사용된 두 영상의 매칭점 X_{t-1} 와 X_t 은 모든 알고리즘에 동일한 값이 사용되었다. 표 1은 다양한 알고리즘에 대한 호모그래피의 에러를 추정한 결과이다.

표 1. 다양한 특징점 추출 알고리즘을 이용하여 추정된 호모그래피의 에러 값 비교

	Video1	Video2	Video3
제안된 파티클 키포인트	6.2987		
SIFT [31]	6.3490		
SURF [32]	6.5318		
Harris [30]+BRIEF [36]	6.6736		
Harris [30]+FREAK [39]	6.7344		
FAST [34]+BRIEF [36]	6.4990		
FAST [34]+FREAK [39]	6.3569		
BRISK [38]	6.9978		

그림 6은 다양한 특징점 추출 방법을 이용하여 인접 프레임간의 매칭점을 추정한 결과이다. 각 알고리즘에 대한 실험 결과는 RANSAC 알고리즘을 이용하여 아웃라이어를 제거하여 나타내었다. 그림 6(a)-(g)에서 보이는 기존의 특징점 추출 방법과 달리, 제안하는 파티클 키포인트 알고리즘을 이용한 매칭점은 평평한 영역을 포함한 영상 전체에 대해 추출된 것을 그림 6(f)를 통해 확인할 수 있다.

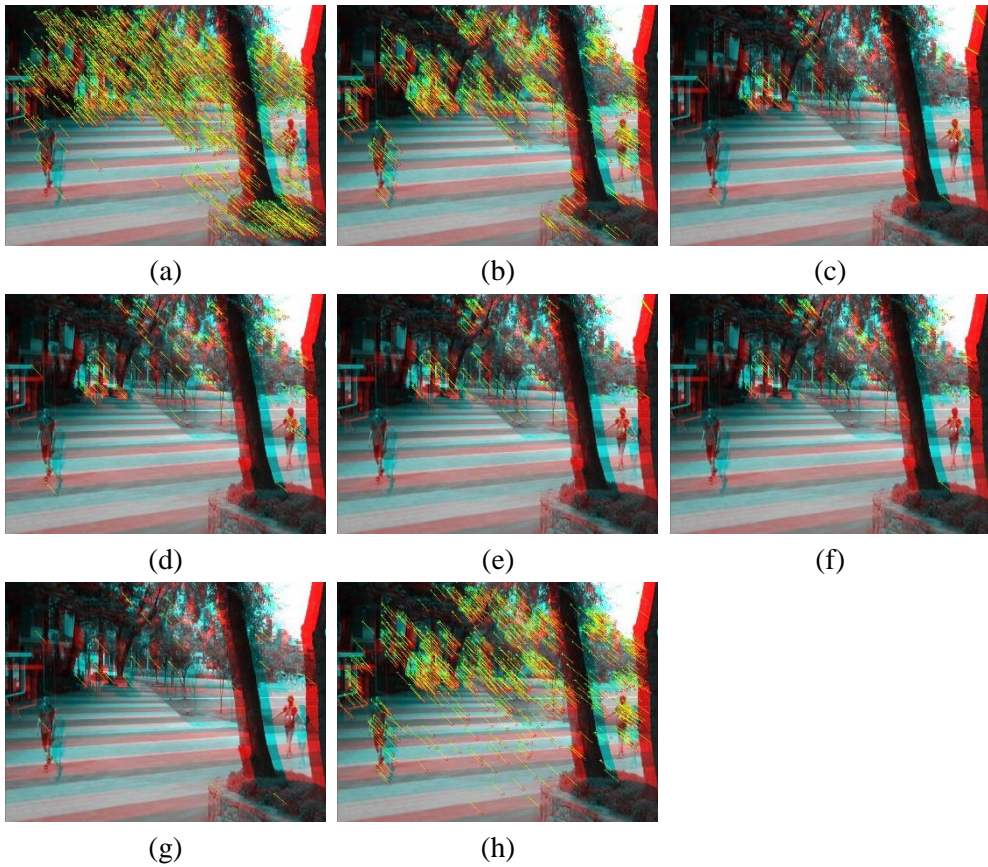
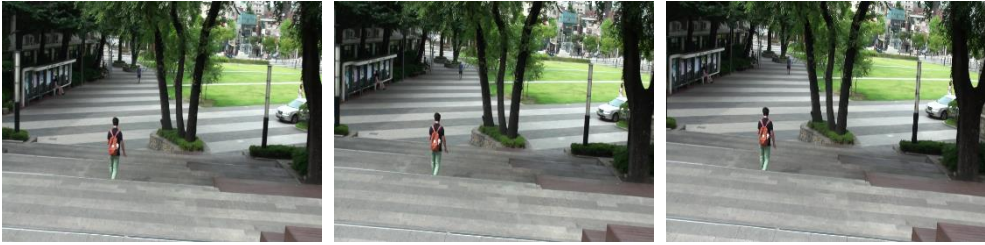


