

특징창과 특징링크를 이용한 스테레오 특징점의 정합 성능 향상

김 창 일[†] · 박 순 용^{††}

요 약

스테레오 정합(stereo matching) 기술은 주어진 두 영상에서 동일한 물체의 영상점이 어떤 위치 관계를 가지고 있는지를 결정하는 기술이다. 본 논문에서는 영상 특징점에 대해 스테레오 위치관계를 결정하는 새로운 스테레오 특징점 정합(stereo feature matching) 방법을 제시한다. 제안하는 방법은 주어진 스테레오 영상에서 FAST 추출기를 이용하여 특징점을 추출하고, 특징점 벡터들의 정보들을 내부에 포함하는 특징창(feature window)이라는 공간을 정의하여 스테레오 정합의 성능을 향상한다. 제안하는 방법은 표준 영상에 추출된 특징점들에 대해 특징창을 생성하고, 참조 영상에서 표준 영상의 특징창과 가장 유사한 특징창을 탐색 및 결정한다. 결정된 두 개의 특징창 내부의 특징점들의 시차관계는 특징링크(feature link)를 생성하여 시차를 결정한다. 만약, 이 과정에서 시차가 결정되지 않은 특징점들이 있다면, 특징창 내의 결정된 시차 정보를 이용하여 시차 값을 보간한다. 마지막으로, 제안하는 방법의 성능을 검증하기 위해 결과 영상과 정답 영상의 시차를 비교하여 정합 정확성과 수행시간을 비교하였다. 또한, 기존의 특징점 기반 스테레오 정합 방법들과 제안하는 방법의 성능을 비교 및 분석하였다.

키워드 : 스테레오 정합, 특징점 정합, 특징창, 특징링크

Enhancement of Stereo Feature Matching using Feature Windows and Feature Links

Chang-Il Kim[†] · Soon-Yong Park^{††}

ABSTRACT

This paper presents a new stereo matching technique which is based on the matching of feature windows and feature links. The proposed method uses the FAST feature detector to find image features in stereo images and determines the correspondences of the detected features in the stereo images. We define a feature window which is an image region containing several image features. The proposed technique consists of two matching steps. First, a feature window is defined in a standard image and its correspondence is found in a reference image. Second, the corresponding features between the matched windows are determined by using the feature link technique. If there is no correspondence for an image feature in the standard image, its disparity is interpolated by neighboring feature sets. We evaluate the accuracy of the proposed technique by comparing our results with the ground truth of in a stereo image database. We also compare the matching accuracy and computation time with two conventional feature-based stereo matching techniques.

Keywords : Stereo Matching, Feature Matching, Feature Window, Feature Link

1. 서 론

스테레오 정합(stereo matching) 기술은 인간이 양안 시차를 가지고 있는 것에서 착안되어 컴퓨터 비전 연구에서

많은 연구가 진행되고 있는 분야이다. 스테레오 정합 기술은 3차원 장면 복원, 3차원 물체 복원, 로봇 비전 등 여러 분야에 응용되는 기반 기술이다. 스테레오 정합은 스테레오 카메라로 획득한 두 영상에서 동일한 물체의 영상점을 찾는 것을 말한다. 이 때, 두 동일한 영상점 간의 대응 관계의 차이를 시차(disparity)라고 한다. 그리고 이 정보를 영상화한 것을 시차 지도(disparity map)라고 한다. 스테레오 정합을 위해 흔히 스테레오 좌, 우 영상 중 정합의 기준이 되는 영상을 표준 영상(standard image)이라 하고, 표준 영상과 비교를 위해 사용되는 영상을 참조 영상(reference image)으로

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 로봇 비전 연구의 한계 상황 돌파를 위한 핵심 기술 개발 지원사업의 연구결과로 수행되었음(NIPA-2011-C7000-1101-0006).

† 준 회원 : LIG넥스원 연구원

†† 정 회원 : 경북대학교 컴퓨터학부 부교수

논문접수 : 2011년 6월 13일

수정일 : 1차 2011년 9월 22일, 2차 2011년 10월 17일

심사완료 : 2011년 10월 18일

구분한다. 즉, 스테레오 정합은 표준 영상의 동일한 영상점을 참조 영상에서 찾는 것을 의미한다.

스테레오 정합은 시차 지도의 획득 방법에 따라 일반적으로 두 가지로 나눌 수 있다. 첫째는 영역 기반 방법(area based method)으로 영상 전체의 정보를 이용하여 시차 값을 구하는 방법이다[1]. 따라서 이 방법의 결과로 조밀한 시차 지도(dense disparity map)를 얻을 수 있다. 하지만, 이 방법은 전체 영상에 대해 대응되는 시차를 탐색하므로 기본적으로 많은 수행 시간을 요구한다.

두 번째 방법은 특징 기반 방법(feature based method)으로 위의 방법과는 달리 영상에서 특징이 두드러진 영역을 기반으로 시차를 구하는 방법이다[2,3]. 이런 이유로 영상에서 특징점을 추출하는 특징 추출기(feature detector)의 사용이 추가적으로 필요하다. 이 때, 주로 많이 이용되는 영상의 특징은 에지(edge)나 코너(corner) 등이 있다. 일반적으로 특징 기반 방법은 추출된 특징점의 시차를 먼저 결정 한 후에 결정되지 않은 영역의 시차들을 특징점들의 시차 값을 이용하여 추정한다[2,3]. 또한, 특징 기반의 스테레오 정합 기술에서는 특징점 간의 거리 및 각도 등 기하학적인 정보를 추가적인 제약조건으로 이용할 수 있다[4,5,6].

스테레오 정합에서는 표준 영상과 참조 영상의 상관관계(correlation)를 판단하는 것 역시 중요하다. 이 때, 상관관계를 판단하는 함수를 상관 함수(correlation function)이라 하며, 이를 구하는 방법 역시 크게 두 가지 형태로 구분된다. 첫째로, 창 함수 방법(window method)[1,3,4]은 두 영상의 일부분을 특정 크기의 창의 크기만큼 가져와서 이를 SAD(Sum Absolute Difference), SSD(Sum Squared Difference), NCC(Normalized Cross Correlation) 등의 방법을 이용하여 상관관계 값을 계산한다[1]. 이 방법은 간단하면서 수행 시간은 오래 소요되지 않는다.

두 번째 방법은 영상을 데이터 항(data term)과 연속성(continuity)으로 구성된 에너지 함수를 최소화 하는 방법이 있다[7,8]. 이 방법은 에너지 최소화를 위해 영상의 전역 정보를 반복적으로 탐색한다. 이런 이유로, 반복 횟수가 높을수록 정확성이 높아지지만, 그에 따른 추가적인 수행 시간이 소요된다. 이 방법을 이용한 예로는 그래프 컷(graph-cut)[7], 신뢰 전파(belief propagation)[8] 방법 등이 있다.

