논문 2010-47CI-3-4

깊이 추정을 위한 PCA기반의 특징 축소

(PCA-Based Feature Reduction for Depth Estimation)

신 성 식*, 권 오 봉**

(Sung-Sik Shin and Ou-Bong Gwun)

요 약

본 논문에서는 한 장의 정지 영상에서 학습을 통한 방법으로 깊이 정보를 추정하는데 사용되어지는 특징 정보를 PCA(Principal Component Analaysis)기반으로 축소하여 깊이 정보의 정확성을 향상시키는 방법에 대하여 기술한다. 정지 영상에서 깊이 정보를 추정하기 위하여 이미지의 에너지 값과 기울기와 같은 특징을 추출하며 특징들의 관계를 이용하여 각 영역의 깊이 정보를 추정한다. 이 때 영상 필터를 사용하여 많은 특징을 추출하지만 특징의 중요성을 판단하지 않고 모두 사용하면 오히려 성능에 좋지 않은 영향을 미친다. 본 논문에서는 한 장의 정지 영상의 깊이 추정을 위해 PCA를 기반으로 중요도를 판단하여 특징 벡터의 차원을 줄이고 깊이를 정확하게 추정할 수 있는 방법에 대하여 제안한다. 제안한 방법을 스탠포드 대학의 평가 데이터로 실험한 결과, 깊이를 추정하는데 있어서 전체 특징 벡터의 30%만을 이용하여 평균 0.4%에서 최대 2.5%의 정확도가 향상되었다.

Abstract

This paper discusses a method that can enhance the exactness of depth estimation of an image by PCA(Principle Component Analysis) based on feature reduction through learning algorithm. In estimation of the depth of an image, hyphen such as energy of pixels and gradient of them are found, those selves and their relationship are used for depth estimation. In such a case, many features are obtained by various filter operations. If all of the obtained features are equally used without considering their contribution for depth estimation, The efficiency of depth estimation goes down. This paper proposes a method that can enhance the exactness of depth estimation of an image and its processing speed is considered as the contribution factor through PCA. The experiment shows that the proposed method(30% of an feature vector) is more exact(average 0.4%, maximum 2.5%) than using all of an image data in depth estimation.

Keywords: PCA, Principal Component Analysis, Feature Reduction, Depth Estimation

Ⅰ. 서 론

본 논문은 학습을 통하여 정지 영상의 깊이 정보를 추정하는데 사용되는 특징 벡터의 차원을 PCA[1]기반으로 축소하여 학습시간을 줄이고 성능을 향상 시키는 방법에 대하여 기술한다.

정지 영상에서 깊이 정보를 추정하는 연구는 로봇 비전, 컴퓨터 그래픽스[2], 지능형 자동차[3] 등 여러 분야에서 필요로 하는 기술이며 많은 연구가 이루어지고 있다.

깊이 정보를 추정하는 방법은 영상에서 물체의 꼭지점, 외부 육관선 등의 정보를 수학적으로 계산[10]하는 방법을 비롯하여 영상의 변화, 기울기와 같은 특징 정보를 이용하는 방법 등이 있다. 그리고 최근에는 영상의 특징정보와 실제 깊이 정보를 이용하여 학습을 통한 깊이 정보 추정 방법[4]이 연구된다.

그러나 다양한 방법을 이용하여 추출된 많은 특징 정보들은 중요한 정보인지 아니면 중요하지 않은 정보인지 판단되지 않고 깊이 정보를 추정하는데 모두 사용된다. 일반적으로 특징의 개수가 많아지면 정보가 많아

* 학생회원, ** 정회원, 전북대학교 컴퓨터공학부
(Division of Computer Science and Engineering, Chonbuk National University)
접수일자: 2010년4월5일, 수정완료일: 2010년4월30일

성능이 좋아진다고 생각 할 수 있다. 하지만 오히려 성능도 좋지 않은 영향을 미칠 수 있다. 특징의 개수가 많아지면 불필요한 특징 정보들이 포함되고 부작용이 발생할 수 있기 때문이다.

이러한 문제를 해결하기 위하여 패턴인식 분야에서는 많은 연구가 이루어졌으며 최적화 분석 방법 중 대표적으로 PCA와 LDA(Linear Discriminate Analysis)가 있다.

PCA[5]는 고차원의 정보를 최대한 유지하면서 저차원으로 차원을 축소하며 LDA[6~7]는 분류 클래스간의 분산과 클래스내의 분산 비율이 최대화하는 방식으로 차원을 축소한다.

