

## 스케일 공간에서 동적 계획을 이용한 스테레오 정합

## (Stereo Matching Using Dynamic Programming in Scale-Space)

崔祐榮\*, 朴來弘\*\*

(Woo Young Choi and Rae-Hong Park)

## 要 約

본 논문에서는 스케일 공간상의 영교차점 지문 영상을 특징으로 사용하여 스테레오 정합을 하는데 있어서 올바른 정합률을 높이기 위한 정합방법을 제안한다. 정합 방법으로는 지문영상에 적합한 동적 계획 기법을 이용한 방법을 제안하며 정합과정에서 잘못 정합된 경우를 교정하기 위한 후처리과정을 이용하여 정합률을 높인다. 실험에서는 합성영상과 자연영상에 대해 제안한 알고리즘을 적용하여 올바르게 정합된 결과를 얻을 수 있었다.

## Abstract

In this paper, a matching method is proposed to improve the correct matching rate in stereo correspondence matching in which the fingerprint of zero-crossing points on the scale-space is used as the robust matching feature. The dynamic programming, which is appropriate for the fingerprint feature, is introduced for correspondence matching. We also improve the matching rate by using the post-processing for correcting mismatched points. In simulation, we apply the proposed algorithm to the synthetic and real images and obtain good matching results.

## I. 서론

2차원 영상들로부터 3차원 깊이 정보를 취득하는 데는 스테레오(stereo), 움직임(motion), 명암(shading), 텍스춰(texture) 혹은 선화(line drawing)를 이용한 방법들이 있다. 이 중에서 스테레오를 이용한 방법은 능동감지소자를 필요로 하지 않는 수동적 방법이라는 장점을 가지고 있으므로 실험환경의 특별한 제약이 없기 때문에 일반적인 scene을 처리대상으로

할 수 있는 장점이 있으며 넓은 응용분야를 가지고 있다.

스테레오 정합은 공간상의 한점이 좌우 두 영상면에 각각 투영되었을 때 이 두 점을 서로 같은 점으로 정합하는 것이 목적이다. 정합된 두 점은 각 영상의 좌표계에서 상대적으로 다른 위치에 존재하게 되는데 이들 위치의 차이를 변위(disparity)라고 한다. 이 변위와 두 영상 좌표계간의 관계를 이용하면 3차원상의 물체에 대한 깊이 정보를 구할 수 있다.<sup>[1]</sup> 변위를 구하기 위해서는 2차원 영상의 모든 점들간의 정합여부를 검사하여야 하지만 탐색영역을 줄이기 위해 epipolar 구속조건을 가정한다. 이 구속조건을 가정하면 탐색영역은 2차원 영상의 한 행(1차원)으로 범위를 줄일 수 있다. 또한 영상의 모든 화소들에 대해서 정합을 하지 않고 영상에서 특징(feature) 들에

\*正會員, 明知大學校 電子工學科  
(Dept. of Elec. Eng., Myongji Univ.)

\*\*正會員, 西江大學校 電子工學科  
(Dept. of Elec. Eng., Sogang Univ.)

接受日字: 1991年 8月 13日

대해서만 정합을 하게 되면 정합 문제를 더욱 간단히 할 수 있다. 이를 특징 베이스 정합(feature-based matching)이라 한다. 특징베이스 정합은 영상에서 모서리, 선소, 에지(edge), 영교차점(zero-crossing) 및 특별한 연산자에 의하여 선택된 점들을 특징점들로 이용하여 정합한다.<sup>[1-4]</sup> 이러한 특징들은 영상의 빛의 밝기보다는 그변화에 관련됨으로써 물체표면의 불연속을 의미하므로 광학적 변화에 덜 민감하며 고려될 특징들의 수가 화소의 수보다는 적기 때문에 계산처리 시간도 빠르다. 그러나 영상의 특징점들에 대해서만 정합되어 모든 화소에 대해서 깊이 정보를 구할 수 없으므로 나머지 화소들에 대해서는 보간법을 이용하여 깊이 정보를 구하여야 하는 단점이 있다. 이러한 단점에도 불구하고 장점이 갖는 특성이 훨씬 두드러지므로 많은 연구가 이루어져 왔다.

본 논문에서는 특징베이스 정합 방법에 대하여 연구하였다. 정확한 정합점을 구하기 위하여 Gaussian-Laplacian 연산자에 의해 구한 영교차점을 사용하였으며 Gauss 함수의 스케일(scale) 값 크기에 따라 영교차점이 변하게 되는 단점을 보완하기 위해 각 스케일의 값과 영교차점 에지와의 관계를 나타낸 Witkin<sup>[5]</sup>에 의해 제안된 스케일 공간(scale-space) 상의 지문 영상을 정합을 위한 특징으로서 사용하였다. 스케일 공간상의 영교차점 지문 영상을 특징으로 사용하여 스테레오 정합을 하는데 있어서 올바른 정합률을 높이고 지문 영상의 특징을 이용하는데 적합한 정합방법으로 동적 계획(dynamic programming) 기법을 이용한 방법을 제안하였다. 또한 정합과정에서 잘못 정합된 경우를 교정하기 위하여 영상의 각 행에 존재하는 특징점들간의 연속성 관계를 이용한 후처리 과정을 적용하여 정합률을 높였다.

먼저 II장에서는 스테레오 정합에서 정합에 사용될 특징으로서 스케일 공간상의 영교차점 지문 영상에 대해 논하고 III장에서는 이를 특징점으로 사용한 특징베이스 정합방법에 대해 논한다. 이 때 정합 방법으로는 지문영상을 특징점으로 이용하였을 때 정합률을 높이기 위해 이 특징점에 적합한 동적 계획 기법을 이용한다. IV장에서는 보다 높은 정합률을 얻기 위하여 동적 계획 정합과정에서 잘못 정합된 부분을 영상의 각 라인간의 관계를 이용하여 수정하는 후처리 과정에 대해 논한다. V장에서는 실험 결과에 대하여 검토하며 IV장에서 결론을 맺는다.

