

FAST를 이용한 파노라마 영상 생성 기법

*김종호, **박시영, ***유지상

광운대학교

*jjong3927@kw.ac.kr, **pksiyoung@kw.ac.kr, ***jsyoo@kw.ac.kr

Panorama image generation using FAST algorithm

*Jongho Kim, **Siyoung Park, ***Jisang Yoo

Kwangwoon University

요 약

본 논문에서는 자연스러운 파노라마 영상 생성을 위해 FAST(features from accelerated segment test)를 이용한 특징점 기반의 파노라마 영상 생성 기법을 제안한다. 다수의 영상을 이용해 자연스러운 파노라마 영상을 만들기 위해 실린더 투영을 수행 한 후 추출된 특징점들을 RANSAC(random sample consensus)을 이용해 정합 시 오차율을 최소화한다. 서로 다른 방향에서 얻은 다수의 영상을 합성할 때 정합 경계 주변의 이질감을 보완하기 위해 블렌딩 기법을 사용함으로써 자연스러운 파노라마 영상을 생성한다. 다수의 영상으로 실험을 한 결과 왜곡이 보정되고 자연스러운 파노라마 영상을 생성할 수 있었다.

1. 서 론

영상처리를 이용한 파노라마 영상 생성은 여러 방향과 각도에서 촬영한 다수의 영상을 하나의 영상으로 정합하는 것을 의미한다. 기존의 파노라마 영상 생성 기법 중에는 SIFT(scale invariant feature transform)와 SURF(speed up robust feature)기법을 이용한 특징점 기반 알고리즘이 있었다[1][2][3]. 기존의 기법은 많은 계산량 때문에 수행속도가 느린 단점이 있으며 많은 영상을 정합할 때 왜곡현상이 발생한다. 본 논문에서는 FAST기법을 이용한 특징점 기반의 파노라마 영상을 생성하는 기법을 제안한다. 다수의 영상을 정합할 때 생기는 왜곡현상을 보정하기 위해 영상을 실린더 좌표계로 변환한다 [2]. 변환된 영상에서 FAST(features from accelerated segment test)기법을 통해 특징점을 검출한 뒤 SURF 기법의 서술자(descriptor) 벡터를 특징점마다 각각 계산한다. RANSAC(random sample consensus)기법을 이용해 특징점 중 가장 정확하게 정합된 네 점을 찾고 영상간의 호모그래피를 계산한다[4]. 호모그래피를 이용하여 영상을 좌표변환한 뒤 다른 영상과 정합하게 되면 파노라마 영상이 생성된다.

본 논문의 2장에서는 파노라마 영상을 생성하기 위한 알고리즘과 제안하는 방법을 설명한다. 3장은 다수의 영상으로 실험을 한 결과와 성능비교를 확인하고 4장에서 결론을 내린다.

2. 본 론

본 논문에서는 여러 영상에서 특징점을 추출한 뒤 호모그래피를 계산해 자연스러운 파노라마 영상을 만드는 방법을 제안한다. 파노라마 영상을 생성하는 기법의 흐름도는 그림1과 같다.

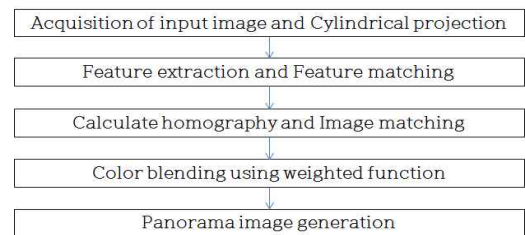


그림 1. 제안하는 파노라마 영상 생성 기법
Fig 1. Proposed panorama image generation method

영상을 디지털 카메라, 웹캠 등을 이용하여 적당한 간격으로 겹치게 촬영한다. 파노라마 영상을 생성할 때 촬영 각도와 초점거리의 차이가 있기 때문에 영상이 늘어지는 현상이 발생한다. 이를 해결하기 위해 기존의 평면 좌표계에 있던 영상을 실린더 모양의 좌표계에 투영한다 [5]. 정방향 실린더 투영을 할 때 홀이 생기는 문제점을 해결하기 위해 식(1)의 역방향 실린더 투영을 수행한다.