그림 6. 다양한 알고리즘을 이용한 특징점 추출 및 매칭 결과: (a) SIFT [31], (b) SURF [32], (c) Harris [30]+BRIEF [36], (d) Harris [30] +FREAK [39], (e) FAST [34] +BRIEF [36], (f) FAST [34] +FREAK [39], (g) BRISK [38], 그리고 (h) 제안하는 파티클 키포인트

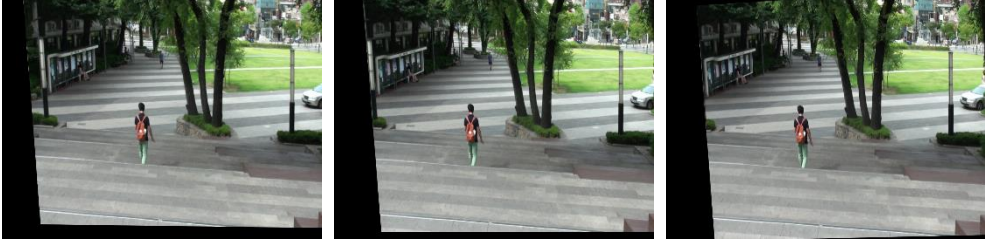
4.2 제안하는 알고리즘을 이용한 비디오 안정화 결과

본 장에서는 기존 방법과 제안된 방법의 비교를 통해 제안하는 비디오 안정화 방법의 성능을 증명한다. 본 실험에서는 입력 영상 3개를 이용하여 제안하는 비디오 안정화 알고리즘의 성능을 평가하였다.

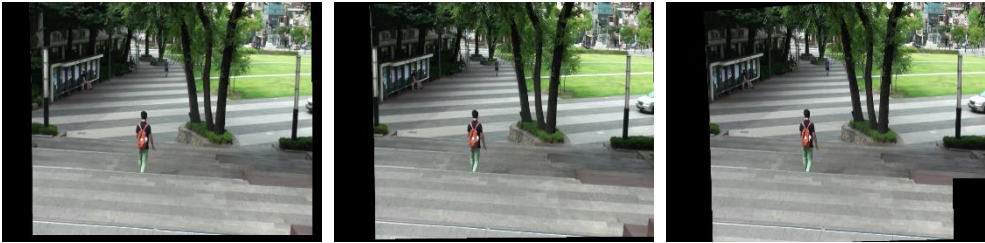
그림 7은 흔들리는 카메라로부터 얻은 원본 비디오를 기존 방법 및 제안하는 방법을 이용하여 안정화 한 결과 영상을 보인다. 그림 7(b)에서 보듯, 그림 7(a)의 입력영상에 대해 기존의 특징점 추출 알고리즘 기반의 글로벌 카메라 패스를 이용하여 안정화하였을 때 잘못 추정된 호모그래피로 인하여 가장자리에 기하학적 왜곡이 발생한 것을 확인할 수 있다 [7]. 영상의 중심에 위치해 있는 객체를 보면, 글로벌 카메라 패스를 이용하여 안정화 한 결과상에서 객체가 휘어짐으로써 기하학적 왜곡을 쉽게 확인할 수 있다. 또한 그림 7(c)에서 번들 패스(bundled path) 알고리즘을 이용하여 지역적 카메라 패스를 추정함으로써 안정화 한 결과를 보이며, 무늬가 없는 블록에서의 호모그래피 추정의 오류로 인하여 비디오 안정화를 위한 워핑 시 오류가 발생한 것을 확인할 수 있다 [23]. 반면에, 제안하는 방법을 이용하여 안정화를 수행한 결과 그림 7(d)와 같이 적은 기하학적 왜곡을 가지며 안정화의 성능이 향상된 것을 확인할 수 있다.



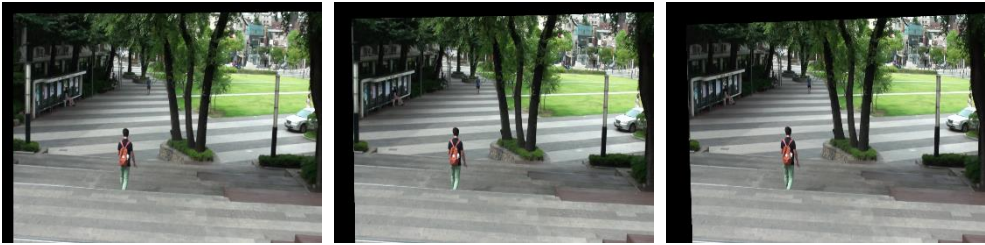
(a)



(b)



(c)



(d)

그림 7. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) 65, 66, 67번째 흔들리는 입력 비디오 영상, (b) Matsushita 방법을 이용한 안정화 영상 [7], (c) Liu 방법을 이용한 안정화 영상 [22], 그리고 (d) 제안하는 방법을 이용한 안정화 영상

