상관 윈도우사이의 왜곡을 보정한 특징점 정합 기법

(Feature Point Matching Technique using Adjustment of Distortion between Correlation Windows)

하 승 태 † 한 준 희 ††
(Seung-Tae Ha) (Joon Hee Han)

요 약 본 논문은 영상과 연관된 3차원 정보로부터 초기 3차원 변환을 유추, 상관윈도우를 변환시켜 정합에 이용하는 새로운 정합기법을 제안한다. 즉, 초기 스테레오 정합 등을 통한 3차원 정보를 추출하고, 인위적인 초기 특징점의 대응을 통해 3차원 변환을 얻으며, 이를 이용해 상관 윈도우의 3차원 변환을 가능하게 한다. 상관 윈도우의 3차원 변환은 기존의 방법이 가지는 영상 흐름의 2차원적인 제한을 이용한 정합방법에 비해 실제 카메라의 변환 유추에 합당하다. 또한 3차원 변환을 통해 정합 대상 점의 탐색범위를 최소화하고 정합의 결과에 신뢰성을 더한다. 실험에서는 다양한 영상의 정합 결과와 기존 방법과의 상관 계수 비교를 통해 본 논문이 제안하는 정합방법의 우월성을 보인다.

Abstract This paper proposes a new feature point matching algorithm that uses 3-dimensional transformed correlation windows. We use an initial stereo matching method to find 3-dimensional information which corresponds to the image. Applying initial feature point-matching with human-computer interaction, a 3-dimensional transformation can be obtained. We can also calculate 3-dimensional transformed correlation windows. This algorithm is more reliable than previous matching algorithms which use 2-dimensional transformation because the use of 3-dimensional transformation is closer to a real camera transformation. We show various image-matching examples and graphs which represent higher correlation coefficients than those of previous matching algorithms. Experimental result show that our proposed matching algorithm is more accurate than previous matching algorithm for matching image pairs having large disparities.

1. 서 론

서로 다른 두 영상에서의 정합은 한 영상에서의 특정한 한 점과 이에 대응되는 다른 영상의 점을 찾는 문제이다. 정합을 수행하고자 하는 물체에 대해 카메라로써 서로 다른 위치에서 두 장의 영상을 취득한다면, 각각의 카메라에 투사되는 영상도 달라지게 된다. 이 때, 대상 물체가 갖는 특정 지점이 두 영상에서 갖는 위치를 파악하여 두 영상간의 대응 관계를 알아내는 것은 컴퓨터 비전 분야에서 오랫동안 연구해 온 중요한 문제로서, 이를 일치문제(correspondence problem)라 한다. 일치문제는 모델에 대한 깊이(depth)정보나 모델의 움직임(motion), 카메라의 움직임을 추정하는 기초가 된다.

이에 대한 기존의 연구를 살펴보면, Deriche[1]는 초기 정합으로 찾은 에피폴라 기하(epipolar geometry)를 이용하여 기초 행렬(fundamental matrix)을 유추, 개선된 특징점 정합 결과를 도출하고자 하였다. Zhang[2]은 Deriche [1]와 마찬가지로 에피폴라 기하를 도입하였으나 특징점의 초기 정합의 결과로 발생하는 애매성(ambiguity)의 해결을 위해 이완화(relaxation) 방법을 도입하였다. Cheng [3]은 2차원 affine 변환 제한을 둔 영상의 흐름을 가정한 특징점 정합방법을 제안하였고, 또한 Li[4]는 2차원 변환에 불변적인 특징 성분과 이들간의 관계를 사용하여 ARS라는 가중치 그래프를 구성하여 정합을 시도하였다.

그러나 카메라의 위치 이동이 클 경우 위의 방법들은 좋은 정합결과를 내지 못한다. 이는 기존의 방법들이 정합의 대상을 찾을 때 특징점의 이동 거리에 제한을 둠으로

† 정 회 원 : (주)싸이버뱅크 PDA 개발부 연구원
hast@postech.ac.kr
†† 종신회원 : 포항공과대학교 컴퓨터공학과 교수
joonhan@postech.ac.kr
논문접수 : 2000년 6월 22일
심사완료 : 2001년 2월 8일

서 정합의 대상을 한정하거나, 영상 전체의 흐름을 2차원 변환으로 근사함으로서 정합쌍을 구하려 하기 때문이며, 이는 두 영상을 취득하는 카메라의 위치 이동이 클 경우에 그 제한을 적용하기 어려우므로 정합 결과의 신뢰성이 떨어진다.

이에 대해 본 논문에서는 카메라간의 3차원 변환을 직접 유추함으로서 정합의 신뢰성을 획득한다. 즉, 인위적인 초기 특징점의 대응과 스테레오 정합을 이용한 3차원 정보를 통해 모델간의 초기 3차원 변환관계를 얻고, 추출해낸 각 특징점에 이를 적용, 3차원 변환(3-dimensional transformation)된 상관 윈도우를 얻는다. 3차원 변환된 상관 윈도우는 카메라간의 변환에 의해 달라지는 상관 윈도우의 효과(그림 1)를 보정해주며, 이에 따라 영상의 특징점에 대한 정합 결과도 정확하게 된다.

본 논문의 차례는 다음과 같다. 2장에서는 기존의 정합 방법을 살펴보고, 3장에서는 기존 정합방법의 문제점과 제안하는 정합방법의 이론과 알고리즘을 살펴본다. 4장에서는 다양한 실험 결과를 통하여 기존 정합방법과 제안하는 정합방법의 상관 계수를 비교, 제안하는 방법의 우월성을 보이며, 마지막으로 5장에서는 결론 및 향후 연구 방향을 토론한다.

