

論文97-34S-2-6

복수의 거리영상 간의 변환계수의 추출

(Registration of Multiple Range Views)

鄭徒賢*, 尹逸東*, 李商郁*

(Do Hyun Chung, Il Dong Yun, and Sang Uk Lee)

요 약

거리영상으로부터 물체의 전체적인 형상을 복원하기 위해서는 여러 방향에서 취득한 물체의 거리영상들간의 변환을 알아야 하는데, 이처럼 물체의 전체적인 형상을 복원하는 과정에서 복수의 거리 영상 간의 변환 계수를 추출하는 과정을 거리 영상의 레지스트레이션(registration)이라고 한다. 본 논문에서는 시점이 서로 다른 복수의 거리 영상간의 변환을 추출하는 새로운 방법을 제안한다. 제안하는 기법은 초기 변환 추정과 반복 변환 개선의 2 단계로 구성된다. 초기 변환 추정은 물체 표면의 한 영역이 시각의 변화에 따라서 영상내에 다르게 투영되는 효과를 반영하도록 대표축(principal axis)을 개선시킨다. 다음, 거리 측정기의 역보정으로 찾아진 대응점 주위에서 가장 가까운 거리에 있는 점에 대응되도록 변환을 반복적으로 갱신하여 두 거리 영상 간의 변환 계수들을 추출한다. 제안하는 변환 계수 추정 알고리즘의 성능을 평가하기 위하여 공간 부호 거리 측정기로 실측한 거리 데이터로 실험을 수행한다. 실험 결과 제안하는 초기치 추정이 반복 추정의 수렴을 가속시킴을 알 수 있다.

Abstract

To reconstruct the complete 3-D shape of an object, several range images from different viewpoints should be merged into a single model. The process of extraction of the transformation parameters between multiple range views is called registration. In this paper, we propose a new algorithm to find the transformation parameters between multiple range views. The proposed algorithm consists of two step: initial estimation and iteratively update the transformation. To guess the initial transformation, we modify the principal axes by considering the projection effect, due to the difference of viewpoints. Then, the following process is iterated: in order to extract the exact transformation parameters between the range views: For every point of the common region, find the nearest point among the neighborhood of the current corresponding point whose correspondency is defined by the reverse calibration of the range finder. Then, update the transformation to satisfy the new correspondencies. In order to evaluate the performance the proposed registration algorithm, some experiments are performed on real range data, acquired by space encoding range finder. The experimental results show that the proposed initial estimation accelerate the following iterative registration step.

I. 서 론

컴퓨터 시각(computer vision)에서 최근 가장 활발히 연구되고 있는 문제 중의 하나가 바로 거리 영상으

로부터 물체의 전체적인 3차원의 형상을 복원하는 문제라고 할 수 있다^[6,7,8,11,12,13,14]. 거리 영상으로부터 물체의 전체적인 형상을 복원하기 위해서는 먼저 물체 표면의 모든 부분에 대한 거리 데이터가 취득되어야 한다. 그러나 거리 측정기의 시각 범위(field of view)의 한계, 또는 물체의 표면의 특성 등으로 인하여 한번의 계측만으로는 물체 표면의 모든 부분에 대한 거

* 正會員, 서울大學校 電氣工學部

(School of Elec. Eng., Seoul Nat'l University)

接受日字: 1996年7月12日, 수정완료일: 1997年2月12日

리 데이터를 취득하는 것은 사실상 불가능하다. 그러므로 물체 표면의 모든 부분의 거리 데이터를 취득하기 위해서, 일반적으로 서로 다른 복수의 시각에서 각각 거리 영상을 취득하여 이들을 합성하는 방법을 사용한다. 이를 위해서는 각 거리 영상들간의 변환을 반드시 알아야 합성을 할 수 있다. 이처럼 물체의 전체적인 형상을 복원하는 과정에서 필수적인 복수의 거리 영상간의 변환 계수 추출의 문제를 거리 영상의 레지스트레이션 (registration)이라고 한다^[7].

물체의 전체적인 3차원 형상의 복원을 위해, 한 물체에 대한 복수의 거리 영상을 취득하는 방법은 다양하게 연구되어 왔다. 먼저, 턴테이블이나 로봇 팔 등의 사전에 잘 보정된 기구를 사용하는 경우^[11]에는, 대부분 위치와 자세 등의 정보가 일정한 오차범위 내에서 주어지므로, 각 데이터간의 변환을 구하는 과정은 크게 문제시 되지 않는다. 한편, 주어진 데이터들 간의 상호 관계만을 이용하여 그 변환을 추정해야하는 경우, 각각의 거리 영상을 영역 분할하거나 특징을 추출하여 이들을 정합함으로써 그 변환을 추출하는 방법^[6], 또는 각각의 거리 영상을 적당하게 모델링한 다음 각 표면간의 거리를 최소화하는 변환을 반복적으로 계산하는 방법 등으로 분류할 수 있다. 전자는 거리 영상에서 특징을 추출하고 이를 정합하는 과정에서 많은 계산량을 요구한다. 또한 이러한 방법은 어떠한 특징을 사용하느냐에 따라 표면의 형태가 제한될 수 있고, 주로 미분 연산을 이용하기때문에 잡음에 민감한 단점이 있다. 후자는 반복계산의 수렴성만 보장된다면 매우 좋은 결과를 기대할 수 있다. 이러한 방법은 대부분 부정방정식의 최적해를 구하는 문제로 해석되고 있는데, 오차의 척도가 되는 함수를 어떻게 정의하느냐와 이것을 어떠한 방법으로 최소화하느냐가 주된 관건이라 할 수 있다. 많은 연구^{[7]-[14]}에서 한 거리 영상의 제어점들과 다른 거리 영상의 표면간의 거리를 최소화하는 ICP (iterated closed point) 알고리즘을 사용하는데, 점과 표면간의 거리를 계산하기 위해서 그 표면상에서 그 점과 가장 가까운 점을 찾기 위한 다양한 방법들이 제시되어있다.

본 논문에서는 물체의 전체적인 3차원 형상을 복원하기 위해 공간 부호 거리 측정기 (space-encoding range finder)로 취득한 시점이 서로 다른 복수의 거리 영상들을 합성하는 과정에서, 각 거리 영상간의 변환을 추출하는 새로운 방법을 제안한다. 일반적으로 물체의

전체적인 3차원 형상을 복원하는 과정은 그림 1에 도시한 바와 같이 먼저 거리 영상을 취득하고, 전처리 (preprocessing)를 거친 다음, 각각의 거리 영상간의 변환 (registration)을 구한다. 그런 다음에는 구해진 변환을 이용하여 모든 거리 데이터를 하나의 완전한 모델로 합성하는 과정 (integration and modeling)을 거친다. 본 논문에서는 그림 1에 표시되어있는 것처럼, 거리 영상의 취득, 전처리, 변환 계수의 추출 단계까지의 내용을 다루었다.