본 논문에서는 고차원 정보를 유지하면서 저차원으로 특징을 축소하는 PCA를 기반으로 특징의 차원을 축소하여 깊이 정보를 추정 하였으며 특징의 차원을 축소함으로써 성능을 향상시키는 방법에 대하여 다음과 같이 설명한다.

Ⅱ장에서는 깊이 추정에서 사용되는 특징의 추출 방법과 특징 벡터의 생성 방법에 대하여 기술하고 Ⅲ장에서 PCA를 이용한 특징 축소 방법을 기술한다. Ⅳ장에서는 실험 결과에 대하여 기술하고 마지막으로 Ⅴ장에서 본 논문의 결론 및 향후 연구에 대하여 기술한다.

Ⅱ. 특징 벡터

사람은 한 장의 정지 영상을 보고 거의 완벽한 깊이 정보를 판단 할 수 있다. 이것은 영상의 변화, 기울기, 조합[8]과 미리 선 학습되어 알고 있는 물체의 크기, 흐릿한 형태 등의 정보를 이용하여 가능하다.

하지만 컴퓨터는 사람과 같이 정보를 판단하기 힘들며 모든 영역에서 찾을 수 있는 특징 정보를 이용해야 한다. 물체의 윤곽선 등은 영상을 작은 영역으로 분할하였을 때 존재하지 않는 영역이 발생 할 수 있으므로 사용하지 않는다. 예를 들면 영상에서 하늘 영역에서는 아무런 정보도 찾을 수 없다.

본 논문에서는 작은 단위의 영역에서도 공통적으로 에너지 정보를 취할 수 있도록 영상의 에너지 값을 이용하여 특징을 추출하며 특징을 추출하기 위해 Laws' 마스크와 텍스쳐의 기울기 필터를 이용[11]한다.

1. 특징의 추출

영상에서 특징을 추출하기 위해 색상 공간을 RGB에서 Y(Intensity Channel)$Cb Cr$(Color Channel)로 변환한다. 그 이유는 영상의 영역에서 특징을 추출 할 때 색상 정보는 거의 사용하지 않기 때문이다.

그림 1. Laws' 마스크와 텍스쳐 기울기 필터
Fig. 1. Laws' Masks and texture gradient filter.

$L_3' \times L_3$	$L_3' \times E_3$	$L_3' \times S_3$
$E_3' \times L_3$	$E_3' \times E_3$	$E_3' \times S_3$
$S_3' \times L_3$	$S_3' \times E_3$	$S_3' \times S_3$

그림 2. 9개의 2차원 Laws' 마스크
Fig. 2. Nine 2dimension Laws' Masks.

이렇게 변환 된 영상의 Y영역에서 Laws' 마스크 필터와 텍스쳐의 기울기 필터를 이용하여 에너지 값들을 계산하여 특징으로 선택 한다.

Laws 마스크는 영상에서 특정 정보를 추출하기 위해 만들어진 Laws' 벡터의 조합으로 생성 하였으며 본 논문에서는 L(Level), E(Edge), S(Spot), R(Ripple), W(Wave)중에 L, E, S만을 사용하였으며 3차원 Laws' 벡터를 이용하였다. ($L_3 = [1\ 2\ 1]$, $E_3 = [-1\ 0\ 1]$, $S_3 = [-1\ 2\ -1]$)

또한, 첫 번째 Laws' 마스크 필터를 이용하여 $Cb Cr$ 영역의 에너지 값을 추가 정보로 추출하여 분할 된 영역별로 총17개(Y영역 : Laws' 마스크 9개, 기울기 필터 6개, Cb영역 : 첫 번째 Laws' 마스크 1개, Cb영역 : 첫 번째 Laws' 마스크 1개)의 특징 정보를 가진다.

작은 단위로 분할 된 영역(patch)의 에너지를 구하는 방법은 식(1)과 같다.

$$E_i(n,k) = \sum_{(x,y)\in patch(i)} \left| F_n(x,y) \right|^k \qquad (1)$$

$F_n(x,y)$는 Laws' 마스크와 텍스쳐 기울기 필터를 통해 구해진 에너지 값을 이야기하며, $k = 1,2$로 순수 에너지 값과 에너지 값의 제곱을 의미한다. 즉, 한 영역의 에너지 값은 필터를 적용하여 나온 결과에서 영역에 포함되는 픽셀 값들의 합이고 이 에너지 값이 특징이 된다.