하지만, 앞서 설명한 두 스테레오 정합 방법은 전체 영상에 대해 시차를 결정하므로 많은 수행 시간이 요구되어 로봇의 3차원 지도 생성과 위치 인식 등과 같이 실시간으로 3차원 정보를 계산하고 처리하는 경우에 사용하기 어렵다. 이런 이유로 본 논문에서는 영상에서 추출된 특징점만의 시차 정보를 추정하는 스테레오 특징점 정합(stereo feature matching) 기술을 제안한다[6]. 기존의 스테레오 정합이 조밀한 시차정보를 획득하는데 주목적이 있었다면 본 논문에서는 영상 특징점에 대해서만 시차를 계산하는 것을 주목적으로 한다. 이러한 스테레오 정합 기술은 영상의 모든 영역에 대한 거리 정보를 필요로 하지 않고 특징점의 거리 정보

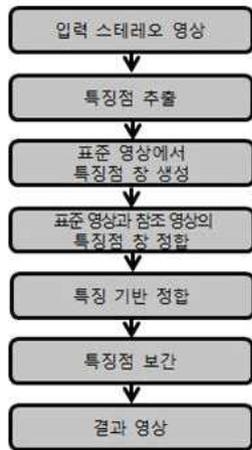
만으로도 충분한 응용분야에 사용될 수 있다. 예를 들어, 스테레오 비전 기술을 이용한 로봇의 자율 주행의 경우, 스테레오 영상의 모든 점에 대한 거리 정보보다는 특징점까지의 거리를 빠른 시간에 계산하고 로봇의 위치를 추정 한 다음 특징점을 마일스톤 점(milestone point)으로 저장하는 기술이 필요하다. 또한, 로봇의 주행 경로에 대한 3차원 지도 정보와 획득된 스테레오 영상에서의 3차원 정보를 실시간으로 비교하여 현재 위치를 추정하는 기술에 적용할 수 있다.

이전에 연구된 단순 스테레오 특징점 정합 방법은 표준 영상에서 추출된 특징점들과 참조 영상의 추출된 특징점들의 대응관계를 상관 함수 값을 이용하여 판단한다. 하지만, 이 방법은 표준 영상과 참조 영상의 특징 추출 지점이 완전히 동일하지 않거나 스테레오 영상에서 추출된 특징점 수가 다르게 추출된다면 정합 오류가 발생할 위험이 크다. 따라서 이런 오류를 개선한 특징링크(feature link) 방법은 수평 위치가 동일한 특징점 간의 거리 정보와 컬러 정보를 이용하여 단순 특징점 정합에서 나타나는 문제를 개선하여 높은 정확성을 보장한다[8]. 하지만, 이 방법은 빠른 속도로 특징점의 시차를 정확히 결정하지만, 스테레오 영상에서 추출된 특징점의 개수에 비해 정합되는 특징점의 수가 많지 않은 단점이 있다.

본 논문에서는 이전에 연구된 스테레오 특징점 정합 방법에서 특징점의 정합되는 수를 증가시키면서 정확성을 개선한 새로운 스테레오 특징점 정합 알고리즘을 제시한다. 제안하는 방법은 특징창(feature window)이라는 새로운 공간을 정의하여, 그 내부에 있는 특징점들의 수직 및 수평 위치 정보들을 이용하여 특징점들을 정합한다.

제안하는 스테레오 정합 방법에 사용되는 영상은 정류화(rectification)가 되어 있다고 가정한다. 먼저, 스테레오 표준 및 참조 영상에서 특징점을 특징 추출기를 이용하여 추출한다. 다양한 특징 추출기들 중에 FAST 추출기를 사용하여 빠른 속도로 영상에서 특징점들을 추출한다[9]. 특징점이 추출된 표준 영상에서 특징창을 생성하고, 참조 영상에서 표준 영상의 생성된 특징창과 가장 유사한 특징창을 최소 시차에서 최대 시차 사이에서 검색한다. 다음으로 이 과정으로 결정된 두 개의 특징창 내부의 대응되는 위치의 특징점들의 시차 값을 추정한다. 이 때, 특징링크(feature link) 방법을 이용하되 수직 성분도 추가적으로 고려하여 특징점들의 시차를 결정한다[6]. 마지막으로 표준 영상의 특징창 내부의 시차가 결정되지 않은 특징점들의 시차를 보간한다. 또한, 스테레오 정합의 정확성 향상을 위해 추가적으로 전역 특징점 지도 정보를 사용한다. (그림 1)은 제안하는 알고리즘의 전체 순서를 나타낸다.

본 논문의 구성은 다음과 같다. 2절에서는 제안하는 알고리즘의 과정의 과정을 설명하고, 3절에서는 본 논문에서 실험을 통해 제안한 방법의 성능과 예전에 연구되었던 방법들을 비교하였다. 마지막 4절에서는 앞의 내용들을 정리하여 결론을 맺는다.



(그림 1) 제안하는 방법의 순서도

2. 제안하는 새로운 특징 기반 스테레오 정합 방법

본 연구에서는 스테레오 영상 정합을 위해 특징점을 이용하므로, 먼저 주어진 스테레오 영상의 특징점을 추출해야 한다. 다음으로, 추출된 특징점들의 기하학적인 정보를 이용하기 위해 특징점 정보를 특징창이란 불리는 공간 내에서 특징점들의 위치 정보를 판단한다. 표준 영상에서 특징창을 생성한 후, 참조 영상에서 표준 영상에서 생성된 특징창과 서로 대응되는 특징링크를 사용하여 스테레오 특징점 정합을 수행한다. 이 후에 시차를 결정하지 못한 특징점은 보간을 통해 시차 값을 결정한다. 추가적으로, 제안하는 방법의 정확성을 향상시키기 위해 전체 영상에 대해 특징점 위치 정보를 가지고 있는 전역 특징점 지도를 함께 이용한다.

2.1 스테레오 정합의 제약 조건

스테레오 영상에서 정합의 정확성을 높이고, 정합 수행 시간을 단축하기 위해 몇 가지 정합의 제약을 이용할 수 있다. 대표적으로 많이 이용되는 제약 중 하나는 에피폴라 제약(Epipolar constraint)이다[1,2,3,4]. 에피폴라 제약은 스테레오 영상에서 대응하는 두 점을 동일 선상에 위치하도록 해준다. 이를 이용하여 스테레오 영상의 대응 영역을 탐색할 경우, 동일한 선상에 해당하는 영역만 비교하여 연산 수행 시간을 줄일 수 있으며, 대응되는 위치가 보다 에피폴라 제약을 이용하기 전보다 정확히 유지된다.

다음으로 많이 사용되는 제약은 순서 제약(ordering constraint)로 표준 영상에서 특징이 나타나는 순서가 참조 영상에서도 동일한 순서로 나타나는 것을 의미한다[11]. 따라서 특징점을 스테레오 정합할 때, 낮은 인덱스 값에서 높은 인덱스 값 방향으로 정합의 순서를 고려하면 된다.

마지막으로 사용되는 시차 제약(disparity constraint)는 스테레오 영상의 시차 값의 범위는 특정 범위 내에 존재한다는 것으로, 보통 최소 시차 값과 최대 시차 값을 지정하여 그 범위 안에 해당하는 값만을 해당 스테레오 영상에서 유효한 시차로 간주한다[10].