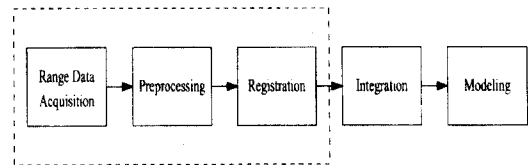


그림 1. 3차원 형상 복원의 블록 다이어그램
Fig. 1. The blockdiagram of 3D shape reconstruction process.

II. 공간 부호화 거리 측정기

1. 공간 부호화 거리 측정기의 기본 원리

공간부호화 기법^[1,2,3,4,5] 공간 부호화 기법이란 계속 대상이 될 공간을 작은 영역들로 나누어 각 영역에 특정한 부호값을 부여하고, 이를 이용하여 2차원 영상으로부터 물체의 3차원 정보를 얻어내는 방법을 일컫는다. 공간을 부호화하기 위하여 액정 소자의 광 투과율을 전자적으로 제어하여 %줄무늬 패턴을 생성하고 여기에 슬라이드 프로젝터의 빛을 비추어 패턴광을 %물체에 투영하는 방법으로 공간을 부호화하고 이로부터 물체의 3차원 정보를 얻어낸다. 이때, n 개의 패턴 마스크를 사용하면 측정 공간은 2^n 개의 작은 영역들로 분할되는데, n 이 충분히 크면 각 영역들은 아주 얇은 췌기 모양의 영역이 되어 평면으로 근사해도 무방하다. 각 점의 공간부호값이 결정되면, 그 값으로부터 2^n 개의 평면 중에 어느 평면에 해당되는지가 결정되고, 카메라에 입력된 영상 내에서의 2차원 좌표로부터 카메라에서 바라보는 시선이 결정되므로, 한 평면과 한 직선의 교점으로 보아 그 점의 3차원 좌표를 결정할 수 있다^[11]. 3차원 좌표 (x_w, y_w, z_w) , 2차원 영상내에서의 좌표 (x_c, y_c) , 그리고 공간 부호값 x_p 와의 관계는 각각 다음과 같이 주어진다.

$$\begin{bmatrix} h_c x_c \\ h_c y_c \\ h_c \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (1)$$

$$\begin{bmatrix} h_p x_p \\ h_p \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (2)$$

여기서 c_{ij} 가 이루는 3×4 행렬을 카메라 계수 행렬 **C**라 하고, p_{ij} 가 이루는 2×4 행렬을 프로젝터 계수 행렬 **P**라 한다. 보정 (calibration)을 통하여 **C**와 **P**를 구하면 x_c, y_c, x_p 로부터 3차원 좌표 (x_w, y_w, z_w) 를 계산할 수 있다.

2. 레인지 파인더의 역보정

본 절에서는 공간부호화 레인지 파인더의 역보정에 대해서 설명한다. 역보정이란 카메라 계수 **C**와 프로젝터 계수 **P**를 이용하여, 3차원 좌표 (x_w, y_w, z_w) 로부터 2차원 영상내에서의 좌표 (x_c, y_c) 와 공간부호값 x_p 를 역으로 계산해내는 과정을 의미한다.

역보정 방법은 식(1)에서 3행의 h_c 로 1행과 2행을 나누어 주면 바로 x_c, y_c 를 계산할 수 있다. 같은 방법으로 식(2)에서 2행의 h_p 로 1행을 나누어 주면 x_p 가 계산된다. 그러나 이 경우에 $h_c=0$, 또는 $h_p=0$ 이 되는 점들은 역보정을 할 수 없다. 만약 주어진 거리영상 내에 이러한 점들이 많이 존재한다면 이 방법은 사용할 수 없으므로 $h_c=0$, 또는 $h_p=0$ 이 되는 점들의 존재 조건을 살펴보는 것은 매우 중요한 일이 될 것이다.

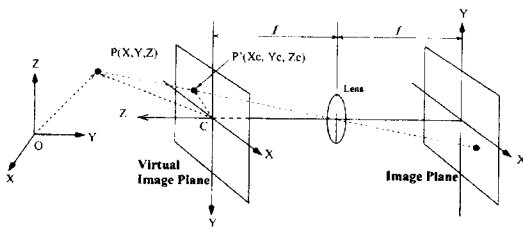


그림 2. 카메라 계수의 의미
Fig. 2. Meaning of the camera parameter.

먼저 $h_c=0$ 이 되는 점들에 대해서 살펴보면, 식(1)의 카메라 계수 행렬 **C**는 물체 좌표계와 카메라 좌표계 간의 변환과 투시 변환으로 이루어져 있음을 알 수 있다. 그러나 식(1)은 다음과 같이 두 단계로 나누어 쓸 수 있다. 먼저 물체 좌표계의 좌표를 다음과 같은 회전 이동과 평행 이동을 포함한 3차원 변환을 통하여

카메라 좌표계의 3차원 좌표로 변환한 후에

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}, \quad (3)$$

다시 아래와 같은 투시 변환을 통하여 카메라의 가상 이미지 평면 내의 2차원 좌표로 변환한다.

$$\begin{bmatrix} x_{ch} \\ y_{ch} \\ z_{ch} \\ w_{ch} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \quad (4)$$

그림 2에서 그 의미를 살펴보면, 벡터 OP 가 회전 이동과 평행 이동에 의하여 CP 로 변환된 다음, 다시 투시 변환에 의해서 CP' 로 변환되는 셈이다. 식(1), 식(3)과 식(4)를 보면 $w_{ch} = \text{frac } z' f + 1 = 0$, 즉, $z' = -f$ 이 되는 점에서 $h_c=0$ 이 됨을 알 수 있다. 이것은 그림 2로 그 의미를 설명하면 카메라의 렌즈의 중심을 지나고 z' 축, 즉 카메라의 시선에 수직인 평면 위의 점들에서 $h_c=0$ 이 성립함을 의미한다는 것을 알 수 있다. 따라서 물체 표면의 거리 영상 중에는 이론상으로 h_c 가 0이 되는 점이 없다는 것을 알 수 있다.

