

특징점 연결 제약을 이용한 고속 스테레오 특징점 정합

김창일^o, 박순용

경북대학교 전자전기컴퓨터학부, 경북대학교 IT 대학 컴퓨터학부

cikim@vision.knu.ac.kr, sypark@knu.ac.kr

요 약

스테레오 비전 분야에서 특징 기반 스테레오 정합 방법은 빠른 수행 시간과 높은 정확성을 갖는다. 본 논문에서는 특징점 연결을 이용한 새로운 특징점 기반 스테레오 정합 기술을 제시한다. 제안하는 방법은 특징점 간의 거리와 컬러 정보를 이용한다. 특징점 연결 방법은 보다 정확한 변이 값을 결정하는데 효율적으로 스테레오 정합의 정확성을 높일 수 있다. 또한, 연결된 특징점 사이에 결정되지 않은 변이 값은 양선형 보간법을 통해 변이 값을 결정한다. 제안하는 방법은 FAST 특징 추출기와 에피폴라, 순서, 변이 스테레오 제약 조건을 사용하여 특징점 연결 방법의 정확성과 성능을 평가하기 위해 5 개의 정답 영상과 비교하여 이를 검증하였다.

1. 서론

스테레오 정합은 컴퓨터 비전 분야에서 가장 많이 연구된 분야 중 하나이다. 스테레오 비전은 인간의 양안 시차에서 시작된 것으로 3 차원 장면 복원, 물체 복원, 로봇 비전 등의 응용에 사용되는 기반 기술이다. 스테레오 정합은 좌우 영상에서 동일한 물체의 위치가 어떤 관계를 가지고 있는지를 결정하는 방법이다. 스테레오 영상 특징 화소(pixel)에서 관계를 정의할 수 있을 때, 대응 관계(correspondence)에 있다고 하고, 두 영상 간의 떨어져 있는 거리 차이를 변이(disparity)라고 한다. 변이 지도(disparity map)은 스테레오 정합의 결과로 영상을 구성하고 있는 화소의 변이 값의 집합을 말한다.

스테레오 정합 기술은 크게 두 가지로 나눌 수 있다. 첫째는 특징 기반 방법(feature based method)와 영역 기반 방법(area based method)으로 나뉘어진다. 특징 기반 방법은 영상의 에지나 코너와 같은 두드러진 특징이 나타나는 영역의 변이 값을 계산하는 방법이다. 따라서 계산 시간이 얼마 되지 않고 정확한 결과 값을 가진다. 특징이 있는 부분에만 변이 값이 계산되므로 변이 지도가 밀집되어 나타나지 않는다. 이를 결정하는 방법으로는 첫번째 방법은 창 기반 방법(window method)은 좌우 영상의 상관관계(correlation)를 이용하는 방법이다 [1]. 이 방법은 빠르게 수행되며 간단한 것이 장점이다. 이 때 상관 관계를 결정하기 위해 사용되는 알고리즘으로는 SAD(Sum Absolute Difference), SSD(Sum Squared Difference), NCC(Normalized Cross Correlation) 등이 있다. 두번째 방법은 기하학적인 정보를 이용하는 방법이다. 특징점 간의 거리나 각도 등을 이용하여 보다 정확한 변이 값을 결정할 수 있다 [2, 3]. 최근에는 순수한 특징 기반 방법에 대한 연구는 많이

이루어지지 않고 있으며, 위의 방법으로 결정된 변이 값을 전파(propagation)하여 주변 영역의 변이 값을 결정하는 방법이 연구되고 있다 [4].

두 번째로 영역 기반 방법은 전체 영역의 변이 값을 모두 결정하는 방법이다. 따라서, 이 방법은 이전 방법과 달리 밀집된 변이 지도를 얻을 수 있다. 이 방법에서 사용되는 방법은 전역적 방법(global method)과 지역적 방법(local method)가 있다. 전역적 방법은 화소 하나의 변이 값을 결정하기 위해 영상 전체의 화소와 비교한 뒤에 변이 값을 결정한다. 이 때, 정합의 정확도를 향상시키기 위해 반복적 연산(iterative scheme) 과정이 사용되기도 한다. 이와 같은 이유로 전역적 방법은 수행 시간이 많이 요구된다. 이 방법은 최근에 많이 연구되고 있으며, 그래프 컷(graph cut) [5]과 신뢰 기반 전파(belief propagation)[6] 방법 등이 있다. 지역적 방법은 특징점 기반의 창 함수 방법과 동일하다.