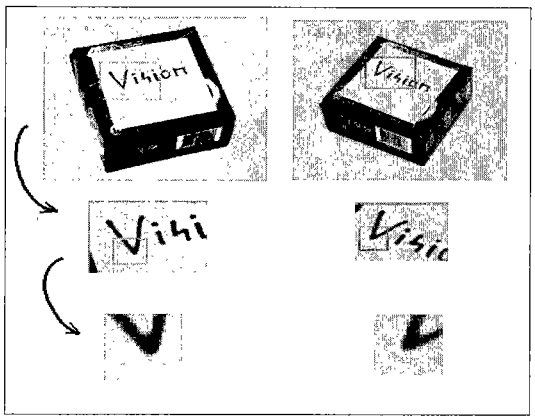

그림 1 시점에 따라 달라지는 상관 윈도우

2. 기존의 정합 방법

기존에 발표된 특징점의 정합 방법은 상관 윈도우를 그대로 이용하기도 하지만, 2차원 affine 변환이나, 영상의 취득시에 발생하는 카메라의 관계인 에피폴라 기하를 이용하기도 한다. Deriche[1]는 초기 특징점을 높은 미분값을 가지는 점들로서 추출하였고, 이에 대해 NCC (normalized cross-correlation)를 이용해 초기 정합을 수행하여 LMedS[5]방법을 통해 잘못된 정합을 바로잡고, 이들 초기 정합쌍으로 두어 에피폴라 기하를 복구하였다. 복구해낸 에피폴라 기하에 의해 각 영상의 대응하는 에피폴라 라인상에 상호 대응하는 특징점들이 놓이게 되므로 추가 정합이 더욱 용이하다. Zhang[2]도 Deriche[1]와 마찬가지로 에피폴라 기하를 정합의 선결조건으로 하고 있다. 다만 초기 정합쌍의 보완에 있어서 에너지 함수를 도입하여 초기 정합쌍의 대응 강도를 계산하는 이완화를 수행, 에너지 함수를 안정된 값으로 수렴시킴으로서 정확한 정합을 가능하게 한다.

Li[4]는 2차원 크기변화, 이동, 회전에 불변인 특성을 가지는 정보들을 이용하였다. 이는 특징점이나 경계선의 특성이나 각 선들이 이루는 각과 같은 관계, 혹은 선들이 가지는 길이에 대한 함수로서, 이들을 이용하여 정합에 적용되는 가중치 그래프를 형성하는데 이를 ARS 그래프라 한다. 이에 대해 이완화 기법을 이용하여 최적의 정합 결과를 얻는다.

Affine 변환을 이용한 점들간의 정합에 관한 연구 결과는 Cheng[3], Sprinzak[6]에 나타나 있다. Cheng[3]은 다음과 같은 2차원 affine 통합함수를 이용하였다.

$$\begin{bmatrix} x_j' \\ y_j' \end{bmatrix} = R\begin{bmatrix} x_j \\ y_j \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} scos\theta - ssin\theta & \\ ssin\theta & scos\theta \end{bmatrix}\begin{bmatrix} x_j \\ y_j \end{bmatrix} \quad (1)$$

이에 대한 최적의 통합 affine 변환은 변환된 점간의 거리가 최소가 되는 변환이다. Sprinzak[6]는 canonical form을 이용한 수학적 증명을 통하여 정합이 빠른 시간내에 수행될 수 있음을 보인다. 정합하고자 하는 두 대상점 집합이 $A=\{a_1, a_2, \ldots, a_n\}$, $B=\{b_1, b_2, \ldots, b_n\}$ 일 때 A, B에 대한 canonical form이 회전 변환 상에서 정합되면, A, B는 선형 변환 상에서 정합된다.

기존 정합 알고리즘은 2차원 물리적 변환이나, affine 변환을 가정으로 하고 있다. 이에 비해 영상의 정보는 실제로 3차원이므로, 영상을 얻은 시점의 차가 클 경우 2차원 변환을 가정으로 하는 기존 알고리즘은 정합의 신뢰성이 떨어지고, 폐색이나 잡음에 약하게 된다.

3. 상관윈도우의 3차원 변환을 통한 정합 방법

이 장에서는 카메라의 이동이 큰 경우에 대해 잘못된 정합 결과를 보이는 원인을 알아보고, 이를 보정할 수 있는 상관 윈도우의 3차원 변환을 이용한 새로운 정합방법을 제안한다.

3.1 기존 정합 방법의 문제점

2장에서 살펴본 기존의 정합방법은 다음과 같이 나눌 수 있다.

· 2차원 affine 변환을 이용한 정합
· 에피폴라 기하를 이용한 정합
· 기타 정합 (그래프 등)

2차원 affine 변환을 이용한 정합은 영상의 흐름을 2차원적으로 제한하여 어떻게 변환되었는가를 가정함으로서 정합을 수행한 것이다. 이는 영상의 변환이 2차원적인 변환으로 근사하기 힘든 경우, 즉, 카메라 시점의 변화에 있어서 그 거리가 크다던가 부분 폐색 같은 문제가 발생할 경우 affine 변환에 의한 신뢰성은 떨어진다고 할 수 있다. 또한, 에피폴라 변환을 이용한 정합은 초기 정합시 고전적인 상관 윈도우의 단순 비교와 영상에서의 정합 거리, 미분 방향에 대한 제한 등을 설정하기 때문에, 카메라의 변환이 큰 경우 정확한 정합쌍을 유추하기가 힘들다. 즉, 기존 정합 방법의 문제점은 영상의 변환이 그리 크지 않다는 가정조건에서 출발한 것이 된다. 따라서, 카메라의 변환이 클 경우, 정합 알고리즘이 가지고 있는 가정조건의 적용이 힘들게 되며, 결과적으로 도출되는 정합쌍도 잘못된 정합쌍이 많아질 가능성이 높아지게 된다.

