논문 2009-46SP-2-10

칼라 불변 기반의 특징점을 이용한 영상 모자이킹

(Image Mosaicking Using Feature Points Based on Color-invariant)

권 오 설*, 이 동 창*, 이 철 희**, 하 영 호***

(Oh-Seol Kwon, Dong-Chang Lee, Cheol-Hee Lee, and Yeong-Ho Ha)

요 약

컴퓨터 비전 분야에서 영상 모자이킹 (Image Mosaicking) 은 제한된 시야각의 카메라를 사용하여 획득한 여러 장의 중첩된 영역을 가지는 영상을 한 장의 영상으로 정합하여 나타내는 기법이다. 최근에는 연속된 영상에서 카메라의 기학학적인 움직임 때문에 발생하는 영상의 왜곡이나 밝기 차에 관계없이 정확한 정합을 수행하기 위해서 특징점을 기반으로 서술자를 구성하는 정합 방법이 많이 연구되고 있다. 그러나 대부분의 특징점 검출 알고리즘들은 영상의 밝기값 기반의 처리 과정을 수행하기 때문에 영상의 칼라 성분은 다르지만 밝기값이 비슷한 경우, 또는 동영상에서 시간의 흐름에 따라 광원이 변화하는 경우에는 광원의 영향에 따라 검출되는 특징점의 수와 각각의 지역 서술자의 특성이 변하여 정확한 대응점을 검출하는데 오류를 유발하게 된다. 이런 문제점을 해결하기 위해서 본 논문은 영상의 칼라 정보를 이용한 특징점 기반의 영상 모자이킹 방법을 제안하였다. 디지털 칼라 카메라로부터 획득한 디지털 값을 좁은 대역을 갖는 가상의 카메라 출력값으로 변환하여 물체의 분광 반사율 기반의 값으로 유도하고 이것을 광원의 변화에 불변하는 칼라 불변 값 (Color-Invariant Value) 으로 정의하였다. 제안된 칼라 불변값의 유효성을 검증하기 위해서 시뮬레이션된 광원들과 Macbeth Color-Checker를 이용하여 확인하였으며, 실험결과에서 제안한 방법과 기존의 SIFT 알고리즘을 비교를 통해 제안된 방법의 정합율의 향상을 확인하였다.

Abstract

In the field of computer vision, image mosaicking is a common method for effectively increasing restricted the field of view of a camera by combining a set of separate images into a single seamless image. Image mosaicking based on feature points has recently been a focus of research because of simple estimation for geometric transformation regardless distortions and differences of intensity generating by motion of a camera in consecutive images. Yet, since most feature-point matching algorithms extract feature points using gray values, identifying corresponding points becomes difficult in the case of changing illumination and images with a similar intensity. Accordingly, to solve these problems, this paper proposes a method of image mosaicking based on feature points using color information of images. Essentially, the digital values acquired from a digital color camera are converted to values of a virtual camera with distinct narrow bands. Values based on the surface reflectance and invariant to the chromaticity of various illuminations are then derived from the virtual camera values and defined as color-invariant values invariant to changing illuminations. The validity of these color-invariant values is verified in a test using a Macbeth Color-Checker under simulated illuminations. The test also compares the proposed method using the color-invariant values with the conventional SIFT algorithm. The accuracy of the matching between the feature points extracted using the proposed method is increased, while image mosaicking using color information is also achieved.

Keywords: image mosaicking, color feature, color-invariant

I. 서 론

영상 모자이킹 (Image Mosaicking) 은 제한된 시야각의 카메라를 사용하여 획득한 여러 장의 중첩된 영역을 가지는 영상을 한 장의 영상으로 정합하여 나타내는 기법이다. 최근 컴퓨터 비전 분야에서 이 기법은 다양

* 학생회원, *** 정회원, 경북대학교 전자전기컴퓨터학부
(School of Electrical Engineering and Computer Science, Kyungpook National University)
** 정회원, 안동대학교 컴퓨터공학 전공
(Major of Computer Engineering, Andong National University)
※ 본 연구는 2008년도 BK21 사업의 지원으로 수행됨.
접수일자: 2008년2월22일, 수정완료일: 2009년2월27일

하게 연구되고 있으며, 파노라믹 영상 (Panoramic Image) 및 동영상을 재현하는데 적용되고 있다[1]. 예를 들면, 정지 영상의 경우 한 장의 영상이 가지는 시야각은 제한되어 있지만 각각의 영상을 하나의 영상으로 정합하면 넓은 시야를 확보 할 수 있기 때문에 광 시야각의 영상을 재현할 수 있고 사실적인 느낌을 부여하는 장점을 가진다.

영상 모자이킹을 수행하는데 있어서 가장 중요한 부분은 인접한 영상사이에 정확한 대응점의 검출을 통해 상호 관계를 도출하는 것이다. 대응점을 결정하는 방법은 미리 정의된 영상의 특정영역을 나타내는 서술자의 구성 유무에 따라서 크게 두 종류로 나뉘어진다.

첫 번째 방법은 서술자 없이 블록 (Block) 기반의 정합 방법으로 각각의 영상 내에 구성된 블록들 간의 상관 (Correlation) 을 이용하는 NCC (Normalized Cross Correlation) 방법과 FFT (Fast Fourier Transform) 의 특성을 적용하여 위상 상관 (Phase Correlation) 을 이용하는 방법 등이 있다[2~3]. NCC 방법[2]의 경우 비교 대상이 되는 블록간의 밝기값의 유사도를 수식적으로 계산하고 이것을 이용하여 대응 유무를 결정하는 방법인데 반하여 위상 상관을 이용하는 방법[3]은 인접한 두 영상 전체에 대해서 크로스 파워 스펙트럼 (Cross Power Spectrum) 을 계산하여 영상간의 변위를 결정하여 적용하는 방법이다. 이러한 블록기반의 정합 방법은 카메라의 기하학적인 움직임에 의한 영상 왜곡에 의해서 도출된 대응점의 정확성이 감소하는 단점을 가진다.