스테레오 정합의 정확도를 높이기 위해 여러 가지 제약조건들을 사용한다. 그 중에 가장 많이 사용되는 것인 에피폴라 제약(epipolar constraint)으로 영상의 대응되는 두 점을 동일한 에피폴라 선상에 두는 방법이다 [1,3,4]. 다음 조건은 순서 제약(ordering constraint)으로 특징점이 좌우 영상에 특징이 나타나는 순서는 동일하게 나타난다는 것이다 [7]. 마지막 방법은 링크 길이 제약(link length constraint)으로 좌우 영상에 대응되는 2 개의 특징점 간의 거리는 동일하게 나타난다는 조건이다. 또한 컬러 정보를 사용하는 것 역시 영상의 차이를 정확하게 구분할 수 있게 하는데 큰 도움을 준다. [4, 6]

본 논문에서는 특징점 연결 제약과 컬러 정보를 이용하여 특징점 기반 스테레오 정합을 연구하였다. 논문의 알고리즘의 정확성을 확인하기 위해 정답 영상(ground truth image)와 비교하여 정확성을 확인해

보았다.

본 연구는 다음과 같이 진행된다. 2장에서 특징점 기반 알고리즘에 대해 설명하고, 3장에서는 본 연구의 결과를 제시하여 이를 통해 4장에서는 결론을 도출한다.

2. 특징점들의 연결 제약을 이용한 방법

이번 장에서는 특징점 연결 제약을 이용한 스테레오 정합 기술에 대해 설명한다. 먼저 특징점을 추출하기 위해 FAST 알고리즘을 사용한다. 단순 특징점 기반 스테레오 정합 알고리즘은 좌우 영상의 동일한 에피폴라 선상에 특징점의 개수가 다르게 추출될 경우 정합의 정확도가 떨어지는 단점이 있다. 이런 이유로 특징점 간의 기하학적인 관계를 이용하여 정확성을 높이는 방법이 고려된다 [2, 3]. 이 논문에서는 특징점 기하학적 특징과 컬러 정보를 사용하여 정합의 정확성을 향상시킨 특징점 연결 (Feature Link) 방법을 제안한다.

2.1 컬러 정보를 이용한 정합 함수

특징점 간의 상관 관계를 분석하는 방법 중에 많이 사용되는 창 기반 방법은 간단하며 수행 시간이 빠르다 [1]. 전에 비슷한 연구에서 창 기반 방법을 사용할 경우, 영상의 명도(intensity) 정보만을 이용한다. 하지만, 명도 정보만을 사용하는 것은 정확한 영상의 차이를 계산하기는 어렵다. 왜냐하면 명도 정보는 단순히 컬러 영상의 합으로 만들어지기 때문이다 [4]. 이런 이유로 영상의 컬러 정보를 사용한다.

많은 컬러 채널 중에 RGB 채널을 사용하였고, 창 기반 방법에서 평균 제곱 오차(Mean Square Error) 알고리즘을 사용하였다. 이 알고리즘은 아래의 수식으로 정의 된다.

$$f(x, y) = \frac{1}{n^2} \sum_{i=-n/2}^{n/2} \sum_{j=-n/2}^{n/2} dist_c(Right(x+i, y+j), Left(x+i+d, y+j))$$

$$dist_c(c_L, c_R) = (R_1 - R_2)^2 + (G_1 - G_2)^2 + (B_1 - B_2)^2$$

$$Left(x, y) = (R_{left}(x, y), G_{left}(x, y), B_{left}(x, y))$$

$$Right(x, y) = (R_{right}(x, y), G_{right}(x, y), B_{right}(x, y))$$

수식 1에서 d 는 우 영상의 특정 화소의 변이 값을 의미하고, n 은 창의 크기를 의미한다. 위에서 정의된 이 수식은 우리 방법에서 초기 추측, 특징점 간 정합, 특징점 연결 정합 세 부분에서 사용된다.

2.2 정합 알고리즘

에피폴라 제약 조건을 이용하므로 제안한 방법은 동일한 에피폴라 선상에 존재하는 특징점만을 고려한다.