$$\begin{aligned} x' &= s\theta + x_c = s \tan^{-1} \frac{x}{f} + x_c \quad (1) \\ y' &= sh + y_c = s \frac{y}{\sqrt{x^2 + f^2}} + y_c \end{aligned}$$

식 (1)에서 (x, y) 는 실린더 좌표를 (x', y') 는 평면 좌표를 나타낸다. f 는 영상의 초점거리를 x_c 와 y_c 는 실린더 좌표계의 중심좌표이며 s 는 이미지의 스케일을 결정하는 변수이다. 그림 2는 초점거리 f 를 변화시키면서 역방향 실린더 투영을 한 결과 영상이다.



(a) 원본 영상 (b) $f = 700$ (c) $f = 400$

그림 2. 역방향 실린더 외평한 결과 영상

Fig 2. Result image using backward warping and cylindrical projection

본 논문에서는 영상 사이의 정합점을 찾기 위해 FAST(features from accelerated segment test)기법으로 특징점을 검출한다[6]. 그림 3은 영상의 특징점 후보중 하나인 p 화소를 중심으로 16개의 주위 화소를 나타낸다. 첫 번째 과정은 특징점 후보를 빠른 속도로 검출하는 것인데 16개의 주위 화소($I_{p \rightarrow x}$)의 밝기값이 p 화소(I_p)의 밝기값에 임계값 t 를 더한 밝기값 보다 더 크거나 임계값 t 를 뺀 밝기값보다 더 작은 화소가 12개 이상 연속으로 있을 때 p 화소는 특징점 후보로 검출된다. 즉, 영상의 모든 특징점 후보들의 주변 화소를 식 (2)을 통해 3가지($S_{dark}, S_{similar}, S_{bright}$)로 정의하고 p 화소와 비교후 특징점 후보를 검출한다.

$$S_{p \rightarrow x} = \begin{cases} I_{p \rightarrow x} \leq I_p - t & (S_{dark}) \\ I_p - t < I_{p \rightarrow x} < I_p + t & (S_{similar}) \\ I_p + t \leq I_{p \rightarrow x} & (S_{bright}) \end{cases} \quad (2)$$

두 번째 단계는 식 (3)을 이용해 특징점 후보로 검출된 화소들에 점수를 부여한다. 식 (3)이 의미하는 것은 특징점 후보로 검출된 화소의 밝기값(I_p)이 인접 화소의 밝기값($I_{p \rightarrow x}$)에 대해 특징이 될 수 있는 상태를 유지할 수 있도록 하는 최대 임계값 t 를 찾고, 그 임계값 t 를 특징점의 점수로 정의한다.

$$V = \max \left(\sum_{x \in S_{bright}} |I_{p \rightarrow x} - I_p| - t, \sum_{x \in S_{dark}} |I_p - I_{p \rightarrow x}| - t \right) \quad (3)$$

마지막 세 번째 단계는 영상에서 인접한 점들은 같이 특징점으로 검출되는 경우가 많은데 그 중 특징이 가장 강한 한점만 특징점 후보로 검출하는 단계이다. Non-maximal suppression 이라고 하는 이 과정은 두 번째 단계에서 계산한 식 (3)의 점수를 이용해 인접한 특징점들 중 자신보다 높은 V 값을 가지는 특징점이 있다면 낮은 점수를 가지는 특징점은 제거하는 방식으로 하나의 특징점만을 검출한다.