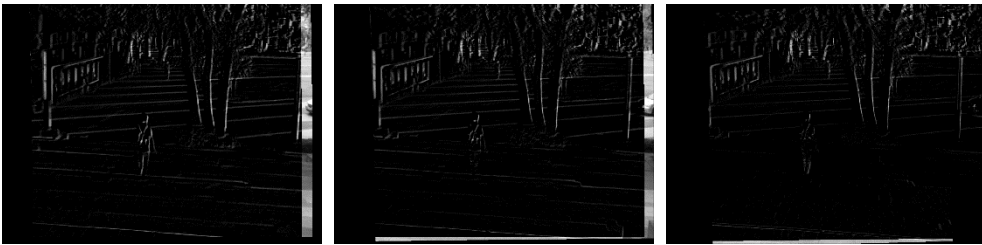
그림 8은 다양한 비디오 안정화 알고리즘에 대해 더 정확한 비교를 위하여 그림 7과 같은 실험 영상에서 인접 프레임간의 차 영상을 나타낸다. 그림 7(a)-(d)는 각각 원본 영상과 안정화 된 영상의 세 가지 프레임 쌍 $\{(64, 65), (65, 66), (66, 67)\}$ 에 대한 차분영상을 나타낸다. 결과적으로 제안하는 알고리즘은 파티클 특징점을 사용하여 평평한 부분의 움직임까지 고려하였기 때문에 바닥부분을 포함하여 영상의 전체적인 흔들림을 효과적으로 제거한 것을 확인할 수 있다.



(a)



(b)



(c)



(d)

그림 8. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) $\{(64, 65), (65, 66), (66, 67)\}$ 에 대한 입력 비디오 차분 영상, (b) Matsushita 방법을 이용한 안정화 차분 영상 [7], (c) Liu 방법을 이용한 안정화 차분 영상 [22], 그리고 (d) 제안하는 방법을 이용한 안정화 차분 영상

그림 9는 다른 실험 영상에서의 제안하는 알고리즘과 기존의 알고리즘을 사용한 안정화 영상의 비교를 나타내며, 가장자리에 발생한 홀을 제거하기 위하여 영상을 크롭하여 나타냈다. 그림 9(a)의 110, 111, 112번째 입력 비디오 프레임에 대하여, 그림 9(b)와 그림 9(c)는 각각 번들 패스 알고리즘 [22]과 서브 스페이스(subspace) 알고리즘 [18]을 이용하여 안정화 한 결과 영상이다. 이때, 3차원 방법을 사용하여 안정화 한 결과는 객체와 배경의 다른 움직임으로 인하여 기하학적 왜곡이 발생했음을 볼 수 있다. 또한 3차원 안정화 알고리즘의 결과는 가장자리의 홀과 왜곡을 보상하기 위하여 원본 비디오의 중요 콘텐츠가 제거되는 문제점이 나타났다. 반면 제안하는 알고리즘을 이용한 안정화 결과는 최적의 카메라 패스 추적을 통하여 고품질의 안정화 영상을 보이며, 이는 그림 9(d)에서 나타난다. 결과적으로, 제안하는 비디오 안정화 방법은 움직임 기반의 영상 분할과 강건한 특징점을 이용하여 효과적으로 흔들림을 제거하였다.

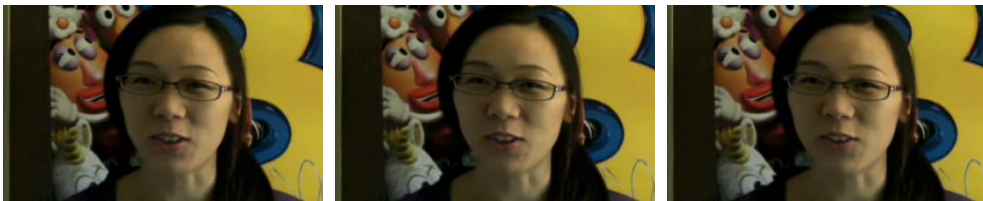
그림 10은 다양한 비디오 안정화 알고리즘에 대해 더 정확한 비교를 위하여 그림 8과 동일한 실험 결과를 나타낸다. 그림 10(a)-(d)는 각각 원본 영상과 안정화 된 영상의 세 가지 프레임 쌍 $\{(109, 110), (110, 111), (111, 112)\}$ 에 대한 차분영상을 나타낸다. 결과적으로 제안하는 알고리즘은 사용자 카메라로부터 얻은 흔들리는 영상에 대해 추가적인 기하학적 왜곡 없이 효과적으로 안정된 비디오를 보인다.



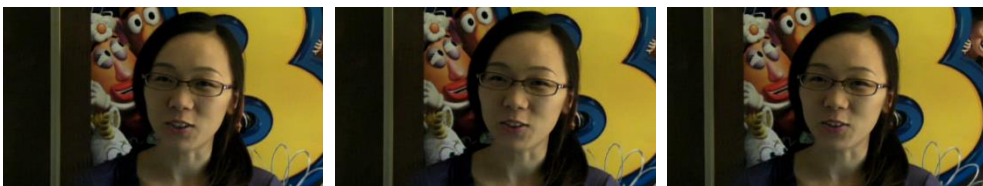
(a)



(b)



(c)