식(2)로부터 위에서와 같이 풀어보면 프로젝터에 대해서도 비슷한 결과를 얻는다. 즉, 물체 표면의 거리 영상 중에는 이론상으로 h_p 가 0이 되는 점이 없다는 것을 알 수 있다. 그러므로 식(1)과 식(2)에 의해서 이론적으로 거리 영상 내의 모든 점에 대해서 x_c, y_c, x_p 를 항상 계산해 낼 수 있다. 즉, 거리 영상로부터 공간부호 영상을 재생성하는 일이 이론적으로 가능하다는 것을 알 수 있다.

여기에서 부수적으로 (x_w, y_w, z_w) 의 형태로 주어진 거리 영상을 기준 평면으로부터의 거리의 형태로 변환하는 방법을 알 수 있다. 평면의 방정식 $h_c=0$ 을 물체 좌표계의 표현으로 고치면 다음과 같다.

$$c_{31}x_w + c_{32}y_w + c_{33}z_w + c_{34} = 0 \quad (5)$$

카메라 렌즈의 중심을 지나고 카메라의 시선에 수직인 평면의 방정식이 주어졌으므로, 각 점으로부터 이 평면까지의 거리를 계산하면, $(x_w, y_w, z_w)'$ 의 3차원 좌표로 주어진 거리 영상을 기준 평면으로부터의 거리로 표현한 깊이 영상 (depth map)으로 변환할 수 있다.

$$d(x_w, y_w, z_w) = \frac{|c_{31}x_w + c_{32}y_w + c_{33}z_w + c_{34}|}{\sqrt{c_{31}^2 + c_{32}^2 + c_{33}^2}} \quad (6)$$

이를 이용하여 거리 영상을 깊이 영상으로 변환한 예를 그림 3에 제시한다.



(a) 공간부호영상 (b) 깊이영상

그림 3. 거리영상을 깊이영상으로 변환한 예
Fig. 3. An example of the conversion of a range image into depth map.

식(5)의 평면 방정식으로부터 카메라가 물체를 바라보고 있는 시각 벡터를 물체 좌표계의 표현으로 다음과 같이 얻을 수 있다.

$$\frac{1}{\sqrt{c_{31}^2 + c_{32}^2 + c_{33}^2}} [c_{31} \ c_{32} \ c_{33}]^T \quad (7)$$

이렇게 구해진 거리 측정기의 시각 벡터는 거리 영상 간의 초기 변환을 계산하는 데 사용된다.

III. 변환 계수 추출 기법

본 장에서는 앞장에서 설명한 공간 부호화 거리 측정기의 역보정을 이용하여 복수의 거리 영상 간의 변환 계수를 추출하는 방법을 제안한다. 복수의 거리 영상 간의 변환을 구하려면 이웃하는 거리 영상들 간의 변환을 차례로 구하면 되므로, 본 논문에서는 2 개의 거리 영상 간의 변환을 구하는 문제에 대해서만 설명하기로 한다.

1. 전처리 과정

효과적인 물체의 형상 복원을 위해서는 무엇보다도 물체 표면의 정확한 거리 데이터를 취득하는 것이 중요하다. 그런데 이렇게 물체 표면의 거리 데이터를 얻는 과정에서 물체 주위에 있는 다른 물체, 또는 배경까지 함께 취득하게 된다. 거리 영상 간의 변환을 구하는 과정에서 이러한 배경의 거리 데이터는 필요없으므로 본 논문에서는 그림 4와 같은 과정을 거쳐서, 공간 영

상에서 물체와 배경을 분리한다.

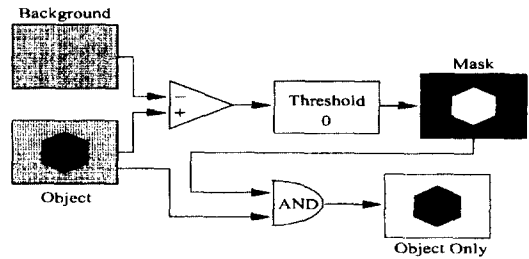


그림 4. 물체와 배경의 분리과정
Fig. 4. Separation of the object and background.

이를 좀더 자세히 설명하면, 물체의 공간부호영상을 생성하기 이전에 물체가 없는 배경만의 공간 부호 영상을 생성한다. 다음, 물체의 공간 부호 영상을 생성하고 이로부터 배경만의 공간 부호 영상을 뺀 것을 적절한 문턱치로 이진화하여 마스크를 만든다. 이렇게 생성한 마스크로 물체의 공간 부호 영상을 마스킹하여 물체만의 공간 부호 영상을 추출한다.

한 가지 더 고려해야할 사항은 바로 계측시에 발생하는 잡음인데, 이러한 잡음은 대상 물체 표면의 거리 데이터 중에 그릇된 정보를 포함시키므로 정확한 변환 계산을 할수가 없다. 그러나 이러한 잡음을 제거하기 위해서 보통의 smoothing 필터를 사용하게 되면 물체의 경계 부분의 정보가 없어지기 때문에 본 논문에서는 중간값 (median) 필터를 사용하여 잡음을 제거하였다.

2. 초기 변환의 추정

본 절에서는 거리 영상 내의 물체 표면의 데이터들의 통계적 특성을 이용하여 초기 변환을 추정하는 방법에 대해서 설명한다. 2개의 거리 영상이 주어졌을 때, 첫번째 거리 영상 의 한 점 p_1 을 두번째 거리 영상 내의 한 점 p_2 로 대응시키는 변환을 다음과 같이 쓸 수 있다.

$$p_2 = R p_1 + t, \quad (8)$$

여기서 R 은 회전이동을 의미하는 (3×3) 행렬이고, t 는 평행 이동을 의미하는 3차원 열벡터이다.

Dorai 등^[14]은 2 개의 거리 영상 간의 변환을 구하는 과정에서 데이터 점들의 3차원 좌표간의 공분산 행렬 (covariance matrix)의 고유벡터 (eigenvector)들이 이루는 세개의 직교축을 대표축 (principal axes)이

라 정의하고 이들 간의 변환을 계산하여 초기 변환을 추정하는 방법을 제안하였다.