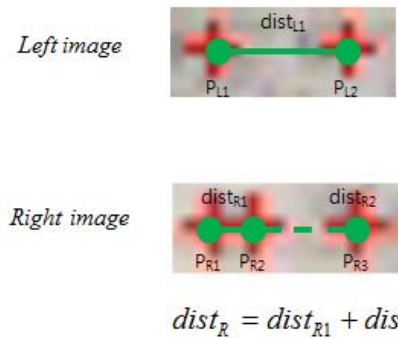
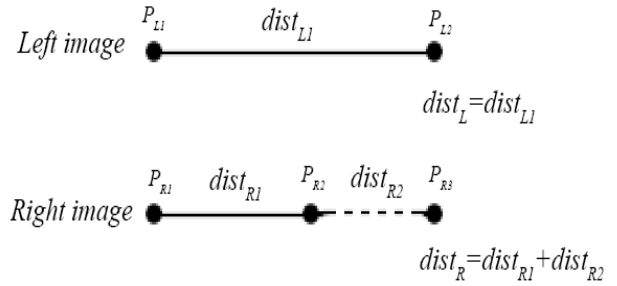


그림 1: 특징점 연결 알고리즘. (a) 거리가 다를 때, 특징점을 추가로 연결. (b) 실제 Teddy [9] 영상에서 특징점 연결이 필요한 경우

제안한 방법에서 특징점을 추출하는 방법으로 FAST 알고리즘을 사용한다. FAST 알고리즘은 영상 주변의 화소들의 밝기 간의 차이와 특정 문턱치(threshold) 간의 밝기를 비교하여 특징점을 결정한다. 이 알고리즘은 빠른 속도로 특징점을 추출한다 [8]. 추출된 특징점은 내림차순으로 정렬된다.

다음으로, 동일한 에피폴라 선상에 추출된 특징점의 수의 개수를 분석한다. 만약 좌우 영상에서 특징점이 하나만 발견된다면, 특징점과 특징점 간의 정합 방법을 사용한다. 왜냐하면 특징점이 하나일 경우에는 특징점 간의 기하학적 특성을 이용할 수 없기 때문이다.

만약 특징점이 좌우 영상 모두 특징점이 2개 이상 나타나면, 최초의 특징점이 정확한 변이 값을 가지게 해야 한다. 최초의 특징점의 변이 값에 오류가 있을 경우 그 뒤에 나타나는 특징점까지 잘못된 변이 값을 갖게 될 확률이 높기 때문이다. 이런 이유로 초기에 추출된 몇 개의 화소에 대해 특징점 주변의 상관값(correlation value)을 계산, 비교하여 가장 정확하다고 판단되는 변이 값을 결정해준다. 이 과정을 초기 추측이라고 한다. 초기 추측의 정확도에 따라 뒤에 나타나는 화소의 변이 값에도 영향을 주기 때문에 이 과정은 매우 중요하다. 왜냐하면 특징점은 내림차순 형태로 정렬되어 있기 때문이다.

초기 추측 후에, 나머지 특징점에 대해 변이 값을 결정해야 한다. 정합의 정확도를 높이기 위해 특징점 간의 기하학적인 특성을 사용하기 위해 특징점 사이의 거리 값을 계산한다. 그리고 이 두 특징점을 서로 연결한다. 이를 특징점 연결 (Feature

link)라 정의한다. 좌우 영상의 특징점 간의 거리가 같고, 수식 1 을 통해 계산된 값이 특정 문턱치보다 작다면, 이 특징점들은 서로 대응관계에 있다고 판단하여 변이 값을 결정한다. 만약 거리가 같지 않다면, 특징점을 더 연결한다. 특징점 연결로 거리가 가까진다면, 연결되어 있는 최초와 마지막에 특징점에 변이 값을 결정한다. 이 과정은 그림 1 에 설명되어 있다.

이 때, 변이 값을 부화소 단위의 값으로 계산하기 위하여 2 차 함수 곡선 일치법을 사용한다. 이 때 사용한 수식은 아래처럼 정의된다.