그림 1은 같은 영상을 서로 다른 카메라 시점에서 취득한 것이다. 영상의 대응하는 특징점에 대하여 설정된 상관 윈도우가 나타나 있는데, 두 특징점은 대응하는 점임에도 불구하고, 카메라의 시점이 다름에 따라 상관 윈도우가 상당한 차이를 가지는 것이 보인다. 이러한 상관 윈도우의 차이는 기본적으로 카메라의 서로 다른 위치에서의 영상취득에 있으며, 이는 궁극적으로 2차원 변환이 아니라 3차원 변환에 기인한다. 따라서 이러한 상관 윈도우의 변화를 영상 전체적으로 유추하기에는 2차원 변환의 가정으로는 부족하며, 3차원 변환을 기초로 한 정합방법이 필요하게 된다.

3.2 상관 윈도우의 3차원 변환을 이용한 정합 방법

(1) 상관 윈도우를 변환하는 이유

상관 윈도우 자체는 3차원 모델을 2차원 카메라로 투사한 영상의 부분영역(patch)이다. 두 영상을 얻을 때의 두 카메라의 다른 위치는 결국 카메라 좌표(camera coordinate)의 3차원 변환을 야기시키게 되며, 이에 따라 정합의 대상인 특징점의 상관윈도우에 해당하는 영역 또한 3차원 변환에 의해 모양이 변하게 된다. 이것이 위에 설명한 상관 윈도우가 뒤틀리는 원인이 된다. 거꾸로, 각각의 카메라 사이의 3차원 변환을 알고 있다고 가정하면 변환에 의해 뒤틀어진 상관윈도우를 보정할 수 있다는 뜻이 된다.

(2) 3차원 변환의 유추

스테레오 정합과 인위적으로 구한 영상간의 초기 정합쌍으로부터 정합하고자 하는 두 영상 정보간의 3차원 변환 행렬을 얻는다. 대응하는 3차원 좌표가 각각 $\vec{x}$, $\vec{y}$이고 3차원 변환 행렬이 M_{3d}일 때, M_{3d}는 4*4의 차원을 가지며, 대응하는 3차원 좌표의 관계는 수식 (2)와 같게 된다.

$$M_{3d}=\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\vec{y}=M_{3d}\vec{x} \qquad (2)$$

따라서 3차원 좌표의 대응 집합으로부터 최적의 M_{3d}를 구할 수 있으며, 이러한 방법에는 Levenberg-Marquadt 최소화 방법, Newton 방법, multiple linear regression등이 있으나 본 논문에서는 그 연산 속도가 빠르고 결과값 수렴에 안정적인 multiple linear regression[7]을 스테레오 정합으로 얻은 3차원 정보에 적용하며, 셀프-캘리브레이션의 경우 3차원 유클리디안 정보이므로 강체 변환의 제한이 가해진 Horn[8]의 방법을 사용한다.

(3) 상관 윈도우의 탐색 위치

3차원 변환을 이용하여 특징점이 정합되는 대략적인 위치를 추측할 수 있다.

첫 번째 영상에서 취득한 특징점이 k개 일때, 각 특징점을 m_i라 하면, 이 m_i가 가지는 3차원 좌표를 $m_{(3d)i}$라 하면 $m_{(3d)i}$를 3차원 변환 행렬에 의해 변환한 좌표는 수식 (3)과 같이 표현된다.

$$m_i=(x_i, y_i) \quad , \quad 1\le i\le k$$

$$m'_{(3d)i}=M_{3d}\cdot m_{(3d)i} \qquad (3)$$

이 $m_{(3d)i}$가 두 번째 영상에서 가지는 좌표를 m'_i라 하면, m'_i는 m_i와 대응되는 정합쌍이 된다. 그러나 초기 정합에서 구한 3차원 변환행렬 M_{3d}가 정확하지 않으므로, m'_i는 m_i와 대응되는 정합점이 아니며, m_i와 대응하는 정합점을 m'_i가 갖는 좌표의 근방에서 재탐색함으로서 정확한 정합점을 구한다.

(4) 상관 윈도우의 3차원 변환 방법

추출한 특징점 m_i에 $d*d$의 상관 윈도우를 설정한다. m_i에 씌운 상관 윈도우에 속한 화소들을 수식 (4)와 같이 정의한다. 곧, $N_{(\alpha\beta)}(mi)$는 상관 윈도우의 좌측 위로부터 가로로 α, 세로로 β의 위치에 있는 좌표값을 나타낸다.

$$N_{(\alpha\beta)}(m_i)=(x_i-d/2+\alpha, y_i-d/2+\beta) \qquad (4)$$

상관 윈도우에 3차원 변환을 적용하는 것은, 각각의 상관 윈도우의 화소 $N_{(\alpha\beta)}(m_i)$에 3차원 변환을 적용하는 것이다. 3차원 변환을 적용하려면 좌표들 또한 3차원 좌표이어야 하므로 스테레오 또는 셀프-캘리브레이션으로 구한 3차원 정보에 2차원 화소를 대응시켜 화소에 대응하