두 번째 방법은 특징점 (Feature Point) 기반의 정합 방법으로 각각의 영상에서 특징점을 검출하고 검출된 특징점 주위 영역을 나타내는 지역 서술자 (Local Descriptor) 기반의 정합 과정을 수행한다. 대표적인 특징점 검출 알고리즘에는 Harris Corner Detector, SUSAN Corner Detector, Morphological Corner Detector, SIFT (Scale Invariant Feature Transform) 등이 있다[4~8]. Harris Corner Detector[4]는 기본적으로 지역적인 신호 변화를 측정할 수 있는 자기 상관 (Auto-Correlation) 함수에 바탕을 두고 있으며 테일러 확장을 적용한 영상의 그레디언트 기반의 헤이시안 (Hessian) 행렬식을 이용하여 해당 화소의 코너의 정도를 결정하고 미리 정한 임계값을 만족하는 화소에 대해서 특징점으로 검출한다. SUSAN Corner Detector[5]는 일정 크기의 원 영역 안에서 밝기 정보의 유사도를 구하여 임계값 이상의 밝기 차가 발생하면 특징점으로 검출한다. 마지막으로 Morphological Corner Detector[6]는 대칭적인 구조의 구성요소를 이용하는데 기본적인 연산은 확장 (Dilation) 과 침식 (Erosion) 을 연속적으로 적용하는 열림 (Open-ing) 과 닫힘 (Closing) 연산을 통하여 영상 내의 특징점을 검출한다. SIFT[7~8] 알고리즘은 설정된 크기 공간 (Scale Space) 에서 가우시안 차 영상을 구하고 이것을 이용하여 좁은 영역내에서 최대값 또는 최소값으로 표현되는 화소를 특징점으로 검출한다.

그 중에서도 SIFT 알고리즘은 정확성에 있어서 우수함이 검증된 알고리즘으로 알려져 있고 물체 인식 (Object Recognition), 영상 정합 (Image Matching), 모션 추정 (Motion Tracking) 등과 같은 다양한 분야에서 응용 되고 있다[4~5]. SIFT 알고리즘의 방법은 먼저 영상으로부터 정합에 이용될 특징점을 검출하고 다음으로 검출된 특징점 주위의 영역을 나타내는 서술자를 구성함으로써 두 영상의 공간적인 위치를 도출함으로써 정합에 사용되는 영상의 크기와 회전, 밝기 변화에 강인한 특성을 가지는 것이 장점이다. 그러나 SIFT 알고리즘을 포함한 대부분의 특징점 검출 알고리즘들은 수행 속도의 이득 및 처리 과정의 복잡성을 줄이기 위해서 영상의 밝기값 기반의 처리 과정을 수행하기 때문에 밝기가 유사하지만 다른 색을 가지는 영상이나 특징점을 포함하는 영상에서 오류를 발생시키는 원인이 된다.

본 논문에서는 이러한 문제점을 해결하기 위해서 표면 칼라 정보를 고려한 특징점 기반의 영상 정합 방법을 제안하였다. 영상획득 과정에서 물체에 영향을 주는 광원의 특성을 배제한 물체의 분광 반사율 특성에 기반한 값을 도출하고 이를 칼라 불변값 (Color Invariant Value) 으로 정의한다. 그리고 획득한 칼라 불변 값을 이용해서 특징점을 검출하고 서술자를 구성한다. 이때 칼라 불변 값을 획득하기 위해서 카메라의 감도와 광원에 대한 가정이 주어진다. 카메라의 감도는 서로 중첩되지 않으며 좁은 대역을 가지며 이때 광원은 플랑키안 로커스 (Planckian Locus) 의 특성을 가진다. 이 두 가정에 기반한 영상의 디지털 값을 로그 색도 좌표계에서 표현하면 광원의 색을 배제한 물체의 분광 반사율 기반의 동일한 값으로 표현할 수 있다. 이를 토대로 카메라의 응답을 변환 행렬을 이용하여 칼라 불변값으로 변환하고, 변환된 칼라 불변값을 SIFT 구조에 적용하여 영상 모자이킹을 수행한다.

Ⅱ. 기존의 방법

2.1 기존의 방법의 특징

영상 정합은 기본적으로 인접 영상간의 중복되는 부분의 정보를 이용해서 이루어지고 크게 블록 기반의 정합 방법과 특징점 기반의 정합 방법으로 구분된다. 그 중에서도 특징점 기반의 알고리즘들이 정확성에서 우수한 성능을 가지며 최근 발표된 Lowe의 SIFT 알고리즘이 대표적이다[7~8].

SIFT 알고리즘은 크게 영상 내의 특징점을 검출하는 단계와 검출된 특징점을 표현하는 지역 서술자를 구성하는 단계로 나눌 수 있다. 우선 영상 내의 다수의 특징점을 검출하는 단계에서는 설정된 크기 공간을 다수의 옥타브 (Octave) 로 구성하고 각각의 옥타브에서는 동일 영상에 대해서 크기를 반으로 축소한 영상을 이용한다. 크기가 다른 입력영상에 대해서 시그마 값을 k(=1.6)배 증가시키면서 가우시안 영상을 여러 장 생성하고 이것을 이용해서 가우시안 차 (Difference of Gaussian) 영상을 생성한다. 다음으로 영상 크기를 변화시키면서 특징점을 검출한다. 각각의 옥타브에서 수행하는 세부적인 과정은 영상에 같은 비율로 가우시안 마스크의 시그마 값을 변화시키면서 여러 장의 블러된 영상을 획득한 후, 각각의 블러된 영상에서 여러 장의 차 영상을 획득한다. 두 번째 차 영상을 기준으로 3 x 3 마스크를 적용하여 마스크 내의 중심값이 첫 번째 차 영상과 세 번째 차 영상의 3 x 3 마스크 내의 값과 비교하여 중심 값이 가장 크거나 가장 작으면 극 값 (Extreme) 으로 정의하고 그 위치를 특징점으로 검출한다. 세 번째 차 영상에 대해서도 동일한 방법으로 두 번째, 네 번째 차 영상과 비교하여 특징점을 검출한다. 이 과정을 하나의 옥타브로 정의하고 그 다음 각각의 옥타브는 영상의 크기를 다운 샘플링한 후 첫 번째 옥타브에 적용한 절차를 반복 수행한다.