II. 정합에 사용할 스케일 공간상의 지문 영상 특징

일반적으로 특징베이스 방법에서는 정합을 위한 특징으로 에지 정보를 많이 사용한다. 그러나 일반 에지 연산자들에 의해 추출된 에지는 한 화소이상의 두께로 표현되므로 정합을 하기 위하여 세선화 과정이 필요하게 된다. 세선화 과정에서 에지들의 위치가 실제 위치와 다른 결과를 얻을 수 있으며 또한 많은 계산량이 필요하므로 한 화소 두께의 에지를 구할 수 있는 영교차 에지 연산자가 스테레오 정합에 많이 이용되고 있다. 이것은 Marr-Poggio 스테레오 모델<sup>[6]</sup>에 근거한 것으로 영교차 에지를 구하기 위하여 Gaussian-Laplacian 연산자가 주로 이용된다. 일반적으로 Gauss 함수의 스케일 값을 크게함에 따라 보다 특성이 강한 영교차 에지를 구할 수 있지만 그 위치가 원영상의 에지 위치와는 다르게 되므로 이를 이용하여 정합을 할 경우 구하게 되는 깊이 정보에는 오차가 발생하게 된다. 따라서 특징으로 사용할 영교차점 에지는 원영상 에지 위치와 일치하면서도 강한 특성을 가져야 한다.

다단계 정합(multistage matching) 방법<sup>[1]</sup>은 영상의 해상도를 줄여서 먼저 영교차점을 이용하여 정합하고, 다음 단계에서 해상도를 높인 후 이전 단계에서 정합된 점들의 주변 영역에 대해 다시 영교차점을 구해 정합을 시도한다. 이러한 과정을 원래 영상의 해상도가 될 때까지 해상도를 높이면서 반복하여 최종 정합 결과를 얻는다. 이 방법은 영교차점을 구하는 측면에서 보면 스케일 값이 큰 값에서 낮은 값으로 변화시키며 정합하는 과정으로 볼 수 있으며 이러한 방법을 취함으로써 정합할 후보 영역에 제한을 둘 수 있고 정합시간도 줄일 수 있다. 그러나 해상도가 낮은 첫단계에서 정합되지 못한 점들은 다음 단계에서도 정합 후보 영역이 될 수 없으므로 최종 단계에서 정합되지 못하는 단점이 있다. 따라서 높은 스케일 값과 낮은 스케일값에서 구한 각각의 영교차 에지의 성질을 동시에 갖출 뿐 아니라 높은 스케일에서 구한 정보와 낮은 스케일에서 구한 정보들 사이의 관계를 잃지 않는 특징이 필요하게 된다. 이러한 각 스케일의 값과 영교차 에지와의 관계를 나타낸 것이 Witkin<sup>[5]</sup>에 의해 제안된 스케일 공간상의 지문 영상이다. 본 논문에서는 이를 정합에 이용되는 특징으로 사용한다. 이때 스케일 상의 지문 영상이 실제 에지에 해당하지 않는 경우가 있다.<sup>[7,8]</sup> 이를 제거하고 실제 에지에 해당하는 부분만을 추출하기 위해서 제안한 지문영상 추출 방법을 사용하였다.<sup>[9]</sup>

지문 영상을 정합의 특징으로 사용하기 위해서 지문 영상의 길이, 상위 기울기, 하위 기울기 및 부호를 특징값으로 사용한다. 스케일 공간을 그림 1에서

와 같이 x축을 영상의 한 단면으로 하고, y축을 스케일값으로 하고 이때 최소 스케일값을  $\sigma_{min}$ , 최대 스케일 값을  $\sigma_{max}$  라고 하자. 길이 특징값은 스케일 공간상에서  $\sigma_{min}$ 에서  $\sigma_{max}$ 까지 스케일값을 점점 크게 하였을 때 영교차점을 구할 수 있을 때까지의 지문 영상의 길이를 말한다. 즉

$$\text{길이} = \sum_{i=\sigma_{min}}^{\sigma_{max}} E(i), \tag{1}$$

여기서 E(i)는 한 지문 영상에 대해 스케일값이 i일 때 영교차점이 존재하면(지문 영상이 사라지지 않았으면) 1의 값을 갖고 아니면 0의 값을 갖는다. 상위 기울기 특징값은 지문 영상에서 기울기의 최대 변곡점을 기준으로 지문 영상의 양끝점을 잇는 두선분중 변곡점의 위에 있는 선분의 기울기를 의미하며 하위 기울기 특징값은 변곡점의 아래에 있는 선분의 기울기를 의미한다. 각 기울기는 그림 1에서와 같이 기울기를 표시하는 화살표 방향이 반시계 방향이면 양의 값을 갖고 시계 방향이면 음의 부호를 갖는다. 두 가지 기울기 특징을 사용한 것은 정합할 두 지문 영상들의 하위 기울기는 서로 같으나 상위 기울기가 다른 경우(또는 그 반대가 되는 경우)가 있기 때문이다. 부호 특징값은 영교차점이 양수값에서 음수값으로 변할 때 생겼으면 +를, 그와 반대 경우면 -를 나타낸다.

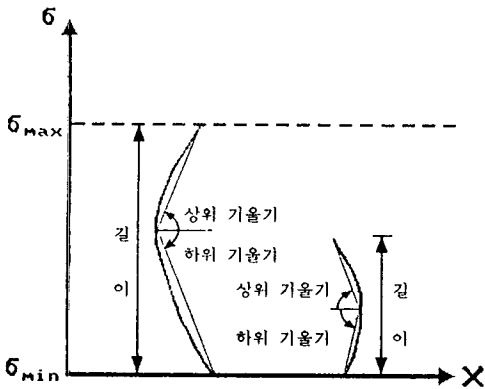


그림 1. 스케일 공간  
Fig. 1. Scale space.

Ⅲ. 스케일 공간상의 지문 영상을 정합하기 위한 동적 계획 정합

먼저 두 스테레오 영상에서 정합할 점들에 대한 탐색 영역을 줄이기 위해 두 카메라가 수평선상에 놓

여있고 서로 평행하다고 가정하여 epipolar 구속조건을 사용한다. Epipolar line은 각 영상의 한 행에 해당하므로 이 구속조건을 이용하면 2차원의 영상의 정합을 각 영상의 행단위의 정합으로, 즉 1차원 신호의 정합으로 단순화할 수 있다. 따라서 일차원 영상에 대해 스케일 공간상의 지문 영상을 구해 정합의 특징으로 사용할 수 있다.

