DOI: 10.3745/KIPSTB.2010.17B.5.333

크기 및 회전 불변 특징점을 이용한 파노라마 영상 합성 알고리즘

권 기 원[†] · 이 해 연[††] · 오 득 환[†††]

요 약

본 논문은 동일한 대상물을 촬영한 영상을 합성하여 파노라마 영상을 생성하는 방법에 대하여 설명한다. 디지털 카메라의 보급으로 파노라마 영상에 대한 관심이 높아지면서 다양한 방법의 파노라마 영상의 제작 방법이 연구되고 있다. 본 논문에서는 크기 및 회전 불변 특징점을 활용하여 파노라마 영상을 합성하는 방법에 대해서 제안한다. 먼저, 입력 영상들에 대해서 특징점을 추출하고, RANSAC 알고리즘을 통해 추출된 특징점을 정합한다. 정합점을 이용하여 투영 변환식을 모델링하고, 모델링된 변환식을 통하여 영상을 정렬하여 파노라마 영상을 생성한다. 제안한 알고리즘은 SURF 특징점 추출 알고리즘을 적용하여 영상의 크기 및 회전 등의 기하학적 변형에 강인하며, 처리 속도도 향상하였다. 실험에서는 기존 Harris corner 검출기나 SIFT 알고리즘을 통해 검출한 특징과 제안한 알고리즘에서 사용된 SURF 알고리즘을 비교 분석 하였고, 640×480 크기의 영상을 이용하여 제안한 알고리즘을 통해 파노라마 영상을 합성하였다. 그 결과 파노라마 영상의 합성에 소요되는 시간은 평균 0.4초로 나타났고, 기존 알고리즘에 비하여 효율적인 것으로 나타났다.

키워드 : 파노라마 영상, 크기 및 회전 불변 특징, SURF, 특징 매칭, 원근 투영

Panoramic Image Composition Algorithm through Scaling and Rotation Invariant Features

Ki-Won Kwon[†] · Hae-Yeoun Lee[††] · Dukhwan Oh[†††]

ABSTRACT

This paper addresses the way to compose paronamic images from images taken the same objects. With the spread of digital camera, the panoramic image has been studied to generate with its interest. In this paper, we propose a panoramic image generation method using scaling and rotation invariant features. First, feature points are extracted from input images and matched with a RANSAC algorithm. Then, after the perspective model is estimated, the input image is registered with this model. Since the SURF feature extraction algorithm is adapted, the proposed method is robust against geometric distortions such as scaling and rotation. Also, the improvement of computational cost is achieved. In the experiment, the SURF feature in the proposed method is compared with features from Harris corner detector or the SIFT algorithm. The proposed method is tested by generating panoramic images using 640×480 images. Results show that it takes 0.4 second in average for computation and is more efficient than other schemes.

Keywords : Panoramic Image, Scale and Rotate Invariant Feature, SURF, Feature Matching, Perspective Projection

1. 서 론

컴퓨터 기술이 발전함에 따라 가상 현실 및 3D 기술에 대한 관심이 높아지고 있다. 가상 현실 기술은 사용자에게 시각적으로 풍부한 원격 현장감을 주거나 가상 체험을 할 수 있도록 해주는 목적으로 이용되기 때문에, 가상 현실에서의 가상세계는 높은 현실감을 사용자에게 제공할 수 있어야 한다. 특히 가상 공간에서의 배경은 현실감의 중요한 요소로, 실세계 환경으로부터 만드는 방법과 3차원 공간을 바탕으로 컴퓨터 그래픽스에 의해 제작된 두 가지의 형태가 있다. 컴퓨터 그래픽스를 통해 만드는 가상 환경은 복잡한 모델링 작업과 실시간으로 가상 환경을 만들어내기 위한 특별한 하드웨어가 필요하고, 영상의 복잡도가 제한되어 있어 현실감이 있는 가상 환경을 표현하기가 힘들다. 반면, 실세

※ This paper was supported by Research Fund, Kumoh National Institute of Technology.
† 준 회 원 : 국립금오공과대학교 컴퓨터공학부 4학년
†† 정 회 원 : 국립금오공과대학교 컴퓨터공학부 교수
††† 정 회 원 : 국립금오공과대학교 컴퓨터공학부 교수(교신저자)
논문접수 : 2010년 3월 9일
수 정 일 : 2010년 5월 31일
심사완료 : 2010년 6월 7일

계 환경으로부터 만드는 가상 환경은 저렴한 하드웨어의 사용으로도 쉽게 표현이 가능하다.

실세계 환경으로부터 표현된 가상 환경은 사용자 주변의 여러 방향에 대하여 영상을 촬영한 후 사용자가 주위의 모든 방향을 바라볼 수 있도록 하기 위해 단일 영상들을 결합하여 합성한 파노라마 영상을 사용함으로써 얻을 수 있다. 이러한 파노라마 영상은 전체 경치 중에서 360° 방향의 모든 경치를 담아내는 환경, 180° 방향의 시야를 담아내는 전경 등으로 말하기도 한다[1]. 이러한 파노라마 영상은 흔히 볼 수 있는 광고, 의료 영상, 지도·내비게이션 서비스 등의 다양한 용도로 활용되고 있다. (그림 1)에는 파노라마 영상을 이용하여 서비스를 제공하고 있는 예를 나타내었다. Google의 Street View나 Daum의 Road View는 지도와 파노라마 영상을 합쳐서 사용자들에게 더 편리하게 정보를 제공해 줄 수 있다. Yellow Bird 카메라의 경우는 동영상을 이용한 파노라마 합성을 통해 재생과 동시에 파노라마로 볼 수 있도록 해준다.

파노라마 영상은 하드웨어적으로 파노라마 기능을 지원하는 카메라를 사용하거나 일반카메라를 이용하여 여러 장의 사진을 촬영한 후에 합성하여 생성할 수 있다. 파노라마용 카메라는 특별한 어안 렌즈와 같은 175° 이상의 넓은 시야각을 가지는 렌즈를 사용하여 광각도를 촬영할 수 있도록 만들어지거나, 긴 한조각의 필름에 영상을 기록할 수 있도록 만들어진다. 그러나 설계 및 기술적인 난이도로 인하여 카메라의 가격이 고가이며 범용성이 떨어진다. 일반카메라의 경우 360° 방향의 모든 장면을 한 번에 촬영할 수 없기 때문에 여러 장의 사진으로 나누어 촬영 후 복잡한 알고리즘을 통해 합성한다. 일반카메라를 활용한 파노라마 영상의 생성 방법은 direct-based 방법과 feature-based 방법의 두 가지로 분류할 수 있다. Direct-based 방법은 픽셀들 사이의 정합을 통한 방법으로 복잡한 계산을 요구한다. 그러나 feature-based 방법은 추출된 특징들 사이의 정합을 통한 방법으로 신속하고 신뢰성 있는 파노라마 영상을 획득 할 수 있다.

본 논문에서는 크기 및 회전 불변 특징을 활용한 파노라마 영상의 생성 알고리즘을 제안한다. 먼저, 입력 영상들에 대해서 SURF(Speed-Up Robust Features) 알고리즘을 통해 특징점을 추출하고, 추출된 특징점을 정합하여 원근 변환식을 모델링한다. 모델링된 변환식을 통하여 영상을 정렬하여 파노라마 영상을 생성한다. 제안하는 방법은 feature-based 방법에 해당되며, 신속하고 신뢰성 있는 특징의 추출을 위하여 SURF 알고리즘을 채택하였다.

