반도체디스플레이기술학회지 제11권 제3호(2012년 9월)
Journal of the Semiconductor & Display Technology, Vol. 11, No. 3. September 2012.

가우시안 프로세스 회귀분석을 이용한 영상초점으로부터의 3차원 형상 재구성

무하마드 타릭 마흐무드 · 최영규†
†한국기술교육대학교 컴퓨터공학부

3D Shape Recovery from Image Focus using Gaussian Process Regression

Muhammad Tariq Mahmood and Young Kyu Choi†
†Korea University of Technology and Education, School of Computer Science and Engineering

ABSTRACT

The accuracy of *Shape From Focus* (SFF) technique depends on the quality of the focus measurements which are computed through a focus measure operator. In this paper, we introduce a new approach to estimate 3D shape of an object based on Gaussian process regression. First, initial depth is estimated by applying a conventional focus measure on image sequence and maximizing it in the optical direction. In second step, input feature vectors consisting of eginvalues are computed from 3D neighborhood around the initial depth. Finally, by utilizing these features, a latent function is developed through Gaussian process regression to estimate accurate depth. The proposed approach takes advantages of the multivariate statistical features and covariance function. The proposed method is tested by using image sequences of various objects. Experimental results demonstrate the efficacy of the proposed scheme.

Key Words : Depth recovery, Shape From Focus; Focus Measure; Gaussian Process; Regression

1. 서 론

2차원의 형태로 추출되는 영상으로부터 물체의 3차원 깊이 정보를 추출하는 것은 다양한 응용분야에서 매우 중요한 과정이 되고 있다. 예를 들어, 반도체 부품의 생산 및 조립 공정에서는 다양한 머신 비젼 검사 시스템(inspection system)이 활용되고 있는데, BGA (Ball Grid Array) 부품의 볼의 높이나 납땜용 합금 (solder paste)의 두께, 납땜 부위의 크기와 상태 등과 같은 다양한 검사에서 물체의 높이 (또는 깊이) 측정이 필수적으로 요구된다. 최근에는 3D TV와 컴퓨터 게임, 보안 감시, 의료 영상 등을 포함한 다양한 산업 분야에서도 물체의 깊이 정보를 추출하고 활용하려는 연구가 활발히 진행되고 있다[1,2]. 본 논문에서는 광학 기반의 다양한 측정 방법들 중에서 레이저와 같은 능동 광원을 사용하지 않고 렌즈의 초점 정보를 이용하여 물체의 깊이를 측정하는 방법에 대해 논한다.

초점 정보를 이용하여 물체의 깊이를 측정하는 방법은 크게 Shape From Defocus (SFD)와 Shape From Focus (SFF) 로 나뉘어진다. SFD에서는 여러 장의 영상과 함께 카메라의 정보가 사용되는데[3,4], 이와 달리 SFF에서는 다양한 초점에서 취득된 영상 시퀀스만을 이용하여 조밀한 깊이 정보를 추출할 수 있다[5,6]. 영상 시퀀스의 각 화소에서 초점이 잘 맞는 정도를 나타내는 초점 측도(focus measure)를 계산하고 이를 합하면 영상 초점 볼륨(image focus volume)이 만들어진다. 이 볼륨에서 가장 초점이 잘 맞는 화소를 이용해 깊이 정보를 구할 수 있는데, 광학 축(optical axis)을 따라 최대 초점 값을 계산하여 초기 깊이 맵(initial depth map)이 만들어진다.

초점 값의 계산시 발생하는 오차에 의해 초기 깊이 맵에는 오차가 발생하게 되는데, 이동이나 확대, 흔들림 등에 의해 초점 측도에 에러가 발생하고, 이러한 초점 측도의 성능이 깊이 맵의 정확도를 결정하게 된다.

†E-mail : ykchoi@kut.ac.kr

이와 함께, 인접한 프레임 사이의 초점 레벨의 비연속성(discreteness)도 초점 정보의 누락을 발생시키게 된다[7]. 대부분의 잡음들은 주로 고주파 영역에 집중하여 나타나는데, 이러한 영역에서는 여러 가지 방법으로 밝기 값을 보완해 주더라도 깊이 값이 부정확하게 계산되는 경우가 많다. 또한 영상에서 텍스쳐가 약하게 나타나는 영역(특별한 무늬가 나타나지 않는 영역)에서도 초점 측정이 어려울 수 밖에 없는데[5], 이러한 초기 깊이 맵의 정확도가 SFF 알고리즘에서는 매우 중요한 요인이 된다.

지금까지 초기 깊이 값의 보다 정확한 예측을 위해 두 가지 접근 방법들이 제안되고 있다. 한 가지 접근 방법은 영상 초점 볼륨을 필터링 하는데 기반을 두고 있다. 이들 방법들은 작은 윈도우 내의 초점 값들을 단순히 더하여 깊이 값의 안정성을 향상하고자 하는 것으로, 이전 프레임과 다음 프레임의 초점 값을 활용하기 위해 3차원 이웃(neighborhood) 연산을 이용하였다. 그러나 이 방법들은 윈도우를 너무 작게 잡으면 잡음을 충분히 제거하기가 어려우며, 윈도우가 너무 크면 깊이 값이 흐릿해지는(blurry) 문제가 발생하게 된다[8,9]. 이를 보완하기 위해 단순한 필터링이 아니라 잡음과 관련이 있는 고주파 성분을 제거하기 위해 도메인을 변환하여 처리하는 알고리즘도 제안되었다.