2.2 특징점 추출

영상에서 두드러진 특징을 가지는 영역을 추출을 위해 많은 알고리즘이 사용된다. 이러한 특징 추출 알고리즘 중에서 FAST 알고리즘을 사용한다[9]. FAST 알고리즘은 현재 픽셀과 주변 영역의 n 개의 픽셀의 밝기 값을 비교하여, 그 밝기 차이가 특정 문턱값(threshold)보다 클 때, 현재 픽셀을 영상에서 특징점으로 결정한다. 이 방법은 다른 알고리즘들보다 매우 빠른 속도로 특징점을 추출할 수 있으며, 또한 특징점 추출 시에 비교하는 픽셀의 수에 따라 FAST 추출기의 성능이 달라진다. 이 방법으로 추출된 특징점들의 결과는 오름차순 형태로 저장되는데, 이런 이유로 FAST 추출기를 이용하면 스테레오 순서 제약을 만족시킨다.

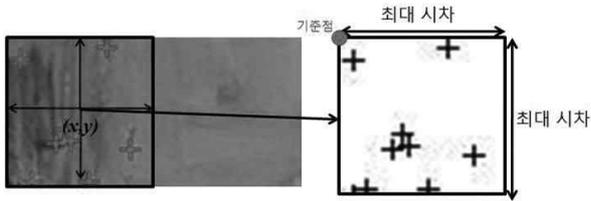
2.3 표준 영상에서 특징창 생성

스테레오 특징점 정합은 스테레오 영상의 특징점에 대해서만 정합하므로 영역 기반 방법보다 기본적으로 높은 정확성을 가진다[5]. 하지만 스테레오 특징점 정합 방법에서 정합 오류가 발생할 수 있다. 정합 오류가 발생하는 원인은 크게 세 가지로 분류할 수 있는데, 첫째로 어떤 추출된 특징점이 표준 영상에서는 나타나지만 참조 영상에서는 나타나지 않는 경우이다. 둘째로, 특정 패턴이 반복되는 영역일 경우에는 특징점들 간의 대응 관계를 정확히 결정하는 것이 어렵다. 마지막으로, 특징 추출기의 오류로 특징점의 위치를 잘못 결정할 경우이다. 이를 해결하기 위해서 추출된 특징점들 간의 길이, 각도 등 기하학적인 정보를 사용하여 이런 오류를 어느 정도 제거할 수 있다[2].

이전 연구들 중에 특징링크 방법은 동일한 에피폴라 선상에 존재하는 특징점 간의 길이 정보를 이용하여 스테레오 특징점 정합의 발생하는 오류를 줄인다[6]. 하지만 이 방법은 추출된 특징점 수에 비해 정합된 유효한 특징점 수가 적다. 그 이유는 특징링크 방법은 오로지 특징점 간의 수평 거리만을 고려하기 때문이다. 또한, 정확하게 스테레오 영상이 정렬되지 않은 영상에서는 정합의 좋은 결과를 기대하기 어렵다.

따라서, 특징링크 방법에서 고려되지 않은 수직 정보를 이용하기 위해 특징점 간의 기하학적인 정보 특히 수직 성분을 추가적으로 고려해야 한다. 이 때, 수직적인 정보와 수평적인 정보의 어느 범위 내에서 고려하는지가 정합에 있어서 중요한 요소가 된다. 따라서 이런 정보를 고려한 특징창(feature window)이라는 불리는 공간을 정의한다. 특징창은 (그림 2)에서와 같이 특정 영상 좌표 (x, y) 를 기준으로 스테레오 영상의 최대 시차 값 크기를 가지는 정사각형 모양으로 생성된다. 특징 창을 최대 시차로 한 것은 특징 창 내의 시차 값들이 시차 제약을 만족하게 하기 위함이다.

(그림 2)에서 생성된 내부에 특징점 정보들은 기준점을 토대로 2차원 벡터 형태로 저장된다. 특징창은 본 논문에서 제안하는 특징점 스테레오 정합의 기본 단위로 사용한다. 즉, 표준 영상에서 생성된 특징창 내부의 특징점 벡터들과 참조 영상에서 표준 영상과 가장 유사한 특징창 내부의 특징점 벡터들을 서로 비교 및 정합하여 특징점들의 시차 값을 결정한다.



(그림 2) 특정 위치 (x, y) 에서 특징창 생성 과정과 생성된 특징 창

2.4 표준 영상과 참조 영상의 특징창 정합

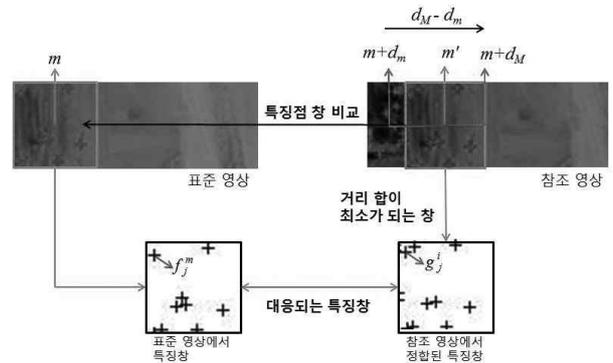
표준 영상의 특징창 내부의 특징점을 정확하게 정합하기 위해서는 참조 영상에서 표준 영상과 가장 유사한 특징창을 찾는 것이 중요하다. 표준 영상에서 특징창이 결정된 후에 참조 영상에서 표준 영상의 최소 시차에서 최대 시차만큼 이동하면서 특징창을 생성한다. 그리고 생성된 표준 영상과 참조 영상 각각에서 내부의 특징창의 특징점 개수를 비교하여 차이가 2 이하로 나타나면서 각 특징창 내부의 대응되는 위치에 특징점의 거리의 차이를 계산하여 그 값을 누적한다. 이 내용은 수식 1로 표현할 수 있다.

$$m' = \underset{m+d_m \leq i < m+d_M}{\operatorname{argmin}} \left(\sum_{j=1}^k \sqrt{(f_j^m - g_j^i) \cdot (f_j^m - g_j^i)} \right) \quad (1)$$

$$k = \operatorname{mim}(n_1, n_2) \quad (2)$$

수식 1에서 m 은 표준 영상에서 특징창을 생성한 x 방향으로 위치를 의미하며, d_m 은 최소 시차를, d_M 은 최대 시차를 의미한다. k 는 비교할 특징점 벡터의 개수이며 수식 2와 같이 표준 영상의 특징창의 특징점 벡터 수 n_1 과 참조 영상의 특징점 벡터 수 n_2 중에 작은 값을 선택한다. f 는 표준 영상의 특징창 내부의 특징 벡터를 의미하고, g 는 참조 영상에서 특징 벡터를 의미한다. 그리고 f^m 와 g^i 는 각각 표준영상과 참조영상에서 x 위치가 m 과 i 인 특징창에서 j 번째 특징 벡터를 말한다. 즉, 수식 1을 이용하면 표준영상의 특징창의 내부에 있는 특징점들의 벡터 거리 차이가 최소가 되는 참조영상의 특징창을 구할 수 있다. 이때 참조영상에서 정합된 특징창의 x 위치는 m' 가 된다. 이 과정은 (그림 3)에 나타내었다.