정합을 하는 방식중에는 가장 단순한 방법은 특징 점들간의 국부적 상관관계(correlation)를 이용하여 순차적으로 적용하는 brute force방식이다.<sup>[2]</sup> 그러나 이 방법은 특징점들간의 국부적인 특성을 이용하여 적용하므로 전체적인 특징점들간의 관계를 고려하지 못할 뿐만 아니라 적용하는 순서에 따라 구해진 결과가 다르게 된다. 이러한 단점을 해결하기 위한 방법으로 적용하는 순서에 무관하며 병렬처리가 가능하고 특징점들간의 상호관계를 이용하는 이완(relaxation) 기법이 있다.<sup>[2]</sup> 이 방법은 특징점들간의 상호관계를 이용하므로 잡음등에 의한 특징점들에 있어서의 오차에 불구하고 올바른 정합 결과를 얻을 수 있다. 그러나 이 방법에서는 해를 구하기 위해 해가 수렴할 때 까지 반복 계산하여야 하므로 많은 계산 시간을 요하게 되고 수렴하지 않을 경우도 있다. 따라서 보다 빠르고 정확한 정합결과를 얻을 수 있고 주어진 특징점에 적합한 정합 방법이 필요하게 된다.

본 논문에서는 스테레오 영상에 대해 epipolar 구속조건을 가정하여 두 영상 행에 존재하는 지문 영상을 특징으로 하여 정합하는 것이 목적이다. 따라서 그림 2에서와 같이 2차원 축에서 x축을 우측 영상의 한 행에 존재하는 지문 영상들로 보고 y축을 좌측 영상의 한 행에 존재하는 지문 영상들로 보았을 때 이들 특징점들 간에 가장 올바른 정합관계를 맺는 최적 정합 경로를 찾으면 된다. 이와 같은 문제는 동적 계획기법을 사용하면 효율적으로 해결할 수 있다. Ohta 등<sup>[10]</sup>과 Lloyd 등<sup>[11]</sup>은 영상의 빛의 밝기를 특징으로 하여 스테레오 정합에 동적계획 기법을 사용하였다.

동적 계획 기법은 어떤 문제에 대해 해를 구하기 위해 연속된 결정을 하여야 하는 경우 다단계 결정 과정을 최적성 원리(principle of optimality)에 기초하여 푸는 방법이다.<sup>[12]</sup> 여기서 최적성 원리는 다단계 결정 과정에서 최적 결정의 계열(sequence)은 어떤 단계에서의 결정으로 정해지는 상태에 관하여 남은 단계에서의 결정이 최적 결정이라고 하는 성질을 지니고 있어야 한다는 원리이다. 따라서 최적성 원리가 성립하는 다단계 결정 과정에서는 그 n단계 부분 과정에 있어서 최초의 상태  $S_0$ 로 부터 출발하여 최적 정책을 채용, 결정할 때 얻게 되는 목적 함수값

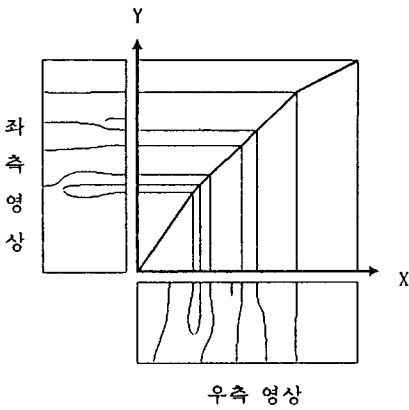


그림 2. 우측, 좌측 영상의 각 행에 존재하는 지문 영상의 정합

Fig. 2. Matching of fingerprints on each row of the right and left images.

을  $fn(s_0)$ 라 하고  $s_0$ 로부터 1회의 결정을 한 후에 도달한 상태를  $s_1$ 이라 하면 최적성 원리에 따라  $fn(s_0)$ 와  $fn-1(s_1)$ 을 결합하여 점화식을 얻게 된다.  $n$ 이 무한일 때는 재귀적 축차 반복 과정으로 해답을 구할 수 있다. 동적 계획 기법의 장점은 최적해가 될 수 없는 다단계 결정 과정들을 피함으로써 계산량을 크게 줄일 수 있다는 점이다.

본 논문에서 최적화될 목적 함수는 비용함수(cost function)를 최대화하는 형태를 사용한다. 즉 정합될 각 쌍마다 음이 아닌 비용(cost)을 부가하고 그 쌍이 정합 불가능한 쌍일 때는 최소값을 부가하며 비용함수는 이들 비용값의 합으로 표시된다. 최적 정합이 가능한 정합쌍의 집합은 전체 비용함수를 최대화하는 집합이 된다. 좌측영상의 한 행에  $M$ 개의 지문 영상이 있고 우측 영상에  $N$ 개가 있다고 가정하면 정합 가능한 각쌍들에 부가되는 초기비용  $c(m, n)$ 은 일반적으로 두 지문영상의 특징값의 차의 선형적 합으로 부가된다. 그러나 지문 영상을 특징으로 사용할 경우 부호정보가 서로 다른 쌍들은 정합될 수 없으므로  $c(m, n)$ 은 다음과 같이 정의할 수 있다.

$$c(m, n) = \begin{cases} K / (1 + \alpha D(1) + \beta D(ac) + \gamma D(bc)), & m \text{과 } n \text{의 부호가 같을 때} \\ 0, m \text{과 } n \text{의 부호가 다르거나 } m=0 & \text{또는 } n=0 \text{ 일 때,} \end{cases} \quad (2)$$

여기서  $(m, n)$ 은 좌측영상의  $m$ 번째 지문 영상이 우측영상의  $n$ 번째 지문 영상과 정합됨을 의미한다( $0 \leq m \leq M, 0 \leq n \leq N$ ).  $(m, 0)$  또는  $(0, n)$ 은  $m$  또는  $n$ 번

째 지문영상과 정합될 대응 지문 영상이 존재하지 않는 경우를 의미한다. 이것은 한 영상에서는 볼 수 있는 특징점이 다른 영상에서는 물체들간의 겹침이나 영교차 추출을 할 수 없어 정합을 할 수 없는 경우에 해당한다.  $D(1)$ 은 지문 영상  $m$ 과  $n$ 의 길이의 차를,  $D(ac)$ 과  $D(bc)$ 은 각각 상위 기울기와 하위 기울기의 차를 나타낸다.  $\alpha, \beta, \gamma$ 는 가중치이고  $K$ 는 정규화 상수이다. 출발점으로 부터  $(m, n)$ 까지의 경로의 총 비용  $f(m, n)$ 은 비용  $c(m, n)$ 을 이용하여 다음과 같은 순환식의 형태로 표시할 수 있다.