이 방법을 자세히 설명하면, N 개의 관측점 \mathbf{x}_i 로 구성된 1 매의 거리 영상의 무게중심 $\bar{\mathbf{x}}$ 과 공분산 행렬 Σ 을 각각 다음과 같이 계산한다.

$$\bar{\mathbf{x}} = \begin{bmatrix} \mu_x \\ \mu_y \\ \mu_z \end{bmatrix} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad (9)$$

$$\Sigma = \frac{1}{N} \sum_{i=1}^N \{(\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T\}. \quad (10)$$

다음, Σ 의 세개의 정규직교 고유 벡터들을 계산하여 이들로 구성된 세개의 직교축을 그 거리 영상의 대표축이라고 정의한다. 첫번째 거리 영상의 대표축을 이루는 세 벡터를 $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ 일 때에 이들이 이루는 3×3 행렬을 S_1 이라 하고, 두번째 거리 영상의 대표축을 이루는 세 벡터를 $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ 일 때에 이들이 이루는 3×3 행렬을 S_2 라 하면, S_1 을 S_2 로 대응시키는 회전 변환행렬 3×3 R 은 다음과 같이 계산된다.

$$R = S_2 S_1^{-1}. \quad (11)$$

또한 두 거리 영상 무게중심을 각각 $\mathbf{g}_1, \mathbf{g}_2$ 이라 하면 다음 식에 의해서 평행 이동벡터 \mathbf{t} 를 구한다.

$$\mathbf{t} = \mathbf{g}_2 - R \mathbf{g}_1. \quad (12)$$

여기에서 주의해야 할 것은 고유벡터들의 대응관계이다. 각 대표축간의 회전이동을 계산하기 위해서는 \vec{u}_i 중의 어느 벡터가 \vec{v}_i 중의 어느 벡터로 변환된다는 대응관계를 알아야 하는데, $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ 과 $(\vec{v}_1, \vec{v}_2, \vec{v}_3)$ 가 이 순서대로 일치한다는 보장이 없다. 그래서 \vec{u}_i 들과 \vec{v}_i 들을 서로의 내적이, 다른 벡터와의 내적에 비해서 절댓값이 가장 큰 것들끼리 %대응되도록 재배열한 뒤에 사용한다.

이러한 방법은 두 거리 영상이 물체 표면 중에서 같은 부분들만을 포함하고 있어야만 정확한 결과를 얻을 수 있다. 즉, 거리 측정기의 시각 벡터에 수직인 평면 상에서의 회전 및 평행 이동만 했다고 가정했을 때에만 좋은 결과를 얻을 수 있다. 또한, 거리 영상을 취득할 때 물체의 표면에 있는 각 영역들은 거리 측정기의 시각 벡터에 수직인 평면으로 투영된 2차원 영상 내에서의 면적에 비례하는 최소만큼 샘플링된다. 만약 물체가 거리 측정기의 시각과 일치하지 않는 축에 대해서

회전이동을 하게 되면, 각 영역의 샘플링되는 점들의 개수가 변하게 되는데, Dorai 등의 방법은 이러한 변화를 반영하지 못하고 있다.

본 논문에서는 이러한 시각의 변화에 따른 데이터의 변화량을 고려하여 대표축을 계산하는 방법을 개선하였다. 앞에서 설명한 것과 같이 거리 영상 내에서 어떤 영역이 샘플링된 수는 거리 측정기의 시각 벡터에 수직인 평면에 투영된 2차원 영상 내에서의 면적에 비례한다. 이 2차원 영상 내에서의 면적은 원래 3차원 상에서의 면적에다 거리 측정기의 시각 벡터와 그 영역의 표면 법선 벡터가 이루는 각 θ 의 \cos 값을 곱한 것으로 볼 수 있다. 그러므로 역으로 2차원 영상 내에서의 면적에 $\cos \theta$ 의 역수를 곱해주면, 이 값은 그 영역의 3차원 상에서의 실제 면적을 의미한다고 볼 수 있다. 그 관계를 그림 5에 도시한다.

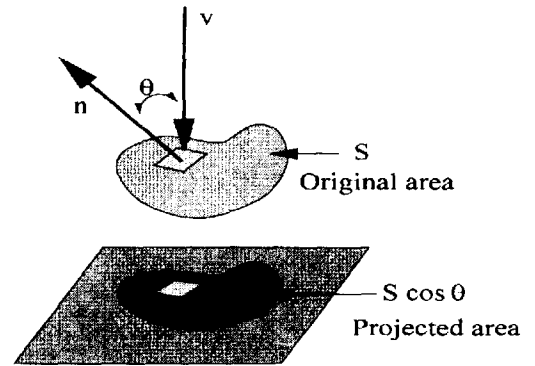


그림 5. 평면상으로 투영될 때의 면적의 변화
Fig. 5. Change of the area when projected to a plain.

그러나 이러한 원리를 이용하여 본 논문에서는 물체 표면의 각 점을 하나의 작은 평면 영역으로 보고, 각 점에서의 표면 법선 벡터와 거리 측정기의 시각 벡터가 이루는 각 θ_i 의 \cos 값의 역수를 각 점에 대한 가중치 w_i 로 정의하여, 무게 중심과 공분산 행렬을 계산한다. 거리 측정기의 시각 벡터는 앞 장에서 설명한 바와 같이 거리 측정기의 카메라 계수 C 로부터 식 (7)과 같이 주어진다.

각 점에서의 표면 법선 벡터는 각 점을 중심으로 5×5 크기의 원도우 내에 존재하는 모든 점들을 대표하는 평면을 최소사승법으로 구해 그 법선 벡터를 사용하며, 이 값들을 이용하여 i 번째 관측점에 대한 가중치 w_i 를 다음과 같이 정의한다.

$$\omega_i = \frac{1}{\cos \theta_i}, \quad (13)$$

여기서

$$\cos \theta_i = \frac{\mathbf{v} \cdot \mathbf{n}_i}{\|\mathbf{v}\| \|\mathbf{n}_i\|}. \quad (14)$$

이 때에 \mathbf{v} 는 거리 측정기의 시각 벡터를 의미하고, \mathbf{n}_i 는 각 점에서의 표면 법선 벡터를 의미한다. 이제 이를 가중치로 이용하여 거리 영상 내의 물체 표면의 점들의 가중치 무게 중심과 공분산 행렬을 각각 다음과 같이 계산한다.

$$\bar{\mathbf{x}} = \begin{bmatrix} \mu_x \\ \mu_y \\ \mu_z \end{bmatrix} = \frac{1}{W} \sum_{i=1}^N \{\omega_i \mathbf{x}_i\}, \quad (14)$$

$$\Sigma = \frac{1}{W} \sum_{i=1}^N \{\omega_i (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T\}, \quad (15)$$

여기서 $W = \sum_{i=1}^N \omega_i$ 이다. 이제 가중치 공분산 행렬의 고유 벡터들을 구하고, 식 (11)과 식 (12)를 이용하여 \mathbf{R} 과 \mathbf{t} 를 계산함으로써 시각의 변화를 고려한 두 거리 영상간의 변환을 추정한다. 제안하는 초기변환의 추정 과정을 그림 6에 도시하였다.