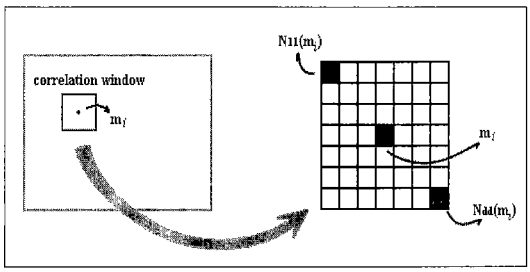

그림 2 특징점에 설정한 상관 윈도우의 각 화소들

는 3차원 좌표를 구한다. 대응된 좌표를 $N_{3d(\alpha\beta)}(m_i) = (X(m_i)_{\alpha\beta}, Y(m_i)_{\alpha\beta}, Z(m_i)_{\alpha\beta}, 1)$이라 하면, 3차원 변환된 상관 윈도우의 좌표들은 수식 (5)와 같다.

$$N'_{3d(\alpha\beta)}(m_i) = M_{3d} \cdot N_{3d(\alpha\beta)}(m_i) = M_{3d} \cdot \begin{bmatrix} X(m_i)_{\alpha\beta} \\ Y(m_i)_{\alpha\beta} \\ Z(m_i)_{\alpha\beta} \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} a_{11}X(m_i)_{\alpha\beta} + a_{12}Y(m_i)_{\alpha\beta} + a_{13}Z(m_i)_{\alpha\beta} + a_{14} \\ a_{21}X(m_i)_{\alpha\beta} + a_{22}Y(m_i)_{\alpha\beta} + a_{23}Z(m_i)_{\alpha\beta} + a_{24} \\ a_{31}X(m_i)_{\alpha\beta} + a_{32}Y(m_i)_{\alpha\beta} + a_{33}Z(m_i)_{\alpha\beta} + a_{34} \\ 1 \end{bmatrix} \quad (5)$$

이 때 변환된 좌표인 $N'_{3d(\alpha\beta)}(m_i)$ 도 또한 3차원 좌표이므로, 이를 2차원 영상의 좌표로 바꾼 것을 $N'_{3d(\alpha\beta)}(m_i)$라 하면, 2차원 좌표인 $N_{(\alpha\beta)}(m_i)$와 $N'_{(\alpha\beta)}(m_i)$는 이론적으로 서로 대응 관계에 놓이는 정합 윈도우가 되고, 이 상관 윈도우의 중심 좌표를 각각 m_i, m'_i라 하면 m_i와 m'_i는 서로 대응하는 특징점이 된다.

그러나 초기 정합쌍에서 유추한 M_{3d}는 구하고자하는 정확한 3차원 변환 행렬이 아니므로 위의 정합윈도우들과 특징점들은 정확한 대응의 관계를 갖지는 않게 되며, 이에 대해 정확하게 대응하는 특징점을 찾는 것이 바로 본 논문의 목적이 된다. 여기서 초기 대응점을 수작업으로 얻는다고 가정하였을 때, M_{3d}는 어느 정도의 정확성을 보장하게 되며, 이에 따라 m_i에 대응하는 정확한 대응점은 m_i가 M_{3d}에 의해 변환된 점인 m_i의 주변에 있게 된다. 따라서 대응점의 탐색은 m'_i의 주변에서 행하게 되며, 이때 대응정도의 결정은 m'_i의 각 주변 점에 $N'_{3d(\alpha\beta)}(m_i)$가 두 번째 영상에서 가지는 2차원 좌표인 $N'_{(\alpha\beta)}(m_i)$로 형성되어 변형된 상관 윈도우의 모양을 적용한다. 변형되기 전의 상관 윈도우의 모양과 이에 대응되는 변형 후의 상관 윈도우 모양이 그림 3에 나타나 있다.

그림 3과 같이 변형된 상관 윈도우를 각 탐색점에 적용함으로서 대응의 기준이 되는 상관 계수를 얻는다. $N_{(\alpha\beta)}(m_i)$와 대응할 새로운 상관 윈도우의 모양(template)을 $W_{(\alpha\beta)}(m_i)$라 하면, $W_{(\alpha\beta)}(m_i)$는 $N'_{(\alpha\beta)}(m_i)$의 좌표에서 $N'_{(\alpha\beta)}(m_i)$의 중심 좌표인 m'_i의 좌표를 뺀 값이 된다.

$$W_{\alpha\beta}(m_i) = N'_{(\alpha\beta)}(m_i) - m'_i \quad (6)$$

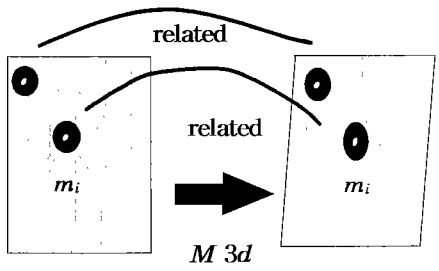

그림 3 초기 변환 행렬에 의해 변형된 상관 윈도우의 예

수식 (6)에 의해 $N'_{3d(\alpha\beta)}(m_i)$의 모양이 m_i에 대한 정합 탐색점 $s(m_i)$에 적용된다. m_i에 대한 정합 탐색점 $s(m_i)$가 m'_i로부터 가로로 a, 세로로 b만큼 떨어진 점일 경우, 이를 $S_{ab}(m_i)$이라 정의하면, $S_{ab}(m_i) = m'_i + (a, b)$가 된다. 따라서 최종적으로 m_i에 씌운 상관 윈도우의 화소 $N_{(\alpha\beta)}(m_i)$와 대응되는 탐색 정합점 $S_{ab}(m_i)$에 씌운 상관 윈도우의 화소 $K_{(\alpha\beta)}(S_{ab}(m_i))$는 수식 (7)과 같이 도출된다.