$$f(m, n) = c(m, n) + \text{MAX}(f(m-1, n-1), c(m-1, 0) + f(m-2, n-1), c(0, n-1) + f(m-1, n-2), c(m-1, 0) + c(m-2, 0) + f(m-3, n-1), c(0, n-1) + c(0, n-2) + f(m-1, n-3)), \quad (3)$$

여기서  $\text{MAX}()$ 는 원소중 최대치를 구하는 함수이고  $\text{MAX}()$ 안의 원소중 첫번째 원소는 이전 단계에서  $m-1$ 과  $n-1$ 번째 지문 영상이 서로 정합된 경우를 의미한다. 둘째, 셋째 원소는  $m-1$ 과  $n-1$ 번째 지문 영상이 정합될 수 없어 건너 뛰고  $m-1$ 과  $n-2$  또는  $m-2$ 와  $n-1$ 이 서로 정합된 경우를 의미한다. 넷째, 다섯째 원소는 연속된 두 지문영상에서 대응 정합쌍을 구할 수 없는 경우를 의미한다. 더 많은 연속된 지문 영상이 정합되지 않는 경우를 생각할 수 있으나 이것은 더 많은 계산량이 요구되므로 적용하는 영상에 따라 고려할 수 있는 범위를 선택하면 된다. 본 논문에서는 최대 정합되지 않고 건너뛰게 되는 갯수를 두 개로 하였다. 최적 경로는 경로의 출발점에서 끝점까지 구한 총비용  $f(m, n)$ 을 최대화하는 경로가 되며 최적비용  $f_{max}$ 는 다음 식과 같다.

$$f_{max} = \text{MAX}(f(m, n)), \text{ for } 0 \leq m \leq M, 0 \leq n \leq N. \quad (4)$$

최적 경로는  $f_{max}$ 로 부터 출발하여 식(3)을 이용하여 역추적(back-tracking)하여 구하면 된다.

#### IV. 정합률을 높이기 위한 후처리 과정

동적 계획 기법을 이용하였을 때 높은 정합률을 얻을 수 있지만 보다 높은 정합률을 얻고 잘못 정합된 부분을 수정하기 위하여 후처리 과정을 한다. 이 과정은 영상의 각 행에 대해서 두 단계로 수행하는데 첫 단계에서는 정합된 각 지문 영상쌍에 대해 행간

의 관계를 이용하여 검증 및 보정을 한다. 둘째 단계에서 정합과정이나 보정과정에서 정합되지 않은 쌍들에 대해 정합가능성을 조사하여 정합한다.

일반적으로 영상에서 한 물체에 속한 에지는 연속성을 갖는다. 이것은 한 물체에 속한 에지점들중 영상의 한 행에서 구한 에지점은 이전 행 또는 다음 행에서 구한 에지점이 위치한 주변에 존재하게 됨을 의미한다. 또한 스테레오 영상의 경우 한 영상에서 에지의 연속성이 만족되면 다른 영상에서도 그 에지는 연속성을 갖는다. 이러한 에지의 연속성을 이용하여 후처리 첫 단계에서는 영상의 현재 행에 존재하는 정합된 쌍들이 올바른지를 판단한다. 현재 행과 이전 행 또는 현재 행과 다음 행에 대해서 연속성 구속조건이 만족되지 않는 경우 현재행과 이전 행 또는 현재 행과 다음 행 사이에 어느 행이 올바른 것인지를 판단하기 위해 식 (2)에서 사용된 지문 영상의 길이 차  $D(1)$ , 상위 기울기차  $D(ac)$ , 하위 기울기 차  $D(bc)$ , 좌측 평균 빛의 밝기 차, 우측 평균 빛의 밝기 차들의 가중치 합을 구하여 그 값이 작은 경우를 올바르게 정합된 행으로 보고 이를 기준 행으로 정한다. 이 때 좌측, 우측 평균 빛의 밝기 차는 지문 영상을 기준으로 좌우 양옆의 5화소의 평균 빛의 밝기를 구하여 비교하는 두 지문 영상에 대해 그 차를 구한 값을 의미한다. 기준 행에서 올바르게 정합된 쌍  $i$ 와  $j$ 에 대해 현재 행에서 정합여부를 판단하기 위한 쌍을  $m$ 과  $n$ 이라고 하였을 때 그림 3과 같은 경우에 오류가 있다고 보고 교정을 한다. 그림 3의 각 항에서 위의 두 지문 영상들은 좌측, 우측 영상의 기준 행에 존재하는 지문 영상들을 의미하고  $i$ 와  $j$ 는 올바르게 정합된 쌍을 나타낸다. 아래 두 지문 영상들은 현재 행에 존재하는 지문 영상으로 정합이 올바르게 되었는지를 판단하려는 지문 영상이다. 직선으로 표현된 것은 정합된 지문 영상을, 점선으로 표시된 것은 정합되지 않은 지문 영상을, X로 표시된 것은 주변에 정합에 사용될 지문 영상이 없음을 의미한다. 그림 3 (a)-(c) 경우는  $i, j$ 가 정합되었을 때  $m$ 과  $n$ 이 정합하려면  $m$ 이  $i$ 와 연속성을 가지므로  $n$ 도  $j$ 와 연속성을 가져야 하는데 그렇지 못하므로 이 때는  $m$ 과  $n$ 이 서로 정합된 것이 아닌 것으로 하여 두번째 후처리 단계에서 다시 정합여부를 판단한다. 그림 3 (d)-(f)에서는  $m$ 이  $i$ 와 연속성을 가질 때  $j$ 와 연속성이 있는 것은  $k$ 이므로  $m$ 은  $k$ 와 정합된 것으로 교정하고  $n$ 은 정합되지 않은 것으로 한다. 그림3(g)-(i)에서는  $m$ 이  $i$ 와 연속성이 있을 때  $j$ 와 연속성을 갖는 것이  $k$ 인데  $k$ 는 이미 다른 지문 영상과 정합이 되어 있으므로 이 때에는  $m, n, k$  모두 정합되지 않은 것

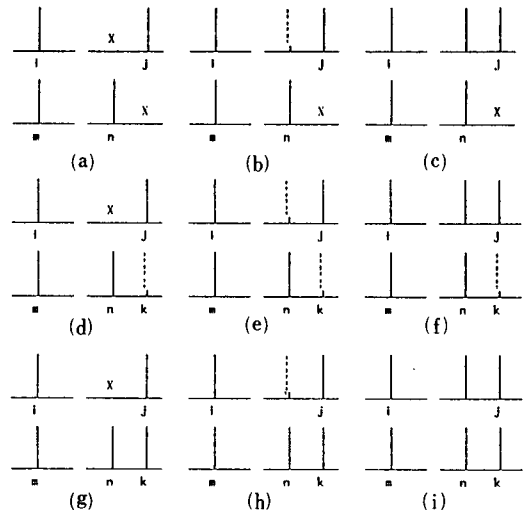


그림 3. 연속성에 의해 교정되어야 할 경우  
 Fig. 3. The case to be corrected due to the continuity constraint.