두 번째 접근 방법들은 먼저 기존의 방법으로 초기 깊이 맵을 구한 후 이를 추가적으로 개선하는 방식으로 정확도를 향상하고자 하였는데, 이를 위한 다양한 근사 기법들이 제안되었다[1, 6]. 그러나 실제 상황에서는 이들 방법에서도 초기 예측치의 에러가 최종적인 깊이 값에 큰 영향을 주게 되는데, 따라서 일반화된 최적의 깊이 연산자(depth estimator)를 필요하게 된다.

본 논문에서는 가우시안 프로세스 회기 분석(Gaussian process regression)을 이용하여 깊이 맵의 정확도를 향상시키는 방법을 제안하는데, 잡음이 있는 초기 깊이 정보와 초점 값들로부터 물체의 형상을 안정적으로 추출하고자 한다. 초기 깊이 값은 기존의 방법을 이용해 계산하며, 입력 특징 벡터(input feature vector)를 만들기 위해 초기 깊이 주변의 3D 이웃들에 대한 고유치들을 계산한다. 정확한 깊이 값의 계산을 위해 특징량들로부터 정보를 통합하는 가우시안 프로세스 회귀 분석을 통해 잠재 함수(latent function)를 유도한다. 제안된 방법은 다변량 통계 특징량(multivariate statistical feature)과 공분산 함수(covariance function)를 사용하는데, 이들은 커널(kernel)과 같이 동작하게 된다. 실험에서 제안된 방법은 3차원 형상 복원 측면에서 기존의 방법들과 비교되었는데, 제안된 방법의 성능이 우월함을 알 수 있었다.

가우시안 프로세스 회귀분석은 비모수 (non-parametric) 알고리즘으로, 통계적인 학습 방법과 기계 학습 방법을 결합하는 커널 기법(kernel method)의 한 종류이며, 최근에 특히 영상 및 신호 처리 분야에서 많은 관심을 모으고 있다[10, 11]. 제안된 방법은 점들의 집합(cloud point)으로부터 3차원 표면을 재구성하기 위해 커널 기법이 사용된 논문[12]에서 동기를 얻었는데, 이 논문에서는 3D 데이터가 커널 함수를 통해 특징 공간으로 변환되었으며, 물체의 표면은 이러한 특징 공간에서의 초월 평면(hyperplane)의 형태로 나타나게 되는데, 이때, 가우시안 프로세스는 공분산 함수의 형태로 커널을 제공해 준다.

2. 초점으로부터의 형상 재구성

Fig. 1은 볼록 렌즈를 통해 영상 평면에 2차원 형상이 생성되는 기하학적인 원리를 보여주고 있다. 물체 표면 S상의 한 점 P에 의해 발사된 빛들은 영상 평면의 한 점 P'에 잘 초점이 맞추어진다. 한편 S상에서 초점이 맞지 않은 한 점 Q에 의해 발사된 빛들은 영상 평면에서 한 점이 아니라 일정한 영역 Q'에 분산된다. 따라서 취득된 영상 I (i, j)은 초점이 잘 맞은 물체 영역과 함께 초점이 맞지 않아 흐릿하게 나타난 영역의 영향을 함께 포함하고 있다. 광학 축(optical axis)을 따라 물체를 움직이면서 Z장 취득한 영상 시퀀스를 I (i, j)라 하자. SFF에서는 영상 내의 화소들 중에서 초점이 잘 맞은 점까지의 거리(카메라로부터 표면)를 측정한다. 일반적으로 초점이 잘 맞은 부분은 경계가 선명하여 그렇지 않은 부분에 비해 훨씬 많은 고주파 성분을 포함하게 된다. 따라서 영상 시퀀스의 모든 점들에 대해 초점이 잘 맞은 정도를 나타내는 기준이 필요한데, 이를 초점 측도(focus measure)라고 한다.

초점 측도를 계산하기 위해 가장 일반적인 접근 방법

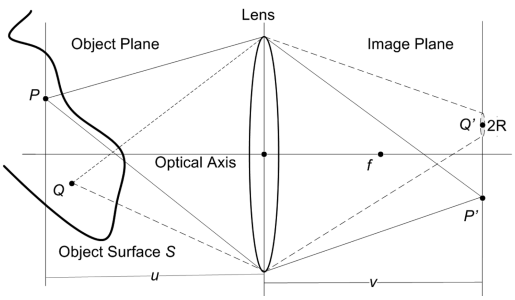

Fig. 1. Basic image formation in convex lens.