특징창을 이용하여 전체 영상의 스테레오 특징점 정합을 위해 특징창은 영상의 왼쪽에서 오른쪽으로, 위쪽에서 아래



(그림 3) 표준 영상에서 생성된 특징창과 참조 영상에서 거리 합 비교를 통한 특징창 정합

쪽으로 이동하면서 진행한다. 이 경우에 특징창의 이동 간격의 크기에 따라 제안하는 방법의 정확성에 영향을 준다. 왜냐하면, 특징창의 이동 간격을 최대 시차 이상으로 할 경우는 특징창의 이동 중에 특징점이 중복적으로 정합될 일이 없기 때문이다. 즉, 이러한 경우에는 특징창이 한번 생성된 영역 내의 특징점 벡터는 이후에 다시 시차를 추정할 수 없다. 하지만, 특징창의 이동 간격이 최대 시차 값보다 작을 경우 그 간격 크기에 따라 특징창의 정합 오류로 인해 계산되지 못한 특징점의 시차 값을 결정될 수 있으며, 또한 결정된 시차 값을 가지고 있는 특징점은 시차 값을 재결정하여 보다 정확히 스테레오 특징점 정합을 수행할 수 있다. 그러나 이 경우에는 이동 간격이 작을수록 스테레오 특징점 정합에 추가적인 수행 시간이 소요된다. 반대의 경우는 수행 시간은 적게 소요되지만, 정합의 정확성을 재검증하지 않으므로 정확성은 낮아지며, 정합되는 특징점들의 수 역시 감소한다.

2.5 대응 특징창 내의 특징점들의 정합

앞의 절에서 결정된 대응하는 두 개의 특징창 내부의 특징점들을 정합해야 한다. 이를 위해 특징점 간의 상관관계를 규명해줄 수 있는 상관 함수의 정의가 필요하다. 이를 위해 제안하는 방법에서는 창 함수 방법을 이용한다[1]. 이 방법을 이용한 예전에 많은 연구에서는 주로 회색 영상을 이용하여 이 방법을 적용했다[4]. 하지만, 회색 영상은 단지 컬러 영상 정보의 평균값으로 구성되기 때문에 대응 관계를 정확히 판별하는데 문제가 있다[3]. 이런 이유로 최근에 컬러 정보를 이용한 연구가 많이 이루어지고 있다[3,8]. 본 논

$$f(x, y) = \frac{1}{n^2} \sum_{i=-n/2}^{n/2} \sum_{j=-n/2}^{n/2} \operatorname{dist}_c(C_R(x+i, y+j), C_L(x+i, y+j)) \quad (3)$$

$$\operatorname{dist}_c(C_R, C_L) = (R_R - R_L)^2 + (G_R - G_L)^2 + (B_R - B_L)^2 \quad (4)$$

$$C_L(x, y) = (R_L(x, y), G_L(x, y), B_L(x, y)) \quad (5)$$

$$C_R(x, y) = (R_R(x, y), G_R(x, y), B_R(x, y)) \quad (6)$$

문에서도 역시 컬러 정보를 이용한 상관 함수를 아래와 같이 정의하여 사용한다. 이 함수는 RGB 컬러 채널을 이용한 최소 제곱 오차(Mean Square Error) 방식이다[3].

식 (3)에서 n 은 상관 함수에 필요한 창 크기를 의미하며, d 는 스테레오 영상에서 결정된 시차 값을 의미하며, C_R 과 C_L 은 표준 영상과 참조 영상의 특정 위치의 RGB 성분 값이다. 이 함수의 값의 결과 값이 특정 문턱값보다 적을 경우 대응 관계에 있다고 판단한다.

제안하는 방법은 앞에서 결정된 두 대응하는 특징창 내의 특징점들의 시차 결정 방법으로 특징링크 방법을 이용한다. 특징링크는 특징점 간의 연결된 형태를 의미하며, 특징링크 길이가 동일할 때 시차를 결정한다. 따라서 특징링크 방법은 특징점 추출기의 수평 방향으로 발생할 수 있는 오류를 줄인다[6]. 하지만, 특징링크 방법은 특징점 추출기 오류로 인한 미세하게 발생하는 수평 및 수직 방향의 특징점 추출 오류를 해결할 수 없다. 이런 이유로, 정합되는 특징점들의 수가 감소한다. 이를 해결하기 위해, 제안하는 방법에서는 수평 및 수직 방향의 특정 오류 허용 범위 내에 존재하는 특징점들 역시 정합될 수 있도록 한다.

특징점들의 정합을 위해 정합하려는 특징점의 현재 수직 위치에서 아래, 위로 v 차이 내에 존재하는 특징점의 개수를 파악한다. 이 때, v 값을 수직 허용 범위라고 정의한다. 만약, 수직 허용 범위 내의 특징점 개수가 1개라면, 해당하는 특징점 간의 상관 함수 값을 이용하여 대응 관계를 파악한다. 하지만, 표준 영상과 참조 영상의 특징점 개수가 2개 이상이 될 경우, 인접 특징점과 연결하여 그 길이를 계산한다. 이는 (그림 4)에 제시하였다.

이 때, 각각의 특징링크 길이를 L_{link} , R_{link} 라고 한다. 이 거리가 동일하지 않고, 그 차이가 h 이하라면, 수식 7을 이용하여 링크 길이가 동일한 값을 가지도록 재조정한다. 이 때, h 값을 수평 허용 범위라고 정의한다.

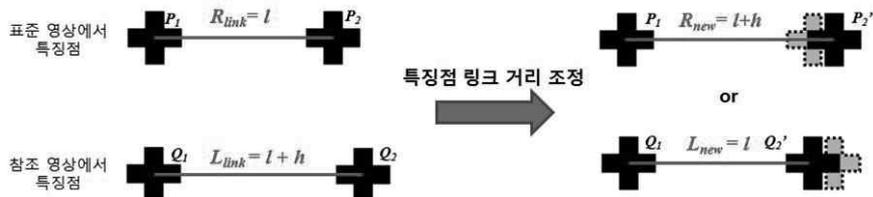
$$\begin{aligned} \theta_1 &= f(P_2', Q_2) \\ \theta_2 &= f(P_2, Q_2') \end{aligned} \tag{7}$$

여기서 f 는 수식 3에 정의된 상관 함수를 의미하며, θ_1 는 특징링크를 확장한 후에 대응되는 영역과의 상관 함수 값을 의미한다. 반대로, θ_2 는 특징링크를 축소한 후의 대응되는 영역과의 상관 함수 값을 의미한다. θ_1 이 θ_2 보다 작으면서 특정 문턱값보다 작을 경우, 특징링크의 길이는 강제로 확장되고, 그 반대의 경우는 특징링크의 길이를 강제로 축소시킨다. 이 과정을 특징링크의 길이 조정이라고 한다. 이 방법은 특징점 추출의 오류로 유사 특징점의 위치가 미세하게 잘못 나타난 경우의 문제들을 해결할 수 있기 때문에 정합의 정확성과 정합될 특징점의 수를 증가시킨다.