다음으로 안정도 판별을 통해서 검출된 다수의 특징점 중에서 정합에 이용될 특정 특징점들을 선별하는 과정이다. 영상의 균일한 부분의 특징점에 대해서는 여러 장의 테스트 영상을 통해서 가우시안 차분 영상에 임계값을 설정하여 제거가 가능하고 강한 에지부분에서 나타나는 특징점들에 대해서는 헤이시안 (Hessian) 행렬식을 이용하여 제거가 가능하다.

인접한 영상들 사이에 특징점을 검출하고 안정도 판별 과정을 거친 후, 선별된 특징점들의 정확한 대응점을 결정하기 위해서 각각의 특징점에 대한 지역 서술자를 구성한다. 특징점을 중심으로 16×16의 윈도우를 설정하고 윈도우 내의 화소들의 방향에 대한 히스토그램을 계산하는 과정으로 히스토그램은 각각의 화소들의 방향에 대한 크기 값을 누적함으로써 계산된다. 히스토그램의 최대값을 나타내는 방향을 현재 특징점의 주방향으로 설정하고 주방향이 결정되면 히스토그램내의 다른 방향값은 주방향을 기준으로 정렬한다. 특징점 전체의 서술자에 대해서 각각의 주방향을 결정하는데 이것은 영상의 회전 변환에 대해서 동일한 위치의 특징점 정합이 가능한 특성을 가진다. 방향 히스토그램은 각각의 특징점들 주위의 16×16 윈도우 안에서 계산되고 360도의 방향을 10도 간격으로 즉, 36개의 방향으로 양자화 한다. 특징점의 주위 화소들에 대한 크기와 방향은 다음 식을 이용한다.

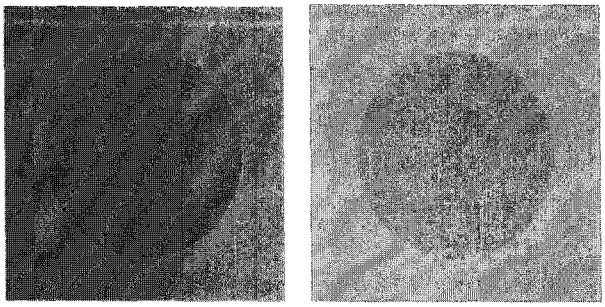
$$m(x,y)=\sqrt{(L(x+1,y)-L(x-1,y))^2+(L(x,y+1)-L(x,y-1))^2} \quad (1)$$

$$\theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \quad (2)$$

여기서 $L(x,y)$은 크기값이고 $\theta(x,y)$은 방향값이다. 그리고 $m(x,y)$은 특징점 위치에서의 주변 화소의 밝기 값이다. 각각의 특징점의 주방향이 결정되면 지역 서술자를 128차원의 특징 벡터로 표현한다. 앞서 특징점 주변 영역에 대해서 각 화소들의 방향 히스토그램을 구한 후 36방향으로 양자화 한 방향을 각 서브 블록 (Sub-Block) 4×4에 대해서 8방향으로 양자화 한다. 이때 각 방향에 대해서 크기값을 더하기 전에 16×16 윈도우 안의 화소들의 크기값은 가우시안 커널에 의해서 가중치 (Weighting) 가 결정한다. 가우시안 커널을 적용하는 것은 윈도우의 중심근처 즉, 특징점 위치에서 멀어질수록 낮은 가중치를 설정하여 특징 벡터를 계산하는 것으로 특징점 근처의 화소의 크기 값이 특징 벡터 정합 시 정확한 정합에 영향을 준다는 것을 의미한다. 128차원으로 계산된 특징점의 벡터는 가장 큰 값을 가지는 차원에 대해서 정규화한다. 특징점을 중심으로 생성된 128 차원의 벡터값을 비교 및 분석함으로써 영상간의 상관관계를 도출할 수 있다. 이때, 특징점들이 외부 환경변화에 강인할 때 정합률이 높아진다.

2.2 기존 방법의 문제점

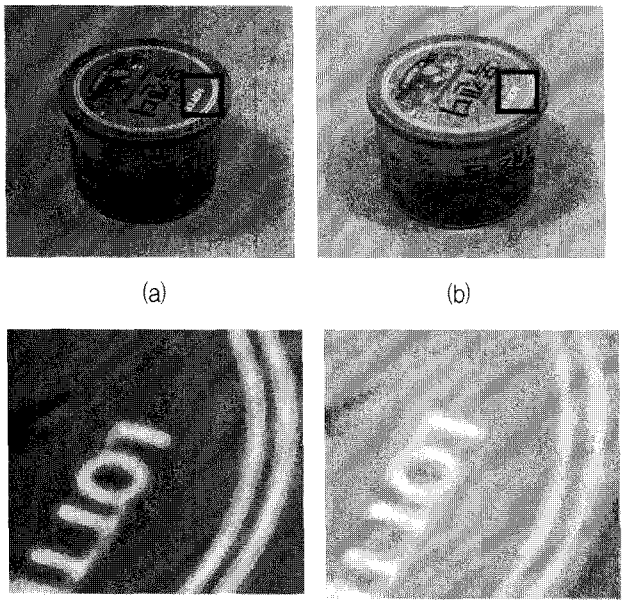
대부분의 특징점 기반의 정합 알고리즘들은 영상의 3차원 정보를 처리하지 않고 수행 속도 및 처리 과정의 복잡성을 줄이기 위해서 일차원의 밝기값 기반의 처리 과정을 수행한다. 이때, 물체의 표면 색 정보는 다르지만 밝기값으로 변환했을 때 인접하는 영역이 유사한 밝기값을 가진다면 기존의 SIFT 알고리즘에서는 특징점이 검출되지 않을 것이다. 왜냐하면 밝기값으로 변환했을 때 유사한 밝기값 속에서 특징점을 추출하지 못하기 때문이다. 그림 1에서와 같이 밝기값은 유사하나 색상의 차이가 있는 경우, 기존의 SIFT 알고리즘을 이용하여 특징점을 검출한 결과 특징점이 검출되지 않았다. 이것은 영상의 밝기값만을 고려한 기존의 알고리즘에서 발생할 수 있는 문제점이다. 실제로 인접한 화소간 값의 차이가 밝기값 기준으로 13 이하의 값의 차이를 가지게 되면 SIFT 알고리즘을 적용하였을 때 특징점을 검출하지 못함을 확인하였다. 또한 광원의 변화가 영상내의 물체에 영향을 미치는 경우, 영상 정합에서 오류를 발생시키는 원인이 된다. 그림 2(a)의 경우 표시된 영역내의 D65 광원하에서 붉은 색과 흰색은 밝기차가 확연하여 특징점 추출에 문제가 없다. 그러나 2(b)의 경우 붉은 색을 많이 포함하는 A 광원하에서 물체는 광원의 영향을 받아서 2(a)에 비해 경계영역이 사라지는 것을 확인할 수 있다. 이러한 경우 특징점의 검출 확률은 더욱 낮아지게 된다.