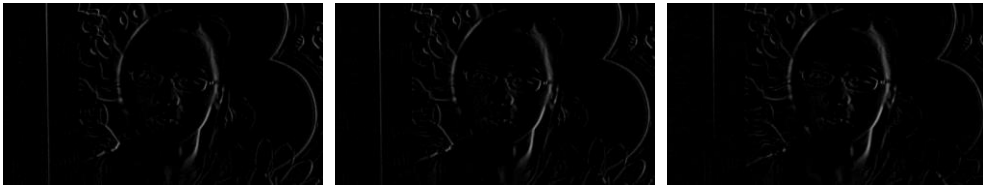


(d)

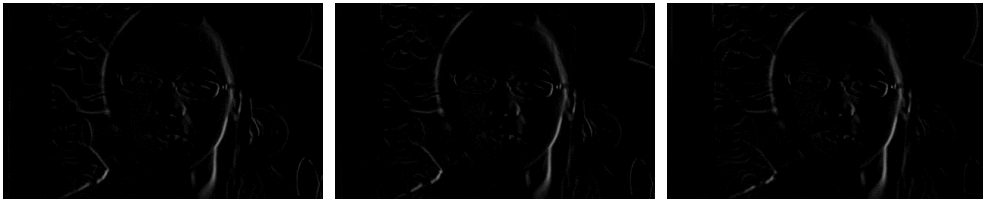
그림 9. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) 110, 111, 112번째 흔들리는 입력 비디오 영상, (b) Liu 방법을 이용한 안정화 영상 [22], (c) Liu 방법을 이용한 안정화 영상 [18], 그리고 (d) 제안하는 방법을 이용한 안정화 영상



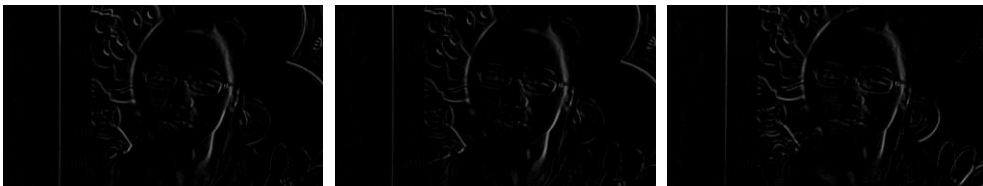
(a)



(b)



(c)



(d)

그림 10. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) $\{(109, 110), (110, 111), (111, 112)\}$ 에 대한 입력 비디오 차분 영상, (b) Liu 방법을 이용한 안정화 차분 영상 [22], (c) Liu 방법을 이용한 안정화 차분 영상 [18], 그리고 (d) 제안하는 방법을 이용한 안정화 차분 영상

그림 11는 다른 실험 영상에서의 제안하는 알고리즘과 두 개의 기존 알고리즘을 사용하여 안정화 한 결과를 비교한다. 마찬가지로 공정한 비교를 위하여 가장자리에 발생한 홀을 제거하기 위하여 안정화 된 영상을 크롭하여 나타냈다. 그림 11(a)에서 보이는 입력 비디오에서 선택된 세 개의 입력 프레임에 대하여, Grundmann [12]와 Zhou [19]이 제안한 방법을 이용하여 안정화 한 영상은 각각 그림 11(b)와 그림 11(c)에서 보인다. 그림 11(d)에서 보듯, 제안하는 비디오 안정화 방법은 영상의 내용을 보존하며 원치 않았던 움직임도 효과적으로 제거함으로써 비디오의 품질을 향상시켰다.

그림 12는 제안하는 알고리즘의 성능을 증명하기 위하여 인접 프레임간의 차분 영상을 보인다. 그림 12(a)-(d)는 각각 원본영상, Grundmann의 방법 [12]을 이용한 안정화 영상, Zhou의 방법 [19]을 이용한 안정화 영상, 그리고 제안하는 방법을 이용한 안정화 영상에 대한 차분 영상이다. 결과적으로, 인접 프레임간의 흔들림을 최소화함으로써 안정화 된 영상이 생성된 것을 확인 할 수 있다.



(a)



(b)

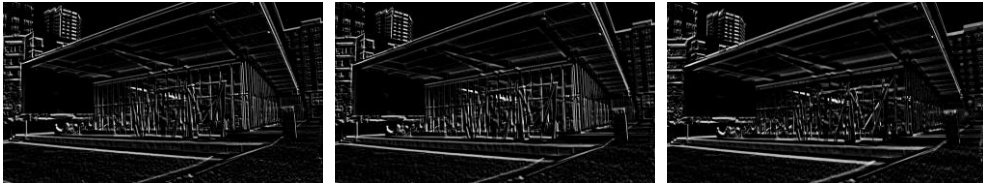


(c)



(d)

그림 11. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) 334, 335, 336번째 흔들리는 입력 비디오 영상, (b) Grundmann 방법을 이용한 안정화 영상 [12], (c) Zhou 방법을 이용한 안정화 영상 [19], 그리고 (d) 제안하는 방법을 이용한 안정화 영상



(a)



(b)



(c)



(d)

그림 12. 다양한 알고리즘을 이용한 비디오 안정화 결과: (a) $\{(333, 334), (334, 335), (335, 336)\}$ 에 대한 입력 비디오 차분 영상, (b) Grundmann 방법을 이용한 안정화 차분 영상 [12], (c) Zhou 방법을 이용한 안정화 차분 영상 [19], 그리고 (d) 제안하는 방법을 이용한 안정화 차분 영상