2. 특징 벡터의 생성

하나의 영역에는 2.1절의 특징 추출과 같이 기본적으로 34개의 특징이 존재한다. 이러한 특징 정보는 작은 영역의 정보만을 포함하고 있기 때문에 전역적인 정보를 포함하기 위해 그림 3과 같이 이웃하는 영역(상,하,좌,우)의 정보와 영역의 크기(1배,3배,9배)를 달리한 이웃의 정보를 추가적인 특징 정보[9]로 사용한다.

전체 특징 정보는 그림 3과 같이 A0영역의 특징 정보 34개와 A1(상),A2(하),A3(좌),A4(우)영역의 특징 정보 136개와 영역의 크기를 변경하여 다시 B0의 특징 정보 및 B1, B2, B3, B4의 특징 정보를 가진다. 이와 같은 방법으로 한 번 더 크기를 변경하여 C0, C1, C2, C3, C4의 정보를 가지며 최종적으로 한 영역의 특징 정보는 자신의 영역과 이웃영역 그리고 크기별 특징을 포함하여 총 510개의 특징을 가지게 된다.

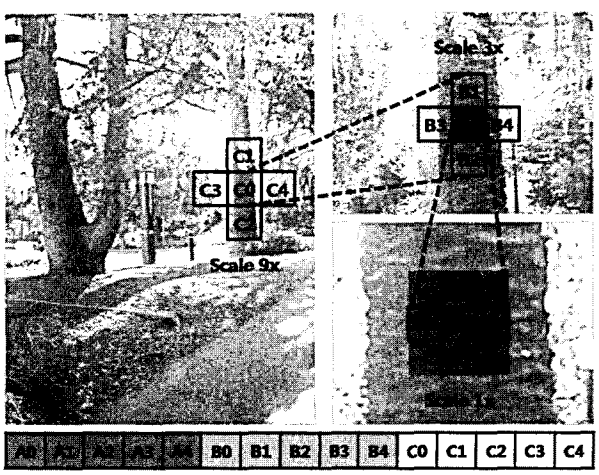


그림 3. 한 영역의 특징은 크기 별로 이웃특징을 포함
Fig. 3. A feature of a patch includes neighbors features from each scales.

Ⅲ. 특징 공간 축소

1. 특징 벡터 정규화

Ⅱ장에서 추출한 특징들은 어떤 필터를 이용하여 추출되었는지에 따라서 특징 공간의 분포가 서로 다르다. PCA에서는 특징 공간의 분포에 따라 주성분을 분석하기 때문에 추출된 특징들을 정규화하지 않고 그대로 사용하게 되면 의도되지 않은 주성분을 찾게 된다.

예를 들어 그림 4와 같이 데이터의 분포가 [0~1]의 범위 중 [0.1~0.9]를 가지는 특징 분포와 [0~10]의 범위 중 [2~4]을 가지는 특징분포의 중요도는 [0.1~0.9] 범위 분포를 가지는 특징이 높으나 PCA에서는 [2~4]의 범위 분포를 가지는 특징을 주성분으로 찾는다. 후자의 경우 절대적 분포의 크기가 전자에 비해 더 크기 때문이다.

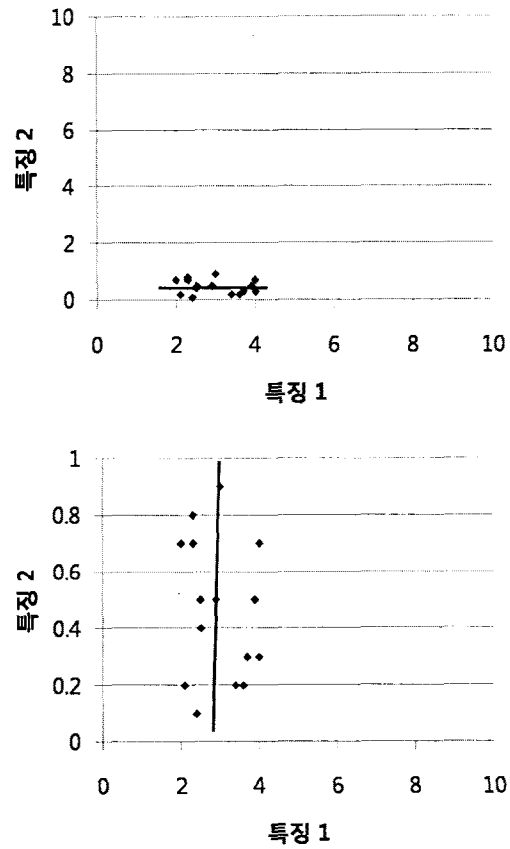


그림 4. 비정규화 PCA결과(위)와 정규화 PCA결과(아래)
Fig. 4. Non-normalization PCA result(Top) and normalization PCA result(Bottom).