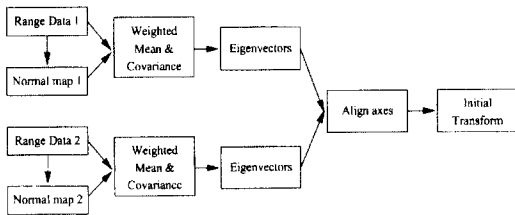


그림 6. 초기 변환의 추정 과정의 블록 선도
Fig. 6. Blockdiagram of the initial guess process.

3. 반복 계산에 의한 변환 계수의 추출

본 절에서는 공간 부호화 거리 측정기의 역보정을 이용하여 오차 함수를 정의하는 방법과 이를 최소화하는 방향으로 반복계산을 진행시키는 과정에 대해서 설명한다.

서로 다른 시각에서 취득된 2 개의 거리 영상 간의 변환 계수를 추출하기 위해서는 양쪽 거리 영상에 공통으로 계측된 영역이 상당 부분 존재해야만 한다. 이러한 공통영역의 데이터 점들간의 1 대 1 대응관계가 성립한다면, 두 데이터 간의 변환은 쉽게 구할수 있겠지만, 실제 경우에는 그 대응 관계를 쉽게 알 수가

없다.

본 논문에서는 공간 부호화 거리 측정기의 역보정을 이용하여 공통 영역의 데이터 점들 간의 대응 관계를 구한다. 거리 측정기의 역보정은 Blais 등 [15]이 레이저 거리 측정기에 적용하였는데, 각 화소의 거리 데이터로부터 그 점이 어느 위치의 레이저 광선으로 계측된 것인가를 결정하는 알고리즘을 제안하였고, 이 방법을 이용하여 2 개의 거리 영상에서 같은 레이저에 의해 계측된 점들을 대응시키는 방법으로 대응점 관계를 해결하였다.

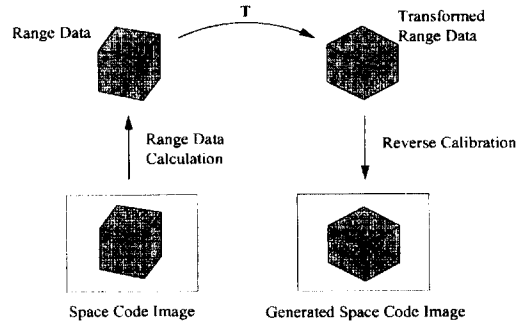


그림 7. 변환계수 추출과정에서의 역보정
Fig. 7. Reverse calibration in the registration process.

본 논문에서도 이들과 비슷한 방법을 공간 부호화 거리 측정기에 적용한다. 각 화소의 거리 데이터로부터 계측 당시의 입력 영상 내에서의 좌표 및 공간 부호값을 계산하는 방법은 앞에서 이미 설명하였으므로, 이를 이용하여 시각이 서로 다른 거리 영상에서 주어진 변환에 대한 대응점을 바로 찾아낼 수 있는 방법을 제안한다. 이렇게 구해진 대응 관계의 정확성의 여부는 주어지는 변환의 정확도에 기인한다. 즉, 정확한 변환이 알려져 있는 상황에서는 역보정에 의해서 결정된 공통 영역의 점들간의 대응관계도 정확하다고 볼 수 있는 것이다. 하나의 물체에 대해서 서로 다른 시각에서 취득한 두 벌의 거리 영상을 각각 V_1 , V_2 라고 하자. V_2 의 점들을 V_1 의 좌표계로 변환해주는 변환 행렬을 \mathbf{T} 라고 하자. 그림 7에서와 같이 \mathbf{T} 를 이용하여 V_2 의 점들을 V_1 과 같은 좌표계 표현으로 바꾼 다음 변환된 데이터에 앞에서 설명한 역보정을 적용한다.

이때 V_1 또는 V_2 어느 한 쪽에만 속한 점들은 계산된 영상 내의 좌표 (x_c, y_c) 나 공간 부호 x_p 의 값이 주어진 범위를 벗어나게 되어 쉽게 고려 대상에서 제

외시킬 수 있다. 이렇게 함으로써 V_2 에 속한 점들 중 공통 영역에 속하는 점들에 대해서는, 이 점들이 V_1 의 시각에서 보면 어느 위치에 해당하는가 내지는 V_1 의 점들 중에서 어느 점과 대응하는가를 쉽게 알 수 있게 된다. 그림 7과 같은 단계를 반복 적용할 때, 대응점들의 3차원 좌표 간의 차를 이용하여 k 번째 반복에서의 오차함수 E_k 를 다음과 같이 정의한다.

$$E_k = \frac{1}{N_k} \sum_{i=1}^{N_k} d(T_k \mathbf{p}_i, C_k(\mathbf{p}_i)), \quad (16)$$

여기서 \mathbf{p}_i 는 역보정 과정에서 제외되지 않고 오차함수 계산에 포함된 공통 영역의 점들을, N_k 는 \mathbf{p}_i 의 갯수를, $C_k(\cdot)$ 는 역보정에 의해서 성립된 대응 관계를, $d(\cdot)$ 는 3차원 공간상에서의 점들 간의 거리를 각각 의미한다. 즉, 이 식은 역보정에 의해 대응한다고 알려진 공통 영역의 점들 간의 3차원 거리의 평균값을 의미한다. 만일 \mathbf{T}_k 가 참값과 같다면 E_k 는 거의 0에 가까운 값을 갖게 될 것이다. 그러나 서로 다른 시각에서 얻어진 거리 영상의 점들 간에는 시각 방향에 따른 투영 효과나 샘플링에 의한 양자화 효과 등에 의해 정확한 1 대 1 대응관계는 기대하기 힘들기 때문에 \mathbf{T}_k 가 참값을 갖는다 하더라도 E_k 가 정확히 0이 되지는 않을 것이다. 따라서 E_k 는 각 단계에서 계산된 변환이 얼마나 참값에 가까운가를 판별하는 척도로는 사용될 수 있겠지만, 이것을 반복 계산에서 최소화하려는 목적함수로 사용하기는 적절하지 않다. 결국 현재의 변환이 잘못된 변환이라면, 그에 근거해서 결정된 대응 관계를 만족시키는 최적의 변환을 찾더라도 이는 결국, 계속해서 잘못된 변환을 찾아가는 결과가 되기 때문이다. 그러므로 그 대안으로서, V_2 의 어떤 점 \mathbf{p} 에 대해서 역보정에 의해 결정된 대응점 $C_k(b\mathbf{p}_i)$ 의 주위에 적당한 크기의 원도구를 띄워서 그 안에 존재하는 점들 중 변환된 \mathbf{p}_i 와 가장 가까운 점을 구한 다음 이 점과의 거리를 이용하여 오차 함수를 정의하고 이를 최소화시키는 방향으로 반복 계산을 진행시킨다. 그림 8로 이러한 개념을 다시 설명하면, 앞에서 정의한 오차 함수 E_k 는 그림에서 a 와 b , d 와 e 등의 거리의 평균을 의미하므로, 이를 최소화하는 변환을 다음 단계의 변환으로 갱신하는 것은 현재의 변환이 잘못된 변환임에도 불구하고 그 변환이 지시하는 잘못된 대응점으로 더욱 접근시키는 것