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_{j-1}^2 & x_{j-1} & 1 \\ x_j^2 & x_j & 1 \\ x_{j+1}^2 & x_{j+1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} f(x_{j-1}, x_i) \\ f(x_j, x_i) \\ f(x_{j+1}, x_i) \end{bmatrix} \quad (5)$$

$$x_j' = -\frac{b}{2a}, \quad d = x_j' - x_i \quad (6)$$

x_i 는 우영상의 특징점 위치를 말하고, x_j 는 좌영상의 특징점 위치를 의미하며, x_j' 는 2 차 함수 곡선 일치법을 사용하여 계산된 부화소 단위의 화소 값이다.

2.3 변이 보간

위의 방법을 이용하면 특징점 연결의 최초와 마지막 특징점에 대해서만 변이 값을 가지게 되므로 최초와 마지막 특징점 사이의 존재하는 특징점 모두에 대해 변이 값을 보간(interpolation)한다. 이를 위해 양선형 보간 방법(bilinear interpolation)을 사용한다.

	단순 특징점 정합	보간 전 특징점 연결 정합	보간 후 특징점 연결 정합
Teddy	13.05	13.50	13.70
Aloe2	20.32	20.63	20.72
Cloth1	19.08	19.88	19.90
Lampshade	19.05	20.08	20.02
Rocks2	19.61	20.76	20.76
Average	18.22	18.97	19.02

표 2. 제안된 알고리즘의 수행 시간 분석 (ms)

3. 실험 결과 및 분석

실험은 Middlebury 에서 제공되는 스테레오 영상과 정답 영상을 사용하여 논문에서 제안하는 알고리즘의 정확성을 비교해보았다. 이를 위 해 Teddy (2002), Aloe2, Cloth1, Lampshade, Rocks2 (2006) 5 개의 영상을 사용하였다. 창 의 크기는 3 으로 정의하였고, 우 영상에서 좌 영상 방향으로 스테레오 정합을 수행하였다.

실험은 알고리즘으로 수행하여 얻어진 변이 값과 정답 영상의 변이 값을 비교하여 정확도를 계산하였다. 계산된 변이 값이 부화소 단위이므로 오차 범위를 0.5, 0.75, 1.0 에 대해 비교해보았다. 또한 제안한 알고리즘의 성능을 비교하기 위해 단순한 특징점 정합과 결과를 비교하였다. 이 결과는 표 1 과 그림 2 에 나타나 있다.

결과를 보면 알 수 있듯이, 제안한 방법은 단순한 정합 방법보다 정합의 높은 정확성을 가진다. 그림 2 을 통해 단순 특징점 정합 방법에서 나타나는

	좌영상 특징점 개수	우영상 특징점 개수	단순 특징점 정합			보간 전 특징점 연결 정합			보간 후 특징점 연결 정합		
			오차 범위			오차 범위			오차 범위		
			0.5	0.75	1.0	0.5	0.75	1.0	0.5	0.75	1.0
Teddy	1071	1202	85.43%	90.07%	92.05%	90.08%	92.56%	94.21%	90.98%	93.23%	94.70%
Aloe2	995	1029	99.67%	99.67%	99.67%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Cloth1	928	875	99.25%	99.25%	99.62%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Lamp	1132	980	91.84%	91.84%	98.64%	95.27%	95.27%	99.32%	94.67%	94.67%	99.33%
Rocks2	1025	989	97.66%	97.66%	98.83%	97.93%	97.93%	99.31%	97.93%	97.93%	99.31%
평균	1030.2	1015	94.77%	95.70%	97.76%	96.66%	97.15%	98.57%	96.72%	97.17%	98.67%

표 1: 단순 특징점 정합과 특징점 연결 정합의 정확도 분석


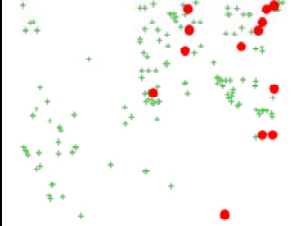
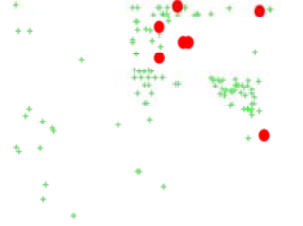
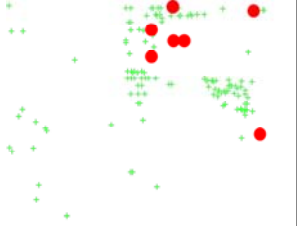