이러한 문제를 해결하기 위하여 동일한 공간 분포를 가지도록 식(2)와 같이 최소 최대 정규화를 이용하여 각 특징 벡터를 정규화 한다.

$$Normalization(x_{ij}) = \frac{x_{ij} - \min x_i}{\max x_i - \min x_i} \qquad (2)$$

x는 특징 벡터를 의미하고 i는 특징 벡터의 차원의 색인 번호를 j는 i번째 특징 차원에서 특징 데이터의 색인 번호를 의미한다.

2. PCA(Principal Component Analysis)

통계학에서 PCA는 데이터 집합을 분석하는 기법 가운데 하나이다. PCA는 데이터를 하나의 축으로 사상시

켰을 때 그 분산이 가장 커지는 축이 주성분이 되고, 두 번째로 커지는 축이 그 다음 주성분이 된다. 이와 같은 방법으로 차례로 놓이도록 새로운 좌표계로 데이터를 선형 변환한다.

데이터 집합의 전체 평균이 0이라고 가정하면, 데이터 집합 x의주성분 w_1은 다음과 같이 정의된다.

$$w_1 = \arg\max_{\|w\|=1} E\left\{\left(w^T x\right)^2\right\} \tag{3}$$

k-1개의 주성분이 이미 주어져 있을 때 k번째 주성분은 앞의 k-1개 주성분을 제거함으로써 찾을 수 있다.

$$\hat{x}_{k-1} = x - \sum_{i=1}^{k-1} w_i w_i^T x \tag{4}$$

그리고 이 값을 데이터 집합에서 제거하면 다음 주성분을 찾을 수 있다.

$$w_k = \arg\max_{\|w\|=1} E\left\{\left(w^T \hat{x}_{k-1}\right)^2\right\} \tag{5}$$

3. 특징 축소

Laws' 마스크와 텍스쳐 기울기 필터를 이용하여 구해진 특징 벡터는 마스크에 따라 그 특징의 최소값과 최대값이 서로 다르다. 따라서 PCA를 통하여 특징 벡터의 차원을 축소하기 전에 3.1절과 같이 최소 최대 정규화를 통한 정규화가 필요하다. 정규화 과정을 거친 후에 3.2절의 내용과 같이 PCA를 이용하여 주성분을 구한 후 특징을 축소한다. 적절한 특징의 차원 축소를 통하여 성능 향상을 기대 할 수 있다.

본 논문에서는 깊이 추정에 적합하도록 특징의 차원을 줄이기 위해 50개 단위로 특징의 차원 축소 실험을 하고 그 결과를 제시 한다.

Ⅳ. 실 험

1. 데이터 수집

실험 데이터는 3차원 레이저 스캐너를 이용하여 생성된 미국 스탠포드대학에서 제공하는 데이터를 사용하였다. 데이터는 총 400여장의 1704x2272해상도를 가지는 영상과 305x55해상도를 가지는 깊이 지도의 쌍으로 이루어져 있다. 이 중 300여장의 영상 데이터를 학습 데이터로 사용하고 100여장의 영상을 테스트 데이터로 사용하였다.

2. 실험

Ⅱ장의 설명과 같이 이미지를 여러 개의 작은 영역으로 분할하였으며, 각 영역마다 Laws' 마스크와 텍스쳐 기울기 필터를 사용하여 특징을 추출한다. 또한, 이웃하는 영역의 특징 정보를 추가하여 영역 당 510개의 특징 정보를 생성한다.

이렇게 생성된 특징 벡터는 각 차원마다 서로 다른 특징 공간 분포를 가지고 있어 PCA를 적용하여 원하는 주성분을 구할 수 없다. 그렇기 때문에 정규화 과정을 거치게 된다.