기존의 Harris corner를 사용한 방법에서는 영상의 크기가 변하거나 회전된 영상이 입력될 경우 특징을 잘 추출하지 못하였다. 하지만 SURF나 SIFT(Scale Invariant Feature Transform) 알고리즘을 사용하면 영상의 크기 변화와 회전에 강인한 특징을 추출 할 수 있다. 특히 SURF 알고리즘은 SIFT 알고리즘의 속도를 개선하였고, 흔들림에도 강하다. 보통 사진을 찍을 경우 삼각대나 모노포드 등을 사용하지 않고 손에 들고 찍기 때문에 손떨림이나 기타 파노라마 영상 합성에 방해되는 요소가 생길 경우가 많지만 알고리즘적으로 해결 가능하기 때문에 파노라마 영상을 생성하는데 SURF 알고리즘이 더 효율적이다.

본 논문의 구성은 다음과 같다. 2장에서는 기존 파노라마 영상의 합성을 위한 방법들에 대해 소개하고, 3장에서는 크기 및 회전 불변 특징을 이용하는 파노라마 영상 합성 알고리즘을 제안한다. 4장에서는 특징 추출 알고리즘의 성능을 비교 분석한 후에 제안한 알고리즘으로 합성한 파노라마 영상을 제시하고, 5장에서는 결론을 짓도록 하겠다.

2. 관련연구

동일한 대상물을 촬영한 영상의 합성을 통한 파노라마 영상 생성 방법은 direct-based 방법과 feature-based 방법으로 분류할 수 있다.

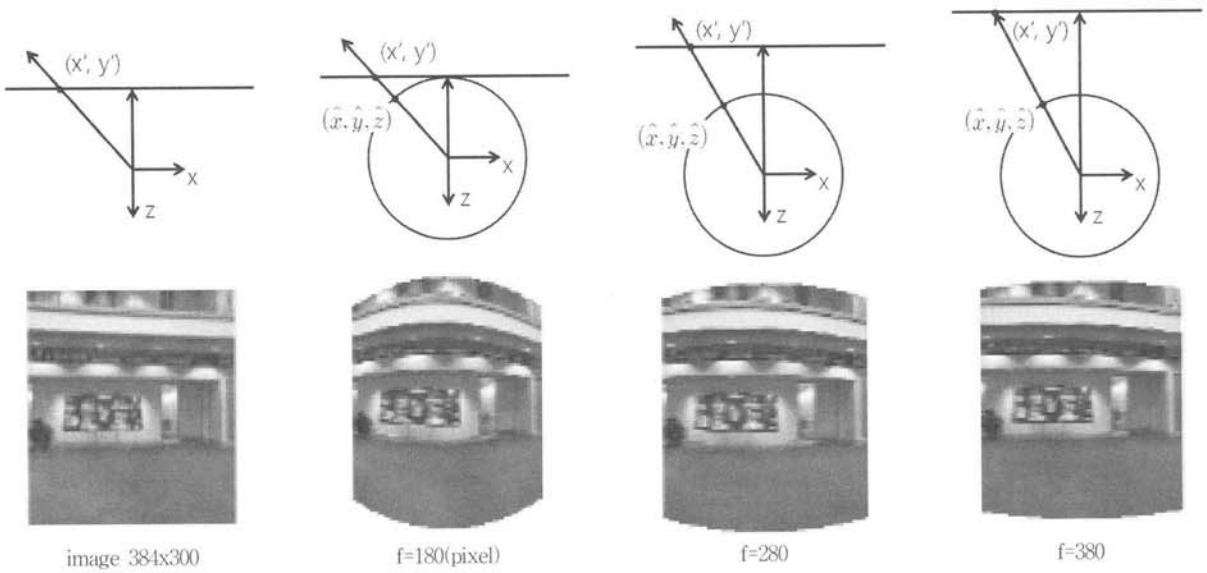
2.1 Direct-based 방법을 통한 파노라마 영상 합성

Direct-based 방법은 영상의 픽셀들 사이의 정합을 통한 파노라마 영상 합성 방법이다[2, 3]. Szeliski는 고정된 시점에서 연속적인 영상을 촬영한 후에, 촬영된 영상을 3차원 원통형 좌표에 맞추어 2차원 영상으로 변환하고, hierarchical lucas-kanade 알고리즘을 사용하여 영상을 정렬하여 파노라마 영상을 생성하였다[2]. Szeliski의 알고리즘에서는 (그림 2)와 같이 렌즈 초점거리에 따라 영상의 왜곡 정도가 변화되므로, 영상 취득에 사용하는 카메라에 맞는 올바른 초점거리값을 알아야 한다.

Suen et al.은 direct-based 방법의 처리 속도를 개선한 파노라마 영상 합성 알고리즘을 제안하였다[3]. 색상과 위치

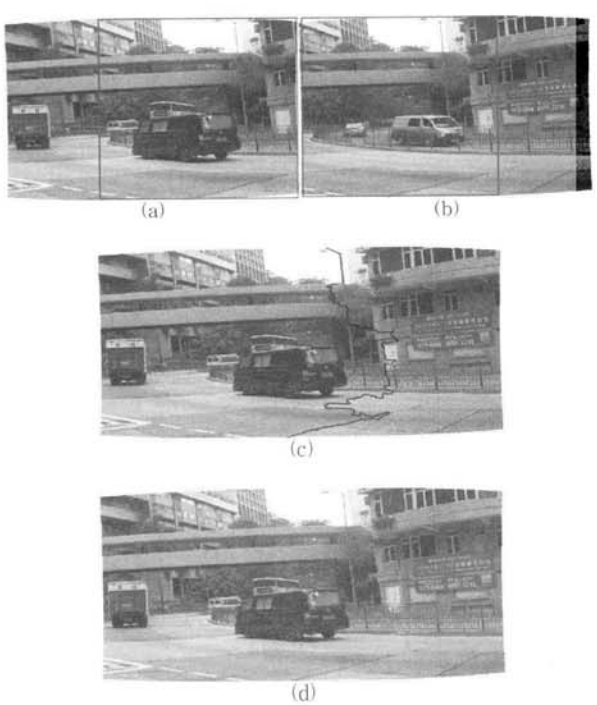
(a) Google Street View (b) Daum Road View (c) Yellow Bird 서비스

(그림 1) 파노라마 영상을 활용한 서비스 예



(그림 2) Szeliski 알고리즘의 초점 거리값에 따른 영상 왜곡

가 불일치하는 영상에 대해 합성이 가능하고, 움직이는 물체 등에 의해서 영상의 연결이 절단되지 않도록 방지하기 위해서 커브 강도와 밝기값 변화율의 일관성을 활용한다. 또한 합성 결과물의 밝기값 변화 등을 조절해서 합성된 영상의 절개 부분의 픽셀값 불연속성을 개선했으며, 다중 해상도 기법을 통해서 원본 영상의 크기에 독립적으로 처리 시간을 단축하였다. Suen et al.의 영상 합성의 과정은 (그림 3)과 같다. (그림 3a)와 (그림 3b)는 합성을 위해 입력된 영상으로 서로 겹쳐지는 부분은 빨간색 사각형 영역이다. (그림 3c)는 각각의 영상에서 겹치는 부분에서 커브 강도에 따라 잘라내고 영상을 붙인 것이다. (그림 3d)는 합성한 후



(그림 3) Suen et al.의 파노라마 영상 합성 과정

에 영상들의 경계선을 블렌딩 처리한 결과로 입력 영상의 중간값을 계산해서 서로 색상을 혼합하여 어색한 경계선을 제거하게 된다.

이와 같이 direct-based 방법을 사용하여 파노라마 영상을 합성하는 알고리즘은 쉽게 구현이 가능하고, 합성된 파노라마 영상의 품질도 적절하지만, 촬영되는 영상의 수평이나 수직이 고정되어야 하며, 정해진 순서에 따라서 영상이 입력되어야 한다. 따라서 잘못된 영상이 입력되거나 영상의 기하학적인 왜곡이 존재할 경우 파노라마 합성이 되지 않거나 성능이 급격히 저하된다. 또한 모든 픽셀들을 고려하기 때문에 잘못된 밝기 항상성으로 인하여 전체 알고리즘 수행 결과가 다를 수 있으며, 매번 반복해서 접근할 때마다 초기화가 필요하므로 속도가 느리다. 특히, 입력되는 영상의 조명 차이나 잡음에 약하다는 단점이 있다.