이 된다. 그러나 만약 V_2 의 표면에서 b 주위에 있는 점들 중, a 와 가장 가까운 점이 c 라 하면, 새로운 오차함수 η_k 는 이러한 거리들의 평균을 의미하므로, 이 값을 최소로 하는 변환을 계산한다면, 현재의 변환에 비해서 V_1 의 표면과 V_2 의 표면을 더 가깝게 만드는 역할을 할 수 있다는 것이다.

이러한 개념을 적용시켜 새로운 오차 함수 η_k 를 정의해 보면 다음과 같다.

$$\eta_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \min_{\mathbf{q}_i \in \mathcal{N}\{C_k(\mathbf{p}_i)\}} d(T_k \mathbf{p}_i, \mathbf{q}_i), \quad (17)$$

여기서 $\mathcal{N}(x)$ 는 x 주변 점들의 집합이다. 이러한 방법은 기존의 많은 연구^{[7]-[14]}에서 사용한 ICP 알고리즘에서 대응점을 결정하는 방법을 새롭게 제안한 것이 된다. 제안하는 방법은 그 대응점을 찾는 과정이 매우 간단하다는 장점이 있다. 즉, 기존의 많은 연구들은 이 대응관계를 결정하기 위하여 반복적인 방법을 사용하는 반면, 본 논문에서는 역보정으로 주어진 대응점들의 주위만 탐색해서 가장 가까운 점을 새로운 대응점으로 선택하므로 그 과정이 훨씬 간단하다.

마지막으로 새로운 대응관계를 만족시키는 최적의 변환을 어떻게 계산하는가에 대해서 설명한다. 3차원 공간에서의 회전 이동과 평행 이동을 나타내는 임의의 변환 행렬 \mathbf{T} 는 동차 좌표계 표현으로 다음과 같은 형태를 갖는다.

$$\mathbf{T} = \begin{bmatrix} C_2 C_y & C_2 S_y S_x - S_2 C_x & C_2 S_y C_x + S_2 S_x & t_x \\ S_2 C_y & S_2 S_y S_x + C_2 C_x & S_2 S_y C_x - C_2 S_x & t_y \\ -S_y & C_y S_x & C_y C_x & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (18)$$

여기서 C_i 와 S_i 는 각각 $\cos \theta_i$ 와 $\sin \theta_i$ 를 의미한다. 이제 N 쌍의 대응점들이 주어질 때 이들 간의 대응관계를 만족시키는 최적의 변환은 각 원소들간의 의존관계때문에 단순한 최소 사승법의 적용으로는 해를 구할 수가 없다. 이 경우 변환 계수 $\theta_x, \theta_y, \theta_z, t_x, t_y, t_z$ 들의 비선형 방정식을 풀어야 하기 때문에 상당히 어려운 문제가 된다. 이것은 회전 이동에 의해, \cos, \sin 함수들이 식에 나타나기 때문이다. 이때 각 회전각들이 1에 비해 아주 작은 값이라고 가정한다면, $\cos \theta_i = 1, \sin \theta_i = \theta_i$ 가 되어서 위의 변환을 다음과 같

이 선형화할 수 있으며, 이로부터 최소 자승법에 의해 해를 구할 수 있다. (x_i, y_i, z_i) 을 (x'_i, y'_i, z'_i) 에 대응시키는 근사적인 선형 변환 ΔT 는 다음과 같다.

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -\theta_z & \theta_y & t_x \\ \theta_z & 1 & -\theta_x & t_y \\ -\theta_y & \theta_x & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (19)$$

이렇게 해서 갱신된 변환은 새로운 대응 관계를 만족시키는 최적의 변환이 된다. 즉 η_k 를 최소화하는 변환이 되는 것이다.

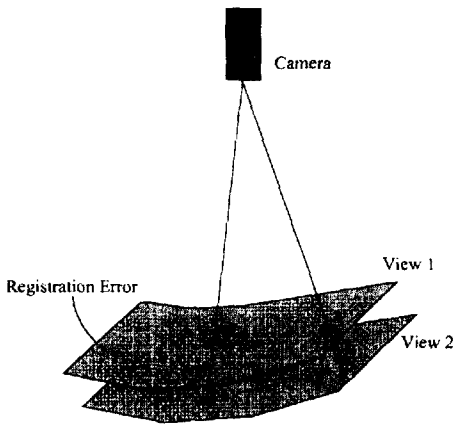


그림 8. 레지스트레이션 오차
Fig. 8. The registration error.

IV. 실험 결과

1. 초기 변환의 추정

제한한 알고리즘의 성능을 평가하기 위하여 공간 부호화 거리 측정기로 취득한 회전 이동과 평행 이동이 알려진 거리 데이터들을 이용하였다. 그림 9에 정육면체를 15도씩 돌려가면서 취득한 거리 영상을 제시하였다. 그림 10에 Pot를 15도 돌려서 취득한 거리 영상을 제시하였다.

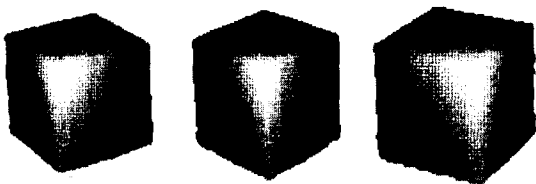


그림 9. 정육면체의 거리영상
Fig. 9. Range Images of a Cubic



그림 10. Pot의 거리영상
Fig. 10. Range Images of a Pot.