정규화 과정을 거친 특징 벡터들은 서로 같은 크기의 공간 분포를 가지게 되며 PCA를 통하여 차원을 축소한다. 학습과정에서는 영상의 윗부분의 정보와 영상의 아랫부분의 정보는 같은 값을 가져도 서로 다른 깊이를 가질 수 있기 때문에 각 *row*별 학습을 한다. (영상에서 파란색은 윗부분에서는 하늘이고 아랫부분에서는 물일 확률이 높다.)

본 논문에서는 50개 단위로 차원을 축소 후 반복 가중 최소자승법[12](Iteratively reweighted least squares)을 이용하여 학습을 최종 결과를 확인 했다.

3. 오차율 비교

오차율은 실제 깊이 지도에서 추정 된 깊이 정보를 비교하여 측정 하였으며, 모든 영역에서의 깊이 차이의 합을 이미지 전체의 오차율로 계산 한다.

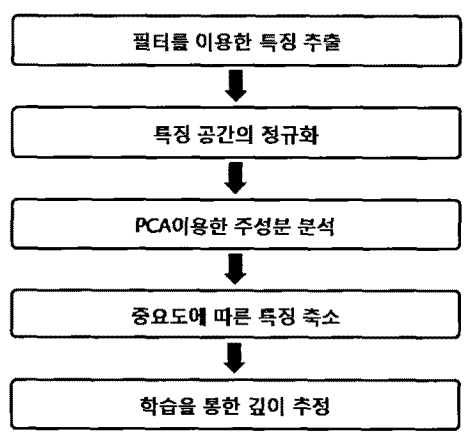


그림 5. 깊이 추정 실험 과정
Fig. 5. A experiment process of depth estimation.