으로 판정하여 두 번째 후처리 단계에서 정합 여부를 결정한다.

두번째 후처리 단계에서는 순서 구속조건 (ordering constraint)에 근거하여 이미 정합된 지문 영상을 이용하여 동적계획 정합에서 또는 첫번째 후처리 과정에서 정합되지 않은 지문 영상들의 정합가능성을 조사하여 정합한다. 순서 구속조건은 좌, 우 영상의 한 행에서 정합된 어느 한 쌍을 기준으로 한 영상에서 그 쌍의 좌측(우측)에 있는 지문 영상은 다른 영상에서도 그 정합된 쌍의 좌측(우측)에 있는 지문 영상과 정합된다는 조건이다. 실제 영상에서 이 조건이 위배되는 경우는 극히 드물며 위배된 경우에는 사람의 시각으로도 스테레오 정합을 하기 어렵다. 따라서 이미 정합된 쌍들 사이에 정합되지 않은 지문 영상들이 있으면 이들에 대해 정합을 시도한다. 먼저 정합되지 않은 지문 영상들이 있으면 이들에 대해 정합을 시도한다. 먼저 정합되지 않은 지문 영상 쌍중에서 같은 부호를 갖는 경우에만 정합가능성이 있는 것으로 보고 다음 과정을 수행한다. 그림4에서와 같이 좌측, 우측 영상의 한 행에 정합되지 않은 지문 영상  $m, n$ 과 정합된 쌍  $a-c, b-d$ 가 있을 때 영상의 한 행에서 지문 영상들 사이의 떨어진 거리를 계산하는 함수를  $d(\ )$ 라고 하자. 이 때  $|d(m, a)-d(n, c)|$ 와  $|d(m-b)-d(n, d)|$  모두 임계치 보다 작으면  $m$ 과  $n$ 은 정합된 것으로 판정한다. 이상의 두 단계를 영상의 각 행마다 적용하여 전체 영상의 모든 행에 대하여 수행한다.

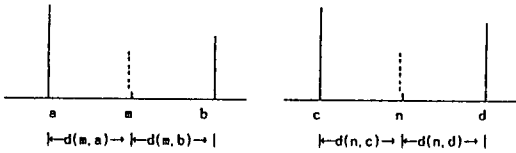


그림 4. 좌, 우측 영상에서 정합되지 않은 지문 영상쌍 (m, n)  
 Fig. 4. Unmatched fingerprint pair (m, n) in the left and right images.

V. 실험결과 및 검토

실험에서는 epipolar 구속 조건을 적용하여 두 영상의 수평선이 일직선 상에 놓이고 수직선이 평행한 영상 취득 구조로 하여 2차원 영상의 정합을 각 영상의 행단위의 정합으로 단순화하였다. 0.1mm이하의 정밀도를 갖는 산업용 로봇 FANUC을 사용하여 취득한 256×256의 크기의 256단계의 밝기를 갖는 영상들을 취득하였으며 이 중 한 예를 그림 5에 보였다. 그림에서 두 영상들 사이에 빛의 밝기에 차가 약간 있는데 이것은 스테레오 영상 취득시 카메라 이동에 의해 조명과의 각도가 달라지게 되어 생기는 현상이다.

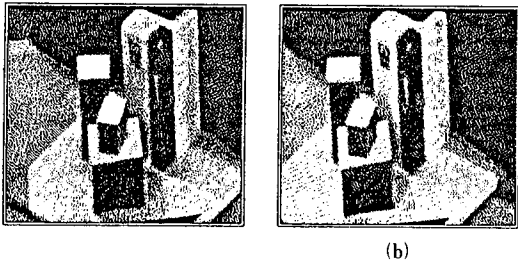


그림 5. 실제 스테레오 영상  
 Fig. 5. Real stereo image.

그림 6 (a)는 그림 5의 스테레오 영상에서 168번째 행의 영상 강도(intensity) 단면을 보인 것으로 약간의 빛의 밝기에 차이가 있음을 알 수 있다. 따라서 Ohta등<sup>[10]</sup>와 Lloyd<sup>[11]</sup> 방법에서와 같이 단순히 영상의 빛의 밝기만을 특징으로 하여 스테레오 정합을 할 경우 어려움이 있음을 알 수 있다. 그림 6 (a)의 단면에 대해 특징으로 사용될 실제 예지에 해당하는 스케일 공간상의 지문영상을 구한 결과를 그림 6 (b)에 보였다. 이 지문 영상들로부터 길이, 상위 기울기, 하위 기울기 및 부호등의 특징값들을 구한다. 동적

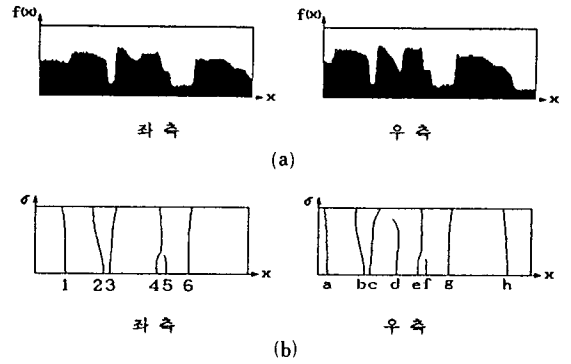


그림 6. (a) 영상 강도의 단면 (b) 추출된 지문 영상  
 Fig. 6. (a) Intensity profile, (b) Extracted fingerprint.