$$K_{(\alpha\beta)}(S_{ab}(m_i)) = S_{ab}(m_i) - W_{\alpha\beta}(m_i) \quad (7)$$

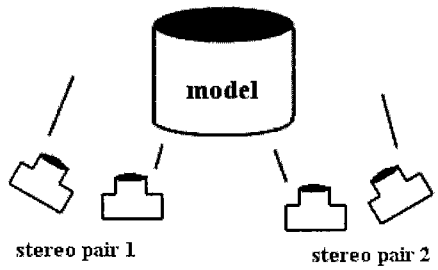

그림 4 모델로부터 영상 취득

3.3 알고리즘의 흐름

(1) 모델로부터 영상 취득

정합하고자 하는 두 위치로부터 그림 4와 같이 각각 2장씩의 스테레오 영상을 얻는다. 각 위치에서 두장씩의 영상을 취득하므로 필요한 영상은 네 장이 된다.

(2) 초기 3차원 모델 정보 취득

취득한 각 쌍의 영상에 대하여 2장에서 설명한 스테레오 정합을 이용해 3차원 정보를 구한다. 이런 방법으로

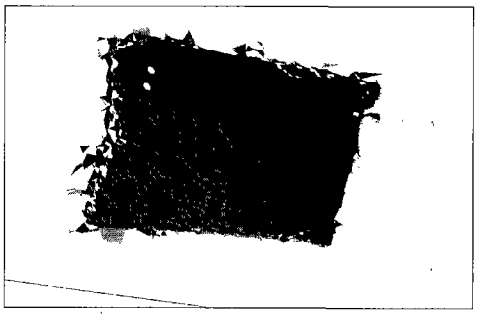

그림 5 좌측 스테레오로 얻은 3차원 정보

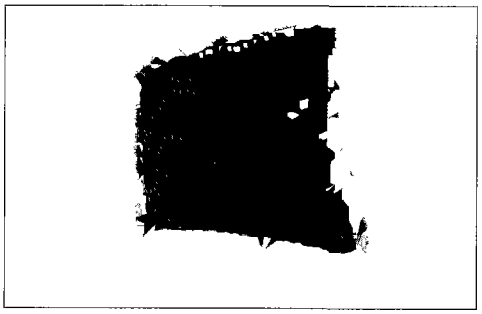

그림 6 우측 스테레오로 얻은 3차원 정보

얻은 3차원 모델은 세부적 정확도가 떨어지는 단점이 있으나 기존의 거리센서를 이용한 방법과는 달리 영상만으로도 어느정도 모델의 모양을 추정할 수 있다는 장점이 있다. 또한 그레이 레벨 영상과 대응되는 3차원 정보이므로 우리가 원하는 특징점 정합 방법에 유용하다고 할 수 있다.

(3) 초기 정합쌍 구하기

정합할 대상 영상의 쌍으로부터 10개 이상의 초기 특징점 정합쌍을 선택한다. 정합쌍은 모델의 전체 영역에서 골고루 얻는 것이 좋으며 정합의 대상은 특징이 확실한 코너 등으로 찾는 것이 정확한 초기 정합쌍을 구하는데 좋은 결과를 가져온다. 또한 얻어낸 영상에서의 초기 정합쌍으로 그에 해당하는 3차원 정합쌍을 구한다. 스테레오 방법으로 얻은 3차원 정보를 이용할 경우 렉티피케이션 된 영상의 2차원 좌표와, 구해진 3차원 정보의 (x, y)좌표는 서로 대응하므로 3차원 정보의 (x, y)좌표에 해당하는 z좌표를 찾아 점의 3차원 좌표를 구한다.

(4) 3차원 변환 행렬 얻기

구한 3차원 좌표의 대응으로부터 카메라의 이동 행렬을 얻는다. 대응하는 3차원 좌표가 각각 $\vec{x}$, $\vec{y}$일때 두 좌표의 관계는 $\vec{y} = M_{3d}\vec{x}$이며, 이때 M_{3d}는 4*4의 크기를 가지는 행렬이다. 3차원 좌표의 대응 집합으로부터 최적의 M_{3d}를 구할 수 있으며, 이러한 방법에는 Levenberg-Marquadt minimization, multiple linear regression[7], Horn[8]등이 있고, 본 논문에서는 스테레오 정합으로 3차원 정보를 얻었을때 multiple linear regression[7]을, 셀프-캘리브레이션으로 3차원 정보를 얻었을 때 Horn[8]의 방법을 이용한다. 또한 초기 정합쌍이 많은 에러를 포함하고 있을 경우를 생각해, 첫 번째 정합점을 M_{3d}를 이용해 변환한 좌표와 두 번째 정합점과의 차이가 클 경우 이를 정합쌍에서 제거하고 다시 변환행렬 M_{3d}를 유츄한다.