제안되는 방법은 카메라의 흔들리는 움직임을 제거하였기 때문에 안정화 된 영상의 연속적인 프레임의 차가 최소가 된다. 따라서 제안하는 방법의 실험결과에 대한 객관적인 평가를 위해 인접한 프레임의 최대신호 대잡음비(peak signal to noise ratio; PSNR)의 수치를 비교하며, 이는 다음과 같이 정의된다.

$$PSNR = 10 \log \frac{MAX_f^2}{MSE}, \quad (15)$$

여기서 $MSE = \|f_{i-1} - f_i\|^2$ 는 인접 프레임간의 평균제곱오차(mean square error)를 나타내며, MAX_f 는 프레임 내의 가장 큰 밝기 값을 나타낸다. 표 2는 안정화 된 비디오의 인접 프레임에 대한 PSNR 값을 나타낸다. 결과적으로, 제안하는 비디오 안정화 방법을 이용하여 인접 프레임간 화소들의 잘못된 위치를 바로 잡아준 것을 확인할 수 있다.

표 2. 원영상과 안정화 영상의 PSNR 값 비교

	프레임	원영상	안정화영상
Video1	65	13.8732	17.2121
	66	14.2882	17.6466
	67	13.8210	20.5377
	평균 (300 프레임)	14.5765	17.0475
Video2	110	18.2937	25.0486
	111	19.4547	25.1494
	112	20.0965	23.8199
	평균 (446 프레임)	21.6836	23.6703
Video3	334	20.5033	21.9704
	335	20.4595	22.3818
	336	18.3253	21.1192
	평균 (450 프레임)	14.9423	17.7095

마지막으로, 제안하는 방법과 기존 방법들의 객관적인 비교를 위하여 Liu의 방법을 이용하여 사영 왜곡을 측정한다 [22]. 2장에서 언급 한 바와 같이, 일반적으로 사영 왜곡은 실세계가 이미지 센서로 사영될 때 발생한다. 따라서 잘못 추정된 호모그래피를 사용하여 영상을 안정화 한다면, 사영 왜곡에 의해 비디오의 기하학적인 품질이 저하되는 문제점이 있다. 따라서 우리는 흔들리는 원본 카메라 패스와 안정화된 카메라 패스 사이의 변환 관계를 통해 사영 왜곡을 측정하며, 이는 다음과 같이 정의된다.

$$P_i = B_i Q_i \quad (16)$$

여기서 Q_i 와 P_i 는 각각 원본 카메라 패스와 안정화 된 카메라 패스를 의미하며, B_i 는 둘 사이의 변환관계를 나타내는 행렬이다. 사영 왜곡은 고품질의 비디오 획득을 위해 중요한 지표이기 때문에, 본 논문에서는 각 프레임에서 측정된 B_i 의 사영 요소의 평균을 이용하여 이를 측정하였다. 결과적으로, 제안하는 알고리즘은 사영 왜곡이 감소하는 방향으로 비디오 안정화가 수행되었으며, 표 3에서 제안된 알고리즘과 기존 알고리즘의 비교를 통해 제안된 알고리즘의 우수성을 보인다.

표 3. 기존 비디오 안정화 방법과 제안된 방법에 대한 사영 왜곡 결과

	Video1
제안된 방법	1.03×10^{-4}
Matsushita [7]	0.98×10^{-4}
Liu [22]	2.31×10^{-4}
	Video2
제안된 방법	2.56×10^{-4}
Liu [22]	2.72×10^{-4}
Liu [18]	5.99×10^{-4}
	Video3
제안된 방법	2.7×10^{-5}
Grundmann [12]	1.53×10^{-4}
Zhou [19]	3.8×10^{-5}

제 5장 결 론

본 장에서는 제안하는 비디오 안정화 방법의 장단점을 간략히 요약하고, 응용 분야와 향후 진행될 보안 과제 및 연구 방향에 대해 기술한다.

5.1 결론

제안하는 알고리즘은 움직임에 따른 영상 분할과 강건한 특징점 추정을 통해 각 영역에 적합한 최적의 호모그래피를 추정하고 분할된 영역에 적합한 카메라 패스를 스무딩 함으로써 비디오의 흔들림을 효과적으로 제거하였다. 기존의 안정화 방법은 객체와 배경의 불연속적인 움직임에 따른 기하학적 왜곡이 발생하는 문제가 있었다. 본 논문에서는 최적화 방법을 사용하여 분할된 영역에 따른 적응적 카메라 패스를 추정함으로써 객체와 배경의 움직임이 상이한 비디오에서의 안정화 성능을 향상시켰다. 또한 영상 전체에 분포된 파티클 특징점 기반의 강건한 특징점을 사용하여 최적의 카메라 패스를 추정함으로써, 영상의 평평한 영역에서의 특징점 오검출로 인해 호모그래피 추정 성능이 저하되는 기존 안정화 방법의 문제점을 해결하였다. 결과적으로 제안하는 방법은 영상에 따른 최적의 카메라 패스를 추정함으로써 2차원 안정화 방법의 한계를 보완하여 사용자에게 보다 자연스러운 안정화 영상을 제공할 수 있다.

