

VR 영상 화질 개선을 위한 계층적 탐색 기반 매칭 방법

남다운, *한종기

세종대학교, *세종대학교

ndy9709@sju.ac.kr, *hjk@sju.ac.kr

Hierarchical Searching-based matching method for improving VR quality

Da-yun Nam *Jong-ki Han

Sejong University *Sejong University

요 약

기존의 스티칭 사용되어 온 특징점 추출 방법들은 영상의 고주파 영역과 저주파 영역에 따라 빈도수가 다르다는 특징이 있고, 이러한 특징점 빈도수의 불균일성 없이, 일정한 간격으로 분포하는 특징점 매칭하는 계층적 탐색 기반 매칭 방법을 제안한다. 이는 스티칭 영상의 화질 개선뿐만 아니라 3D VR 영상의 화질개선에서도 효과를 줄 수 있다.

1. 서론¹

VR (Virtual Reality) 기술이 다양한 분야로 활용도가 높아지고 있다. 특히, VR 기술을 통해 비대면 방식의 교육, 공연 등 서비스가 확대되고 있다. 이와 같은 실사 영상을 기반으로 제작된 VR 영상은 ERP(Equi-Rectangular Projection) 영상, 3D 콘텐츠 등으로 보여질 수 있다. 2D VR ERP(Equi-Rectangular Projection) 영상은 스티칭 (stitching) 시스템을 기반으로 제작되고, 6DoF VR 콘텐츠는 가상 시점 합성 프로그램 또는 포인트 클라우드(point cloud) 생성 프로그램 등을 통해 제작될 수 있다. 실사영상 기반 VR 영상을 제작하기 위해서는 카메라 배열의 위치정보가 제공되지 않을 경우, 카메라 캘리브레이션(calibration) 과정을 통해 카메라의 위치를 추정한 후 시작된다. 이 때 보편화된 캘리브레이션 방법들은 특징점 추출 및 매칭(keypoint extraction)에서부터 시작된다. 특징점 추출은 꽤 오랫동안 연구되어 온 주제이고, 현재까지 가장 널리 사용되는 특징점 추출 및 매칭 알고리즘으로는 SIFT[1], SURF[2], ORB[3] 등이 있다. 일반적인 특징점 추출 알고리즘들은 전체 동작 시간을 고려하면서 영상에 포함된 노이즈의 영향을 적게 받기 위해 영상

안에 선명한 코너점, Blob, 엣지와 같이 영상 내 공간주파수가 높은 지점에 위치하는 픽셀을 특징점으로 둔다. 그리고 해당 특징점에 대한 기술자(descriptor)를 정의해서 기술자 간의 유사성 비교를 통해 중첩되는 영상간 특징점 매칭을 할 수 있다. 특징점 매칭을 기반으로 전체적으로 이미지 매칭을 할 수 있다.

현재까지 특징점 추출 및 매칭 기법들은 점차 고숙화되고 높은 매칭 정확도를 가지도록 발전해왔지만, 현재 VR 기술에 적용하기 적절하지 않는 문제점이 존재한다. 특징점 추출 기법은 고주파 영역에서는 특징점의 빈도가 높아지고, 저주파 영역에서는 특징점의 빈도가 낮아 공간적으로 균일한 특징점을 매칭시키는 것이 어렵다는 점이다. 실제 일반 스티칭, 스테레오스코픽 스티칭(stereoscopic stitching)에서는 영상 와핑[4](image warping)의 개선을 위해 APAP[5], SPHP[6]을 포함한 여러 매쉬 단위 와핑(mesh-based warping) 기술이 등장하였다. 이는 스티칭 과정에서 영상을 정합시킬 때 전역적으로 하나의 호모그래피를 통한 와핑을 하는 것이 아니라, 영상을 매쉬 그리드로 쪼개, 매쉬 폴리곤 단위로 정합시키는 것을 의미한다. 이러한 매쉬 기반 와핑을 위해서는 영상 내 균일한 위치에 분포하는 특징점이 적합하다. 또한 깊이 지도를 기반으로 생성되는 6DoF VR 영상에서 보면, 카메라 캘리브레이션을 정확하게 해야 깊이 지도를 생성 및 정제(refinement)과정의

¹ 연락저자 : 한종기

성능이 높아진다. 높은 정확도의 캘리브레이션을 위해서는 영상 내 균일한 위치에 분포하는 특징점이 적합하다. 그리고 여기서 균일한 위치의 특징점은 특징점 추출 및 매칭 프로그램에 의해서 정해진 예측 불가능한 위치가 아닌 사용자에게 의해 정해진 균일하게 위치한 특징점을 의미한다.