은 영상 도함수(image derivative)를 사용하는 것으로, 선명한 에지가 있는 경우에는 인접 화소의 값의 차이가 크다는 개념에 바탕을 둔다. 이들은 다시 1차 미분과 2차미분을 사용하는 방법으로 나눌 수 있다. Tenenbaum[13]은 수평 및 수직 기울기 연산자(gradient operator)를 사용하는 방법 F_{TEN}을 제안하였다. Sum modified Laplacian F_{SML} 역시 매우 유명한 초점 측도 연산자이다[5]. 2차 미분 방식은 1차 미분 방식에 비하여 더 정확한 결과를 보일 수 있지만, 잡음에 더 민감하다는 단점이 있다. 통계를 바탕으로 한 초점 측도 방법 중에서는 회색조 분산법(gray level variance) F_{GLV}이 많이 사용되고 있는데, 작은 윈도우 내에서 밝기 값의 분산이 크면 영상이 선명하다는 가정을 바탕으로 한다. 영역(domain) 변환을 바탕으로 한 초점 측도법도 제안되었는데, 저주파에 비해 고주파 성분이 주로 선명도를 결정하고 데이터에서의 대부분의 변화를 포함한다는 것을 바탕으로 한다. 이 방법은 정확도는 비교적 좋으나 계산 량이 많다는 단점이 있는데, [14]에 선명도 측정에 대한 비교가 잘 나타나 있다.

초점 측도는 초점의 품질이 높을수록 증가하며, 가장 잘 초점이 맞는 프레임에서 최대값을 갖는다. 지금까지 공간 영역과 주파수 영역에서 많은 초점 측도 방법이 제시되었는데, 보통 작은 2차원 이웃을 바탕으로 영상 시퀀스 $I_z(i, j)$의 각 화소 (i, j)에서 초점의 품질을 계산하게 된다. 이들을 모으면 결과적으로 초점 측도를 나타내는 새로운 3차원 볼륨 $I_z'(i, j)$ 이 다음과 같이 구해진다.

$$I_z'(i, j) = F(I_z(i, j)),\ z = 1, 2, \ldots, Z. \tag{1}$$

이러한 초기 깊이 값을 개선하기 위해 Nayar 등은 가우시안 보간법(Gaussian interpolation)에 바탕을 둔 근사 기법을 제안하였다[5]. 이때 초점 값은 F_{SML}을 사용하여 구하였으며, 피크(peak) 주위의 세 초점 값에 가우시안 모델을 적용하고, 평균값을 최적의 깊이 값으로 사용하였다. 그러나 초점 값이 항상 특정한 분포를 따른다는 보장이 없으며, 곡선 근사 또한 최적의 깊이 값을 구하지 못할 수도 있다. 또한 하나의 프레임에서 계산된 초점 값이 인접한 프레임의 영향을 반영하지 못할 수도 있으므로 복잡한 형태를 나타내는 자연적인 물체들에 대해서는 적합하지 않다. 이런 문제를 해결하기 위해 Subbarao[6]는 FIS(Focused Image Surface)의 개념을 도입하였는데, 물체의 FIS는 카메라 렌즈에 의해 잘 초점이 잡혀진 물체 표면의 점들의 집합을 의미한다. 이들은 초점 측도를 위해 한 장의 영상이 아니라 영상 볼륨에서의 평면 윈도우에서 계산하는 방법을 제안하였는데, 모든 가능한 평면들 중에서 최적의 평면을 찾기 위한 처리 시간이 많이 걸린다는 단점이 있다. 일반적인 물체의 FIS가 항상 평면은 아니라는 점에 착안하여 Choi등은 평면 윈도우 대신에 구간적 곡선 근사를 통해 FIS를 예측하는 방법을 제안하였다. 이 방법도 보다 나은 결과를 보일 수는 있지만 계산 량이 더욱 증가하게 된다[15]. Ahmad 등은 정확한 FIS를 얻기 위해 동적 프로그래밍(Dynamic Programming, DP) 기법을 적용하였다[1]. 이러한 방법들의 잡음에 민감한 초점 측도를 사용하여 잡음이 많은 경우 성능이 떨어지는 문제가 있다. 따라서 본 논문에서는 잡음에 강하고 보다 정확한 깊이 측정을 위해 가우시안 프로세스 회귀분석 기반의 방법을 제안한다.

3. 가우시안 프로세스 회귀분석 방법

3.1. 데이터 준비

잠재 함수의 도출을 위해 효율적인 특징량의 선택이 매우 중요한데, 본 논문에서는, 초기 깊이와 초점 측도에 대응되는 고유치(eigenvalue)를 이용하여 특징 벡터(feature vector)를 만든다. 각 화소 (i, j)에 대해 초기 깊이 $D(i, j)$는 가장 초점이 잘 맞은 프레임의 번호를 나타낸다. 이웃(neighborhood) $H(i, j)$는 다음과 같이 정의하는데, 이것은 $D(i, j)$에 대해 몇 개의 앞뒤 프레임을 포함하게 된다.