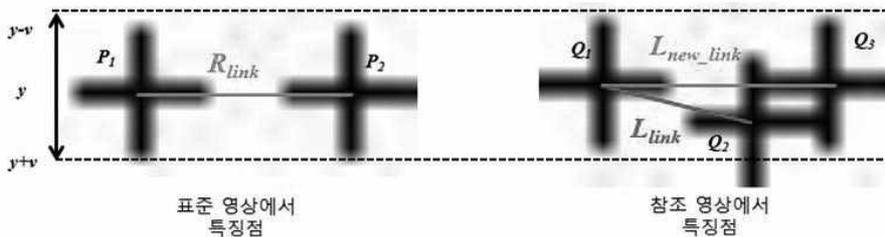
만약, 특징링크 길이의 차이가 수평 허용 범위보다 크게 나타나는 경우에는 (그림 5)에서 같이 P_1 과 P_2 간의 특징링크 R_{link} 와 Q_1 과 Q_2 간의 특징링크 L_{link} 간의 길이 차이를 비교한다. 이 때, 길이 차이가 수평 허용 범위보다 클 때, 링크 길이 값이 작은 L_{link} 를 특징점 Q_1 을 Q_2 가 아닌 Q_3 와 연결시켜 길이를 확장하여 특징링크 L_{new_link} 가 되도록 길이를 조정한다[6]. 이 방법을 특징링크 확장이라고 하며, 이와 같은 방법으로 특징링크 간의 길이 차이가 수평 허용 범위 이하가 될 때까지 특징링크 길이를 확장시킨다.

앞의 방법을 통해 특징링크의 길이가 동일하고, 두 특징링크의 처음과 마지막 특징점들의 상관 함수 값이 특정 문턱값보다 작게 나타날 경우에만 처음과 마지막 특징점들의 시차를 결정한다. 따라서 제안하는 방법을 사용하면 대응 특징점 간의 위치가 동일하게 조정된 후에 시차를 결정하므로 보다 정확한 특징점의 시차를 결정할 수 있다.

위에서 결정된 특징링크의 시차를 부화소(sub-pixel) 단위로 계산하기 위해 2차 함수 곡선 일치법 (2nd order curve fitting)을 수식 (8)을 이용하여 계산한다.



(그림 4) 특징링크의 길이가 수평 허용 범위 내에 있을 때, 특징링크 거리 조정



(그림 5) 특징링크 길이 차이가 수평 허용 범위보다 클 때, 특징링크의 확장

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} x_{j-1}^2 & x_{j-1} & 1 \\ x_j^2 & x_j & 1 \\ x_{j+1}^2 & x_{j+1} & 1 \end{bmatrix}^{-1} \begin{bmatrix} f(x_{j-1}, x_i) \\ f(x_j, x_i) \\ f(x_{j+1}, x_i) \end{bmatrix} \quad (8)$$

$$x'_j = -\frac{b}{2a} \quad (9)$$

$$d = x'_j - x_i \quad (10)$$

x_i 는 우영상의 특징점 위치를 의미하며, x'_j 는 좌영상의 특징점 위치를 의미한다. 이를 통해 계산된, x'_j 는 2차 함수 곡선 일치법을 사용하여 계산된 부화소 단위의 값이다. 이를 통해 시차를 계산하게 되면 일반 화소 단위보다 정확한 시차 값을 얻을 수 있다.

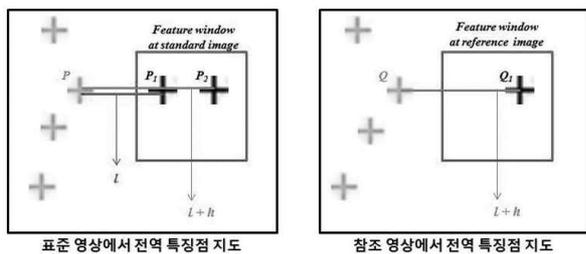
위의 방법으로 특징창 내의 특징점들을 정합 할 경우에 상관 함수 값에 의해 또는 특징링크의 길이 등으로 인해 시차가 결정되지 않은 특징점은 보간을 통해 시차 값을 추정해준다.

2.6 전역 특징점 지도

스테레오 특징점 정합의 정확성을 높이기 위해 추가적으로 전역 특징점 지도(global feature map)를 이용한다. 이는 주어진 스테레오 영상과 동일한 크기를 가지며, 최초로 스테레오 영상의 FAST 추출기를 이용하여 특징점이 추출될 때 생성한다. 생성된 전역 특징점 지도는 FAST 추출기로 추출된 특징점들 모두에 대한 위치 정보를 가진다. 이 전역 특징점 지도 정보는 특징점들이 시차가 결정될 때마다 표준 및 참조 전역 특징점 지도 각각에 대응된 특징점 위치 정보를 기록한다.

제안하는 방법에서 전역 특징점 지도 정보는 정합을 하려고 하는 특징점을 선택하는 과정에서 이용한다. 예를 들면, 스테레오 영상이 (그림 6)과 같이 유사 영역이 반복되는 지점이 있다면, 인근 영역의 영상 정보 값이 모두 유사하게 나타난다. 이런 이유로 상관 함수만을 이용한다면, 대응되는 특징점의 정확한 위치를 판단하는데 어려움이 있다. 이러한 경우에 전역 특징점 지도의 정보를 이용하면 이러한 문제를 해결할 수 있다.

(그림 6)과 같이 현재 정합하려는 특징점 P_1 와 Q_1 이 있을 때, 현재 생성된 특징창 밖에 위치하면서 이전 유사 영역에서 정합된 특징점 P 와 Q 가 있는지를 확인한다. 만약, P 와 Q 가 존재한다면, 특징창 내에 존재하는 동일선 상에서 수직 허용 범위 내에 존재하는 특징점들 각각과 거리를 비교하여,



(그림 6) 전역 특징점 지도 정보를 이용하여 대응 특징점 판단

거리의 차이가 수평 허용 범위 내에 해당하는 특징점을 선택한다. 이 때, 거리의 차이가 있다면, 수식 7의 특징링크 길이 조정이나 특징링크 확장 방법을 통해 동일한 거리 값을 갖도록 한다. 따라서 이 방법을 통해 표준 영상에서 정합하려고 하는 특징점을 참조 영상에서 정확히 대응되는 지점의 특징점을 결정할 수 있다.

2.7 특징창 내의 시차 보간

특징창을 이용한 스테레오 특징점 정합 과정을 통해 시차가 결정되지 않은 특징점들의 시차를 결정하면 보다 많은 수의 시차를 가진 특징점들을 획득할 수 있다. 앞에서 설명했듯이 특징창 내의 시차 값은 시차 제약에 의해 최소 시차에서 최대 시차 사이의 값을 가진다. 따라서 특징창 내의 결정되지 않은 다른 특징점들 역시 최소 시차에서 최대 시차 범위 안에 있는 특징창 내의 특징점들의 결정된 시차 값을 가진다. 이런 이유로 특징창 내의 시차 보간을 위해 현재 특징창에서 결정된 모든 시차의 정보를 가지고 있는 특징점 후보 집합을 생성한다.

다음으로, 전역 특징점 지도의 내용을 확인하여 보간이 필요한 특징점으로부터 최소에서 최대 시차 사이의 특징점이 존재 여부를 검사한다. 만약, 범위 내에 특징점이 존재한다면, 이 시차 값을 특징점 후보 집합에 추가한다. 이는 특징창의 정합 오류로 인해 정합되지 않은 점이 있을 수 있기 때문이다. 이렇게 현재 보간 하려는 특징점과 특징점 후보 집합의 내용을 각각 적용한 시차 값을 모두 적용하여 상관 함수의 값을 모두 계산한다. 이렇게 계산된 상관 함수의 값이 최소가 되면서 문턱값보다 작은 값의 시차 값으로 특징점의 시차를 결정한다.