본 논문에서는 실험결과에서 주관적인 평가뿐만 아니라 객관적인 평가를 통해 기존의 방법보다 향상된 안정화 방법을 제안한다. 하지만 여전히 영상 안정화 시 홀이 발생하기 때문에 영상의 가장자리를 크롭 함으로써, 안정화 된 결과를 나타내게 된다. 또한 호모그래피의 누적 에러를 방

지하기 위하여 특정 프레임마다 호모그래피를 초기화함으로써, 프레임 사이에 불연속성이 발생하게 된다. 따라서 이를 해결하기 위해 추후에는 인접 프레임의 정보를 이용하여 영상 안정화 시 생성되는 홀을 보상하는 연구와 초기화 된 호모그래피를 이전 프레임과 연결하는 동시에 연산의 복잡도를 감소시키는 연구가 필요하다.

제안하는 알고리즘은 흔들림 없이 높은 해상도의 영상을 제공하며 이는 흔들림이 발생할 수 있는 다양한 환경에서 적용될 수 있다. 제안하는 알고리즘은 높은 수준의 계산 처리가 가능한 서버를 기반으로 클라우드 서비스를 사용하여 업로드 된 영상의 왜곡을 개선해 주는데 적합하다. 또한 서버 기반으로 개인 영상을 저속도 촬영 비디오(time lapse video) 형태로 영상을 요약하여 제공함으로써 개인용 디바이스의 용량 한계나 무선 인터넷의 속도의 한계를 극복할 수 있으며, 이때 전처리 과정으로 비디오 안정화를 사용함으로써 고품질의 영상을 제공할 수 있다 [43][44]. 또한 비전 기반 모바일 로봇으로부터 얻은 영상의 흔들림을 제거하기 위하여, 최근에는 SURF 특징점 기반의 비디오 안정화 방법이 모바일 로봇 기반 시스템에 적용되고 있다 [17]. 제안하는 알고리즘은 이와 같은 시스템에 적용되어 평평한 영역에서의 적절한 특징점 추정을 통해 비디오 안정화의 성능을 향상시킬 수 있다. 또한 최근에는 항공 감시 시스템에서 넓은 지역에서의 물체 추적을 위하여 비디오 안정화를 사용하고 있다 [45]. 항공 감시 시스템은 움직이는 카메라로부터 얻기 때문에 인접 프레임간의 흔들

림이 존재하며, 강건한 물체 추적을 위해 비디오 안정화의 전처리 과정이 필수적이다. 제안하는 알고리즘은 항공 영상으로부터 얻은 평평하거나 저 해상도의 영상에서 중요 특징점을 정의해줌으로써 기존 비디오 안정화의 성능을 향상시킬 수 있다.

휴대가 용이한 카메라를 사용하는 사람들은 사이클링, 하이킹 등 다양한 활동들을 비디오로 함께 공유함으로써 새로운 커뮤니케이션을 하며, 이때 흔들림에 의해 왜곡된 비디오에 대한 비디오 안정화 방법은 중요하다. 제안하는 알고리즘은 카메라 패스를 추정할 때 충분한 길이의 프레임을 누적하여 추정한 후 스무딩을 하게 되어 서버 기반의 안정화에 적합하다. 그럼에도 모바일 카메라와 같은 휴대용 기기에 실시간으로 적용하기 위하여 근사화를 통하여 계산량을 줄일 수 있다. 또한 움직임에 따라 분할된 영역을 병렬적으로 처리하게 된다면 그에 따른 속도 향상 결과를 가져올 수 있다. 이를 통해 모바일 카메라뿐만 아니라 움직이는 가상 현실 (virtual reality; VR) 기기로부터 얻은 흔들리는 비디오의 안정화를 통한 고품질의 콘텐츠 제공 및 비전 기반의 운전자 보조 시스템 등에 활용 할 수 있다.

참고 문헌

- [1] B. Golik, "Development of a test method for image stabilizing systems," *PhD thesis, Diploma Thesis at the Department of Imaging Sciences and Media Technology Cologne University of Applied Sciences*, 2006.
- [2] X. Cheng, Q. Hao, and M. Xie, "A comprehensive motion estimation technique for the improvement of EIS methods based on the SURF algorithm and kalman filter," *Sensors*, vol. 16, no. 4, pp. 486, April 2016.
- [3] C. Buehler, M. Bosse, and L. McMillan, "Non-metric image-based rendering for video stabilization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, no. 52, pp. II-609-614, December 2001.
- [4] S. Jang, M. Pomplun, G. Kim, and H. Choi, "Adaptive robust estimation of affine parameters from block motion vectors," *Image and Vision Computing*, vol. 23, no. 14, pp. 1250-1263, December 2005.
- [5] L. Xu and X. Lin, "Digital image stabilization based on circular block matching," *IEEE Transactions on Consumer Electronics*, vol. 52, no. 2, pp. 556-574, May 2006.
- [6] H. Chang, H. Lai, and R. Lu, "A robust and efficient video stabilization algorithm," *IEEE International Conference on Multimedia and Expo*, vol. 1, June 2004.
- [7] W. Xu, X. Lai, D. Xu, and A. Tsoligkas, "An integrated new scheme for digital video stabilization," *Advances in Multimedia*, vol. 2013, no. 8, January 2013.