본 논문에서는 사용자가 지정하는 특정 위치의 점을 특징점으로 두었을 때, 해당 특징점에 대한 매칭점을 탐색 방법이다. 매칭 시키려는 두 영상이 존재할 때, 한 쪽 영상에 대해 먼저 특징점의 위치가 정해지면 그 점들을 기반으로 다른 영상 안의 매칭점을 찾는 것이다. 제안하는 알고리즘은 특징점에 대한 기술자를 계층적으로 정의하고, 계층적 기술자를 통해 계층적 탐색을 하여 최적의 매칭점을 찾는 방안을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 기존 특징점 추출 알고리즘의 원리와 특징점 추출 기술의 동향을 살펴본 후, 3 절에서는 제안하는 기법을 설명하고, 4 절에서는 제안한 기법의 성능을 실험을 통해서 확인한다. 마지막으로 5 절에서는 본 논문에 대한 결론을 맺는다.

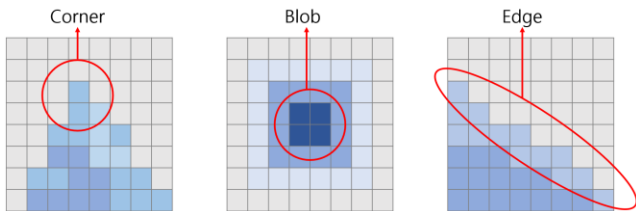


그림 1. 특징점의 성질을 가진 코너, Blob, 엣지영역

2. 관련 연구

2.1 특징점 추출 알고리즘

대표적인 특징점 추출 알고리즘 SIFT[1], SURF[2], ORB[3] 등이 있다. 비교적 오래된 기술임에도 불구하고 모두 시간 대비 높은 매칭 정확도를 보인다. SIFT 코너 검출기와 LoG(Laplacian of Gaussian)을 근사화한 개념인 DoG(Difference of Gaussian) 피라미드를 기반으로 스케일, 회전, 노이즈에 불변하는 Blob 을 특징점으로 지정하였다. SIFT 는 추출된 특징점 중심의 윈도우 내의 정보를 이용해 128 차원의 기술자가 정의한다. SIFT 는 높은 매칭 정확도를 갖는 알고리즘이지만, 다른 알고리즘 대비 긴 동작시간을 가진다. 이 후, SIFT 의 정확도에 근접한 성능을 보이면서 고속 처리가 가능한 SURF 가 제안되었다. SURF 는 LoG 를 근사화한 hessian 검출기

기반으로 blob 을 찾고, 특징점으로 지정하였다. 그리고 특징점 중심의 윈도우 정보를 이용해 64 차원의 기술자를 정의한다. 가장 최근 기술인 ORB 는 FAST[7]의 특징점과 BRIEF[8] 의 기술자를 혼합한 기술이다. FAST 는 고속 처리를 위해 설계된 특징점 추출 기술이며, BRIEF 역시 메모리 절약에 효과적인 이진 기술자(descriptor) 정의 기법이다. ORB 는 FAST 와 회전에 강인하도록 BRIEF 를 개선한 rBRIEF 을 융합한 기술이다. 따라서, 고속 동작에 큰 효과를 보이고, SIFT 에 근접한 매칭 정확도를 보인다. 위와 같은 알고리즘을 통해 얻은 특징점들을 매칭할 때에는 보편적으로 기술자 간 유클리드 거리를 구하고, 가장 거리가 가까운 특징점과 매칭하는 방법을 사용한다. 기존의 사용되는 특징점 추출 방법들은 특징을 찾고 특징점의 위치를 결정하고 특징점끼리 매칭하는 방식이었다면, 제안하는 방법은 특징점 위치를 결정하고, 기술자를 정의하고, 매칭되는 점을 찾는 방식이다. 따라서 사용자는 영상에 특징점이 균일하게 분포하도록 위치 지정을 할 수 있으며, 균일 분포하는 특징점에 대해 매칭 과정을 수행할 수 있다. 제안 알고리즘은 특징점의 위치가 사용자가 의해 가장 먼저 고정되기 때문에, 매칭 정확도를 높이기 위해 계층적으로 탐색영역을 좁혀가면서 매칭점을 찾는 방식을 사용하였다.