$$H(i, j) = \{z | D(i, j) - r \le z \le D(i, j) + r\} \tag{2}$$

이때, r은 $H(i, j)$의 반지름을 나타낸다. Fig. 2는 이웃의 선택을 보여주고 있는데, 첫 프레임 번호 초과이고 마지막 프레임 번호 미만인 이웃들을 고려하여 새로운 영상 볼륨 $J_z(i, j)$을 다음과 같이 만든다.

$$J_l(i, j) = I_{H(i,j)}(i, j),\ 1 < l < Z \tag{3}$$

이러한 새로운 영상 볼륨 $J_z(i, j)$은 L장의 영상을 포함하는데, 중앙 프레임은 가장 초점이 잘 맞는 영상을 나타내고, 이웃 프레임들은 비슷한 초점 레벨을 나타낸다. 이제, 새로운 영상 볼륨을 이용하여 각각의 객체점 (i, j)에서 행렬 $M \in R^{l^2 \times l}$ 을 만들 수 있는데, 이 행렬의 열 $M = \{m, \ldots m_l\}$ 은 $l \times l$크기의 l장의 각각의 작은 영상에서 값들을 이용하여 만들어 진다. 따라서

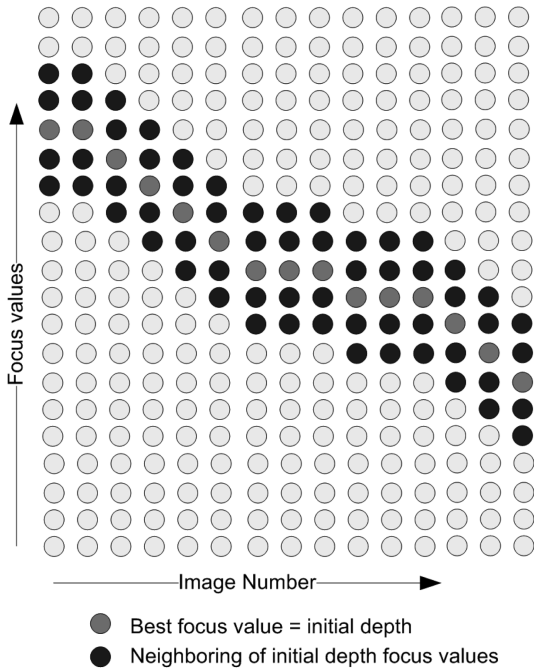

Fig. 2. 3D neighborhood selection: vertical lines represent images, central curve (red color) is the initial depth map, and left and right curve shows lower and upper bound in z-direction.

M의 공분산 행렬은 다음과 같이 구해진다.

$$C_{ij} = E[(m_i - E[m_i])(m_j - E[m_j])] \tag{4}$$

이 행렬은 정방형이며 $l \times l$ 크기의 대칭 행렬이다. 이 행렬에서 ij^{th}번째 성분은 i번째 열 벡터와 j번째 열 벡터의 내적(dot product)으로 구해진다. 대각선 요소 C_{ii}는 i번째 열 벡터들의 분산(variance)을 의미하며, 그 외의 요소들 C_{ii}은 열 벡터들 간의 공분산(covariance)을 나타낸다. 공분산 행렬의 고유치 분해(eigenvalue decomposition)는 다음과 같이 표현된다.

$$\mathbf{CG} = \mathbf{EG} \tag{5}$$

이때, $\mathbf{E} \in R^{l \times l}$ 은 대각선 행렬을 나타내고, $\mathbf{G} \in R^{l \times l}$ 는 정방형 행렬을 나타낸다. $E = diag(\lambda_1, ..., \lambda_l)$의 대각선 성분들은 공분산 행렬 C의 고유치가 되는데, 크기 순으로 정렬되며, $\mathbf{G} = \{g_1, \ldots g_l\}$ 의 열은 고유치와 관련된 고유 벡터(eigenvector)가 된다. 분산과 고유치는 서로 관련이 깊은데, 데이터 대부분의 분산들은 처음 몇 개의 고유치에 의해 결정된다. 그 이후의 고유치들은 크게 많지 않은 분산을 나타내게 되는데, 이들은 보통 잡음과 관련이 된다. 따라서 본 논문에서는 처음 몇 개의 고유치 $p < l$ 들 만을 특징 량으로 사용한다. 결국, 초기 깊이 D(i, j)에 대한 특징 벡터 $x = \{\lambda_1, \ldots \lambda_l, d_1, \ldots d_5\}$ 가 만들어지는데, $\{\lambda_1, \ldots \lambda_p\}$ 는 고유치를 나타내고, $\{d_1, \ldots d_5\}$ 는 D(i, j)의 4-방향 (4-connected) 이웃으로부터 추출된 깊이 값을 의미한다.