3. 실험 결과

이번 절에서는 제안하는 방법의 성능을 실험을 통해 검증하였다. 먼저, 제안하는 방법의 정확성을 평가하기 위해 스테레오 특징점 정합 결과에 대한 정답이 제시된 Middlebury에서 제공하는 스테레오 데이터와 정답 영상을 이용한다 [12]. 즉, 결과로 얻어진 영상의 시차 값과 정답 영상의 시차 값을 비교하여 그 차이가 오차 허용 범위(ϵ) 이내라면 정확한 결과라고 판단한다[1]. Middlebury에서 제공하는 스테레오 데이터 중 공개된 연도 순서대로 5가지 Venus, Teddy, Cone, Aloe, Art 영상을 사용하였다[12].

실험 조건은 스테레오 영상의 특징점 수가 1000개 이상 추출되도록 FAST 추출기의 문턱값을 설정하고, 정합의 방향은 표준 영상을 우 영상으로 참조 영상을 좌 영상으로 하였다. 이 때, 참조 영상의 추출된 특징점 수가 표준 영상보다 조금 많게 나타나게 하여 표준 영상에서 추출된 특징점이 참조 영상에서도 나타날 수 있게 하였다. 또한 영상마다 최소 시차와 최대 시차의 값의 차이가 있기 때문에 이를 각 영상의 형태에 따라 달리하였다. 해당 실험 조건은 <표 1>에 제시하였다.

<표 1> 실험 조건

	Venus (2001)	Teddy (2003)	Cone (2003)	Aloe (2006)	Art (2006)
영상 너비	434	450	450	641	695
영상 높이	383	375	375	555	555
표준 영상에서 추출된 특징점 수	1025	1000	1118	1029	1096
참조 영상에서 추출된 특징점 수	1066	1016	1189	1156	1102
최대 시차	20	55	55	110	110
최소 시차	1	14	16	20	40

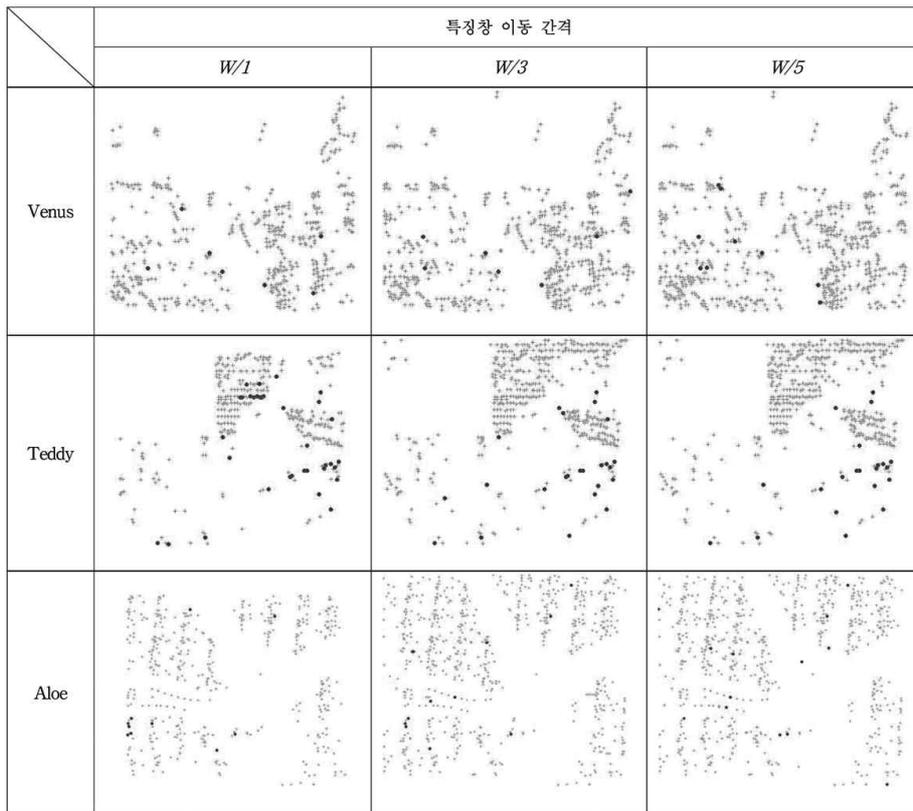
<표 1>에 제시된 것 외에 공통적으로 이용한 조건은 다음과 같다. FAST 추출기의 비교 픽셀 수는 9, 특징링크 연결에서 허용한 수직 허용 범위와 수평 허용 범위는 2, 상관함수의 창 크기 7, 판별 함수의 문턱값은 500이다. 또한, 실험에서 사용한 시스템의 조건은 Intel i5 750 프로세서, 4.00GB 메모리를 가지고, Windows 7 Enterprise K OS에 Microsoft Visual Studio 2008에서 수행하였으며, 추가적으로 OpenCV 2.0 라이브러리를 이용하였다.

제안하는 방법에서 정확성에 가장 큰 영향을 주는 요소인

특징창의 이동 간격을 변화시키면서 정합을 수행하였다. 이때, 시차 값의 오차 허용 범위는 1.0과 2.0 이하로 설정하였다. 제안하는 방법의 결과로 획득된 시차 영상은 (그림 7)에 나타냈고, 정합된 특징점의 수와 정확성, 수행 시간을 <표 2>와 <표 3>에 나타내었다.

<표 2>를 보면, 특징창의 이동 간격이 적을수록, 정합되는 특징점의 수가 늘어났다. 이는 특징창이 이동할 때, 특징창의 정합 오류 등으로 인해 특징창에서 대응되지 않은 특징점 정보가 특징창이 미세하게 이동하면서, 대응되는 특징점을 찾아 정합 또는 보간 되었기 때문이다. 또한, 약간의 차이가 있지만, 이전에 결정되었던 시차를 재계산하여 추출된 특징점들에 대해 정확하게 정합될 확률 역시 향상되었다. 하지만, 이동 거리가 적을수록 추가적인 수행 시간이 필요한 것을 확인 할 수 있다. 그리고 <표 3>에 제시한 수행 시간은 주어진 스테레오 영상의 크기와 최소 및 최대 시차의 범위 차이로 인해 제안하는 방법의 수행 시간의 차이가 발생함을 또한 확인할 수 있다.

제안하는 방법의 정확성을 두 번째로 검증하는 방법으로 시차가 결정된 특징점들을 3차원 점들의 집합으로 표현하였다. 특징점들을 3차원 형태로 나타내는 것이 수치로 표현되는 것보다 직관적이며 신뢰도 있는 정보를 가지기 때문이다. 이를 위해, 결정된 특징점들을 수식 (11)을 이용하여 3차원 정보로 변환한다.