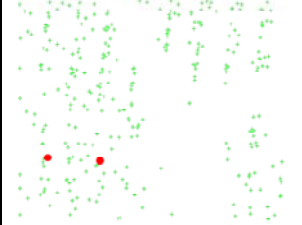


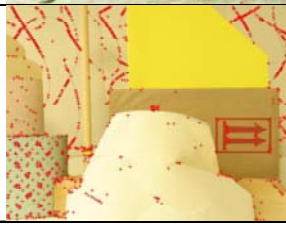
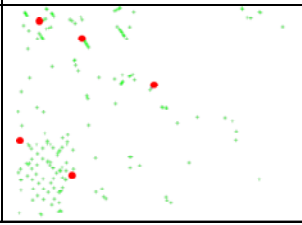
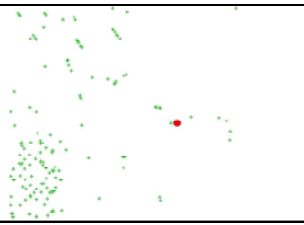
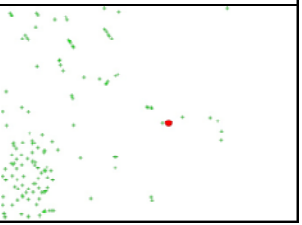
특징점과 원영상	단순 특징점 정합	보간 전 특징점 연결 정합	보간 후 특징점 연결 정합
			
			
			

그림 2: Teddy, Aloe2, Lampshade [9] 영상에서 오차 범위가 1.0일 때, 단순 특징점 정합과 특징점 연결 정합의 결과. 녹색은 정답 영상과 비교하였을 때, 정확한 지점을 의미하고, 빨간원은 잘못된 값을 가지는 위치를 의미한다.

세밀한 오류를 제거하는데 효과적인 것을 알 수 있다. 또한, 정합 수행 시간은 고속으로 가능하다는 것이 표 2 에 나타나있다. 이는 실시간 스테레오 정합 시스템에 적합하다.

4. 결론

본 논문은 고속으로 수행되고, 높은 정확성을 갖는 특징점 연결 스테레오 정합 방법을 제안하였다. 제안한 방법은 세 가지 제약 조건을 사용하여 정확성을 높일 수 있었으며, 세밀한 영역의 오차를 줄일 수 있었다. 또한, 보간법을 사용하여 유효한 변이 값을 증가하였다. 제안한 방법은 빠른 수행 시간과 높은 정확성을 가지고 있으므로 실시간으로 3 차원 정보를 계산해야 하는 시스템에 응용될 수 있다.

감사의 글

본 연구는 지식경제부 및 정보통신산업진흥원의 로봇 비전 연구의 한계 상황 돌파를 위한 핵심 기술 개발 지원사업의 연구결과로 수행되었음.

참고문헌

[1] A Saaidi, H. Tairi and K. Satori, "Fast Stereo Matching using Rectification and Correlation Techniques", ISCCSP 2006.
 [2] Gerard Medioni and Ramakant Nevatia, "Matching

Images Using Linear Features", IEEE Transaction on Pattern Analysis and Machine Intelligence, Nov. 1984.
 [3] Sekhavat D. Sharghi and Farhad A. Kamangar, "Geometric Feature-Based Matching in Stereo Images", IDC 1999.
 [4] Hajar Sadeghi, Payman Moallem, and S. Amirhassan Monadjemi, "Feature Based Dense Stereo Matching using Dynamic Programming and Color", International Journal of Computational Intelligence Vol. 4, 2008.
 [5] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts", IEEE Tran. Pattern Analysis and Machine Intelligence, vol. 23, pp. 1222-1239, 2001
 [6] Qingxion Yang, Liang Wang, Ruigang Yang and Henrik Stewenius and David Nister, "Stereo Matching with Color-weighted Correlation, Hierarchical Belief Propagation, and Occlusion Handling", IEEE Transaction on Pattern Analysis and Machine Intelligence, March 2009.
 [7] I. Cox, S. Hingorani and S. Rao, "A Maximum Likelihood Stereo Algorithm", Computer Vision and Image Understanding, Vol. 63, No. 3, May 1996.
 [8] Edward Rosten and Tom Drummond "Machine Learning for High-speed Corner Detection", European Conference on Computer Vision, May 2006.
 [9] Stereo datasets with ground truth image, Middlebury college, <http://cat.middlebury.edu/stereo/data.html/>