(a) (b)

그림 1. 밝기값은 유사하나 색상의 차이가 있는 영상의 SIFT 알고리즘 적용 결과; (a) 테스트 영상, (b) SIFT 특징점 추출 결과.

Fig. 1. Result images using SIFT algorithm in the images with similar intensity but different hue; (a) test image and (b) feature result by SIFT algorithm.



(a) (b)

(c) (d)

그림 2. 다른 광원에 의한 동일 물체의 변화; (a) D65광원하의 물체, (b) A 광원하의 물체, (c) (a)의 7배 확대, (d) (b)의 7배 확대

Fig. 2. The local descriptor under different illuminations; (a) Macbeth Color-Checker, (b) Tendency of patches under changing illuminations.

III. 제안한 방법

제안한 방법은 영상의 밝기 정보만을 이용하는 기존의 방법과는 달리 영상의 칼라 정보를 고려하여 특징점 검출 및 서술자를 구성하고 이것을 이용하여 영상 모자이킹 결과를 획득하였다. 디지털 칼라 카메라와 같은 입력 장치는 단지 출력값을 기록하기 때문에 광원의 영향을 많이 받기 때문에 일반 카메라로부터 획득된 영상의 물체의 정보는 광원이 변함에 따라서 변하게 된다. 기존의 알고리즘은 광원에 의존적인 밝기값에 기반한 처리과정을 수행하므로 동일한 물체일지라도 다른 광원하에서 획득하면 정합율이 감소한다. 따라서 본 논문에서는 광원에 독립적인 값을 도출하고 이 값을 이용하여 특징점 검출과 서술자 구성 및 영상 모자이킹 결과를 획득하고자 한다. 그림 3은 위에서 설명한 제안한 칼라 불변값 획득 과정의 흐름도이다. 먼저 광원에 독립적인 값을 획득하기 위해서 카메라의 감도와 광원에 대한 두 가지 가정을 한다. 이 가정에 기반하여 유도된 수식과 로그 좌표계를 이용하면 광원에 독립적인 물체 표면의 분광 반사율의 특성을 가지는 값을 획득할 수 있으며 이를 칼라 불변 값으로 정의한다.

3.1 칼라 불변 값 획득 방법

영상에서 칼라 불변 값을 획득하기 위해서 다음과 같은 두 가지 가정을 기반으로 테스트 한다. 첫째, 카메라의 감도는 서로 중첩되지 않으며 좁은 대역을 가지며, 이때 광원은 Planckian Locus의 특성을 가진다. 상용되는 대부분의 카메라는 RGB채널의 대역이 중첩되는 특징이 있다. 이때, 카메라 앞에 중첩되지 않은 대역을 가지는 필터를 부착함으로써 물리적으로 독립된 채널에 대한 영상을 획득할 수 있다. 본 논문에서는 실제 카메라의 응답을 특정 채널을 가지는 가상카메라 응답으로 변환하기 위하여 다음과 같은 수식을 이용하여 대역 중복이 없는 3 채널을 구현하였다.

$$S_i(\lambda) = \frac{1}{\sqrt{2\pi\sigma_i}} \exp\left\{-\frac{(\lambda-\mu_i)^2}{2\sigma_i^2}\right\} \tag{3}$$

여기서 μ 는 최고 피크 파장이며, σ 는 표준편차이다. 다음으로 광원을 수식적으로 모델링하기 위한 방법으로 사용되는 Planckian의 방정식은 색온도와 파장의 함수로 다음과 같이 표현된다.

$$E(\lambda,T) = c_1\lambda^{-5}[\exp(\frac{c_2}{T\lambda}) - 1]^{-1} \tag{4}$$

여기서 $c_1 = 3.74183\times10^{-16} Wm^2$ 이고 $c_2 = 1.4388\times10^{-2} mK$ 으로 상수 값이다.

디지털 칼라 카메라의 출력값은 입력되는 광원과 물체의 분광 반사율과 카메라의 감도의 곱의 적분값으로 정의할 수 있고 다음과 같이 표현된다.

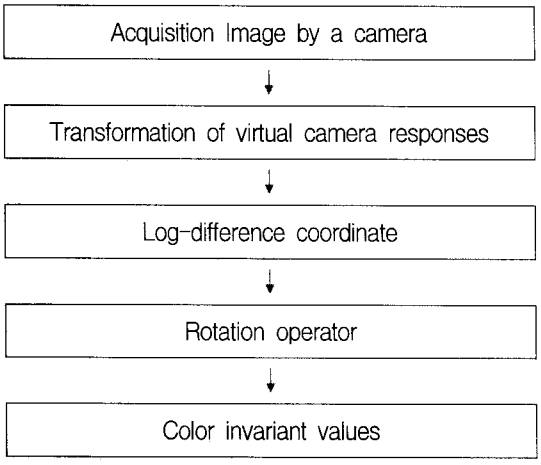


그림 3. 칼라 불변 값 획득 과정
Fig. 3. The flowchart for acquiring color invariant values.

$$C_i = \int_w E(\lambda)R(\lambda)S_i(\lambda)d\lambda, \quad (i = R,G,B) \tag{5}$$

여기서 C_i 는 카메라 출력값이고 $E(\lambda)$ 는 광원 $R(\lambda)$ 는 표면의 분광 반사율이고 $S_i(\lambda)$ 는 카메라의 감도이다. 그러나 첫 번째 가정을 적용하면 카메라의 출력값은 다음과 같다.

$$C_i = \int_w E(\lambda)R(\lambda)\delta(\lambda-\lambda_i)d\lambda = E(\lambda_i)R(\lambda_i) \tag{6}$$

여기에 광원의 식을 대입하면

$$C_i = c_1\lambda_i^{-5}\exp(\frac{-c_2}{T\lambda_i})R(\lambda_i) \tag{7}$$

윗 식은 양변에 자연 로그를 적용하면 표면의 분광 반사율과 광원의 색의 합으로 표현된다.