2.2 Feature-based 방법을 이용한 파노라마 영상 합성

Feature-based 방법은 direct-based 방법과는 달리 픽셀 전체를 고려하지 않고, 입력된 영상에서 특징점를 추출한 후에, 추출된 특징을 기하학적인 변형에 대한 모델의 값을 추정하는 파라미터로 활용하여 파노라마 영상을 생성하는 방법이다[4-6]. Direct-based 방법과 차이는 여러 개의 영상을 놓고, 각각의 영상에서 특징을 추출하고 이 특징을 비교해서 영상 사이에서 연관성이 있는지 식별해서 파노라마를 생성한다는 점이다. 특징을 비교해서 분석하기 때문에 연관이 없는 영상에 대해서는 오류로 처리를 할 수 있게 된다.

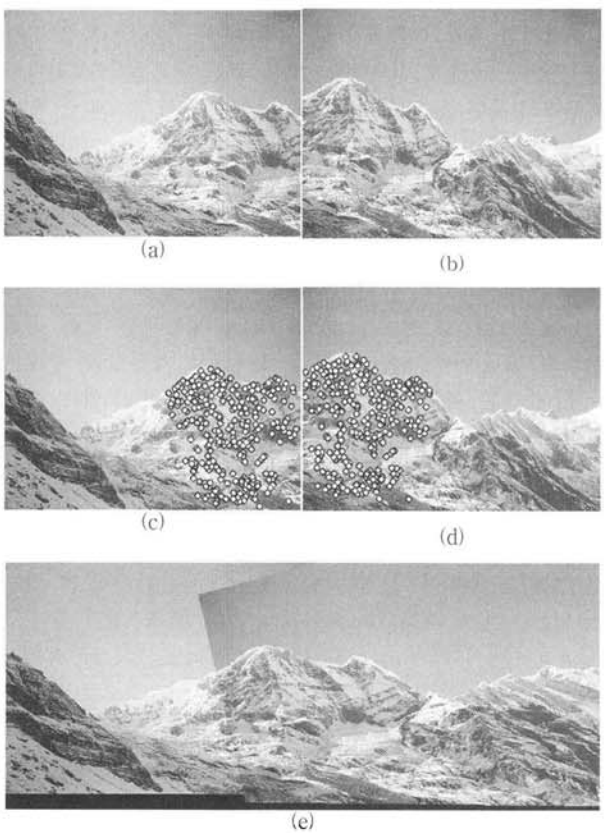
영상처리 및 컴퓨터 비젼 분야에서 Harris corner 검출기는 신뢰성 있는 특징을 검출하기 위하여 범용적으로 활용된다. 그러나 Harris corner 검출기는 영상의 크기 및 회전에 불변하는 특징을 추출할 수가 없다. 파노라마 영상의 합성을 위해서는 입력되는 영상의 경우 크기나 회전 등의 영향으로 인하여 왜곡이 발생하므로, Harris corner 검출기의 특징을 활용하는 것은 파노라마 영상 생성을 위한 알고리즘의

성능 저하를 야기할 수 있다.

Brown and Lowe는 SIFT(Scale Invariant Feature Transform) 특징 추출을 이용한 파노라마 영상 합성 방법을 제안하였다[4, 5, 7]. (그림 4)는 Brown and Lowe의 파노라마 영상의 합성 과정을 나타낸다.

입력되는 영상에 대하여 SIFT 특징을 추출하고(그림 4c 및 4d 참조), 추출된 특징들을 이용해서 가장 인접한 특징점끼리 정합한다. 이때 n개의 특징을 정합하려면 $O(n^2)$의 복잡도를 갖게 되는데, nearest-neighbor 알고리즘을 통하여 O(nlogn)의 복잡도를 갖도록 성능을 향상하였다[8]. 특징들 사이의 정합관계를 수립한 후에, bundle adjustment를 통해서 기하학적인 변환모델의 파라미터를 추정하고, 영상을 정렬하여 변환함으로써 (그림 4e)와 같이 파노라마 영상을 생성하였다. 사용된 SIFT 알고리즘은 입력된 영상의 크기나 회전에 불변하는 특징을 추출할 수 있기 때문에, 이 특징을 이용함으로써 신뢰성 있는 파노라마 영상을 생성할 수 있다. 그러나 SIFT 특징은 계산 복잡도가 높아서 처리 속도가 느린 단점이 있다.

Brown et al. 및 Byrod et al.은 Brown and Lowe가 제안한 알고리즘을 개선한 feature-based 방법을 이용한 파노라마 영상 생성 알고리즘을 제안하였다[6, 9]. SIFT 알고리즘을 사용하여 특징점을 추출하고, 특징 정합 및 bundle adjustment를 통해 기하학적 변환 파라미터를 추정한 후 파노라마 영상을 생성하였다. Brown and Lowe의 알고리즘과



(그림 4) Brown and Lowe의 SIFT 특징을 이용한 파노라마 영상 합성 과정

의 차이점은 추출된 특징점을 이용하여 특징 정합 과정에서 outliers과 inliers을 구별해서 제거해야 하는데, 이전 방식에서는 RANSAC 알고리즘을 사용하였다[10]. 그러나 Brown et al. 방법에서는 MLESAC 알고리즘을 사용하였다[11]. MLESAC은 MLE(maximum likelihood estimation)을 사용해서 주어진 특징점들에서 inliers를 선택할 확률을 최대로 만들어 inliers 추정에 걸리는 시간을 단축한 RANSAC 알고리즘의 파생 알고리즘이다.

이와 같이 feature-based 방법을 이용한 파노라마 영상의 생성 알고리즘은 잘못된 영상이 들어오더라도 추출된 특징의 비교를 통하여 제외 할 수 있기 때문에 자동적으로 파노라마 영상 합성하는 방법으로 적절하다고 할 수 있다. 그러나 입력되는 영상의 크기와 회전 등의 왜곡에 강인하도록 특징을 추출하는 것이 파노라마 영상 생성 알고리즘의 성능과 관련된다. 설명된 Brown and Lowe나 Brown et al. 알고리즘은 특징 추출을 위해 SIFT 알고리즘을 사용하기 때문에, 특징 추출의 결과가 전체 성능을 좌우한다. SIFT 알고리즘이 특징 추출에 있어서 우수한 결과를 나타내지만, 알고리즘의 복잡도가 높아서, 계산량이 많고, 그로 인하여 특징 추출의 속도가 느리다. 따라서 효율적인 파노라마 영상을 생성하기 위해서는 크기 및 회전 변환에 강인하며, 처리 속도가 향상된 특징 추출 방법이 필수적이다.