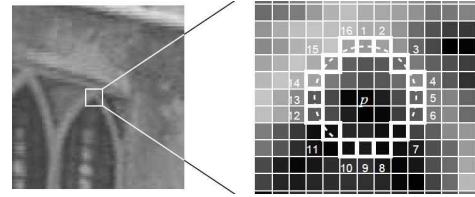


그림 3. 특징점 후보의 중심인 p 화소

Fig 3. The pixel at p is the center of a candidate feature

각 영상에서 특징점을 검출한 다음에는 특징점의 정보를 서술자(descriptor) 벡터로 표현한다. 본 논문에서는 특징점을 검출한뒤 SURF기법의 서술자벡터를 사용하여 특징점을 표현한다. SURF(speed up robust feature)기법은 SIFT(scale invariant feature transform)기법을 개선한 것으로서 크기와 회전에 강인한 특징점을 추출하기 위해 많이 사용되는 기법이다[3]. 먼저 검출된 특징점 주변 화소값에 하 웨이블릿 필터(haar wavelet filter)를 적용해 특징점의 주방향을 계산한다. 주방향을 기준으로 4×4 영역으로 분할한 뒤 각 영역마다 네 가지 종류의 특징벡터($\sum dx, \sum dy, \sum |dx|, \sum |dy|$)를 계산하여 64차원의 벡터를 생성한다. 그림 6은 영상에서 검출한 특징점을 빨간색 점으로 표시한 결과 영상이다.



그림 4. SURF 기법으로 특징점을 추출한 영상

Fig 4. Feature extraction image using SURF algorithm

특징점을 매칭하기 위해 RANSAC(random sample consensus)기법을 이용한다. RANSAC기법은 에러와 잡음이 섞여있는 데이터들로부터 적절한 모델을 예측하고, 예측된 모델과 오차를 최소화하는 기법이다. 모델을 결정하기 위해 전체 데이터 중 최소의 데이터를 랜덤하게 선택하고 가상의 모델로 결정한다. 선택된 특징점이 예측

모델과 잘 맞는지 반복적으로 해를 계산함으로써 최적의 해를 구한다. 기본적으로 특징점 정합 방법은 SURF기법의 64차원 서술자 벡터들 사이의 유클리디안 거리를 계산하여 최소가 되는 두점을 정합점으로 결정한다. 영상 사이의 검출된 많은 특징점 중 가장 정확하게 정합된 4개의 특징점을 결정한다.

두 영상에서 비교적 정확하게 정합된 4개의 좌표를 알면 DLT(direct linear transformation)기법을 이용하여 두 영상 사이의 호모그래피(homography) 행렬을 계산할 수 있다. 호모그래피란 기준영상의 좌표와 그에 대응되는 대상 영상의 좌표 사이의 관계를 행렬로 나타낸 것이다. 호모그래피 행렬의 변수 9개를 구하고 한 영상을 변환하면 기준 영상의 좌표에 정합된다.

다수의 영상을 정합하다 보면 영상마다 조명과 시점이 다르기 때문에 정합 후 부자연스러운 경계선이 생길 수 있다. 이런 불연속적인 경계선을 보정하기 위해 선형가중치함수(bilinear weighted function)를 사용한다. 겹치는 영역의 화소마다 가중치를 다르게 부여하는 방법으로 자연스러운 파노라마 영상을 생성할 수 있다.

3. 실험결과

실험환경은 i5 인텔 CPU, 8GB 램, GeForce GTX460 그래픽카드이며 비주얼스튜디오2010으로 구현하였고 영상 내 특징점 추출을 위해 OpenCV 2.4.6을 사용하였다. 실험에 사용한 영상은 http://mpac.ee.ntu.edu.tw/~sutony/vfx_stitching/pano.htm#Download 에서 다운로드 받았다. 실험은 실린더 투영의 효과 그리고 특징점 추출방법에 따른 과정별 속도를 측정 했다. 그림 5는 파노라마 영상을 생성하기 위해 입력받은 800x600크기의 두 영상이다. 그림 6은 입력 받은 두 영상을 이용해 제안하는 기법으로 파노라마 영상을 생성한 결과이다. 그림7은 각각 324x484 크기를 가지는 4장의 영상을 기존 방법대로 실린더 투영을 하지 않고 정합한 결과 영상이다. 그림8은 실린더 투영 후의 파노라마 결과 영상이다. 실린더 투영을 한 결과 다수의 영상을 정합할 때 왜곡이 사라지며 자연스러운 영상이 생성되었다. 그림9는 총 10장의 영상을 입력받아 제안한 기법을 이용하여 생성한 파노라마 영상이다. 표1은 기존의 특징점 검출기법인 SIFT, SURF 기법과 제안하는 FAST기법을 이용하여 324x484크기를 가지는 10장 영상을 정합할 때 과정별 처리속도를 비교한 것이다. FAST기법의 속도가 가장 빠르며 SIFT기법에 비해 약 10배 정도 수행 시간이 줄었다. 본 논문에서 제안하는 기법을 사용하면 기존 기법보다 수행속도가 빠르면서 자연스러운 파노라마 영상을 얻을 수 있다.