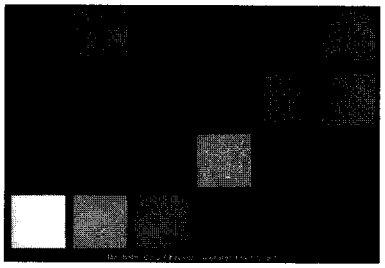
$$\ln C_i = \ln[c_1\lambda_i^{-5}R(\lambda_i)] - \frac{c_2}{T\lambda_i} \tag{8}$$

다음으로 RGB 세 채널의 카메라 응답에 대해서 R채널과 G채널의 로그 색차와 B채널과 G채널의 로그 색차를 각각 계산하고 이것을 2차원 로그 색도 좌표계에서 표현하면 광원에 불변하는 값을 도출할 수 있다.

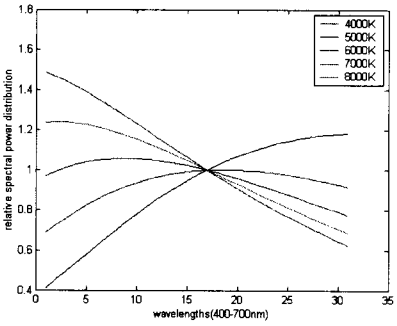
$$\begin{aligned}\ln\frac{R}{G} &= \ln C_R - \ln C_G = R_R' - R_G' + \frac{1}{T}(E_R - E_G)\\ \ln\frac{B}{G} &= \ln C_B - \ln C_G = R_B' - R_G' + \frac{1}{T}(E_B - E_G)\end{aligned} \tag{9}$$

여기서 $R_i' = \ln[c_1\lambda_i^{-5}R(\lambda_i)]$ 이고 $E_i = -c_2/\lambda_i$ 이다.

칼라 불변 값을 유도하는 과정에서 보듯이 카메라의 감도는 이론적으로 델타 함수를 가정한다. 그러나 테스트 과정에서는 RGB 채널간의 중복되지 않는 유사 함수를 사용하였다. 왜냐하면 실제 디지털 칼라 카메라의 감도가 가정한 델타 함수와는 차이가 너무 크기 때문에 적당한 카메라의 감도를 설정하는 것이 중요하기 때문이다. 따라서 적당한 상관 관계를 고려하여 카메라 감도를 설정하였다. 테스트 실험에서는 멕베스 칼라 체커를 이용하여 가상의 카메라 감도와 색온도에 대해서 시뮬레이션된 5개의 광원을 이용하였다. 그림 4는 실험에서 사용된 멕베스 칼라 샘플과 5개 광원의 파장별 분포를 나타내고 있다. 이를 제안한 방법으로 계산하여 로그 색도 좌표계상에서 표현하면 그림 5(a)와 같이 분포하게 된다. 여기서 각각의 샘플이 광원의 색온도의 변



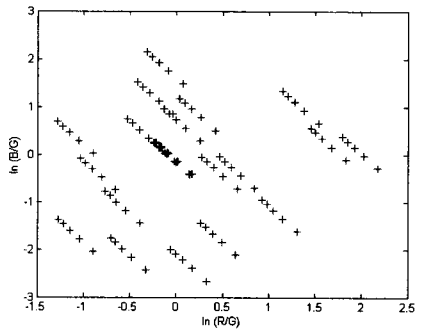
(a)



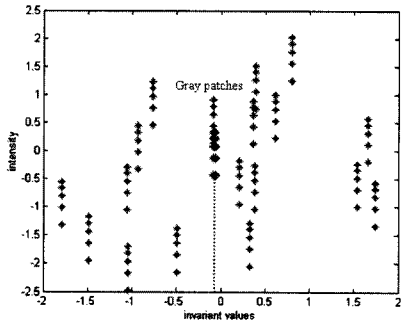
(b)

그림 4. 사용된 샘플과 광원;
(a) Macbeth Color-Checker, (b) 5 가지 광원

Fig. 4. Used samples and illuminations; (a) Macbeth ColorChecker and (b) five illuminations.



(a)



(b)

그림 5. 멕베스 칼라 차트에 대한 칼라 불변 값의 예; (a) 로그 색좌표에서의 결과 (b) 회전 변화 후 결과

Fig. 5. Examples of color-invariant values for Macbeth ColorChecker; (a) the result in Log-coodinates and (b) the result after rotation transform.

화에 따라서 일정한 기울기를 가지며 하나의 직선을 형성함을 확인할 수 있다. 24개의 칼라 패치는 서로 다른 광원 하에서 동일한 기울기를 가지기 때문에 이 결과에 회전 변환을 적용하면 그림 5(b)와 같이 변환된다. 그림 5(b)에서 회색 계통의 패치들은 모두 같은 값으로 도출되고 나머지 색을 가지는 패치들은 각각 다른 값으로 도출되는 것을 확인할 수 있다. 이 값을 칼라 불변 값으로 정의할 수 있으며 실제 카메라로 획득한 칼라 영상에서 각 RGB의 3차원 값을 1차원의 칼라 불변 값으로 변환 후 특징점 검출하고 서술자를 구성한다.