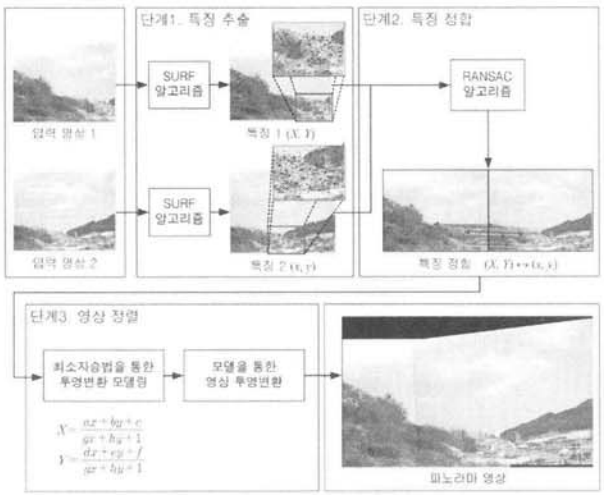
3. 크기 및 회전 불변 특징을 이용한 파노라마 영상 생성 알고리즘

본 절에서는 크기 및 회전 변환에 강인하며, 처리 속도가 향상된 특징 추출 방법을 이용한 파노라마 영상 생성 알고리즘에 대해서 제안한다. 파노라마 영상의 생성을 위해서 2장에서 설명한 것과 같이 일반적으로 SIFT 알고리즘을 많이 사용하고 있다[5, 7]. 하지만, 성능 상의 제약으로 인하여, 본 논문에서는 SURF(Speeded Up Robust Features) 알고리즘을 이용하여 특징 추출의 신뢰성 및 처리 성능을 향상하였다[12]. SURF 알고리즘은 SIFT 알고리즘에 비해서 빠르고 특징점 추출이 정확하기 때문에 파노라마 합성에 더 적합하다.

본 논문에서 제안하는 크기 및 회전 변환에 강인한 특징점을 이용한 파노라마 영상의 합성 과정은 (그림 5)에 도시하였다. 제안하는 알고리즘은 (1) 특징추출, (2) 특징정합, (3) 영상정렬의 3과정으로 구성되어 있으며, 각 과정에 대하여 다음 절에서 자세히 설명하겠다.

3.1 크기 및 회전 불변 특징 추출

입력 영상의 크기 및 회전 등의 기하학적 왜곡에 강인한 특징 추출은 파노라마 영상 생성을 위하여 필수적이다. 기존 논문에서 SIFT 알고리즘을 이용한 방법들은 정확도는 높지만, 대체로 속도가 느렸다. 따라서, 제안한 알고리즘에서는 입력 영상에서 강인하고, 계산비용이 적은 특징점을 추



(그림 5) 제안하는 파노라마 영상 합성 과정

출하기 위해 SURF 알고리즘을 적용하여 효율적인 파노라마 영상의 합성을 모색하였다.

Bay et al.에 의해 제안된 SURF 알고리즘은 입력된 영상의 특징 추출, 특징의 기준 방향과 크기 설정, 디스크립터 생성의 순서로 작업이 이루어진다[12]. SIFT 알고리즘과 다르게 SURF 알고리즘에서는 적분(integral) 영상을 이용하여 밝기값 변화율 계산 속도를 향상하였다. <표 1>에 1024×768 해상도의 영상에 대한 SURF 알고리즘과 SIFT알고리즘에 대한 성능을 정리하였다. SURF 알고리즘은 SIFT 알고리즘에 비하여 추출된 특징점의 수는 적지만, 수행속도 및 메모리 활용에 있어서, 50% 이상 향상된 것을 알 수 있다[12].

SURF 알고리즘은 먼저 입력된 영상을 적분 영상을 생성하게 되는데 이 적분 영상을 생성하게 되면 이후 이 영상을 이용하면 어떤 크기의 영상 영역을 선택하더라도 4번의 연산을 통해 선택 영역의 모든 픽셀의 합을 구할 수 있다. 적분 영상은 다음 식(1)을 통해 구할 수 있다. 주어진 영상 I의 (x,y)좌표를 이용해서 적분영상 I_{Σ}를 구할 수 있다.

$$I_{\Sigma}(X) = \sum_{i=0}^{i<xi<y} \sum_{j=0}^{} I(i,j) \tag{1}$$

SURF 알고리즘에서는 영상의 특징점을 추출하기 위하여 Hessian 행렬에 기반하는 검출기를 사용한다. Hessian 행렬은 다음 식(2)과 같이 표현되며, X는 영상에서 점의 좌표 (x, y)이며, σ는 적분영상의 스케일 값을 의미한다. $L_{xx}(X,\sigma), L_{xy}(X,\sigma), L_{yy}(X,\sigma)$는 x, y, xy 방향으로의 가우시안 2차 미분함수를 통한 Filter와 적분영상과의 콘볼루션을 통해서 구할 수 있다.

〈표 1〉 SURF 알고리즘과 SIFT 알고리즘의 성능 비교

	SURF 알고리즘	SIFT 알고리즘
메모리 요구량	64 bytes	128 bytes
계산 속도	2.4 sec	6 sec
특징점 수	~1000	> 3000

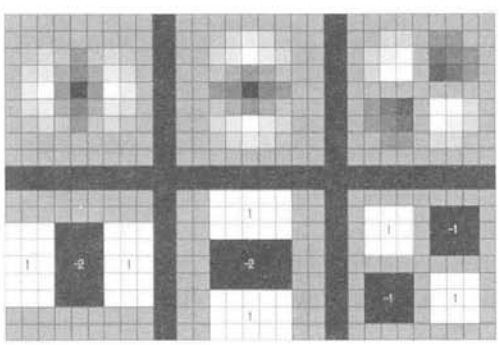
$$H_{(X,\sigma)} = \begin{bmatrix} L_{xx}(X,\sigma)L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma)L_{yy}(X,\sigma) \end{bmatrix} \tag{2}$$

Hessian 행렬을 사용할 때, SIFT와 다르게 가우시안 필터를 사용하지 않고, 결과는 유사하지만, 연산 속도가 빠른 박스 필터를 사용하여 결정자를 구한다. (그림 6)에는 SIFT 알고리즘에서 사용되는 가우시안 필터의 모습과 이에 대응하는 SURF 알고리즘에서의 박스 필터를 나타낸다[12]. 여기서 박스 필터의 검은 부분은 -2, 흰 부분은 1의 가중치에 해당한다. 또한 필터의 가중치 합은 0에 해당한다.

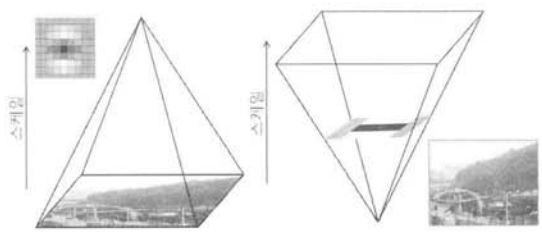
가우시안은 실수값을 갖지만, 박스필터는 정수형 연산이 가능하므로 속도의 향상이 가능하다. 필터의 적용 대상은 원본 입력 영상이 아닌 적분 영상이며, 크기에 불변한 특징을 얻기 위해서 크기 변화가 다양한 영상을 이용하지 않고 박스 필터의 크기를 점점 증가시키며 적분 영상에 적용하게 된다. x, y, xy 방향을 가우시안 필터 결과를 Lxx, Lyy, Lxy 라고 하면, 대응되는 박스 필터의 x, y, xy 방향 결과를 Dxx, Dyy, Dxy 라고 하자. 이를 이용해서 Hessian 행렬의 결정자는 다음 식(3)을 통해서 계산할 수 있다. 계산된 값 중에서 임계치를 초과하는 점이 후보점에 해당한다.

$$\det(H_{\approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \tag{3}$$

SIFT 알고리즘 및 SURF 알고리즘에서 크기 불변 특징의 추출을 위해서 스케일 공간에서 영상을 처리한다. SIFT 알고리즘에서는 스케일 공간은 (그림 7a)와 같이 이미지 피라미드로 표현되며, 가우시안 필터를 반복적으로 영상에 적용함으로서 다양한 스케일의 영상을 생성한다. 그러나 SURF 알고리즘에서는 적분 영상을 이용해서 원래 영상은



(그림 6) 가우시안 x, y, xy 방향 필터(상단)와 대응되는 x, y, xy 방향 박스필터(하단)



(그림 7) 필터 피라미드의 예 (좌측 SIFT 방법, 우측 SURF 방법)

변하지 않고, (그림 7b)와 같이 박스 필터의 크기를 변경하면서 스케일 공간을 생성한다[12]. 이와 같이 스케일 공간에서 반복적으로 검출된 후보점이 크기에 불변하는 특징점으로 선정된다.