계획 기법을 적용하기 위해서 좌우 영상의 각 쌍들에 대해서 특징값들을 이용하여 식(2)로부터 초기 비용  $c(m, n)$ 을 구한다. 이 때 가중치  $\alpha, \beta, \gamma$ 는 1로 하였으며  $K$ 는

$$\text{Max}_{m,n} (c(m, n)) \equiv 1$$

이 되도록 정하였다. 그림 6 (b) 경우 좌측 영상에 존재한 지문 영상들에 대해 좌측에서 우측으로 차례로 1에서 6까지 번호를 붙이고 우측 영상에서는 a에서 h까지 문자를 붙이면 이들 사이의 초기 비용은 표 1과 같다. 이 때 좌측 영상의 지문 영상을 기준으로 하여 이들 지문 영상이 우측 영상의 어느 지문 영상에 해당하는지를 알아보기 위해 순차적 정합 방법을 적용한다.<sup>[2]</sup> 한 영상에서 지문 영상의 특성값이 다른 영상에서 지문 영상의 특성값과 차이가 가장 작은 경우를 정합된 것으로 판정하면 표 1에서의 숫자는 각 쌍들간의 유사도를 나타내므로 다음과 같은 정합 결과를 얻게 된다.

표 1. 초기 비용  
 Table 1. Initial cost.

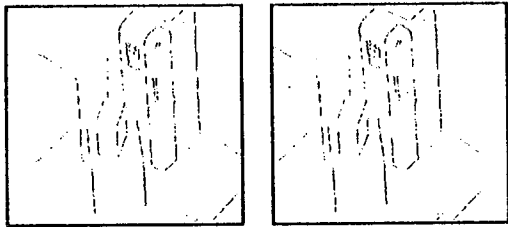
	a	b	c	d	e	f	g	h
1	0.9937	0.0000	0.7385	0.8870	0.0000	0.0000	0.7430	0.0000
2	0.0000	0.9973	0.0000	0.0000	0.7523	0.8932	0.0000	0.9945
3	0.8024	0.0000	0.9961	0.8625	0.0000	0.0000	0.9995	0.0000
4	0.0000	0.8057	0.0000	0.0000	0.9918	0.7377	0.0000	0.8036
5	0.0000	0.9314	0.0000	0.0000	0.7790	0.9968	0.0000	0.9296
6	0.8023	0.0000	0.9966	0.8622	0.0000	0.0000	1.0000	0.0000

순서 구속조건을 무시하였을 경우  
 {(1, a), (2, b), (3, g), (4, e), (5, f), (6, g)}

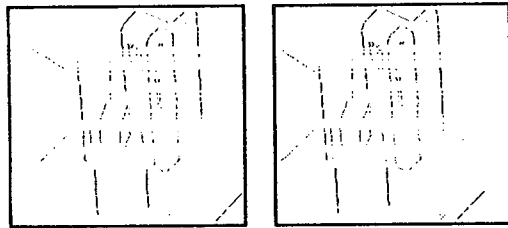
순서 구속조건을 고려한 경우

{(1, a), (2, b), (3, g), (4, h)}

두 경우 모두 올바른 정합 결과를 얻을 수 없는데 이것은 지문 영상들간의 국부적인 특성값에만 근거하여 정합하였기 때문이다. 그림 7 (a)는 그림5의 스테레오 영상에 대해 순서 구속조건을 가정하여 순차적 정합 방법을 적용하였을 때의 결과로서 829쌍의 지문 영상이 올바르게 정합되었다. 그림 7 (b)는 순서 구속조건을 적용하지 않았을 때 결과이며 667쌍이 올바르게 정합되었다.



(a)



(b)

그림 7. 순차적 정합방법을 적용한 경우 정합 결과  
(a) 순서구속조건을 가정한 경우  
(b) 순서구속조건을 가정하지 않은 경우

Fig. 7. Matching results obtained by applying the sequential matching method.

- (a) With ordering constraint assumption,
- (b) Without ordering constraint assumption.

특징들간의 상호관계를 이용하기 위해 이완 정합 방법을 적용했을 때 다음과 같은 정합 결과를 구할 수 있었다.

{(1, h), (2, b), (3, c), (4, e), (5, f), (6, g)}

이것은 순차적 정합 방법에 비해서는 올바른 정합 결과를 구할 수 있었지만 순환식을 이용하여 반복계산하여야 하므로 많은 계산량이 필요하고 또한 단일해로 수렴하지 않는 경우가 있으므로 보다 간단하며 최적해를 구할 수 있는 방법이 필요하다. 이완 정합방법을 그림 5의 영상에 적용한 결과를 그림 8에 도시

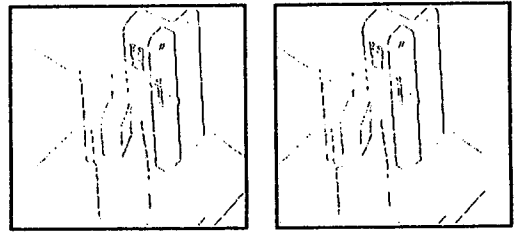


그림 8. 이완 정합방법을 적용한 경우 정합 결과  
Fig. 8. Matching results obtained by applying the relaxation matching method.

하였으며 1036쌍이 올바르게 정합되었다.

동적 계획 기법을 적용하여 정합하기 위하여 식(3)을 이용하여 총비용  $f(m, n)$ 을 계산하였으며 이를 표 2에 제시하였다. 이 표에서 최적 비용을 구하면 5.9757로서 좌측 영상에서 지문 영상 6과 우측 영상에서 지문 영상 g가 정합되었음을 알 수 있다. 따라서 식(3)을 이용하여 역추적하면 표 2의 밑줄친 부분이 되며 그 결과는 다음과 같다.

{(1, a), (2, b), (3, c), (4, e), (5, f), (6, g)}

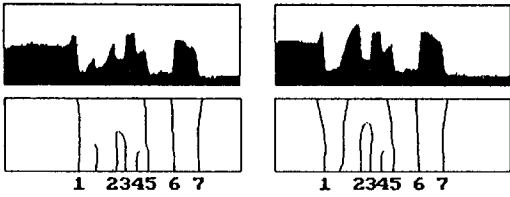
따라서 동적 계획 기법을 사용하면 전체를 최적화하는 결과를 얻게 되므로 올바른 결과를 얻을 수 있다.