(그림 7) 오차 허용 범위가 1.0일 때, Venus, Teddy, Aloe 세 영상에 대한 특징창 이동 간격 변화에 따른 특징점 스테레오 특징점 정합 결과 영상. 녹색 점은 정확한 지점, 적색 점은 부정확한 지점. (W는 특징창 크기)

<표 2> 특징창 이동 간격 변화에 따른 결과

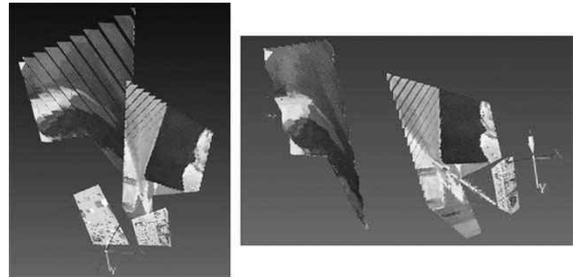
영상 이름			특징창 이동 간격					
			W/1	W/2	W/3	W/4	W/5	W/6
Venus	S+F		492	574	565	570	599	585
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	47.3	55.4	54.4	55.0	57.5	56.2
		$\epsilon=2.0$	47.5	55.6	54.7	55.2	57.7	56.4
Teddy	S+F		315	452	450	444	442	449
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	28.1	42.8	42.6	42.1	42.1	42.7
		$\epsilon=2.0$	28.8	43.4	43.3	42.7	42.6	43.4
Cone	S+F		371	498	494	488	480	508
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	34.9	45.7	45.0	44.4	43.8	45.9
		$\epsilon=2.0$	34.9	45.7	45.1	44.7	43.9	46.3
Aloe	S+F		456	650	689	708	687	708
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	43.3	62.1	65.8	67.8	65.4	67.5
		$\epsilon=2.0$	43.6	62.5	66.3	68.4	66.2	68.4
Art	S+F		320	459	456	469	474	525
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	25.4	36.4	34.4	36.5	35.9	39.5
		$\epsilon=2.0$	26.4	37.2	35.8	37.8	37.9	41.4

[W: 특징창 크기, ϵ : 오차 허용 범위, T: 전체 정합된 특징점 수, S: 정합된 특징점들 중에 오차 허용 범위 내의 특징점 수, F: 정합된 특징점들 중 오차 허용 범위 내에 속하지 않는 특징점 수]

<표 3> 특징창 이동 간격 변화에 따른 수행 시간

영상 이름	특징창 이동 간격					
	W/1	W/2	W/3	W/4	W/5	W/6
Venus	28.5	52.2	76.8	88.3	103.5	128.5
Teddy	67.7	193.9	366.6	508.7	665.3	861.8
Cone	73.9	230.5	378.9	535.1	635.4	785.4
Aloe	195.5	913.3	1946.1	2890.6	3769.4	4940.5
Art	251.4	783.7	1520.8	2421.9	3115.7	4192.6

[W: 특징창 크기, 수행시간 단위: ms]



(그림 8) Venus 정답 영상의 3차원 변환 결과와 제안하는 방법의 3차원 변환 결과. 회색 영역은 정답 영상, 적색 영역은 제안한 방법으로 획득된 정합된 특징점

$$z = \frac{f \cdot B}{d} \tag{11}$$

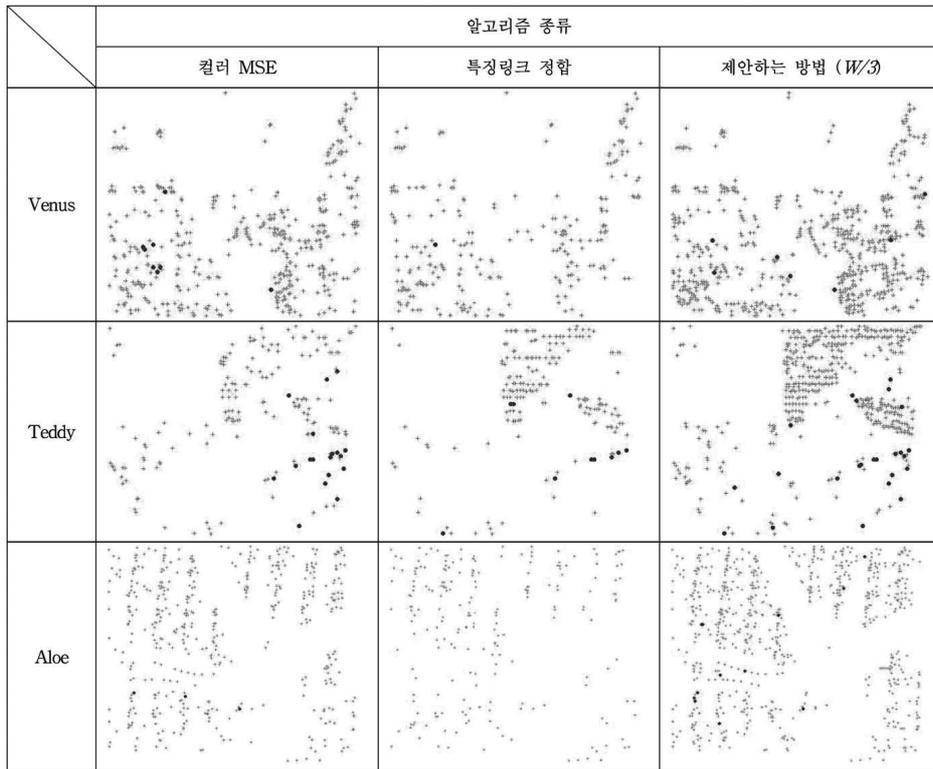
수식 (11)은 삼각화(triangulation)에 기반을 둔 기본적인 3차원 깊이 정보(z)를 획득 방법으로, f는 스테레오 카메라의 초점 거리(focal length)를 말하며, B는 스테레오 카메라의 위치 차이, d는 시차 정보를 의미한다. 본 논문에서는 f=8mm, B=100mm로 가정하였다.

결과 비교를 위해 Venus 영상과 Venus의 정답 영상 전체를 3차원 정보로 변환한다. 그리고 제안하는 방법의 결과를 3차원 정보로 변환한 뒤, 두 정보를 함께 표현하였다. 이 결과는 (그림 8)에 나타내었다. 제안하는 방법의 결과가 정확하다면, 정답 영상의 3차원 좌표에 제안하는 방법의 결과가 겹쳐진 형태로 나타날 것이다. 이런 이유로, (그림 8)을 통해 제안하는 방법의 정합 정확성이 높다는 것을 확인할 수 있다.

다음으로 제안하는 방법이 이전에 연구되었던 방법과 제안하는 방법의 결과가 어떤 차이가 있는지를 비교하였다. 이를 위해 앞에서 사용한 5개의 스테레오 영상을 사용하여 이전에 연구된 두 가지 방법 - 컬러 정보를 이용한 MSE 상관 함수

만을 이용하여 정합한 단순한 방법(이하, 컬러 MSE), 특징링크 정합 방법과 제안하는 방법의 스테레오 특징점 정합을 수행한 결과를 (그림 9)에 나타냈다. 실험의 조건은 <표 1>과 동일하게 하였고, 세 가지 방법 각각 정합된 특징점 수와 정확성, 수행 시간을 <표 4>와 <표 5>에 제시하였다.