SURF 알고리즘에서 디스크립터는 필터 크기가 2인 Haar wavelet 결과를 통하여 계산된 64차원의 불변하는 정보로 표현된다. 디스크립터 계산을 위해 먼저 추출된 후보점에 대하여 6σ 반경 내의 x 방향 Haar wavelet 응답과 y 방향 Haar wavelet 응답을 더하여 벡터를 구함으로 주 방향을 계산한다. 주 방향을 기준으로 4x4 크기의 세부 영역으로 나누고, 각 영역마다 5x5 크기의 x 방향 및 y 방향 Haar wavelet을 적용시켜 dx, dy를 계산한다. dx, dy를 사용하여 sum(dx), sum(dy), sum($|dx|$), sum($|dy|$)의 값을 계산함으로서 각 세부 영역마다 4개의 특징 값을 구할 수 있다. 총 16개 영역에 대해여 4개의 특징 값을 계산할 수 있으므로, 총 64차원을 갖는 디스크립터를 얻을 수 있다[12]. 이와 같은 방법을 통하여 SURF 알고리즘을 통하여 추출한 특징점들은 회전, 크기 변화, 영상의 밝기, 대비에 강한 특징을 갖게 된다.

3.2 연관 특징 정합

입력 영상에 대하여 SURF 알고리즘을 통해 추출한 특징점들을 RANSAC알고리즘을 이용해서 inliers과 outliers을 구별해서 outliers들을 제거하여 연관 특징을 정합한다[10]. 일반적으로 사용되는 최소자승법의 경우 inliers와 outliers를 모두 고려해서 오차가 최소가 되도록 구하는 접근법이다. 그러나 입력되는 데이터에 오차가 많을 경우 추정 결과의 신뢰성에 심각한 저하가 발생한다. 일반적으로 입력되는 영상에 대하여 특징을 추출하면, 영상의 잡음이나 촬영한 대상물의 차이로 인하여 잡음의 역할을 하는 특징들이 다수 검출되므로 최소자승법을 통한 방법은 오차 및 성능의 저하가 발생한다. 반면에 RANSAC 알고리즘은 inliers와 outliers를 구분지어 inliers 만 선정하여 수학적 모델을 추정할 수 있는 알고리즘이다. (그림 8)의 점들을 보게 되면, 직선을 추출해 내는데 파란점이 inliers, 빨간점이 outliers라고 한다면 (그림 8)의 inliers, outliers가 혼재되어 있기 때문에 outliers를 모두 고려해서 inliers를 구별하는 최소자승법으로는 위와 같은 직선을 추출해 내기 어렵다. 하지만 RANSAC 알고리즘에서는 inliers만 선정하여 활용함으로써 직선을 쉽게 추정할 수가 있다.

(그림 8) RANSAC 알고리즘을 통한 직선 추정 예

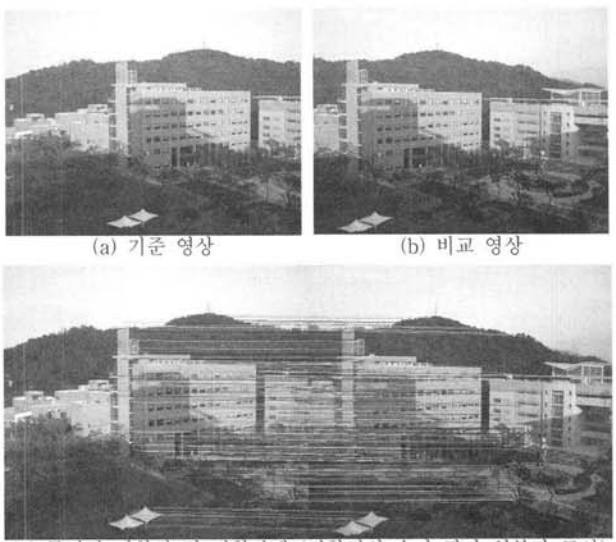
RANSAC 알고리즘의 수행 과정은 다음과 같다. 입력된 특징점에서 N개의 서브데이터를 선택하고, 이 데이터를 가상의 inliers로 생각하고 모델을 예측한다. 선택된 특징점에서 예측한 모델이 잘 맞는지 체크하고 이때 잘 맞는 데이터의 수를 K라고 하자. 만약에 K가 충분하면 이 모델을 확정하고 종료하지만, 모델이 좋지 않으면 계속해서 L번 반복하게 된다. 이때 잘 맞는 데이터의 수 K는 사용자가 지정한 수치로 본 논문에서는 20으로 설정하였다.

최대 반복횟수 L은 특징점에서 뽑은 한 개의 서브데이터가 좋은 inliers 일 확률이 P_g인 경우에 L번의 모델 예측시도가 실패할 확률을 계산해서 이것이 주어진 설정값 P_{fail} 보다 작은경우 모델 예측 실패로 처리하므로 다음 식(4)와 같이 정리할 수 있다.

$$L=\frac{\log(P_{fail})}{\log(1-(P_g)^N)} \tag{4}$$

$$\begin{aligned} P_{fail} &= L\text{번의 모델 예측 실패 확률} \\ &= (\text{한번 시도가 실패할 확률})^L \\ &= (1-\text{한번 시도가 성공할 확률}))^L \\ &= (1-(\text{랜덤 데이터가 모델을 예측할 확률})^N)^L \\ &= (1-(P_g)^N)^L \end{aligned}$$

SURF 알고리즘을 특징점을 추출하고 RANSAC 알고리즘을 통한 특징 정합을 통한 결과를 (그림 9)에 나타내었다. 각 입력 영상은 1024×768 해상도이며, 추출된 특징점 수는 (그림 9a)가 2257개 (그림 9b)가 2522개 이다. (그림 9c)는 특징점을 정합시켜 정합된 점을 서로 연결한 선을 도시하였다. 정합된 점의 수가 많아서 샘플링을 통하여 정합 관계를 가시적으로 확인할 수 있도록 정합점의 수를 조정하였다. 제안한 알고리즘을 통하여 정합된 특징점의 수는 1075개이



(a) 기준 영상 (b) 비교 영상

(c) 특징점 정합점 및 정합관계 (정합점의 수가 많아 일부만 도시)

(그림 9) 각각의 영상에 대한 특징점 추출 및 특징점 정합 예제

며, 약 48%의 정합률을 보였다. 입력된 영상의 중첩이 70% 정도인 것을 고려하면 48% 정합률은 다음 단계인 영상 정렬을 위하여 충분한 정합점 수에 해당된다.

3.3 영상 정렬

파노라마 영상의 생성을 위한 마지막 과정은 정합된 특징점을 사용하여 homography 관계를 모델링하고, 이를 이용하여 각 영상을 원근 변환하고 영상을 정렬하는 단계이다. 2차원 영상에서 기하학적인 변형에 대한 연산의 효율성을 위하여 3차원 동차좌표계를 활용한다. 2차원 기하학적 변환인 이동, 크기변환, 회전에 대한 3×3 행렬로 표현되는 동차좌표계에서의 변환은 식(5), 식(6), 식(7)로 표현된다. t_x, t_y는 x, y 축으로의 변이를 나타내며, s_x, s_y는 x, y 축으로 스케일링 상수값이며, θ는 회전 각도를 나타낸다.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{6}$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{7}$$

상기의 2차원 기하학적인 변환에 대한 3차원 동차좌표계를 활용하여 파노라마 영상의 생성을 위한 homography를 조합을 통하여 투영 변환을 모델링한다. 투영 변환은 식(8)과 같이 모델링된다[13]. 식에서 행렬의 원소 a, b, c, d, e, f, g, h는 이동, 크기변환, 회전에 대한 행렬에서의 요소들의 곱으로 표현된다.

$$\begin{bmatrix} XW \\ YW \\ W \end{bmatrix} = H \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{8}$$

(X, Y)는 기준 영상 좌표계에서의 좌표값이고, (x, y)는 비교 영상 좌표계에서의 좌표값이며, H는 변환행렬이다. 여기서 $W = gx + hy + 1$ 이므로 상기의 식은 식(9)과 같이 표현할 수 있다.