그림 5. 파노라마 영상 생성을 위한 입력 영상
Fig 5. Input image for panorama image generation



그림 6. 생성된 파노라마 영상
Fig 6. Generated panorama image



그림 7. 다수의 영상을 정합할 때 왜곡 현상
Fig 7. Phenomenon of distortion of numerous panorama image



그림 8. 실린더 투영 후 영상 정합 결과
Fig 8. Result image after cylindrical projection



그림 9. 다수의 영상을 정합한 파노라마 영상

Fig 9. Panorama image using numerous image

Process	SIFT (ms)	SURF (ms)	Proposed method (ms)
Acquisition image and Cylindrical projection	281	281	281
Feature extraction and Order decision	30157	6012	824
Feature matching and RANSAC algorithm	5112	2644	954
Calculate homography and Image matching	891	881	717
Color blending using Weighted function	846	762	734
Total	37287	10580	3510

표 1. 10장의 영상 정합시 특징점 추출 방법에 따른 과정별 처리 속도 비교

Table1. Comparisons of processing speed according to feature extraction method with 10 images

4. 결론

본 논문은 기존의 기법 보다 빠르게 특징점을 추출할 수 있는 FAST기법을 이용하여 파노라마 영상 생성 기법을 제안하였다. 또 실린더 투영과 색상 블렌딩 기법을 통한 보정과정을 추가하여 기존보다 자연스러운 파노라마 영상을 생성할 수 있었다. 다수의 영상을 실시간으로 정합하는 것은 한계가 있으나 병렬처리를 이용하여 시간을 줄일 수 있을 것으로 예상된다. 앞으로 더 많은 영상을 정합 할 때 속도 향상과 왜곡을 최소화 하는 연구가 필요하다.

감사의 글

본 연구는 미래창조과학부의 정보통신방송연구개발사업의 일환으로 수행하였음. [라우드니스 기반의 방송음향 기술 및 실내 환경 소음의 스트레스 평가 기술 개발 (2014-044-055-002)]

참 고 문 헌

[1] M. Brown, and D. G. LOWE, "Automatic panoramic image stitching using invariant features," *International journal of computer vision*, Vol. 74, No.1, Dec. 2006.

[2] R. Szeliski, "Image alignment and stitching: a tutorial," *Computer graphics and vision*, Vol. 2, No.1, pp.15~16, Jan. 2006

[3] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool "Speeded-Up Robust Features (SURF)," *Computer vision and image understanding (CVIU)*, Vol. 110, No. 3, pp. 2-8, Jun. 2008

[4] L. Moisan, P. Moulon, and Pascal Monasse, "Automatic homographic registration of a pair of images, with a contrario elimination of outliers," *Image processing on line(IPOL)*, pp. 2-3, May, 2012

[5] R.Szeliski, H.Y.Shum, "Creating Full View Panoramic Image Mosaics and Environment Maps", *Computer Graphics*, pp.251-258, August 1997,

[6] E.Rosten, T.Drummond, "Machine learning for high-speed corner detection", *European conference on computer vision*, Vol.1, pp.430-443, 2006

[7] P. J. Burt and E. H. Adelson, "A multi resolution spline with application to image mosaics," *ACM Transaction on graphics*, Vol. 2, No. 4, pp. 2-5, Oct. 1983.