학습 과정을 위해서는 정확한 깊이 값을 알고 있는 객체로부터 N번 관찰한 학습 데이터 집합 T를 $T = \{(\mathbf{x}_n, y_n)\}_{n=1}^{N}$ 과 같이 준비한다. 이때, $\mathbf{x}$는 차원이 $d = p + 5$인 입력 특징 벡터이며, $\mathbf{y}$는 스칼라 출력 값 또는 목표(target)를 의미한다. 본 연구에서의 출력 값은 깊이 값이 되고, 입력은 와 같은 행렬의 형태가 된다. 이때 $\mathbf{X}$는 열이 공분산(covariates)을 나타내는 $d \times N$ 차원 행렬이며, $\mathbf{y}$벡터는 목표의 집합이다. 관측되는 깊이 값인 $\mathbf{y}$는 잡음이 포함된 입력 $\mathbf{X}$에 잠재 함수 $f(\mathbf{x})$가 적용되어 만들진 값으로 가정한다. 또한 입력 데이터는 평균이 0이고 분산이 σ_n^2 인 가우시안 잡음 $\varepsilon \sim (\mathrm{N}(0, \sigma_n^2))$ 에 의해 오염되어 있다고 가정하였다. 따라서 최종 목표는 가우시안 프로세스 회귀 분석을 통해 이러한 깊이 연산 함수 $f(\mathbf{x})$를 찾아내는 것으로 정형화 된다.

3.2. 가우시안 프로세스 회귀 분석

가우시안 프로세스는 다변량 가우시안 분포 (multivariate Gaussian distribution)를 무한히 많은 변수로 확장한 것으로, 평균 함수 $\mathrm{m}(\mathbf{x}) = \mathrm{E}[f(\mathbf{x})]$와 공분산 함수 $\mathrm{K}(\mathbf{x}, \mathbf{x}') = E[f(\mathbf{x})f(\mathbf{x})']$ 에 의해 표현되는데[10], 먼저 깊이 연산 함수 $f(\mathbf{x})$를 다음과 같이 가우시안 프로세스로 정의한다.

$$f(\mathbf{x}) \cong \mathrm{GP}(m(\mathbf{x}), K(\mathbf{x}, \mathbf{x}')) \tag{6}$$

기본적인 선형 회귀분석 모델에 따라 이것은 $y = f(\mathbf{x}) + \varepsilon$의 형태를 가진다. 잡음 모델의 가정을 바탕으로 회귀분석에 대한 Bayesian 접근법에 따른 공산(likelihood)이 정해지는데, 다음과 같은 관계가 있다.

$$\mathrm{P}(f|T) \propto P(f)P(T|f) \tag{7}$$

이때, $P(T|f)$ 와 $P(f)$ 는 $f(\mathbf{x})$의 사전 확률(priori probability)을 나타내는데, $P(f)$ 는 평탄도(smoothness)에 대한 가정을 나타내고, $P(T|f)$ 는 측정치와 함수 값이 얼마나 관련이 있는지를 나타낸다. 결국 $f(\mathbf{x})$에 대한 가우시안 프로세스 분포는 다음과 같이 표현된다.

$$\mathrm{P}(f\,|T)\ \propto N(f\,|\,m(\mathbf{x}),\,\mathrm{K}(\mathbf{x},\mathbf{x}'))\prod_{n=1}^{N}\mathrm{P}_n(\mathrm{y}\,|\,f(\mathbf{x})\,) \qquad (8)$$

관측된 목표 값 $\mathbf{y}\in R^{N\times d}$ 과 예측 함수 값 f_*의 결합 분포(joint distribution)는 다음과 같이 표현된다.

$$\begin{bmatrix}\mathrm{y}\\ f_*\end{bmatrix}\sim\mathrm{N}\left(0,\begin{bmatrix}(\mathbf{K}_{\mathrm{x,x}}+\sigma_n^2\mathbf{I}) & \mathbf{K}_{\mathbf{x},\mathbf{x}}\\ \mathbf{K}_{\mathbf{x},\mathbf{x}} & \mathbf{K}_{\mathbf{x},\mathbf{x}}\end{bmatrix}\right) \qquad (9)$$

이때, $\mathbf{K}$는 공분산 행렬이고, $\mathbf{X}_*$는 실험 데이터이다. 예측 분포 $\mathrm{P}(\mathbf{f}_*\,|\,T,\mathbf{X}_*\,)$ 는 결합 분포에 비례하는데, 평균 예측 함수는 다음과 같다.

$$\mathbf{f}_*=\mathbf{K}_{\mathrm{x,x}}[\mathbf{K}_{\mathrm{x,x}}+\sigma_n^2\mathbf{I}]^{-1}\mathbf{y} \qquad (10)$$

단일 특징량 벡터 $\mathbf{x}$에 대해서 매핑 함수 $f(\mathbf{x})$의 예측은 다음과 같이 단순화 된다.

$$f(\mathbf{x})=\mathbf{k}_*^T(\mathbf{K}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y} \qquad (11)$$

이때, $\mathbf{k}_*^T$ 는 샘플 포인트와 N개의 학습 데이터 포인트 사이의 공분산 벡터를 의미한다. 평균 예측은 관측 값 $\mathbf{y}$의 선형 조합이며, 따라서 N개의 커널 함수의 선형 조합으로 다음과 같이 나타낼 수 있다.

$$f(\mathbf{x}_*)=\sum_{n=1}^{N}\alpha_n K(\mathbf{x}_n,\,\mathbf{x}_*) \qquad (12)$$

이때, $\alpha=(\mathbf{K}+\sigma_n^2\mathbf{I})^{-1}\mathbf{y}$ 는 가중치 벡터이다.