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}{\begin{bmatrix} g & h & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}} \tag{9}$$

식(9)을 원근 변환의 non-vector 형에 해당하도록 식을 재구성하면 식(10)과 같다.

$$X = \frac{ax + by + c}{gx + hy + 1},\ Y = \frac{dx + ey + f}{gx + hy + 1} \tag{10}$$

이렇게 구성된 식을 투영 변환 행렬의 원소들 a, b, c, d, e, f, g, h 에 대하여 정리하면 식(11)와 같이 투영 변환을 위한 모델을 수립할 수 있다.

$$\begin{aligned} xa + yb + 1c + 0d + 0e + 0f - xXg - yXh = X \\ 0a + 0b + 0c + xd + ye + 1f - xYg - yYh = Y \end{aligned} \tag{11}$$

연관 특징 추출의 과정을 통해 추출된 기준 영상의 연관 특징점 (X, Y)와 비교 영상의 연관 특징점 (x, y)를 사용하여 식(11)에 입력하여 수식을 식(12)과 같이 정리할 수 있고, 최소 자승법을 통하여 a, b, c, d, e, f, g, h 에 대한 해를 계산하면 투영 변환 행렬식을 계산할 수 있다. 이렇게 계산된 투영 변환 행렬식을 통하여 영상을 정렬하면 (그림 10)과 같이 두 입력 영상을 합성한 파노라마 영상을 생성할 수 있다.

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -X_1x_1 & -X_1y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -Y_1x_1 & -Y_1y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -X_2x_2 & -X_2y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -Y_2x_2 & -Y_2y_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n & y_n & 1 & 0 & 0 & 0 & -X_nx_n & -X_ny_n \\ 0 & 0 & 0 & x_n & y_n & 1 & -Y_nx_n & -Y_ny_n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \\ g \\ h \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ X_2 \\ Y_2 \\ \vdots \\ X_n \\ Y_n \end{bmatrix} \tag{12}$$

(그림 10) 투영 변환을 이용하여 합성한 파노라마 영상 결과

4. 실험 결과

본 논문에서 제안한 알고리즘을 이용하여 다양한 영상 그룹에 대하여 실험하였다. 입력된 영상의 사이즈는 모두 640x480으로 통일하였고, AMD Athlon 64 X2 Dual Core 3600+ 2.7GHz CPU와 2GB RAM 사양의 컴퓨터를 이용하였다. 먼저 제안한 알고리즘에서 사용한 특징점 추출 및 정합 성능을 분석한 후에, 파노라마 영상 생성 결과에 대해서 제시하겠다.

4.1 특징점 추출 및 정합 성능 비교

먼저 SURF 특징점 추출의 성능을 분석하기 위하여 동일

한 영상에 대하여 SURF 특징점을 추출한 후에 임의로 특징점을 추가하거나 제거하면서 정합된 특징점의 수를 계산하였고, 이를 통하여 파노라마 영상을 합성한 결과를 <표 2>에 정리하였다. 그 결과에 따르면 각각의 영상에서 추출되는 특징점의 수가 많거나 적더라도 SURF 알고리즘에서는 충분한 정합점을 추출할 수 있고 파노라마 영상의 생성에 문제가 없음을 확인할 수 있다.

제안한 알고리즘에서 사용한 SURF 방법과 기존 논문에서 활용되는 Harris corner 검출기 및 SIFT 알고리즘과 특징점 추출 및 정합 성능에 대한 비교를 수행하였다. (그림 11)와 같이 동일한 대상물에 대하여 대략 15도, 45도, 90도 회전 및 크기 변화를 고려하여 촬영한 영상에 대하여 특징점 추출을 수행한 후에 영상 정합의 결과를 분석하였다. 영상 정합의 경우 0도 영상을 고정하여 0도와 15도, 0도와 45도, 0도와 90도 영상의 비교를 수행하였다. 성능에 대한 평가는 특징점 수, 정합점 수, 수행 시간이다. 특징점의 수가 많아도 정확한 정합점의 수가 많지 않다면 파노라마 영상의 생성 알고리즘의 성능 및 신뢰성이 저하가 된다.

〈표 2〉 같은 영상에서 특징점을 추가, 제거 후 합성 결과

	실험 1	실험 2	실험 3
영상1 특징점	703	503	980
영상2 특징점	586	399	870
정합점	66	59	54
결과			

(그림 11) 특징점 추출 및 정합 성능 분석을 위한 영상

<표 3>, <표 4> 및 <표 5>에는 영상셋1, 영상셋2 및 영상셋3에 대하여 Harris corner 검출기, SIFT 알고리즘과 SURF 알고리즘을 사용하여 분석한 특징점 수, 정합점 수 및 정합률에 대하여 정리를 하였고, <표 6>에는 처리속도에 대하여 정리하였다.

<표 3>, <표 4> 및 <표 5>에 정리된 것과 같이 추출된 특징점 수는 SIFT 알고리즘 > SURF 알고리즘 >> Harris corner 검출기의 순서이며, 정합점 수 및 정합률도 SIFT 알고리즘 > SURF 알고리즘 >> Harris corner 검출기의 순서이다. 추출된 특징점 수는 파노라마 영상의 합성을 위한 정합에 큰 영향은 없으며, 정합점 수 및 정합률이 파노라마 영상의 성공적인 수행을 위해서 중요하며 값이 클수록 유리

〈표 3〉 영상셋 1에 대한 특징 추출 알고리즘 성능 분석

		15도	45도	90도
Harris corner 검출기	특징점수	614	545	558
	정합점수	50	33	34
	정합률	8.1%	6.1%	6.1%
SIFT 알고리즘	특징점수	977	1082	976
	정합점수	298	280	242
	정합률	30.5%	25.9%	24.8%
SURF 알고리즘	특징점수	668	710	757
	정합점수	155	118	268
	정합률	23.2%	16.6%	35.4%

〈표 4〉 영상셋 2에 대한 특징 추출 알고리즘 성능 분석

		15도	45도	90도
Harris corner 검출기	특징점수	1021	1200	1086
	정합점수	94	79	63
	정합률	9.2%	6.6%	5.8%
SIFT 알고리즘	특징점수	1155	1318	984
	정합점수	493	426	369
	정합률	42.7%	32.3%	37.5%
SURF 알고리즘	특징점수	690	840	764
	정합점수	190	198	268
	정합률	27.5%	23.6%	35.1%

〈표 5〉 영상셋 3에 대한 특징 추출 알고리즘 성능 분석

		15도	45도	90도
Harris corner 검출기	특징점수	374	459	966
	정합점수	37	32	33
	정합률	9.9%	7.0%	3.4%
SIFT 알고리즘	특징점수	721	927	1204
	정합점수	201	175	45
	정합률	27.9%	18.9%	3.7%
SURF 알고리즘	특징점수	471	569	550
	정합점수	129	48	17
	정합률	27.4%	8.4%	3.1%

〈표 6〉 영상셋 1, 2 및 3에 대한 각 특징추출 알고리즘의 평균 수행시간

	영상셋 1	영상셋 2	영상셋 3	평균
Harris corner 검출기	3.680sec	5.047sec	3.432sec	4.053sec
SIFT 알고리즘	4.786sec	5.349sec	4.260sec	4.798sec
SURF 알고리즘	0.450sec	0.417sec	0.453sec	0.440sec

하다. <표 6>에 나타난 것과 같이 평균 수행시간은 알고리즘의 복잡도를 의미하며 SIFT 알고리즘 > Harris corner 검출기 >> SURF 알고리즘의 순서로 값이 작을수록 좋은 알고리즘이다.