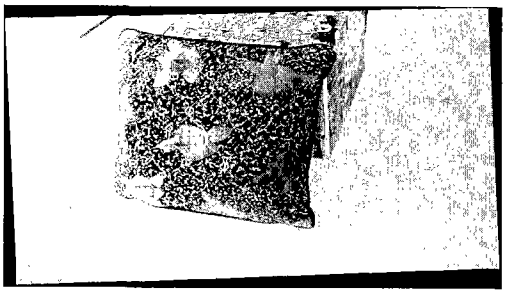

그림 7 방석의 우측 렉티피케이션 된 영상에 초기 정합쌍을 설정한 모습

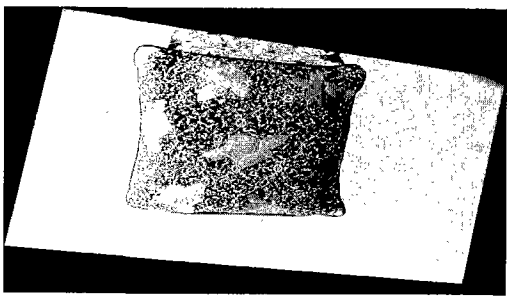

그림 8 방석의 좌측 렉티피케이션 된 영상에 초기 정합쌍을 설정한 모습

(5) 영상에서 특징점 추출 및 상관 윈도우의 변환을 이용한 정합

스테레오 방법등을 통해 얻은 3차원 정보와, 구한 영상의 2차원 정보를 대응시키는 것이 가능하므로 각각의 상관 윈도우에 해당하는 화소의 3차원 좌표를 얻을 수 있다. 또한 윗 단계에서 카메라의 이동행렬을 얻었으므로 구한 특징점 주위의 상관 윈도우를 3차원 변환시킨다. 변환 행렬이 완전히 정확하지 않으므로 대응되는 대상

화소와 완전히 일치된다고 볼 수 없다. 따라서 대상화소의 주변 화소에 대해 상관 윈도우를 적용하여 상관 계수를 비교, 가장 높은 점을 찾는다. 탐색의 범위는 영상의 해상도에 따라 다르나 보통 상하좌우 10화소정도에서 찾는다. 구해진 정합점의 상관계수가 너무 낮은 정합쌍은 임계값을 이용하여 정합쌍에서 제거한다. 상관 임계값의 경우 두 3차원 정보를 얻은 시점에 있어서 차이가 큰 경우나 영상의 특징이 좀더 조밀한 경우 낮은 값(0.6 정도)을 택하는 것이 일반적이며, 그렇지 않은 경우는 그보다 높은 값을 택하는 것이 좋은 결과를 낸다.

(6) 유일 정합쌍 실현

찾은 정합쌍 중에 첫 번째 영상의 점들이 두 번째 영상의 점들에 일치하지만, 반대로 두 번째 영상의 점이 첫 번째 영상의 점과 일치하는가 하는 문제가 있다. 이는 정합 연산자의 일관성에 관한 문제로서, 일관성이 좋은 연산자로 인한 정합쌍은 두 방법으로 얻은 정합쌍이 일치할 것이다. 두 번째 영상의 점과 첫 번째 영상의 점 주위를 탐색하여, 원래 대응점과 상관계수가 가장 높은 점이 일치한다면 정확한 정합으로 간주하고, 일치하지 않는다면 잘못된 정합으로 인정하여 정합쌍에서 제거한다.

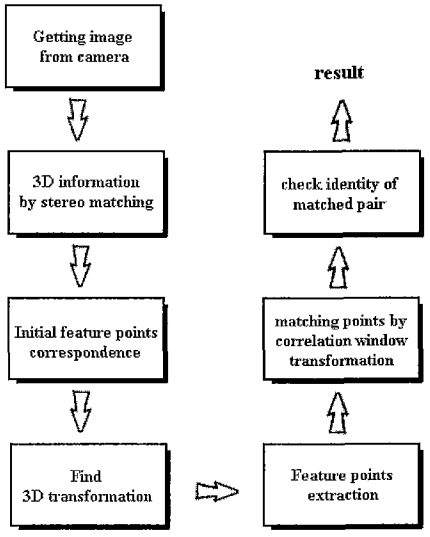

그림 9 정합 알고리즘의 블록 다이어그램

4. 실험 결과

4.1 특징점 정합 결과

실험에 사용된 모델은 모자이크 상자, 커피 상자, 쿠션, 컵등 4개의 물체이며, 각 모델에 대해 4개의 영상을 얻음으로서 2개의 스테레오 정합을 수행하였다. 스테레오 정합의 알고리즘은 Zhang[2]보다 성능이 우수한 권혁민[9]의 알고리즘을 이용한 초기 정합을 이용하였다. 이에 대해 조밀한 정합을 수행함으로서 구해낸 3차원 다각형 모델이 그림 5, 6에 나타나있다. 이에 대해 좌우 스테레오 정합 결과 영상에서 10개 정도의 초기 정합쌍을 그림 7, 8과 같이 인위적인 작업을 통하여 구하였다. 영상으로부터의 특징점 추출 알고리즘은 SUSAN 특징점 추출 알고리즘을 사용하였고, 이때의 밝기 임계값은 영상에 존재하는 무늬의 조밀도에 따라서 다르게 설정하였다. 이때는 조밀도가 높을수록 높은 임계값을, 조밀도가 낮을수록 낮은 임계값을 설정하여 적당한 숫자의 특징점을 추출하도록 하였으며, 실정에 사용된 수치는 50에서 130 이었다. 3차원 변환을 이용한 초기 탐색점으로부터의 탐색 범위는 상하좌우 10픽셀 이내이며, 이 중 변환된 상관 윈도우와의 상관 계수가 가장 높은 특징점을 골라 임계값과의 비교를 통해 정합쌍인지의 여부를 판별한다. 여기서 상관 계수의 임계값은 영상의 조밀도에 따라 0.6에서 0.7을 채택하였다.