3.3. 모델의 선택

최적의 공분산 함수와 파라메터의 선택이 깊이 값이나 영상 복원을 위해 매우 중요하다. 다양한 공분산 함수로는 선형 함수, 구간적 다항 함수, radial basis function, exponential, rational quadratic, 그리고 내적 커널(dot product kernel) 등이 많이 사용되고 있다. 이들 중에서 제곱 지수 함수(squared exponential function, SEF)는 무한한 미분이 가능하여 이를 사용하는 경우 가우시안 프로세스에서 모든 차원의 미분이 가능해 진다. 따라서 이 함수는 매우 평탄(smooth)하며, 또한 이동과 등방성(isotropic)에 불변하므로 본 논문에서는 다음과 같이 이것을 공분산 함수로 사용하였다.

$$\mathrm{K}(\mathbf{x},\mathbf{x}')=\sigma^2\exp\left(-\frac{1}{2}(\mathbf{x}-\mathbf{x}')^T\mathrm{A}(\mathbf{x}-\mathbf{x}')\right)+\sigma_n^2\delta_{\mathrm{xx}'} \qquad (13)$$

이때, $\delta_{\mathrm{xx}'}$ 는 Kronecher delta를 나타내고, $\theta=(\{\mathrm{A}\},\ \sigma_f^2,\ \sigma_n^f)$는 모든 초월 파라메터(hyperparameter)를 포함하는 벡터이고, $\{\mathrm{A}\}$ 는 대칭 행렬 A의 파라메터 집합을 나타낸다. 정확한 모델링을 위해서는 최적의 초월 파라메터 설정이 매우 중요한데, 일반적인 기계 학습 방법에서와 같이 교차 검증(cross validation)을 통해 처리할 수 있다. 여유 공산(marginal likelihood) 는 다음과 같이 주어진다.

$$\log\mathrm{P}(\mathbf{y}|\mathbf{X},\theta)=-\frac{1}{2}\mathbf{y}^T\mathbf{K}_y^{-1}\mathbf{y}-\frac{1}{2}\log|\mathbf{K}_y|-\frac{N}{2}\log 2\pi \qquad (14)$$

이때, $\mathbf{K}_y=\mathbf{K}_f+\sigma_n^2\mathbf{I}$ 는 $\mathbf{y}$에 대한 공분산 행렬이며, K_f는 무잡음 잠재 함수 $\mathbf{f}$에 대한 공분산 함수이다. 최적의 초월 파라메터는 복소 경사법(conjugate gradient method)을 적용하여 이러한 여유 공산을 최대화 하는 방법으로 구해진다.

3.4. 3D 형상 계산

학습을 위해 먼저 실제 원뿔 객체(4장 참조)를촬영한 영상 시퀀스를 준비하였다. 초기 깊이 맵을 구하기 위해 다양한 초점 측도 연산자를 사용할 수 있지만, 제안된 방법에서는 라플라시안 연산자를 사용하였다.

특징 벡터들은 3.1절에 설명한 바와 같이 계산하였으며, 이웃 연산을 위해 5×5×5의 윈도우를 사용하였다. 각각의 물체 점들에 대해 5개의 깊이 값과 6개의 고유치를 이용하여 11차원의 특징 벡터를 구성하였다. 고유치의 개수 선택도 매우 중요한데, 실험 결과 총 25개의 고유치 중에서 6번째 이하에서는 분산이 매우 적게 나타났고, 따라서 6개만을 선택하였다. 원뿔 영상의 해상도가 200×200이므로 총 40000개의 점들이 있는데, 이 중에서 2000개를 학습(training)에 이용하였고, 나머지를 만들어진 모델의 확인과 검증을 위해 사용하였다. 제곱 지수 함수(SEF)를 공분산 함수로 선택하였으며, 최적의 잡음 변화량, 길이 스케일 및 다른 파라메터들은 로그 공산의 음수값(negative log of likelihood)을 최소화하는 방법으로 계산하였다. 실험에서 최적 값으로 $\sigma_n^2=2.07$ 와 $\sigma_f^2=7.25$ 가 선택되었는데, 고유치가 깊이 함수에 영향을 많이 받음을 알 수 있었다.

학습 과정을 통해 모델이 만들어지면, 임의의 객체에 대한 깊이 측정에 이 모델을 사용할 수 있다. 먼저 현미경에 장착된 CCD 카메라를 통해 객체 영상을 취득하고 각 객체 점들에서 3.1절에 설명한 방법으로 특징량을 계산하면 이들을 이용하여 객체의 깊이를 계산하게 된다. 제안된 방법의 정량적 성능 측정을 위해 원뿔 데이터에 대한 RMS(root mean square) 에러를 계

산하였는데, 이는 이상적인 값(ground truth)과 측정값의 차의 제곱을 평균한 값으로 제안된 방법의 성능을 정량적으로 나타내 준다. 제안된 방법의 결과는 F_{SML}, F_{TEN}및 F_{GLV}와 비교하였는데, 이들 알고리즘에서 윈도우의 크기는 5×5를 사용하였다.