Harris corner 검출기의 경우 회전 각도가 높아짐에 따라서 정합률이 떨어지며, 영상의 크기 변화가 있는 경우 정합점의 급격하게 저하되는 경우가 빈번했다. SIFT 알고리즘은 높은 정합률을 가지고 있으나 평균 수행시간이 다른 알고리즘에 비하여 오래 걸리는 문제점이 있다. 그러나 SURF 알고리즘의 경우 알고리즘의 수행시간이 짧으며 정합점의 수나 정합률도 파노라마 영상을 생성하는데 충분하다. 따라서 파노라마 영상의 생성을 위해서 수행시간이 중요한 요소이므로 SURF 알고리즘이 효율적이라고 할 수 있다.

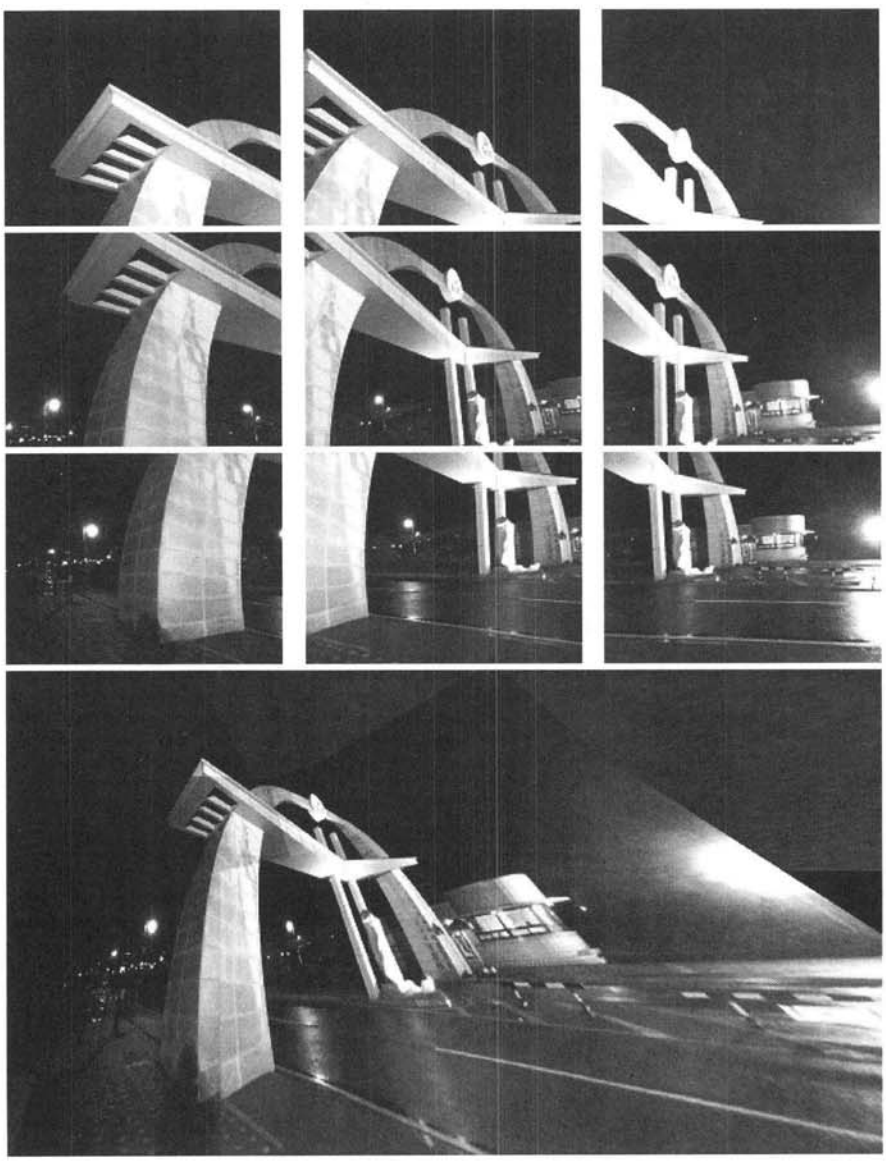
4.2 파노라마 영상 합성 분석

파노라마 영상 합성 성능을 분석하기 위하여 제안한 알고리즘을 Adobe Photoshop CS4의 Photomerge 기능과 수행속도를 비교하였다. Photomerge 기능에서 제안된 방식과 동일한 원근 투영법을 선택하였으며, 블랜딩 기능을 제외한 수행시간을 계산하였다. 영상 2장으로 구성된 10개의 그룹에 대하여 실험을 수행하였고, 평균적인 수행시간은 다음 <표 7>과 같다. 그 결과 Photoshop CS4 보다 약 15배 정도 빠른 것을 확인 할 수 있다.

(그림 12), (그림 13)에는 제안한 알고리즘을 사용하여 다양한 대상물을 포함한 9장의 사진을 촬영하고, 파노라마 영상으로 합성한 결과를 제시하였다. 사진 촬영에 사용된 카메라는 Nikon CoolPix 4500, iPhone 내장 카메라이다. (그림

〈표 7〉 영상 2장으로 구성된 10개의 그룹에 대한 평균 수행시간

	평균 수행시간
Adobe Photoshop CS4	약 6초
제안한 방법	약 0.4초



(그림 12) 야간에 iPhone 내장카메라로 촬영한 영상셋 1 및 합성된 파노라마 영상

12)는 야간에 9장의 사진을 촬영한 결과이며, (그림 13)는 주간에 9장의 사진을 촬영한 결과이다.