- [8] Y. Matsushita, E. Ofek, W. Ge, X. Tang, and H. Shum, "Full-frame video stabilization with motion inpainting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 7, July 2006.
- [9] S. Battiato, G. Gallo, G. Puglisi, and S. Scellato, "SIFT features tracking for video stabilization," *IEEE International Conference on Image Analysis and Processing*, pp.825-830, September 2007.
- [10] K. Lee, Y. Chuang, Y. Chen, and M. Ouhyoung, "Video stabilization using robust feature trajectories," *IEEE International Conference on Computer Vision*, pp. 1397-1404, October 2009.
- [11] J. Xu, H. Chang, S. Yang, and M. Wang, "Fast feature-based video stabilization without accumulative global motion estimation," *IEEE Transaction on Consumer Electronics*, vol. 58, no. 3, pp. 993-999, August 2012.
- [12] B. Pinto and P. Anurenjan, "Video stabilization using speed up robust features," *International Conference Communications and Signal Processing*, pp. 527-531, February 2011.
- [13] M. Grundmann, V. Kwatra, and I. Essa, "Auto-directed video stabilization with robust L1 optimal camera paths," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 225-232, June 2011.
- [14] C. Song, H. Zhao, W. Jing, and H. Zhu, "Robust video stabilization based on particle filtering with weighted feature points," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 2, pp. 570-577, May 2012.

- [15] S. Kim, S. Kang, T. Wang, and S. Ko, "Feature point classification based global motion estimation for video stabilization," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 1, pp. 267-272, February 2013.
- [16] F. Nejadasl and R. Lindenbergh, "Sequential and automatic image-sequence registration of road areas monitored from a hovering helicopter," *Sensors*, vol. 14, no.9, pp. 16630-16650, September 2014.
- [17] T. Shene, K. Sridharan, and N. Sudha, "Real-time SURF-based video stabilization system for an fpga-driven mobile robot," *IEEE Transactions on Industrial Electronics*, vol. 63, no. 8, pp. 5012-5021, April 2016.
- [18] F. Liu and H. Jin, "Content-preserving warps for 3D video stabilization," *ACM Transactions on Graphics (TOG)*, vol. 28, no. 3, pp. 44:1-44:9, August 2009.
- [19] F. Liu, M. Gleicher, J. Wang, H. Jin, and A. Agarwala, "Subspace video stabilization," *ACM Transactions on Graphics (TOG)*, vo. 30, no. 1, p. 4:1-10, 2011.
- [20] Z. Zhou, H. Jin, and Y. Ma, "Plane-based content-preserving warps for video stabilization," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2299-2306, June 2013.
- [21] A. Goldstein and R. Fattal, "Video stabilization using epipolar geometry," *ACM Transactions on Graphics (TOG)*, vol. 31, no. 5, pp. 126:1-10, August 2012.
- [22] Y. Koh, C. Lee, and C. Kim, "Video stabilization based on feature trajectory augmentation and selection and robust mesh grid warping," *IEEE*

Transactions on Image Processing, vol. 24, no. 12, pp. 5260-5273, December 2015.

- [23] S. Liu, L. Yuan, P. Tan, and J. Sun, "Bundled camera paths for video stabilization," *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, pp. 78:1-78:10, July 2013.
- [24] S. Liu, L. Yuan, P. Tan, and J. Sun, "Steadyflow: Spatially smooth optical flow for video stabilization," *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4209-4216, June 2014.
- [25] W. Hong, D. Wei, and A. Batur, "Video stabilization and rolling shutter distortion reduction," in *Proceedings of the IEEE International Conference on Image Processing*, pp. 3501-3504, December 2010.
- [26] C. Liu, "Beyond pixels: exploring new representations and applications for motion analysis," *Ph.D. dissertation, Citeseer*, June 2009.
- [27] D. Hoiem, A. Efros, and M. Hebert, "Geometric context from a single image," in *Proceedings of the IEEE international Conference on Computer Vision*, vol. 1, pp. 654-661, October 2005.
- [28] D. Hoiem, A. Efros, and M. Hebert, "Recovering surface layout from an image," *International Journal of Computer Vision*, vol. 75, no.1, pp. 151-172, February 2007.
- [29] K. Sharma, "A novel approach for object detection in illuminated and occluded video sequences using visual information with object feature estimation," *IEIE Transactions on Smart Processing and Computing*, vol. 4, no. 2, pp. 110-114, April 2015.