4. 실험 및 고찰

제안된 방법을 실험하기 위해 3가지 물체에서 촬영된 영상 시퀀스가 사용되었다. 먼저 정확한 크기를 알고 있는 원뿔에 대해 CCD 카메라를 사용하여 97 프레임의 영상을 촬영하였다. 이때 영상의 해상도는 200×200이며, 원뿔의 실제 깊이 정보는 원뿔 파라메터를 통해 계산하였다. 또한, 두 가지의 작은 물체(microscopic object)에 대해서도 실험하였는데, 1 센트 동전(penny)에 양각으로 새겨져 있는 영문자 P와, 미소제어 시스템(microscopic control system, MCS)을 이용하여 표면에 음각으로 새겨 넣은 영문자 H에 대해 실험하였다. 이때 영상의 해상도는 300×300으로 각각 60프레임의 영상이 사용되었다. Fig. 3은 이러한 물체의 샘플 사진들과 원뿔에 대한 이상적인 깊이 맵을 보여주고 있다.

Table 1은 원뿔 데이터에 대한 처리결과 성능을 RMS 에러로 보여 주고 있다. 제안된 방법을 통해 추출된 깊이 맵이 원뿔의 이론적인 데이터와 유사함을 알 수 있으며, 제안된 방법 적용함에 따른 상당한 정밀도의 향상을 확인할 수 있다. Fig. 4는 원뿔에 대해 다양한 방법으로 복원된 3차원 형상을 보여주고 있다. 실제로 원뿔 데이터에서는 Fig. 3(a)에서 보이는 것과 같이 오른쪽 아랫부분이 조명 상태가 좋지 않고 표면의 텍스처도 다른 부분에 비해 상대적으로 적은(weak texture) 것을 알 수 있다. 기존의 방법들은 Fig. 4의 (a)~(c)와 같이 이러한 부분에서 깊이 정보를 제대로 추출하지 못하였는데, 잘못된 깊이 값이 그림에서 스파이크의 형태로 나타나고 있다. 제안된 방법은 (d)에 나타난 바와 같이 텍스처가 적거나 조도가 약한 영역에서도 좋은 결과를 나타내는 것을 알 수 있다.

Fig. 5는 영문자 H와 P에 대한 실험 결과를 보여주고 있다. 이들 객체에 대한 영상은 이동이나 진동 등 다양한 요인에 따른 잡음을 포함하고 있는데, 기존의 방법들에서는 정확한 초점 품질 의 측정이 어렵고 결

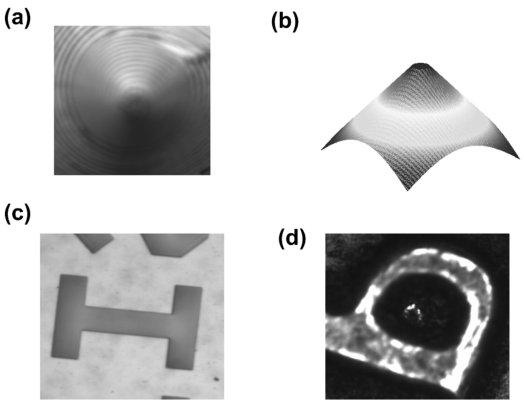

Fig. 3. Sample images of test objects (a) real cone, (b) true depth map of real cone (c) letter H and (d) letter P.

Table 1. Performance comparison in terms of RMSE.

Method	F_{SML}	F_{TEN}	F_{GLV}	f_{GPR}
RMSE	3.2652	3.8586	3.2707	0.7556

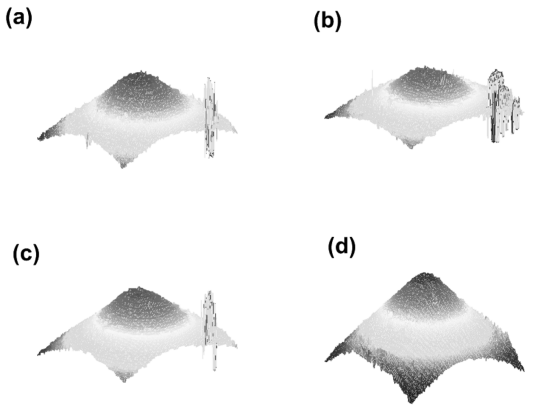

Fig. 4. 3D shapes recovered of real cone using (a) F_{SML}, (b) F_{TEN}, (c) F_{GLV}, (d) f_{GPR}.

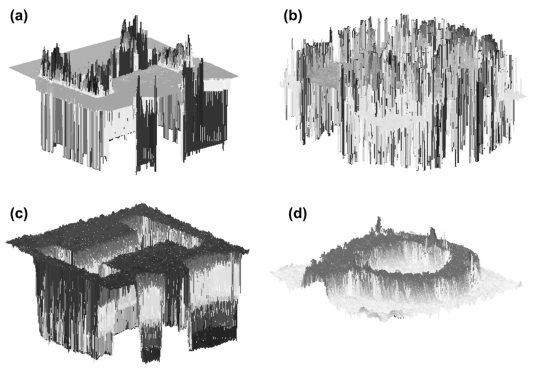

Fig. 5. 3D shapes recovered for test objects H (first column) and P (second column) using (a) F_{SML}, (b) F_{TEN}, (c-d) proposed approach f_{GPR}.