3.2 상용 카메라로부터 가상 카메라 응답 변환

본 논문에서 제시하는 알고리즘은 두 가지의 기본적인 가정으로부터 출발한다. 그 중의 하나가 바로 카메라의 채널별 민감도가 파장에 따라서 서로 중첩이 없어야 한다는 것이다. 그러나 상용되는 대부분의 카메라의 민감도는 채널별로 중첩이 되어있기 때문에 카메라로 획득된 영상을 바로 알고리즘에 적용할 수는 없다. 따라서 카메라로 획득된 응답을 가상 카메라의 응답으로 변환하는 과정이 필요하다. 이를 위해서 표준 광원하에서 표면 반사율을 알고 있는 몇 개의 샘플들을 카메라로 획득하고 다음으로 표면 반사율과 가상 카메라 민감도 및 광원의 분광분포를 계산함으로써 가상 카메라의 응답을 획득한다. 이 두 응답을 다항회귀 방정식으로 계산하면 상용 카메라로 획득된 영상을 가상 카메라의 응답으로 변환할 수 있다.

$$
\begin{aligned}
R'(R,G,B) &= a_{R',0} + a_{R',1}R + a_{R',2}G + a_{R',3}B + a_{R',4}RG + a_{R',5}GB \\
&\quad + a_{R',6}BR + a_{R',7}R^2 + a_{R',8}G^2 + a_{R',9}B^2 + a_{R',10}RGB \\
&\quad + a_{R',11}R^3 + a_{R',12}G^3 + a_{R',13}B^3 + a_{R',14}RG^2 + a_{R',15}R^2G \\
&\quad + a_{R',16}GB^2 + a_{R',17}G^2B + a_{R',18}BR^2 + a_{R',19}B^2R \\
G'(R,G,B) &= a_{G',0} + a_{G',1}R + a_{G',2}G + a_{G',3}B + a_{G',4}RG + a_{G',5}GB \\
&\quad + a_{G',6}BR + a_{G',7}R^2 + a_{G',8}G^2 + a_{G',9}B^2 + a_{G',10}RGB \\
&\quad + a_{G',11}R^3 + a_{G',12}G^3 + a_{G',13}B^3 + a_{G',14}RG^2 + a_{G',15}R^2G \\
&\quad + a_{G',16}GB^2 + a_{G',17}G^2B + a_{G',18}BR^2 + a_{G',19}B^2R \\
B'(R,G,B) &= a_{B',0} + a_{B',1}R + a_{B',2}G + a_{B',3}B + a_{B',4}RG + a_{B',5}GB \\
&\quad + a_{B',6}BR + a_{B',7}R^2 + a_{B',8}G^2 + a_{B',9}B^2 + a_{B',10}RGB \\
&\quad + a_{B',11}R^3 + a_{B',12}G^3 + a_{B',13}B^3 + a_{B',14}RG^2 + a_{B',15}R^2G \\
&\quad + a_{B',16}GB^2 + a_{B',17}G^2B + a_{B',18}BR^2 + a_{B',19}B^2R
\end{aligned}
\tag{10}
$$

일반적으로 다항 회귀 방정식에서 변환 성능은 추정되어야 할 계수가 많을수록 유리한 점이 있지만 변환식의 계산이 복잡해지는 단점이 있다. 본 논문에서는 3, 9, 19개의 계수를 추정하여 각각의 성능 테스트를 통하여 19개의 계수의 사용하였다. 각 항의 계수를 추정하기

위해서는 선형적인 행렬 연산의 수행을 위해 다음과 같이 벡터형식으로 표현할 수 있다.

$$\mathbf{P} = \mathbf{V}^{\mathrm{T}}\mathbf{a} \tag{11}$$

여기서 $\mathbf{V}^{\mathrm{T}}$는 $\mathbf{V}$의 전치 행렬을 나타낸다. 최종적으로 추정하려는 각각의 계수는 가상 역행렬로 표현된다.

$$\mathbf{a} = (\mathbf{V}\mathbf{V}^{\mathrm{T}})^{-1}\mathbf{V}\mathbf{P} \tag{12}$$

결과적으로 실제 카메라로부터 획득된 영상은 추정된 변환 행렬식을 통하면 가상 카메라 출력값으로 변환되고 이 값으로부터 칼라 불변값을 획득할 수 있다.

3.3 칼라 불변 값의 정규화

실제 카메라로부터 획득한 영상의 디지털 값을 변환 행렬식을 통하여 가상 카메라의 출력값으로 변환한 후, 칼라 불변 값을 획득하면 얻어진 값은 각각의 영상에 의존하는 값으로 도출된다. 이것은 영상의 정합 시 정확성을 떨어뜨리고 오류를 유발하는 원인이 된다. 왜냐하면 영상에 따라서 칼라 불변 값의 분포가 다양하기 때문이다. 이런 문제점 때문에 칼라 불변 값을 정규화 과정이 필요하다. 일반적으로 고른 칼라 분포를 가지는 Munsell 패치들에 대해서 칼라 불변 값을 표현하면 그림 6과 같이 나타난다. 그림 6에서 각각의 Munsell 패치는 5 가지의 다른 광원하에서 동일한 값으로 표현된다. 이는 임의의 광원하에서도 각각의 Munsell 패치가 항상 동일한 칼라 불변값을 얻을 수 보여준다.

마지막으로 카메라에 의해 획득된 영상으로부터 계산한 칼라 불변값을 전체 칼라 불변값에 대한 수식으로 정규화함으로써 특징점 추출을 위한 최종적인 서술자 입력값으로 정의한다.

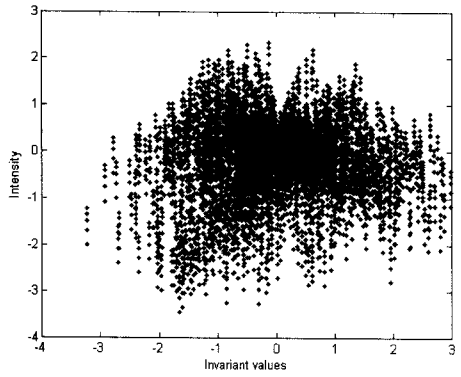


그림 6. 5가지 광원 하에서의 Munsell 패치에 대한 칼라 불변 값

Fig. 6. Color invariant-values for Munsell patches under 5 illuminations.

$$NCIV = \frac{CIV - MP_{\min}}{MP_{\max} - MP_{\min}} \times 255 \tag{13}$$

여기서 NCIV는 정규화 된 칼라 불변 값이고 CIV는 영상의 칼라 불변 값이며 MP는 Munsell 패치의 칼라 불변값이다.