- [30] H. Lee, S. Jeon, I. Yoon, and J. Park, "Recent advances in feature detector and descriptor: a survey," *IEIE Transactions on Smart Processing and Computing*, vol. 5, no.3, pp. 153-163, June 2016.
- [31] C. Harris and M. Stephens, "A combined corner and edge detector," *Alvey Vision Conference*, pp. 147–151, 1988.
- [32] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, November 2004.
- [33] H. Bay, T. Tuytelaars, and L. Gool, "SURF: Speeded up robust features," *European Conference on Computer Vision*, vol. 3951, pp. 404-417, May 2006.
- [34] E. Tola, V. Lepetit, and P. Fua, "DAISY: an efficient dense descriptor applied to wide baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815-530, May 2010.
- [35] E. Rosten and T. Drummond, "Fusing points and lines for high performance tracking," *International Conference on Computer Vision*, vol. 2, pp. 17-21, October 2005.
- [36] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," *European Conference on Computer Vision*, vol. 3951, pp. 430-443, May 2006.
- [37] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BREIF: Binary robust independent elementary features," *European Conference on Computer Vision*, vol. 6314, pp. 778-792, September 2010.

- [38] E. Lublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: an efficient alternative to SIFT or SURF,” *IEEE International Conference on Computer Vision*, pp. 2564-2571, November 2011.
- [39] S. Leutenegger, M. Chli and R. Siegwart, “BRISK: Binary robust invariant scalable keypoints,” *International Conference on Computer Vision*, pp. 2548-2555, November 2011.
- [40] A. Alahi, R. Ortiz, and P. Vandergheynst, “FREAK: Fast retina keypoint,” *International Conference on Computer Vision and Pattern Recognition*, pp. 510-517, June 2012.
- [41] M. Fischler and R. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated artography,” *Magazine Communications of the ACM*, vol. 24, no. 6, pp. 381-395, June 1981.
- [42] A. Litivin, J. Konarad, and W. Karl, “Probabilistic video stabilization using kalman filtering and mosaicing,” in *Proceedings of the SPIE Image Video Commun. Image and Video Communications and Processing 2003*, vol. 5022, pp. 663-674, May 2003.
- [43] J. Kopf, M. Cohen, and R. Szeliski, “First-person hyper-lapse videos,” *ACM Transactions on Graphics (TOG)*, vol. 33, no. 4, pp. 78:1-10, July 2014.
- [44] N. Joshi, W. Kienzle, M. Toelle, M. Uyttendaele, and M. Cohen, “Real-time hyperlapse creation via optimal frame selection,” *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 63:1-9, August 2015.
- [45] G. Chunxian, Z. Zhe, L. Hui, “Hybrid video stabilization for mobile vehicle

detection on SURF in aerial surveillance,” *Discrete Dynamics in Nature and Society*, vol. 2015, 2015.

국문 초록

중앙대학교 첨단영상대학원
영상학과 영상공학 디지털이미징 전공
전 세 미

카메라 기기의 높은 휴대성으로 비 전문가의 비디오 촬영이 증가함에 따라 고품질의 비디오를 얻기 위한 디지털적 안정화 기술이 발달하고 있다. 제안하는 알고리즘은 카메라에서 얻어진 비디오의 흔들림을 효과적으로 제거하기 위해 움직임에 기반한 영역 분할과 강건한 특징점을 이용한 적응적 카메라 패스를 추정하는 방법을 제안한다. 제안하는 방법은 4단계로 구성된다: i) 움직임 기반 영역 분할, ii) 인접한 프레임간의 파티클 키포인트 추출, iii) 적응적 카메라 패스 추정 및 스무딩, 그리고 iv) 비디오 안정화를 위한 렌더링. 결과적으로, 움직임에 기반하여 영상의 영역을 분할함으로써, 불연속적인 움직임에 의해 생기는 기하학적 왜곡을 줄일 수 있다. 또한 파티클 키포인트를 이용하여 영상 내의 평평한 영역의 중요 특징점을 재 정의함으로써, 비디오 안정화를 위한 호모그래피 추정의 정확도를 향상시킬 수 있다. 제안된 알고리즘은 다양한 카메라 기기로부터 얻은 흔들리는 비디오에 적용하여 이미지 프로세스 기술의 중요 전처리 과정으로 활용될 수 있으며, 로봇 비전, 항공영상, 운전자 보조 시스템 등에 응용될 수 있다.

키워드: 비디오 안정화, 카메라 패스 추정, 움직임 추정, 특징점, 호모그래피

Abstract

Semi Jeon

The Graduate School of Advanced Imaging Science, Multimedia & Film
Chung-Ang University

Acquisition of high quality video without jitters becomes more important with the advancement of compact, portable camera with easy handling. This paper presents the video stabilization method using motion-based segmentation and particle keypoints update to adaptively estimate a camera path. The proposed method consists of four steps: i) motion-based segmentation, ii) particle keypoints extraction between adjacent frames, iii) adaptive camera path estimation and smoothing, and iv) rendering for stabilized video reconstruction. As a result, the proposed video stabilization algorithm can efficiently geometrical distortion caused by discontinuity of the motion between dynamic objects and static background. In addition, this paper redefine important feature points in the flat region using particle points and we can optimally estimate the homography for video stabilization. The proposed video stabilization method can significantly improve the visual quality for various video acquisition process and can be applied to vision-based mobile robots, aerial video, and driver assistant systems.

Keywords: video stabilization, camera path estimation, motion estimation, feature extraction, homography