국 (a), (b)와 같이 잡음이 많은 깊이 맵이 만들어진다. 제안된 방법은 이러한 데이터에 대해서도 (c), (d)와 같이 매우 안정적인 깊이 맵을 만들어 주는 것을 알 수 있다.

제안된 방법은 기존의 방법들에 비해 3차원 형상을 더 잘 만들어주는데, 단순히 평활 필터를 사용하여 표면을 완만하게 만드는 방식이 아니다. 기존의 방법들은 인접한 깊이 값의 가중 합을 취하는데, 평활화의 특성상 이웃(윈도우)의 크기가 클수록 너무 평활화되고 형상이 왜곡된다. 최소 윈도우의 크기가 3x3인데, 실험결과 이러한 작은 크기의 윈도우는 잡음을 제거하기 위해 충분하지 않음을 알 수 있었다. 제안된 방법에서는 초기 맵으로부터 4방향 이웃만을 사용하고 나머지 특징량들은 초점 측도로부터 구하는데, 이를 통해 깊이 맵을 단순히 평활화하는 기존 방법에 비해 우월한 성능을 나타냈다.

5. 결론 및 향후 과제

본 논문에서는 초점 정보를 이용하여 3차원 형상을 재구성하는 새로운 방법을 제안하였다. 특히 가우시안 프로세스 회귀분석 기법을 사용하였는데, 형태의 계산 문제를 회귀 문제로 정형화하였으며, 고유치와 초기 깊이 등을 특징량으로 사용하였다. 제안된 방법은 다양한 물체의 깊이 맵 추출에 적용하여 실험해 보았는데, 기존의 방법에 비해 우월한 성능을 보였으며, 특히 잡음이 있는 실제 영상에 대해 탁월한 성능을 나타냈다.

향후 제안된 방법은 서포트 벡터 회귀분석(support vector regression, SVR)이나, 일반화된 회귀분석 신경회로망(generalized regression neural network, GRNN), 및 유전자 프로그래밍(genetic programming) 등 잘 알려진 다양한 방법으로 확장하여 형상 복원이나 영상 복원 응용에 적용해 볼 예정이다. 또한, 특징 량의 추가를 통해 제안된 방법의 정확성을 향상시킬 수 있을 것으로 기대된다.

참고문헌

1. M. B. Ahmad and T. S. Choi, “A heuristic approach for finding best focused shape,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, pp. 566-574, April 2005.
2. C. Hermans, Y. Francken, T. Cuypers, and P. Bekaert, “Depth from sliding projections,” in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 1865-1872.
3. A. P. Pentland, “A new sense for depth of field,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. PAMI-9, pp. 523-531, 1987.
4. P. Favaro, S. Soatto, M. Burger, and S. J. Osher, “Shape from defocus via diffusion,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 30, pp. 518-531, 2008.
5. S. K. Nayar and Y. Nakagawa, “Shape from focus,” Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 16, pp. 824-831, Aug. 1994.
6. M. Subbarao and T. S. Choi, “Accurate recovery of three-dimensional shape from image focus,” IEEE Transactions on pattern analysis and machine intelligence, vol. 17, pp. 266-274, 1995.
7. A. S. Malik and T. S. Choi, “Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery,” Pattern Recognition, vol. 40, pp. 154 -170, 2007.
8. E. Krotkov, “Focusing,” International Journal of Computer Vision, vol. 1, pp. 223-237, 1988.
9. M. T. Mahmood, W. J. Choi, and T. S. Choi, “PCA-based method for 3D shape recovery of microscopic objects from image focus using discrete cosine transform,” Microscopy Research and Technique, vol. 71, pp. 897-907, 2008.
10. C. E. Rasmussen and C. K. I. Williams, Guussian Processes for Machine Learning: MIT Press, 2006.
11. Z. Jianke, S. C. H. Hoi, and M. R. Lyu, “Nonrigid shape recovery by gaussian process regression,” in Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, 2009, pp. 1319-1326.
12. N. Heung-Sun, W. Tai-Pang, and T. Chi-Keung, “Surface-from-Gradients with incomplete data for single view modeling,” in Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, 2007, pp. 1-8.
13. T. M. Tenenbaum, “Accommodation in computer vision,” Stanford University, 1970.
14. R. Ferzli and L. J. Karam, “A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB),” Image Processing, IEEE Transactions on, vol. 18, pp. 717-728, 2009.
15. T. S. Choi and J. Yun, “Three-dimensional shape recovery from the focused-image surface,” Optical Engineering, vol. 39, pp. 1321-1326, 2000.

접수일: 2012년 8월 23일, 심사일: 2012년 9월 5일,
게재확정일: 2012년 9월 17일