3.4 칼라 불변값을 이용한 특징점 검출

특징점 검출은 칼라 불변값을 이용해서 검출된 특징점을 기반으로 재현하였다. 연속된 입력 영상에 대하여 중첩된 영역을 획득하고 각각의 영상에서 특징점을 검출한 후 서술자를 구성한다. 서술자들의 비교를 통해서 특징점들 사이에 정확한 대응점을 결정한다. 이 때 결정된 대응점들 중에서 영상 정합 시 정합 오류를 일으키는 대응점들은 즉, outlier에 해당하는 대응점들을 제거하기 위하여 RANSAC (Random Sample Consensus) 알고리즘을 이용하였다[9]. 다음은 최종적으로 선택된 대응점들을 이용해서 영상의 변환 관계를 도출하고 기하학적인 변환 관계를 이용하여 영상 모자이킹을 재현한다. 영상간의 변환은 투영 변환 (Perspective transform)으로 계산할 수 있다.

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{14}$$

IV. 실험 및 결과

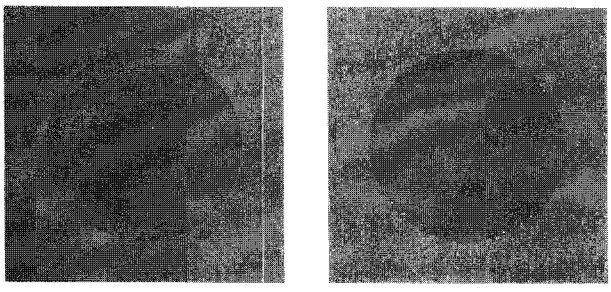
실험 영상으로 사용된 영상은 일반 디지털 카메라 (Kodak 6490), 모바일 폰 카메라 (삼성 SCH-S140), 그리고 캠 카메라 (Sony HDR-SRI) 를 이용하여 획득하였다. 실제 카메라로 획득된 여러 영상을 기존의 SIFT 알고리즘의 밝기값 기반의 특징점 정합과 제안한 칼라 불변값을 이용한 특징점 정합을 비교하였으며 사용된 영상은 동일한 광원하에서 그 외 변수를 변화시킨 영상과 다른 조건은 동일하나, 상이한 광원하에서 획득한 영상으로 구성하였다. 이때 기존의 SIFT 알고리즘의 장점인 크기와 회전 변환에 강인함을 테스트하기 위해서 영상 획득 시 각각의 특성에 맞게 변환하여 영상을

획득하였고 광원이 다른 영상을 획득하기 위해서 영상 부스를 사용하여 D65광원과 A광원 하에서 영상을 획득하였다. 각각의 영상에서 검출된 특징점의 정합율은 대응점의 위치를 고려해서 다음식을 이용하여 계산하였으며, 다만 실험 영상에서 발생하는 모든 특징점을 본 논문에 나타내지 못함을 미리 밝힌다. 다만, 일부의 특징점 검출 결과를 확률적인 근거로 제시하였다.

$$MR(\%) = \frac{Accurate\ \ pairs}{Total\ \ pairs} \times 100 \qquad (15)$$

그림 7은 제안한 방법을 이용하여 특징점을 추출한 결과이다. 이미 앞서 보인 바와 같이 기존의 방법은 이 영상에서 특징점을 검출하지 못하였으나, 제안한 방법은 원 주위에 12개의 특징점이 검출됨을 알 수 있다. 그림 8은 테스트 영상의 특징점 정합 결과로 전체 30개의 대응점에 대한 결과를 보여주고 있다. 왼쪽 영상은 SIFT 알고리즘을 적용하여 얻은 결과이고 오른쪽 영상은 제안한 방법을 적용하여 얻은 결과이다. (a)와 (b)는 동일한 광원하에서 각각 크기 변환과 회전 변환을 시험하기 위한 영상이고 (c)와 (d)는 동일한 광원하에서 여러 가지 물체를 포함하는 영상의 특징점 정합 결과이다. 마지막으로 (e)와 (f)는 다른 광원하에서 획득한 영상의 특징점 결과를 보여주고 있다. 그림 8의 원본 영상은 다음의 홈페이지를 통해 확인할 수 있다. (http://cilab.knu.ac.kr/publication/IEEK2009_images.zip)

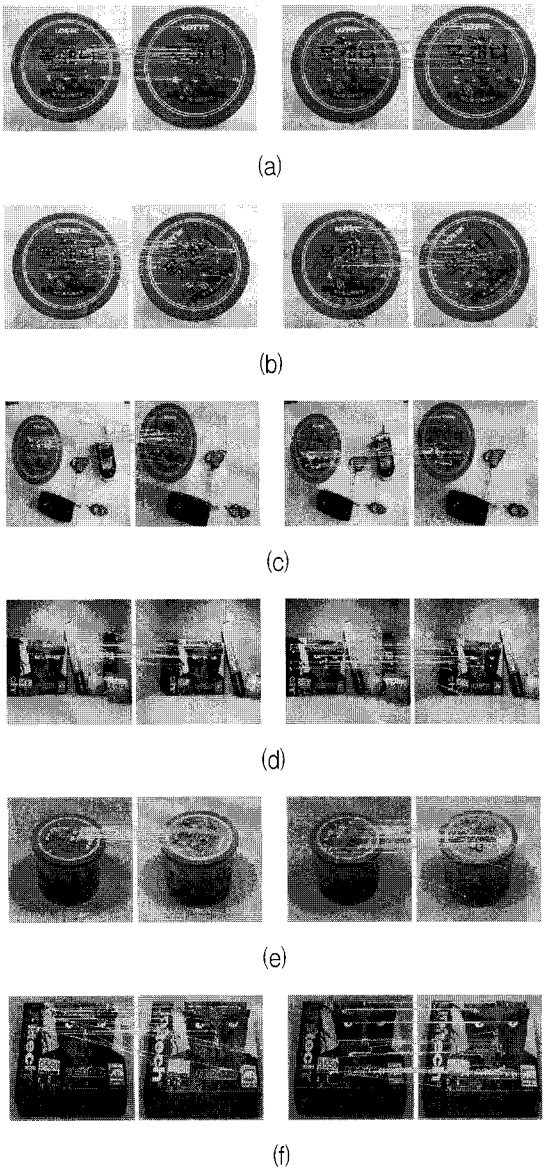
그림 9는 그림 8 영상의 정합율에 대한 결과이다. 정합율을 제시함에 있어서 정확한 정합율을 표현하지는 못하였으나, 동일 조건에서 추출되는 약 50% 특징점에 대한 정합률을 데이터로 제시하였음을 밝힌다. 테스트 결과 대부분의 영상에서 정합률이 향상됨을 확인할 수 있다. 특히 (e)와 (f)에서 정합률의 향상은 급격한 광원의 변화에도 강인한 특징점을 추출할 수 있다는 것을 보여준다. 이 결과 그림 10과 같이 중첩된 부분의 광원이 완전히 다른 상황에서도 별도의 후처리 없이 효율적으로 영상 모자이킹 결과를 획득할 수 있다.