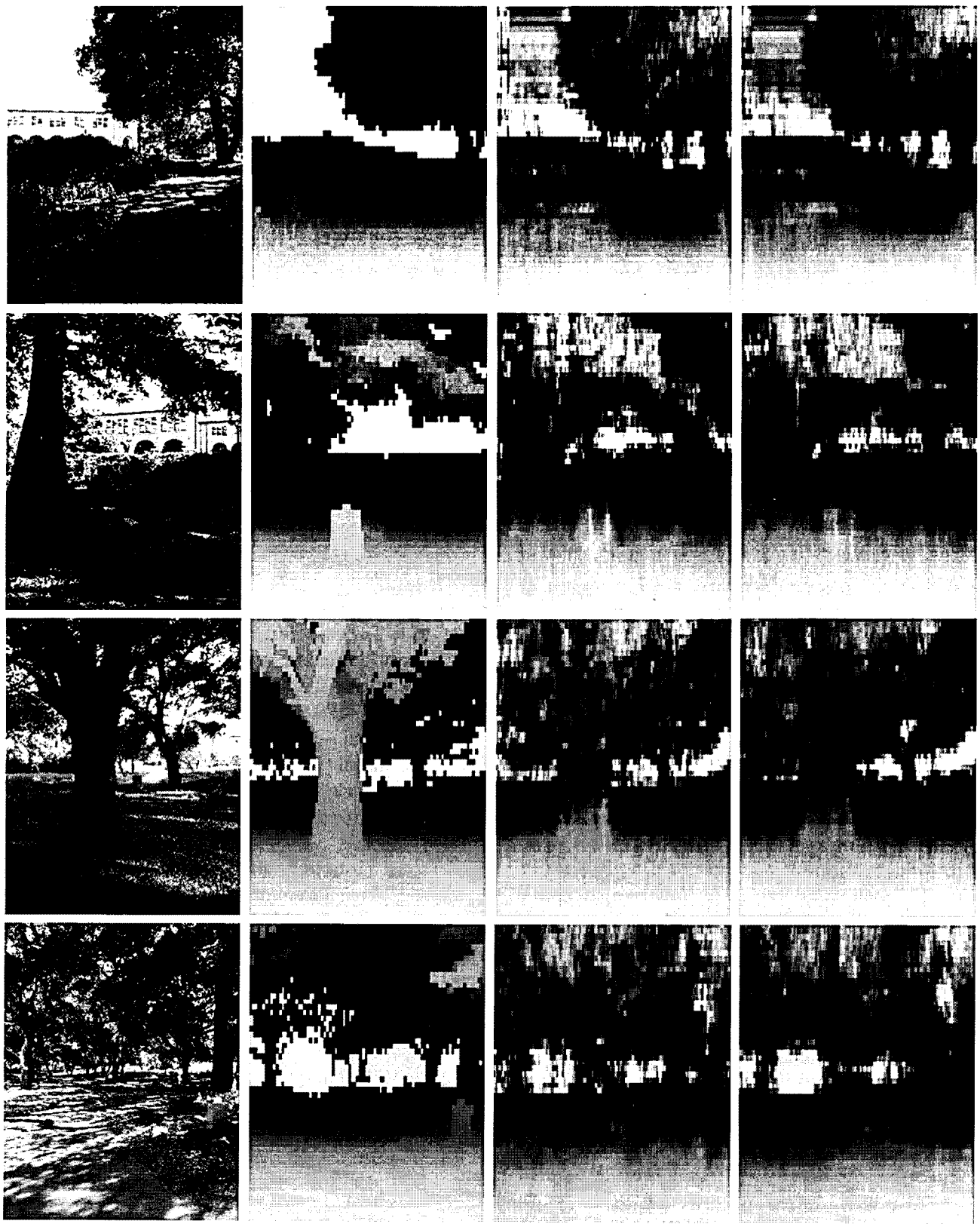


그림 6. 실험 결과 : (첫 번째)원본 이미지, (두 번째)실제 깊이 지도, (세 번째)모든 특징을 이용한 결과, (네 번째)PCA(150차원)를 이용한 결과

Fig. 6. A experiment result : (column 1)original image, (column 2)true-depthmap, (column 3)result using all feature vector, (column 4)result using PCA(150 dimension).

$$Error = \sum_{r=1}^{row} \sum_{c=1}^{col} |d_{rc} - w_r x_{rc}| \quad (11)$$

d_{rc}는 영역의 실제 깊이를 의미하고 w_r은 각 row별로 학습을 통해 구해진 변환 벡터를 의미한다. 그리고 x_{rc}는 해당 영역의 특징 벡터를 의미하며 Ⅲ장에서와 같이 정규화 후에 PCA를 이용하여 특징의 차원이 축소, 변환 되어진 특징 벡터이다.

4. 실험 결과

실험을 통하여 그림 7과 같은 결과를 얻을 수 있었다. X축은 전체 테스트 데이터의 평균 오차율을 나타내고, Y축은 사용한 특징 벡터를 나타낸다. Full-FV는 전체 특징 벡터를 축소하지 않고 원본 그대로 깊이를 추정 했을 때의 오차율이며, PCA-N은 PCA를 통하여 차원을 축소한 특징 벡터로 PCA-50은 축소된 특징 벡터의 차원이 50차원임을 나타낸다.

적은 양의 특징을 사용하거나 전체 특징의 개수와 비슷하게 차원을 축소하는 경우 성능에 영향을 미치지 않거나 정보의 손실에 인한 악영향이 있을 수 있다. 하지만, 대부분의 경우 PCA를 통하여 차원을 줄이고 작은 차원의 특징 벡터를 이용하여 깊이 추정을 할 경우 정확도가 향상됨을 확인 할 수 있다.

본 논문에서 사용한 데이터의 경우 150차원으로 특징 벡터의 차원을 축소하였을 때 원본 데이터에 비해 평균 오차율 0.4%정도 좋은 결과를 보여 주었으며 영상에 따라 최대 2.5%이상 좋은 결과를 보여 주었다.

또한, 성능의 향상과 함께 특징 벡터의 차원이 축소로 학습 시간을 10%정도 줄일 수 있었다.

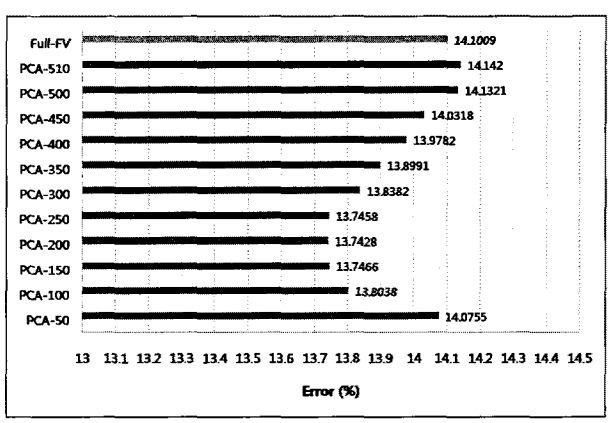


그림 7. 실제 깊이지도와 비교를 통한 PCA를 이용한 특징 개수별 오차율

Fig. 7. Compared each feature vector's error rates using PCA with true-depthmap.