후처리 과정에서는 특징점들의 행간 연속성과 같은 행에 존재하는 특징들간의 순서 구속조건을 이용하여 잘못 정합된 점들을 보정하였다. 그림 9 (a)는 그림 5 영상의 132번째 열의 강도 단면과 이때 올바르게 정합된 지문쌍을 나타낸다. 좌측과 우측에서 같은 번호를 갖는 지문 영상들이 서로 정합되었음을 의미한다. 그림 9 (b)는 133번째 열에 대해 후처리 하기 전의 정합 결과와 후처리 후의 교정된 정합 결과를 보이고 있다. 133번째 열에서 후처리 전에 2번으로 번호매겨진 쌍은 윗 열과 연속성이 없이 새로 정합된 쌍으로 두번째 후처리 단계에서 순서 구속조건을 사용하여 이웃 지문 영상과의 거리차를 계산하였을 때

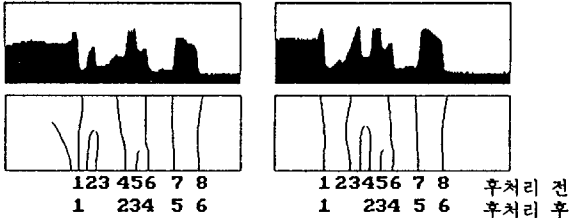
표 2. 총 비용

Table 2. Total cost.

	a	b	c	d	e	f	g	h
1	<u>0.9937</u>	0.0000	0.7385	0.8870	0.0000	0.0000	0.7430	0.0000
2	0.0000	<u>1.9910</u>	0.9937	0.9937	1.6393	1.7802	0.8870	1.7375
3	0.8024	0.9937	<u>2.9871</u>	2.8535	1.9910	1.6393	2.7796	1.7802
4	0.0000	1.7994	1.9910	2.8871	<u>3.9789</u>	3.7248	2.8535	3.5832
5	0.0000	1.7338	1.9910	2.9871	3.6325	<u>4.9757</u>	3.9789	4.9085
6	0.8023	0.8024	2.7960	3.8494	2.8535	3.9789	<u>5.9757</u>	4.9757



(a) 132번째 행



(b) 133번째 행

그림 9. 후처리 결과  
Fig. 9. Postprocessing result.

임계치보다 크게 되므로 이를 제거한다. 또한 후처리 전에 3번으로 매겨진 쌍은 우측 영상을 기준으로 보았을 때 첫번째 후처리 과정에서는 그림3(b)경우에 해당하므로 이들 쌍끼리는 정합되지 않은 것으로 간주하여 두번째 단계에서 정합여부를 판단한다. 이 경우도 거리차를 구했을 때 임계치보다 크게 되므로 제거한다.

그림10은 동적 계획 기법을 적용하고 후처리를 하였을 때의 결과이며 후처리 전 정합된 1348쌍중 올바르게 정합된 것은 1318쌍이며 후처리후 1333쌍이 올바르게 정합되었다. 이 경우 후처리 과정에서 교정되어 새로 정합된 것은 15쌍인데 후처리 전에 대부분 올바르게 정합되어 후처리 과정에서 새로 정합된 수가 적다. 후처리 과정에서 제거된 점은 15쌍으로 주로 기울기가 수평에 가까운 사선 방향의 에지에 속한 점들로 후처리 과정의 두번째 단계에서 오류를 검출하여 제거되었다. 그림들에 표시된 점들은 지문 영상간에 서로 1대1로 올바르게 정합된 경우만을 나타낸 것으로 경계가 끊어진 부분이 생기는 것은 일부의 특징점들이 정합되지 못한것을 의미한다. 또한 수평방향의 에지들에 대해 정합이 되지 않은 것은 epipolar 구속조건을 가정하여 수직 성분의 에지들만 검출하여 정합하였기 때문에 수평방향의 에지들이 검출되지 않았기 때문이다. 그러나 두 스테레오 영상중 좌측 영상에 대해 수직 방향으로 또 다른 영상을 취득하면 사용한다면 즉 trinocular 스테레오구조를

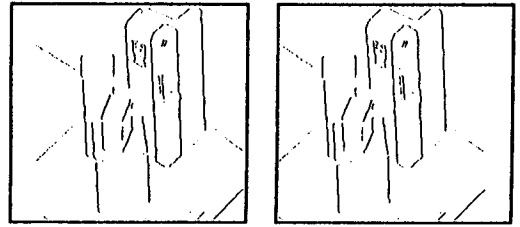


그림10. 동적 계획 기법을 적용하고 후처리를 한 결과  
Fig. 10. Matching result obtained by applying dynamic programming method and postprocessing.

이용한다면 수평 방향의 에지 성분을 검출할 수 있으며 이를 이용하여 정합할 수 있다. 다른 실제 영상들에 대해서도 실험해 본 결과 동적 계획 기법을 사용하였을 때 가장 올바른 정합 결과를 얻을 수 있었으며 정합 결과 및 그 추세들은 위에서 언급한 바와 유사하여 그 결과는 지면상 신지 않았다.

본 논문에서 수행한 정합결과와 비교하기 위해 동적 계획 정합방법을 사용한 Ohta등<sup>[10]</sup>과 Lloyd등<sup>[11]</sup>방법중 Lloyd등<sup>[11]</sup>이 사용한 동적 계획 정합방법에 대한 실험을 하였다. 여기서는 에지점들을 정합하기 위한 특징으로 에지 좌우의 평균 빛의 밝기, 에지 사이의 거리, 각 에지의 방향 등을 사용하였다. 정합된 결과를 그림11에 도시하였으며 1138쌍이 올바르게 정합되었다. 이와 같이 본 논문에서 제안한 방법에 비해 정합률이 떨어지는 이유는 물체들간의 경계가 뚜렷하지 못한 경우 각 물체의 경계에서 에지를 검출하지 못하여 에지와 에지 사이의 영역이 여러 물체에 걸치게 되거나 스테레오 좌우 영상중 한 영상에서만 잡음 등에 의해 에지점이 검출될 경우 에지와 에지사이의 거리와 평균 빛의 밝기가 다르게 된다. 또한 스테레오 영상 취득시 카메라 이동에 의해 조명의 각도가 달라지게 되어 좌우 영상의 전체 빛의 밝기가 다르게 되는 경우도 발생하게 된다. 따라서 빛의 밝기나 에지 사이의 거리 등을 특징으로 사용하여 정합할 경우 올바른 정합 결과를 얻기 힘들어 알 수 있다. 기존의 특징베이스 정합 방법에서 특징으로 선소(line segment) 등을 이용하는 경우 이들 특징을 다각형근사화와 같이 근사화 과정을 이용하여 구하게 된다. 이 때 근사화된 특징들은 원래 화소가 존재하는 위치와 오차가 있게 되므로 정확한 위치와 정합할 수 없게 되고 따라서 깊이 정보를 구하는데 있어서 오차가 발생한다. 본 논문에서 수행한 정합에서는 지문 영상이 존재하는 특징점들에 대해