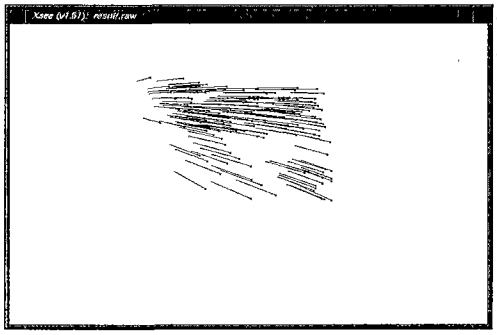

그림 10 영상의 정합 결과

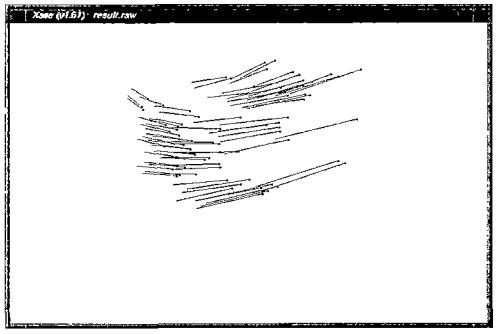

그림 11 커피상자 영상의 정합 결과

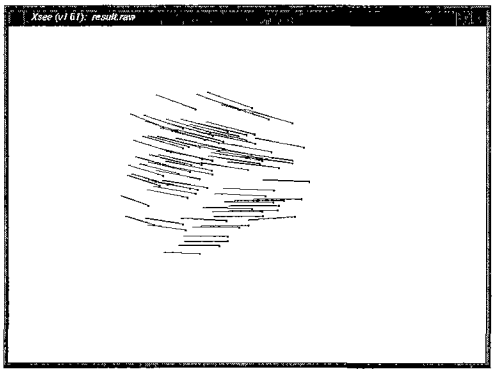

그림 12 방석 영상의 정합 결과

정합 결과에 대한 분석은 상관계수의 향상으로 판별한다. 이 때 비교하는 초기 상관 계수는 정합쌍에 변환을 수행하지 않은 상태에서 상관 윈도우를 씌운 결과이며, 그림 13, 14, 15에서 상관 윈도우의 변화에 따른 상관 계수값의 비교 결과를 확인할 수 있다. 상관 윈도우의 크기는 11에서 21사이로 하였다. 영상의 3차원 변환이 큰 쿠션 영상의 경우에서나 영상의 3차원 뒤틀림이 적은 컵 영상 같은 경우에서도, 상관 윈도우의 변환을 적용한 정합 방법이 적용하지 않은 결과에 비해 좋은 결과를 나타내고 있다.

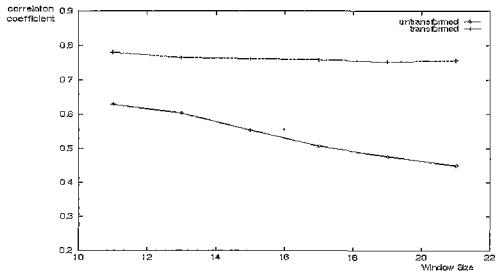

그림 13 컵 영상의 상관계수 변화

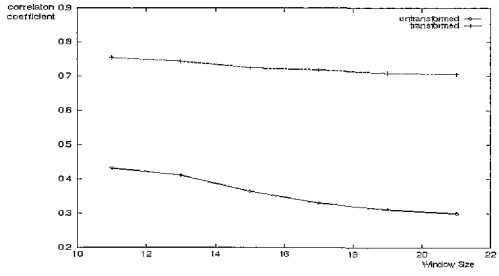

그림 14 커피상자 영상의 상관계수 변화

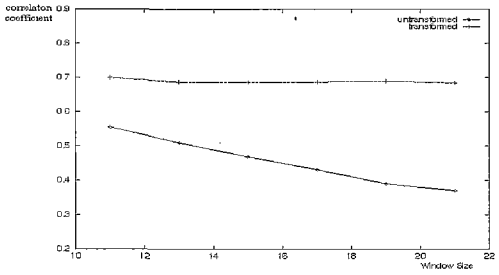

그림 15 방석 영상의 상관계수 변화

표 1 커피상자 영상의 향상된 상관계수 비교

window size	11	13	15	17	19	21
untransformed	0.432	0.411	0.364	0.331	0.312	0.300
transformed	0.754	0.744	0.725	0.719	0.709	0.705

표 2 방석 영상의 향상된 상관계수 비교

window size	11	13	15	17	19	21
untransformed	0.555	0.509	0.468	0.431	0.390	0.369
transformed	0.699	0.686	0.686	0.686	0.689	0.684

4.2 실험 결과에 대한 고찰

실험 결과에서 알 수 있듯이, 제안하는 상관 윈도우의 3차원 정합 방법을 통한 정합 알고리즘이 단순한 상관 윈도우의 비교를 통한 정합 알고리즘보다 뛰어나며, 이는 정합하고자 하는 대상 영상의 시각의 차이나 상관윈도우의 크기에 관계없이 좋은 결과를 보인다.

그러나 정합의 과정에서 적게는 5분에서 많게는 10분 이상의 시간이 소요된다. 이는 초기 스테레오 정합의 선행이 필요한 때문으로 기존의 정합 방법보다는 느리다는 단점을 가지고 있다. 따라서 이는 실시간 정합을 요할시에는 부적절하며, 시간에 관계없이 정확한 정합쌍을 요구하는 곳에 적절하다 할 수 있겠다.

4.3 초기 3차원 정보 및 정합쌍의 오차

초기 스테레오 정합에서 얻은 3차원 정보 또한 대응문제(correspondence problem)로 구해낸 3차원 정보이므로 실제 3차원 정보와는 차이가 있을 수 있다. 이러한 문제점은 초기 정합쌍의 선택시에 문제를 일으키는데, 초기 정합쌍의 인위적 선택시 정합쌍의 3차원 좌표가 실제 좌표와 크게 어긋날 수 있으며, 소수의 잘못된 초기 정합쌍으로 인하여 어긋난 결과가 유추될 가능성이 있다.