3. 제안 방법

그림 2는 제안하는 방법의 전체적인 과정을 도식화한 것이다.

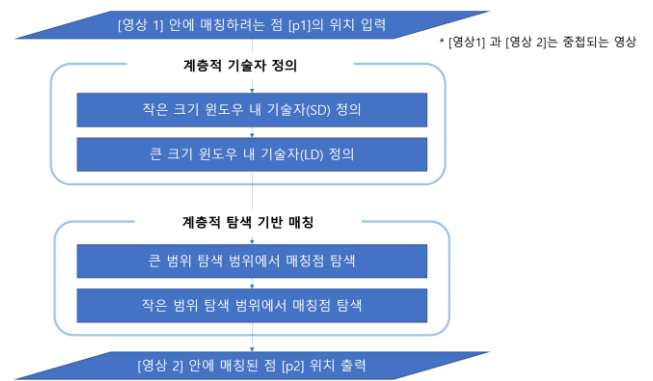


그림 2. 제안하는 방법의 Flowchart

3.1 계층적 기술자 정의

사용자가 정의한 위치를 입력으로 받아 해당 위치의 픽셀 주변으로 작은 크기의 윈도우와 더 큰 크기의 윈도우를 정의한다. 그리고 해당 윈도우들 안에 포함되는 모든 픽셀들의 색상 강도, 그

래디언트, 그라디언트 방향의 값과 거리 관계를 기술자로 정의한다. 본 논문에서 작은 크기 윈도우에 정의된 기술자를 SD(Small-Descriptor)로 표기하고, 큰 크기의 윈도우에서 정의된 기술자를 LD(Large-Descriptor)로 표기하였다. 그림 2 에서 보이는 것처럼, 영상 1 에 대한 특징점 위치를 지정하면, 영상 2 에 대해 해당 특징점과 매칭되는 점을 찾는다. 따라서 영상 1 과 영상 2 를 매칭하기 전, 영상 1 에서는 지정된 픽셀들에 대해서만 SD, LD 를 정의하고, 영상 2 에서는 모든 픽셀에 대해서 SD, LD 를 정의한다.

3.2 계층적 탐색 기반 매칭

본격적인 계층적 탐색에 앞서 위치 지정된 포인트들 중 영상의 가장 중심부에 위치한 10 ~15 개 픽셀과 매칭되는 픽셀들을 찾아야 한다. 이를 통해 영상 1 과 영상 2 간의 대략적인 수평 좌표 차이 Δx 와 수직 좌표 차 Δy 이를 확인할 수 있다. 그리고 영상 2 의 모든 픽셀에 대해 Δx 와 Δy 만큼 좌표를 이동시켜 영상 1 의 좌표계에 대략적으로 일치하도록 한다.

영상 1 과 평행 이동시킨 영상 2 를 가지고, 지정된 특징점들에 대한 매칭점 탐색을 한다. 매칭점 탐색과정 역시 계층적 탐색 과정으로 큰 탐색 영역안에서 대략적인 매칭점 위치를 파악하고 작은 탐색 영역안에서 구체적인 매칭점을 결정한다. 본 논문에서는 작은 탐색 영역의 매칭점 탐색을 SS(Search-in-Small-Range), 큰 탐색 영역의 매칭점 탐색을 SL(Search-in-Large-Range)로 정의하였다. SL 과정에는 LD 에 대해 SSD(Sum-of-Squared-difference)을 기반으로 비용을 계산해, 가장 작은 비용을 가지는 픽셀위치를 찾는다. SL 과정의 최소 비용을 가지는 매칭 점을 식(1)과 같이 나타냈다. 식 (1)의 p_2 과 p_1^{fixed} 는 각각 영상 2 과 영상 1 에 픽셀을 나타내고, 특히 p_1^{fixed} 는 위치 지정된 특징점을 의미하고, p_2 는 큰 탐색 영역 안에 포함된 픽셀을 의미한다 탐색 영역안에서 모든 픽셀 p_2 에 대해 최소 비용을 가지는 점을 p_2^{LS} 로 표기하였다

$$p_2^{LS} = \arg \min_p \sum (p_2.LD - p_1^{fixed}.LD)^2 \quad (1)$$

SL 이 끝나면, 점 p 주위로 작은 탐색 영역을 정의하고, SS 과정을 이어서 한다. 마찬가지로 SS 시에는 SD 에 대해 SSD(Sum-of-Squared-difference)을 기반으로 비용을 계산해, 가장 작은 비용을 가지는 위치의 점 p_2^{SS} 을 찾는다. SS 과정의 최소 비용을 가지는 매칭 점을 찾는 수식은 식(2)와 같다.