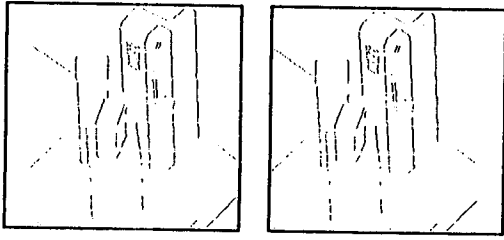


그림 11. Lloyd 등의 방법을 적용한 경우 정합 결과  
Fig. 11. Matching results obtained by applying the matching method proposed by Lloyd et al.

화소단위로 정합을 하게 되므로 보다 정확한 정보를 얻을 수 있다.

## VI. 결론

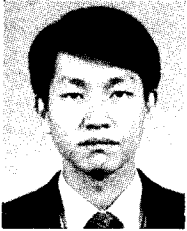
본 논문에서는 특징베이스 스테레오 정합 방법에서 정확한 정합점을 구하기 위하여 Gaussian-Laplacian 연산자에 의한 구한 영교차점과 Gauss 함수의 스케일값과의 관계를 나타낸 스케일 공간상의 지문 영상을 정합 특징으로 사용하였다. 이들 영교차 지문 영상들의 특징을 이용하는데 적합하고 올바른 정합률을 높이기 위하여 스케일 공간상에서 최적성 원리에 근거한 동적 계획 기법을 이용하는 방법을 제안하였다. 동적 계획 기법을 사용하므로써 모든 가능한 정합을 수행한 후 최적 정합을 찾아내는 brute-force 방법에 비해 계산시간을 줄일 수 있는 이점이 있으며, 국부적이 아닌 전체 정합 문제를 최적화하므로 올바른 정합결과를 얻을 수 있다. 보다 높은 정합률을 얻기 위하여 정합과정에서 잘못 정합된 경우가 발생하면 이를 영상의 각 행에 존재하는 특징점들간의 연속성 관계와 순서 구속조건을 이용하여 교정하는 후처리과정을 제안하였다. 실험에서는 제안한 알고리즘을 합성 영상 및 실제 영상에 적용하여 우수한 정합 결과를 얻을 수 있었으며 보다 정확한 깊이 정보를 추출할 수 있다.

앞으로 특징점들에 대해서만 구해진 깊이 정보를 보간하는 방법에 대해 연구가 진행되어야 하겠으며 이 때 단순히 깊이 정보만을 이용하는 것이 아니라 영상의 다른 특징들을 함께 사용한 방법에 대한 연구가 진행되어야 하겠다.

## 參 考 文 獻

- [1] Y. Shirai, *Three-Dimensional Computer Vision*. Springer-Verlag: New York, pp. 122-140, 1987.
- [2] U.R. Dhond and J.K. Aggarwal, "Structure from stereo-a review," *IEEE Trans. syst., Man, Cybern.*, vol. SMC-19, no. 6. pp. 1489-1510, Nov./Dec. 1989.
- [3] 강대갑, 권중장, 김성대, "스테레오 방식에서 일치성 문제를 해결하기 위한 새로운 선소 정합법," 전자공학회논문지, 제27권 제3호, pp. 116-123, 1992년 3월.
- [4] 심영석, 박성한, "3차원 정보를 얻기 위한 rule-based stereo matching algorithm," 전자공학회 논문지, 제 27권 제1호, pp. 151-163, 1990년 1월.
- [5] A. Witkin, "Scale-space filtering" in *Proc. Int. Joint Conf. Artificial Intell.*, pp. 1019-1021, 1983.
- [6] D. Marr and T. Poggio, "A computational theory of human stereo vision." *Proc. Royal soc. London*, vol. B204, pp. 301-328, 1979.
- [7] J.J. Clark, "Singularity theory and phantom edges in scale space," *IEEE Tans. Pattern Anal. Mach. Intell.*, vol. PAMI-10, no. 5, pp. 720-727, Sept. 1988.
- [8] J.J. Clark, "Authenticating edges produced by zero-crossing algorithms," *IEEE Trns. Pattern Anal. Mach. Intell.*, vol. PAMI-11, no. 1, pp. 43-57, Jan. 1989.
- [9] 최우영, 최종수, 이상욱, 박래홍, "스케일 공간상의 지문 영상을 이용한 스테레오 정합," 대한전자공학회 논문집, 제28권 B편 제2호, pp. 53-60, 1991년 2월
- [10] Y. Ohta and T. Kanade, "Stereo by intra-and inter-scanline search using dynamic programming," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-7, no. 2, pp. 139-154, Mar. 1985.
- [11] S.A. Lloyd et al., "A parallel binocular stereo algorithm utilizing dynamic programming and relaxation labelling," *Comput. Vision Graphics Image Process.*, vol. 39, no. 2, pp. 202-225, Aug. 1987.
- [12] E. Horowitz and s. Sahni, *Fundamentals of Computer Algorithms*. Computer Science Press: Maryland, Chap. 5, 1978.

著 者 紹 介



崔 祐 榮 (正會員)

1961年 7月 1日生. 1985年 2月 서강대학교 전자공학과 졸업(공학사). 1987年 2月 서강대학교 대학원 전자공학과 졸업(공학석사). 1992年 2月 서강대학교 대학원 전자공학과 졸업(공학박사). 1990年~1991年 생산기술원 HDTV 연구단 연구원. 1992年 3月~현재 명지대학교 전자공학과 조교수. 주관심 분야는 영상처리, 컴퓨터비전, 영상코딩 등임.

朴 來 弘 (正會員) 第23卷 第6號 參照

현재 서강대학교 전자공학과 부교수.