먼저 위의 데이터로 제안한 초기 변환의 추정 알고리즘과 Doraicite^[14]의 방법을 비교하여 표 1, 표 2, 그리고 표 3에 그 결과를 각각 제시하였다. 정량적인 비교를 위하여 회전 이동 계수들의 상대 오차는 $E_R = \|\hat{R} - R\| / \|R\|$ 로 평행 이동 계수들의 오차는 $E_t = \sqrt{(\hat{t}_x - t_x)^2 + (\hat{t}_y - t_y)^2 + (\hat{t}_z - t_z)^2}$ 로 각각 정의하여 비교하였다. 이러한 비교방법은 Dorai 등^[14]이 레지스트레이션 알고리즘의 성능을 비교하는 데 사용한 것이다. 모든 경우에 있어서 제안하는 방법이 Dorai의 방법에 비해서 우수한 성능을 보이는 것을 확인할 수 있다.

표 1. 정육면체의 초기변환 실험결과: 그림 9 (b) → (a)

Table 1. Result of initial guess: Cubic in Fig. 9 (b) → (a).

변환계수	참값	Dorai의 방법[14]	제안하는 방법
θ_x	0	13.276	-0.858
θ_y	0	13.498	-2.278
θ_z	-15	-12.882	-9.7812
t_x	-16.856	-25.647	-7.6822
t_y	21.967	34.574	10.585
t_z	0	5.249	-0.662
E_R		0.267	0.082
E_t		16.241	14.634

표 2. 정육면체의 초기변환 실험결과: 그림9 (c) → (b)

Table 2. Result of initial guess: Cubic in Fig 9 (c) → (b).

변환계수	참값	Dorai의 방법[14]	제안하는 방법
θ_x	0	12.792	3.946
θ_y	0	12.792	3.517
θ_z	-15	-11.296	-11.212
t_x	-16.856	-22.794	-16.182
t_y	21.967	35.205	23.793
t_z	0	4.434	0.103
E_R		0.265	0.092
E_t		15.171	1.949

표 3. Pot의 초기변환 실험결과: 그림10 (b) →(a)

Table 3. Result of initial guess: Pot in Fig. 10(b) →(a).

변환계수	참값	Dorai의 방법[14]	제안하는 방법
θ_x	0	-1.974	-2.637
θ_y	0	6.637	2.421
θ_z	-15	-18.475	-16.479
t_x	-16.856	-21.641	-18.451
t_y	21.967	20.767	19.026
t_z	0	7.021	3.456
E_R		0.110	0.055
E_t		8.5811	4.810

2. 변환 계수의 추출

앞장에서 설명한 방법대로 반복계산을 수행한 결과, 두 개의 시각이 서로 다른 거리 데이터 간의 변환 계수들을 원하는 오차한계 이내에서 구할 수 있었다. 보통 원하는 오차의 한계란 레인지 파인더의 오차 범위를 의미하며 [8, 13, 11], 본 논문에서는 거리의 오차 한계는 실험적으로 2.08 mm로 정하였으며, 표 11에 정육면체에 대하여 레지스트레이션에 의해 구해진 변환 계수들의 오차를 제시하였다. 이를 통해서 변환 계수들이 초기 추정치에 비하여 그 정밀도가 향상됨을 확인할 수 있다.

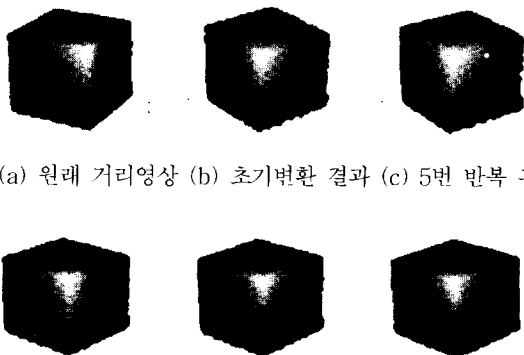


그림 11. 반복계산의 예

Fig. 11. Results of the iteration process.

그림 11에 정육면체의 (c)를 (b)로 변환하는 변환 계수를 반복적으로 구하는 과정을 예로 도시하였다. 반복회

수가 증가하면서 최종 목표 영상에 근접하는 것을 관찰할 수 있다. 이 실험은 19회의 반복 계산으로 최종적인 결과를 얻을 수 있었으며, 참고로 수행 시간은 IBM-PC Pentium-100에서 약 9분 정도가 소요되었다.

표 4. 그림 9의 정육면체 영상간의 레지스트레이션 결과

Table 4. Result of registration: Cubic in Fig. 9.

Parameters	True value	(b) →(a)	(c) →(b)
$\theta_x(\text{deg})$	0	-0.395	0.194
$\theta_y(\text{deg})$	0	-0.010	0.517
$\theta_z(\text{deg})$	-15	-15.013	-14.912
$t_x(\text{mm})$	-16.856	-15.422	-16.182
$t_y(\text{mm})$	21.967	23.382	20.598
$t_z(\text{mm})$	0	-1.218	-0.103
E_R		0.006	0.008
E_t		2.354	1.529

제안한 반복계산 알고리즘이 최소값 (global minimum)으로의 수렴성하는 거의 여부는 고려하지 않고 있다. 사실 거리 영상간의 변환을 구하는 레지스트레이션 문제에서는 변환 계수 공간을 모두 검색하여 최소값을 찾는다 것은 큰 의미가 없으며 속도문제로 인하여 그 실효성도 없다. 물체의 형상을 복원하기 위해서 여러 각도에서 영상을 취득할 때는 필연적으로 어느 정도 계획을 갖고 물체의 자세와 위치를 바꾸어 가면서 취득을 하게 되므로, 각 거리 영상간의 변환을 비교적 정확히 알고있거나, 대략적인 추정치라도 알고 있는 경우가 많다. 또는 전혀 모르는 경우에도 고급 (high level) 특정 정합 (feature matching) 알고리즘 등을 사용하여 어느 정도 비슷한 값을 초기변환은 줄 수도 있기 때문에 대부분의 레지스트레이션 알고리즘은 주어진 초기 변환으로부터 정확한 변환값을 찾아가는 반복 계산의 형태를 취하고 있다. 이에 더하여, 반복 계산의 과정에서 레지스트레이션 오차가 거리 측정기의 계측 오차이내로 수렴한 경우에는 이것이 국부 최소값 (local minimum)일지라도 큰 문제가 되지 않는다. 즉, 최소값이 존재한다 할지라도, 이미 거리 데이터의 취득 단계에서 오차를 포함하고 있으므로 주어진 문턱치보다 더 작은 레지스트레이션 오차를 갖는 변환을 구하는 것은 큰 의미가 없다.