이 경우 해결책으로는 3차원 초기 정합 행렬과 초기 정합쌍과의 관계를 고려한다. 즉, 3차원 초기 정합 행렬은 초기 정합쌍들을 통해 유추되므로, 첫 번째 영상의 정합점이 3차원 초기 정합 행렬을 통해 변환된 좌표와 대응되는 두 번째 영상의 정합점을 비교한다. 이 두 3차원 좌표의 오차가 클 경우 잘못 선택된 초기 정합쌍 혹은 그 정합쌍 부근에서의 초기 스테레오 정합의 오차가 있을 가능성이 높다고 할 수 있으므로 해당하는 정합쌍을 초기 정합쌍에서 삭제하고 정합 행렬을 다시 유추한다면 잘못된 스테레오 정합으로 인한 정합 행렬의 오차를 줄일 수 있게된다.

5. 결 론

지금까지 초기 정합쌍을 이용한 상관 윈도우의 3차원 변환에 의한 개선된 2차원 영상의 특징점 정합방법에 대해 알아보았다. 카메라로부터 영상을 얻을 때 두 카메라 사이의 이동이 큰 경우에 있어서, 기존의 방법을 적용한다면 심한 영상사이의 변위로 인한 상관 윈도우의 심한 뒤틀림으로 인해 정합 결과에 대한 신뢰성을 얻기 어려우며, 특징점 정합시 상관 윈도우를 이용한 비교의 대상이 많아지는 단점이 있다. 이에 대해 본 논문은 초기에 작은 변위를 가진 영상의 쌍으로 얻은 스테레오 모델을 기존의 방법을 이용함으로서 구하고, 이를 영상의 초기 정합쌍과 대응시킴으로서 카메라의 3차원 변환으로 인한 뒤틀려진 상관 윈도우를 보정하여 정확한 상관 계수를 도출하는 동시에 특징점 정합시 정합의 대상이 되는 대응점의 탐색범위를 최소화하였다. 실험에서는 방석과 커피상자, 컵 모델의 정합을 통하여 본 논문이 제안하는 특징점 정합 결과의 우월성을 보였다.

그러나 본 논문에도 단점이 존재한다. 첫번째로 초기 스테레오 정합시 영상의 쌍이 2개 필요하므로

영상을 4개 얻어야 한다는 것이며, 두번째로 3차원 초기 모델의 복원으로 인한 정합 속도가 저하된다는 것이다. 또한 초기 정합의 획득에 있어서 인위적인 대응 작업이 요구된다는 것도 단점으로 지적할 수 있겠다. 따라서 추후 연구되어야 할 방향을 꼽자면 영상의 3차원 변환을 좀더 빠르게 유추하는

문제와 인위적인 대응 작업의 최소화 문제에 대한 것이 되겠다.

본 논문의 응용분야를 볼 때 정확한 정합쌍의 도출은 결국 정확한 모델의 3차원 변환의 도출을 가능하게 한다. 이러한 3차원 변환은 3차원 모델 구축에 필요 불가결한 것이며, 3차원 정보를 얻는 방법에 있어서 영상을 이용한다고 보면, 스테레오나 셀프-캘리브레이션을 이용한 3차원 정보로서 개선된 특징점 정합을 통한 3차원 정합을 수행할 수 있다(하승태[10]).

참 고 문 헌

[1] R. Deriche, Z. Zhang, Q.-T. Luong, and O. Faugeras, "Robust recovery of the epipolar geometry for an uncalibrated stereo rig," ECCV '94, pp.567-575, 1994.

[2] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong, "A robust technique for matching tow uncalibrated images through the recovery of the unknown epipolar geometry," Artificial Intelligence Journal, vol 78, pp.87-119, 1995.

[3] F. H. Cheng, "Point pattern matching algorithm invariant to geometrical transformation and distortion," Pattern Recognition Letters, vol 17, pp.1429-1435, 1996.

[4] S. Z. Li, "Matching : Invariant to translations, rotations and scale changes," Pattern Recognition, vol 25(6), pp.583-594, 1992.

[5] S. Laveau, Geometrie d'un systeme de N cameras. Theirie, estimeation et applications, Ecole polytecg- nique, France : Ph. D. Thesis, 1996.

[6] J. Sprinzak and M. Werman, "Affine point matching," Pattern Recognition Letters, vol 15, pp.337-339, 1994.

[7] S. Shapra and R. Canale, Numerical method of engineers, 2nd Ed., McGraw-Hill Company, 1990.

[8] B. Horn, "Closed-form solution of absolute orientation using unit quaternions," J. Opt. Soc. Amer. : A., vol 4(4), pp.629-642, 1987.

[9] 권혁민, 한준희, "유일 정합쌍을 이용한 효율적인 특징점 정합기법", 한국정보과학회 논문지, 제26권, 제6호, pp. 791-803, 1999.

[10] 하승태, 한준희, "영상정보를 이용한 3차원 정보의 정합", 한국정보과학회 춘계 학술발표논문집(B), pp. 595-597, 2000.

하 승 태

1998년 2월 한국과학기술원 전산학과 학사. 2000년 8월 포항공과대학교 전기전자공학부 졸업(컴퓨터 전공, 공학 석사). 2000년 6월 ~ 현재 (주)싸이버뱅크 PDA 개발부 연구원. 관심분야는 스테레오, 3D Range Data Merging.

한 준 희

정보과학회논문지 : 소프트웨어 및 응용 제 28 권 제 2 호 참조