(a) (b)

그림 7. 특징점 추출 결과;
(a) SIFT 결과, (b) 제안한 방법

Fig. 7. The results of the feature matching;
(a) SIFT algorithm and (b) proposed method.



(a) (b) (c) (d) (e) (f)

그림 8. 실영상의 특징점 정합 결과; (a) 크기 변환에 대한 결과, (b) 회전 변환에 대한 결과, (c),(d) 다양한 물체를 포함하는 영상의 특징점 정합 결과, (e),(f) 다른 광원하에서 얻은 영상의 특징점 정합 결과

Fig. 8. The results of the feature matching for real images; (a) Scale transformation, (b) Rotation transformation, (c),(d) Images including various objects, (e),(f) Images acquired under different illuminations.
(http://cilab.knu.ac.kr/publication /IEEK2009_images.zip)

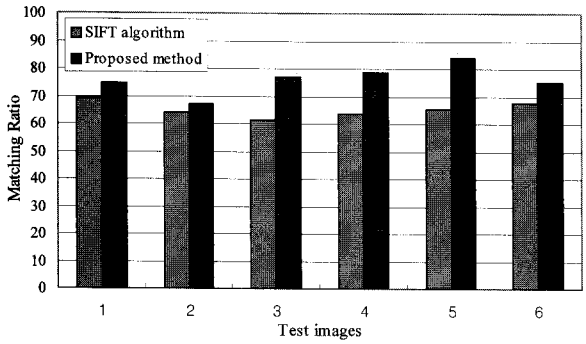
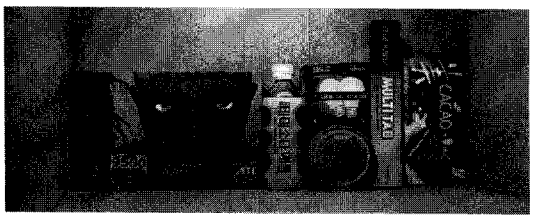


그림 9. 테스트 영상에 대한 정합율 비교
Fig. 9. The comparison of matching rate for test images.

(a)



(b)

그림 10. 급격한 광원의 변화에서 영상 모자이킹
Fig. 10. Image mosaicking when illuminations change suddenly.

V. 결 론

본 논문에서는 칼라 불변값을 이용한 특징점 기반의 영상 모자이킹 방법을 제안하였다. 기존의 알고리즘이 영상의 밝기 위주로 특징점을 추출하기 때문에 광원의 영향을 제거하지 못하는 반면, 제안한 방법은 광원에 불변하는 값을 계산해 냄으로써 조명색의 변화에 상관없이 물체 표면색의 특성을 바탕으로 특징점을 추출하였다. 이를 위해 가상 민감도 함수를 이용한 다항 회귀 방정식을 이용하였으며, 각 화소별로 생성된 광원 불변값을 SIFT 알고리즘에 적용함으로써 기존의 SIFT 알고리즘의 장점을 유지하면서 부가적으로 광원의 변화가 발생시키는 정합 오차를 줄일 수 있었다. 실험을 통하여 광원의 변화가 급격히 발생하는 영상에 대한 정합률을 분석한 결과 기존 방법에 비해 오차가 감소함을 확인하였다.

참 고 문 헌

[1] Alex Rav-Acha, Yael Pritch, Dani Lischinski, Shmuel Pleleg, "Dynamosaics: Video Mosaics with Non-Chronological Time," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

[2] Jun-Wei Hsieh "Fast stitching algorithm for moving object detection and mosaic construction", Image and Vision Computing, vol. 22, pp. 291-306, 2004.

[3] B. Srinivasa Reddy and B. N. Chatterji "An FFT-based technique for translation, rotation, and scale-invariant image registration", IEEE Transactions on Image Processing, vol. 5, no. 8, pp. 1266-1271, Aug. 1996.

[4] C. Harris and M. J. Stephens, "A combined corner and edge detector," In Alvey Vision Conference, pp. 147-152, 1988.

[5] S. M. Smith and J. M. Brady, "SUSAN-a new approach to low level image processing," International Journal of Computer Vision, vol. 23, no. 1, pp. 45-78, May, 1997.

[6] Robert Lagniere, "Morphological Corner Detection," Computer Vision six International Conference, pp. 280-285, 1998.

[7] David G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[8] David G. Lowe, "Object recognition from local scale-invariant features," International Conference on Computer Vision, Corfu, Greece, pp. 1150-1157, 1999.

[9] 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, pp. 381-395, 1981.

저 자 소 개

권 오 설(학생회원)
2004년 2월 경북대학교 전자공학과 석사 졸업
2008년 8월 경북대학교 전자공학과 박사 졸업
2008년 10월 ~ 현재 뉴욕대학교 박사후 연구원

이 철 희(정회원)
대한전자공학회 논문지
제 41 권 SP편 1 호 참조

이 동 창(학생회원)
2006년 2월 경북대학교 전자전기컴퓨터학부 졸업
2008년 2월 경북대학교 전자공학과 석사 졸업
2008년 2월 ~ 현재 삼성전자 통신연구소 연구원

하 영 호(정회원)
대한전자공학회 논문지
제 38 권 SP편 3 호 참조