<표 4>와 <표 5>를 통해 결과를 비교해 보면, 컬러 MSE 정합 방법과 특징링크 정합 방법은 기본적으로 빠른 수행 시간을 가진다. 컬러 MSE 정합 방법에 비해 특징링크 정합은 추출기의 오류로 인한 외치점(outlier)을 제거하여 정합되는 특징점의 시차의 정확성을 개선한 방법이다. 이런 이유로 정합된 특징점들은 더 높은 정확성을 가지고, 빠른 속도로 특징점들의 시차를 결정할 수 있다. 하지만, 특징링크를 이용할 경우 단순 특징점 정합에 비해 실제로 정합되는 특징점의 수는 많지 않다. 또한, <표 4>를 통해 컬러 MSE 특징점 정합과 특징링크 정합 방법 모두 추출된 특징점의 50% 이상의 특징점의 시차를 결정할 수 없는 문제가 있음을 확인할 수 있다.



(그림 9) 오차 허용 범위 1.0일 때, Venus, Teddy, Aloe 세 영상에 대한 세 가지 알고리즘에 대한 스테레오 특징점 정합 결과 영상. 녹색 점은 정확한 지점, 적색 점은 부정확한 지점

<표 4> 세 스테레오 특징점 정합 알고리즘을 적용한 결과

영상이름			컬러 MSE	특징링크 정합	제안하는 방법(W/3)
Venus	S+F		423	241	565
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	40.5	23.4	54.4
		$\epsilon=2.0$	40.5	23.4	54.7
Teddy	S+F		178	180	450
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	15.8	16.7	42.6
		$\epsilon=2.0$	16.5	16.9	43.3
Cone	S+F		289	125	494
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	26.5	11.4	45.0
		$\epsilon=2.0$	26.7	11.5	45.1
Aloe	S+F		455	217	689
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	43.9	21.0	65.8
		$\epsilon=2.0$	44.0	21.1	66.3
Art	S+F		204	102	456
	$\frac{S}{T} \times 100[\%]$	$\epsilon=1.0$	16.7	8.8	34.4
		$\epsilon=2.0$	16.8	8.8	35.8

[W: 특징창 크기, ϵ : 오차 허용 범위, T: 전체 정합된 특징점 수, S: 정합된 특징점들 중에 오차 허용 범위 내의 특징점 수, F: 정합된 특징점들 중 오차 허용 범위 내에 속하지 않는 특징점 수]

이에 비해 제안하는 방법인 특징창을 이용한 방법은 이전 두 방법들보다 정합된 특징점의 수를 약 1.5에서 2배 증가시켰다. 그리고 정합된 특징점 수에 따른 특징점들이 정확한 시차를 가질 확률 역시 증가되었다. 하지만, 정합된 특징점 수가 늘어남과 따른 제안하는 방법의 수행 시간은 이전 방법들

보다 때로는 몇 배 이상 추가적으로 소요되었고, Aloe나 Art 영상과 같이 최대, 최소 시차 범위가 클 경우에는 수행 시간이 다른 영상에 비해 많이 소요되는 것을 알 수 있다. 따라서 제안하는 방법을 실시간 시스템에 적용하기 위해서는 추가적인 속도 개선 알고리즘이 필요하다는 것을 알 수 있다.

〈표 5〉 세 스테레오 특징점 정합 알고리즘의 수행 시간

영상 이름	컬러 MSE	특징링크 정합	제안하는 방법(W/3)
Venus	19.8	10.1	76.8
Teddy	19.0	10.3	366.6
Cone	18.4	9.5	378.9
Aloe	27.4	14.6	1946.1
Art	26.5	13.1	1520.8

[W: 특징창 크기, 수행시간 단위: ms]

4. 결론 및 향후연구

본 논문에서는 새로운 스테레오 특징 기반 정합 방법인 특징창을 이용한 방법을 제안했다. 제안하는 방법을 이용할 경우, 실험을 통해 평균적으로 컬러 MSE 특징점 정합보다 추출된 특징점을 정확히 정합할 확률을 평균 약 1.69배, 특징링크 방법보다 약 2.98배 향상시킨 것을 확인할 수 있었다. 하지만, 표준 영상과 참조 영상의 높은 유사도를 가지는 특징창을 찾는 과정이나 특징창의 이동 거리에 따라 제안하는 방법의 평균 수행시간은 기존 방법보다 추가적으로 약 25배 이상 요구되었다. 이런 이유로, 다음 연구에서는 제안하는 방법의 수행 시간의 단축을 고려하여 실시간으로 특징점들의 시차 값이 결정할 수 있게 할 것이다. 이를 위해 추가적인 병렬 계산 알고리즘 및 GPU와 같은 하드웨어의 사용이 고려될 것이다. 또한, 제안하는 방법을 이용하여 스테레오 정합과 관련된 3차원 물체 복원 등 실제적인 분야에 제안하는 방법을 적용해 볼 것이다.

참고 문헌

[1] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-frame Stereo Correspondence Algorithms", International Journal of Computer Vision, 47(1/2/3): pp.7-42, April-June, 2002.

[2] Y. Ohta and T. Kanade, "Stereo by intra-and inter-scanline search using dynamic Programming," IEEE Transaction, Pattern Analysis and Machine Intelligence, Mar., 1985.

[3] H. Sadeghi, P. Moallem, and S. A. Monadjemi, "Feature Based Dense Stereo Matching using Dynamic Programming and Color," International Journal of Computational Intelligence, Vol.4, No.3, pp.179-186, July, 2008.

[4] S. D. Sharghi and F. A. Kamangar, "Geometric Feature-Based Matching in Stereo Images," Proceeding of Information, Decision and Control (ICDC), Feb., 1999.

[5] B. Tang, D. Ait-Boudaoud, B. J. Matuszewski and L. Shark, "An Efficient Feature Based Matching Algorithm for Stereo Images," Proceedings of the Geometric Modeling and Imaging, July, 2006.

[6] C. Kim and S. Park, "Fast Stereo Matching of Feature Links," The 2011 Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, May, 2011.

[7] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," Transaction Pattern Analysis and Machine Intelligence, Vol.23, pp.1222-1239, Nov., 2001.

[8] Q. Yang, L. Wang, R. Yang, H. Stewenius and D. Nister, "Stereo Matching with Color-weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling," Transaction on Pattern Analysis and Machine Intelligence, March, 2009.

[9] E. Rosten and T. Drummond, "Machine Learning for High-speed Corner Detection," European Conference on Computer Vision, May, 2006.

[10] R. Hartley and A. Zisserman, "Multiple View Geometry in Computer Vision," 2nd Ed, Cambridge University Press, 2004.

[11] J. Sun, Y. Li, S. Bing and K. H. Shum, "Symmetric Stereo Matching for Occlusion Handling," Computer Vision and Pattern Recognition, June, 2005.

[12] Stereo datasets with ground truth image, Middlebury College, "http://cat.middlebury.edu/stereo/data.html"



김 창 일

e-mail : changil.kim@lignex1.com

2005년~2009년 한동대학교 전산전자공학부 (공학사)

2010년~2012년 경북대학교 전자전기컴퓨터 학부(공학석사)

2011년~현 재 LIG넥스원 연구원

관심분야: 컴퓨터비전, 영상처리, 병렬처리 등



박 순 응

e-mail : sypark@knu.ac.kr

1991년 경북대학교 전자공학과(공학사)

1993년 경북대학교 전자공학과(공학석사)

2003년 Electrical & Computer Engineering at State University of New York at Stony Brook(공학박사)

2005년~현 재 경북대학교 컴퓨터학부 부교수

관심분야: 3차원 컴퓨터비전 및 로봇비전