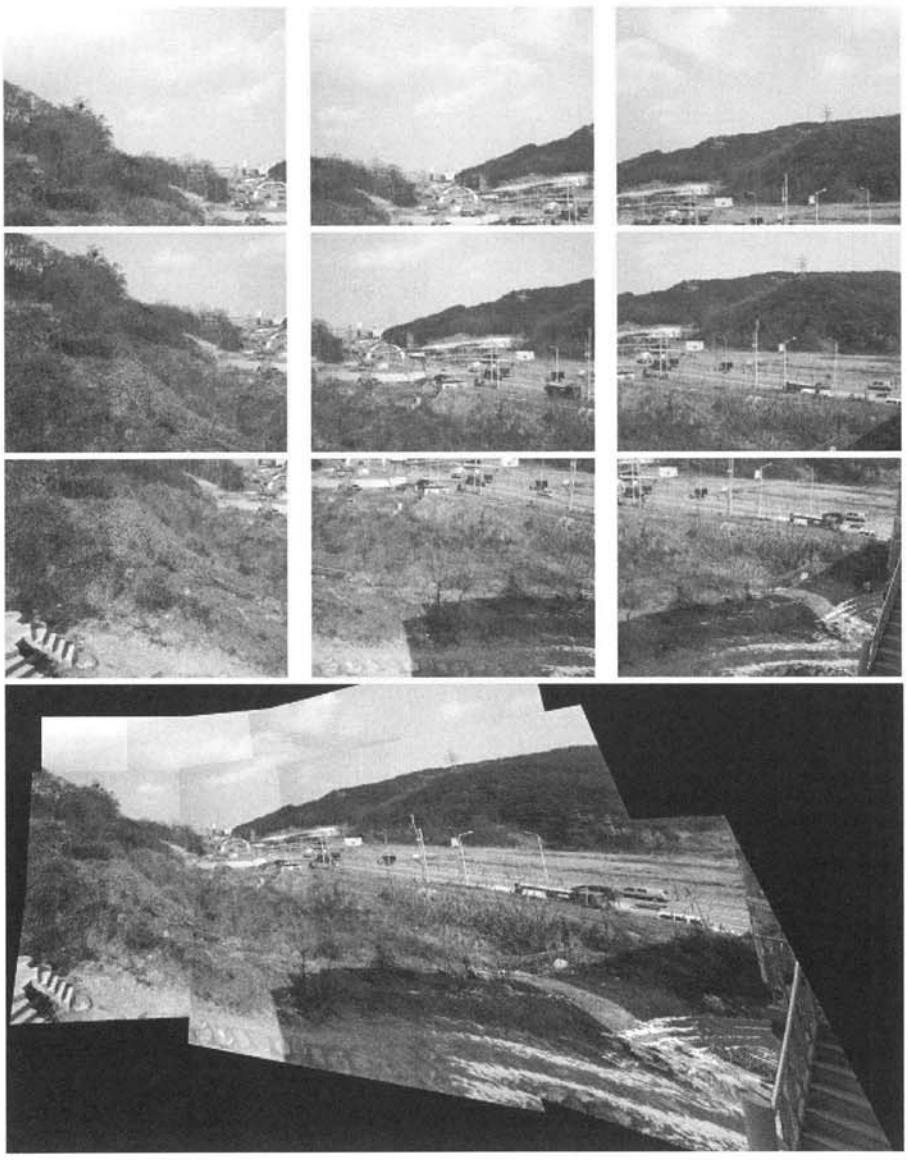
2그룹 영상셋에서 9개의 영상을 파노라마로 만드는데 소요되는 평균시간은 (그림 12)는 12초, (그림 13)는 14초가 소요되었다. 제안된 방법에서는 영상 사이의 관계를 확실하게 찾기 위해서 각각의 영상 사이의 특징점 비교를 전체에 걸쳐 진행한다. 따라서 확실히 중첩되는 영상 2장을 지정해서 입력한 경우 0.4초가 걸렸지만, 다수의 영상을 넣었을 경우 영상의 숫자만큼 시간이 늘어난다.

다양한 대상물을 촬영하여 합성을 수행하였고, 조명 등에 의한 밝기값 및 기하학적인 왜곡 등이 발생하였음에도 불구하고 합성을 올바르게 수행하였음을 확인할 수 있다.

합성된 파노라마 영상에서 영상의 기하학적인 합성은 정상적으로 이루어졌지만 영상의 경계선 부분에서는 밝기값이 부자연스러운 현상이 나타나고 있다. 이는 카메라의 자동 밝기값 조정으로 인하여 촬영된 영상에서 동일 대상물의 위치에서의 픽셀 밝기값이 다르기 때문이다. 이의 해결을 위하여 다중밴드 블랜딩 기술들을 도입할 필요가 있으며, 차후 연구를 통하여 개선할 예정에 있다.

5. 결 론

컴퓨터 기술이 발전함에 따라 가상 현실 및 3D 기술에 대한 관심이 높아지고 있고, 실세계 환경으로부터 만드는 가상 환경은 기술적인 구현의 용이하며 다양한 활용이 가능하다. 일반 카메라를 이용하여 파노라마 영상을 생성하는 경우 360° 방향의 모든 장면을 사진 한번에 촬영할 수 없기 때문에 몇 장의 사진을 나누어 촬영 후 붙여서 파노라마 영상으로 만든다. 본 논문에서는 크기 및 회전 불변 특징점을 활용하여 파노라마 영상을 합성하는 방법을 제안하였다. 입력 영상들에 대해서 특징점을 추출하고 추출된 특징점을 비



(그림 13) 주간에 Nikon 카메라로 촬영한 영상셋 2 및 합성된 파노라마 영상

교 후 정합되는 점을 이용해서 원근 변환식을 모델링한다. 모델링된 변환식을 통하여 영상을 정렬하여 파노라마 영상을 생성하였다.

본 논문에서는 특징점의 추출을 위하여 SURF 특징점 추출 알고리즘을 적용하여 영상의 크기 및 회전 등의 기하학적 변형, 조명 변화 및 영상의 흔들림에도 강인성을 확보하였다. 또한 기존의 SIFT 알고리즘을 사용한 방법들에 비하여 빠른 파노라마 생성 속도의 향상도 달성하였다.

차후의 연구로는 다중밴드 블랜딩 등을 통한 합성 영상의 화질 향상이 필요하다. 실험에서의 합성 결과에서 확인 가능한 것과 같이 입력 영상의 촬영시에 카메라 설정 파라미터의 자동 변화로 인하여 경계선 부분의 밝기값이 차이가 존재하여 밝기값의 불연속성이 발생한다. 다중밴드 블랜딩 기술을 개발하여 적용할 경우 화질의 향상이 기대된다.

참 고 문 헌

[1] Wikipedia, "Panorama," http://en.wikipedia.org/wiki/Panorama.

[2] R. Szeliski, "Image Alignment and Stitching: A Tutorial1," Now Publishers 2006, ISBN 1933019042, 9781933019048, pp.120.

[3] S. T. Suen, E. Y. Lam, and K. K. Wong, "Photographic stitching with optimized object and color matching based on image derivatives," Optics Express, Vol.15, 2007, pp.7689-7696.

[4] M. Brown and D. G. Lowe, "Recognising Panoramas," Proc. of IEEE International Conference on Computer Vision, 2003, pp.1218-1227.

[5] M Brown, and D. G. Lowe, "Automatic panoramic image stitching using invariant features," International Journal of Computer Vision, Vol.74(1), 2007, pp.59-73.

[6] M Brown, R. I. Hartley, and D Nister, "Minimal Solutions for Panoramic Stitching," Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2007.

[7] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60(2), 2004, pp.91-110.

[8] J. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbor search in high-dimensional spaces," Proc. of International Conference on Computer Vision and Pattern Recognition, 1997, pp.1000-1006.

[9] M. Byrod, M. Brown, and K. Atrom, "Minimal Solutions for Paoramic Stitching with Radial Distortion," Proc. of IEEE International Conference on Computer Vision and Pattern Recognition, 2008, pp.1-8.

[10] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, Vol.24(6), 1981, pp.381-395.

[11] P. H. S. Torr, and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," Computer Vision and Image Understanding, Vol.78(1), 2000, pp.138-156.

[12] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "Speeded Up Robust Features (SURF)," Computer Vision and Image Understanding, Vol.110(3), 2008, pp.346-359.

[13] A. Criminisi, I. Reid, and A. Zisserman, "A Plane Measuring Device," http://www.robots.ox.ac.uk/~vgg/presentations/bmvc97/criminispaper/

권 기 원

e-mail : waterdrw@gmail.com

2004년~현 재 국립금오공과대학교 컴퓨터공학부 4학년

관심분야 : 멀티미디어, 모바일시스템 등

이 해 연

e-mail : haeyeoun.lee@kumoh.ac.kr

1997년 성균관대학교 정보공학과(학사)

1999년 한국과학기술원 전산학과(공학석사)

2006년 한국과학기술원 전자전산학과(공학박사)

2001년~2006년 (주)쎄트렉아이 선임연구원

2006년~2007년 코넬대학교 박사후연구원

2008년~현 재 국립금오공과대학교 컴퓨터공학부 교수

관심분야 : 멀티미디어, 영상처리, 콘텐츠보안, 디지털워터마킹 등

오 득 환

e-mail : dhoh@se.kumoh.ac.kr

1982년 경북대학교 전자공학과 전산전공(학사)

1985년 한국과학기술원 전산학과(공학석사)

1994년 한국과학기술원 전산학과(공학박사)

1986년~현 재 국립금오공과대학교 컴퓨터공학부 교수

관심분야 : 인공신경망, 컴퓨터네트워크, 임베디드시스템 등