V. 결 론

본 논문에서는 한 장의 정지 영상에서 깊이를 추정하는데 사용되어지는 특징의 차원을 PCA를 통하여 줄이는 방법에 대하여 기술하였다.

또, 특징의 차원을 단계 별(50개 단위)로 축소하고 실제 깊이를 추정하여 오차율을 확인하고 성능이 향상되었음을 증명하였다.

본 연구에서는 중요도에 따라 PCA를 이용하여 특징의 차원을 축소하여 실험 하였다. 하지만, 차원을 이동하는 과정을 거치므로 어느 특징 차원이 중요한지는 알 수 없기 때문에 많은 연산 시간(모든 특징 정보를 구한 후 PCA를 통한 차원 축소)을 단축하지는 못 하였다.

향후에는 중요하지 않은 특징 차원을 사전에 파악하여 특징 추출에 필요한 사전 연산 시간을 단축하고, 성능을 향상 시키는 방법에 대하여 연구가 필요하다.

참 고 문 헌

[1] M. Tipping and C. Bishop, "Probabilistic principal component analysis," Journal of the Royal Statistical Society, Vol. 61, No. 3, pp. 611-622, 1999.

[2] T. Nagai, T. Naruse, M. Ikehara, and A. Kurematsu, "Hmm-based surface reconstruction from single images," In Proc IEEE Int'l Conf Image Processing., Vol. 2, pp. 561-564, 2002.

[3] J. Michels, A. Saxena, and A. Y. Ng, "High speed obstacle avoidance using monocular vision and reinforcement learning," International conference on Machine learning., Vol. 17, pp. 593-600, Bonn, Germany, August 2005.

[4] A. Saxena, S. H. Chung, and A. Ng, "Learning depth from single monocular images, " Advances in Neural Information Processing Systems., Vol. 18, pp. 1161-1168, 2006.

[5] A. V. D. Linde, "PCA-based dimension reduction for splines," Journal of Nonparametric Statistics., Vol. 15, pp. 77-92, 2003.

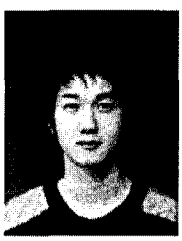
[6] W. Chen, M. Er, and S. Wu, "PCA and LDA in DCT domain," , Pattern Recognition Letters, Vol. 26, No. 15, pp. 2474-2482, 2005.

[7] X. Wei and W. B. Croft, "LDA-based document models for ad-hoc retrieval," Proc. of ACM SIGIR., Vol. 15, pp. 178-185, Washington, USA, 2006.

[8] B. Wu, T. L. Ooi, and Z. J. He, "Perceiving

distance accurately by a directional process of integrating ground information," Letters to Nature., Vol. 428, pp. 73-77, 2004.
[9] X. He, R. S. Zemel, and M. A. Carreira-Perpinan, "Multiscale conditional random fields for image labeling," In Proc. CVPR., Vol. 2, pp. 694-702, 2004.
[10] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," International Journal of Computer Vision., Vol. 47, pp. 7-42, 2002.
[11] S. Das and N. Ahuja, "Performance analysis of stereo, vergence, and focus as depth cues for active vision," IEEE Trans Pattern Analysis & Machine Intelligence., Vol. 17, pp. 1213-1219, 1995.
[12] R. Wolke, "Iteratively reweighted least squares : A comparison of several single step algorithms for linear models," BIT Numerial Mathematics, Vol. 32, No. 3, pp. 506-524, 1992.

저 자 소 개



신 성 식(학생회원)
2006년 전주대학교 컴퓨터공학과 학사 졸업
2008년 전북대학교 컴퓨터공학과 석사 졸업
2008년 전북대학교 컴퓨터공학과 박사 재학
<주관심분야 : 컴퓨터그래픽스, 인공지능, 패턴인식, 소프트웨어공학>

권 오 봉(정회원)
1980년 고려대학교 전기공학과 학사 졸업
1983년 고려대학교 전기공학과 석사 졸업
1993년 일본구주대학교 이공학연구과 박사 졸업
1992년~1993년 일본구주대학교 정보공학과 조수
1994년~현재 전북대학교 전자정보공학부 교수
<주관심분야 : 컴퓨터그래픽스, 사이언티픽 비주얼라이제이션>