$$p_2^{SS} = \arg \min_p \sum (p_2.SD - p_1^{fixed}.SD)^2 \quad (2)$$

4. 실험 결과



그림 3. 위치 지정된 한 개의 특징점



그림 4. 그림 3 에서 제안 알고리즘을 적용한 매칭 결과

그림 3 은 임의로 정의한 위치의 특징점을 파란색 점으로 표현한 것이다. 그림 4 는 그림 3 의 파란색 점에 대해 현재 이미지와 중첩 이미지 간의 매칭 결과를 보여준다. 특징점이 매칭된다는 것을 확인할 수 있다.



그림 5. 균일 분포하는 특징점.

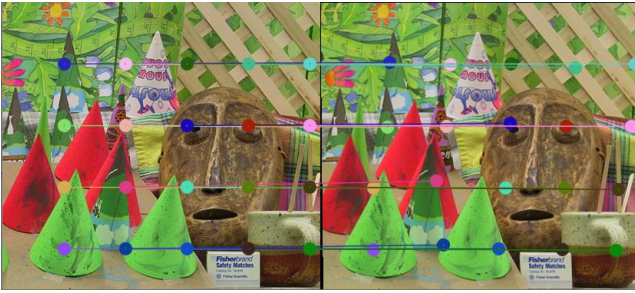


그림 6. 그림 5에서 제안 알고리즘을 적용한 매칭 결과

그림 5는 균일한 격자 구조에 위치하도록 특징점을 지정한 것이다. 그림 6은 그림 5에서 나타난 위치 지정된 특징점에 대한 중첩되는 이미지 간의 매칭 결과를 보여주고, 정확도 높게 매칭된다는 것을 확인할 수 있다.

5. 결론

제안하는 방법은 두 영상이 중첩하는 경우 한쪽 영상 내 지정된 위치의 픽셀과 매칭되는 다른 영상 내 매칭 포인트를 찾는 알고리즘으로, 매칭 과정에서 필요한 기술자를 계층적으로 정의하였으며, 계층적 탐색을 통해 영상 간 매칭되는 픽셀을 찾을 수 있었다. 현재 알고리즘은 매칭 점의 위치가 정해지는 현재 영상에서는 지정된 픽셀에 대해서만 기술자를 정의하면 되지만, 매칭시킬 타겟 영상에서는 모든 픽셀에 고유한 기술자를 정의해야 하기 때문에 연산 복잡도가 크다는 한계점이 있다. 향후 분 알고리즘의 한계점과 매칭 성능을 개선시킬 계획이고, 이를 통해 결과적으로 스티칭 기법의 개선, 더 나아가 VR 영상의 화질 개선을 기대해 볼 수 있다.

참고 문헌

- [1] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, 2004
- [2] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *European conference on computer vision*. Springer, 2006, pp. 404-417
- [3] Ethan Rublee, Vincent Rabaud, Kurt Konolige and Gary Bradski, "ORB: an efficient alternative to SIFT or SURF," *International Conference on Computer Vision*, January 2012
- [4] R. Szeliski, "Image alignment and stitching: a tutorial," *Foundations and Trends in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1-104, 2006.

- [5] Zaragoza, J., Chin, T.J., Brown, M.S., Suter, D, " As-projective-as-possible image stitching with moving dlt" *2013 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2339-2346. CVPR'13 (2013)
- [6] Chang, C.H., Sato, Y., Chuang, Y.Y, " Shape-preserving half-projective warps for image stitching," *IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3254-3261. CVPR'14, 2014
- [7] E. Rosten and T. Drummond. "Machine learning for high-speed corner detection", In *European Conference on Computer Vision*, Vol. 1, 2006.
- [8] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief: Bi-nary robust independent elementary features," In *European Conference on Computer Vision*. 2010