V. 결 론

참 고 문 헌

본 논문에서는 시점이 서로 다른 복수의 거리 영상 간의 변환을 추출하는 새로운 방법을 제안하였다. 이를 위하여 거리 데이터의 취득을 위해 공간 부호화 거리 측정기의 기본원리를 분석하여 3차원 좌표의 형태로만 주어지는 거리 영상을 깊이 영상의 형태로 변환할 수 있는 방법, 거리 측정기의 시각 방향을 물체 좌표계의 벡터로 구하는 방법, 그리고 3차원 좌표로부터, 입력 영상에서의 2차원 좌표와 공간 부호값을 역으로 계산 하는 역보정 기법 등을 제안하였다.

초기변환의 추정 과정에서는, 각 거리 영상 내의 물체표면 데이터의 세 좌표간의 공분산 행렬의 고유벡터들이 이루는 세 축들간의 변환으로부터 두 거리 영상 간의 변환을 추정하여, 이후의 반복계산에서 이를 초기값으로 사용하도록 하는 Dorai의 방법을 수정하여, 물체표면의 한 영역이 시각의 변화에 따라서 영상내에 다르게 투영되는 효과를 반영하는 대표축의 계산 방법을 새롭게 제안하였다. 결과적으로 Dorai의 방법에 비하여 훨씬 더 근사한 초기변환을 구할 수 있음을 실험적으로 확인하였다. 반복 계산에 의한 변환 계수의 추출 과정에서는, 각 단계에서 구한 변환이 얼마나 정확한가를 판단할 수 있도록, 공간 부호화 거리 측정기의 역보정을 이용하여 결정된 대응점들 간의 거리의 평균값으로 레지스트레이션 오차의 척도 함수 (measure function)를 정의하였고, 공통 영역의 점들에 대해서 각 단계에서 역보정으로 찾아진 대응점 주위에서 가장 가까운 점을 찾아 그 점에 대응하도록 대응 관계를 재정의하여 이를 만족시키는 최적의 변환을 찾아 변환을 갱신하도록 하였다. 이때 변환은 비선형 방정식을 풀어 야만 하지만, 초기변환이 충분히 참값에 가깝다는 가정 에 의해서 회전각의 크기를 제한함으로써 선형 방정식으로 근사하여 최소 자승법을 적용하였다.

향후에는 본 논문에서 제안한 방법으로 추출한 변환 계수들을 이용하여 실제로 복수의 거리 데이터를 합성 하여 하나의 모델로 표현하는 인티그레이션 (integration)에 관한 연구가 필요하다고 하겠다. 이 과정에서 정확한 변환계수들이 주어진 상태이고 공간 부호화 거리 측정기의 역보정을 이용하여 바로 정확한 대응점을 찾을 수가 있으므로, 중복되는 데이터의 처리 문제와 모델의 설정 문제 등을 고려하여야 할 것이다.

- [1] S. Inokuchi, K. Sato, and F. Matsuda, "Rangeimaging system utilizing nemetic liquid crystal mask," iccv pp. 657-661, London, June 1987.
- [2] P. Vuylsteke and A. Oosterlinck, pp. 616-619, Osaka, Japan, Nov. 1993. "Range image acquisition with a single binaryencoded light pattern," pami vol. 12, no 2, pp. 148-164, Feb. 1990.
- [3] H. Yamamoto, K. Sato and S. Inokuchi, "Range imaging system based on binary image accumulation," icpr pp. 223-235, Oct. 1986.
- [4] J. Posdamer and M. D. Alschuler, "Surface measurement by space encoded projected beam systems," cgip vol. 18, no. 1, pp. 1-17, 1982.
- [5] Y. Sato, M. Otsuki, "Three Dimensional shape reconstruction by active rangefinder," cvpr pp. 142-147, New York, June 1993.
- [6] Y. Kawai, T. Ueshiba, T. Yoshimi and M. Oshima, "Reconstruction of 3D Objects by Integration of Multiple Range Data," icpr pp. 154-157, Hague, Aug./Sep. 1992.
- [7] Y.Chen, and G. Medioni, "Object Modeling by Registration of Multiple Range Images," icra pp. 2724-2729, Sacramento, Apr. 1991.
- [8] M. Soucy, A. Croteau, D. Laurendeau, "A Multi-Resolution Surface Model for Compact Representation of Range Images," icra pp. 1701-1706, Nice, May, 1992.
- [9] M. Soucy, D. Laurendeau, "Multi-Resolution Surface Modeling from Multiple Range Views," cvpr pp. 348-353, Jun. 1992.
- [10] M. Soucy, D. Laurendeau, "Surface Modeling from Dynamic Integration of Multiple Range Views," icpr pp. 449-452, Hague, Aug./Sep. 1992.
- [11] G. Truk and M.Levoy, "Zippered Polygon from Range Images," Proc. SIGGRAPH'94. In Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH, pp. 311-318, Orlando, Jul. 1994.

- [12] R. Bergevin, D. Laurendeau, and D. Poussart, "Estimating the 3D Rigid Transformation Between Two Range Views of a Complex Object", *icpr* pp. 478-482, Hague, Aug. Sep. 1992.
- [13] H. Gagnon, M. Soucy, R. Bergevin, D. Laurendeau, "Registration of Multiple Range Views for Automatic 3-D Model Building," *cvpr* pp. 581-586, Seattle, June 1994.
- [14] C. Dorai, J. Weng and A. K. Jain, "Optimal Registration of Multiple Range Views," *icpr* pp. 569-571, Los Angeles, Oct. 1994.
- [15] G. Blais, M. D. Levine, "Registering Multiview Range Data to Create 3D Computer Objects," *Technical report of Centre for Intelligent Machines, McGill Univ.*, 1995.

 저 자 소 개

鄭 徒 賢(正會員)

90년 3월 - 94년 2월 서울대학교 제어계측공학과 공학사
 94년 3월 - 96년 2월 서울대학교 제어계측공학과 공학석사
 96년 9월 ~ 현재 미네소타대학교 전기공학과 박사과정

尹 逸 東(正會員)

85년 3월 ~ 89년 2월 서울대학교 제어계측공학과 공학사.
 89년 3월 ~ 91년 2월 서울대학교 제어계측공학과 공학석사.
 91년 3월 ~ 96년 8월 서울대학교 제어계측공학과 공학박사.
 93년 7월 ~ 96년 8월 서울대학교 제어계측신기술센터 연구원
 96년 3월 ~ 97년 2월 대우전자영상연구소 선임연구원
 97년 3월 ~ 현재 한국외국어대학교 전임강사.
 주관심분야는 신호 및 영상처리, 3차원 컴퓨터 시각, 거리 영상의 취득 및 모델링, 컬러 영상처리

李 商 郁(正會員) 第 33卷 B編 第 2號 參照