

특징 점 Appearance Model을 이용한 3차원 물체 인식

주성문, 박재완, 이철우
전남대학교 컴퓨터공학과

e-mail : tjdansld87@naver.com, cyanlip@naver.com, leecw@chonnam.ac.kr

3D Object Recognition Using Appearance Model of Feature Point

Seong-Moon Joo, Jae-Wan Park, Chil-Woo Lee
Chonnam National University Electronics Computer Engineering

요 약

3차원 물체는 카메라의 시선 방향에 따라 다른 영상을 생성하므로 2차원 영상만으로 3차원 물체를 인식하는 것은 쉬운 일이 아니다. 특히 영상생성 시 강한 perspective transformation 이 발생할 경우 2차원 국소 특징을 이용하는 SIFT(Scale-Invariant Feature Transform) 알고리즘은 매칭에 활용하기 어렵다. 본 논문에서는 3차원 물체를 하나의 특징 축 중심으로 회전시키면서 얻은 복수의 영상을 학습 데이터로 활용하여 SIFT 알고리즘을 개선한 물체인식 방법을 제안한다. 이 방법은 복수 영상의 특징점들을 하나의 특징 공간으로 합성하고 그 특징점들 간의 기하학적인 제약조건을 확인하여 3차원 물체를 인식하는 방법이다. 실험에서는 알고리즘의 유용성을 먼저 확인하기 위해 조명조건과 카메라의 위치를 일정하게 유지하였다. 이 방법에 의해 SIFT 알고리즘만으로 인식이 힘들었던 3차원 물체의 다양한 외관(appearance) 인식이 가능하게 되었다.

1. 서론

2차원 영상만을 이용하여 3차원 회전이 수반된 3차원 물체를 인식한다는 것은 어려운 일이다. 그러나 영상만으로 3차원 물체를 인식할 수 있다면 다양한 응용 서비스가 가능하게 되어 다양한 분야에서 이 연구가 널리 진행되고 있다. 예를 들어 이 연구를 증강현실 분야에 적용할 경우 기존의 영상, 사진, 물체 등을 인식하여 그 결과를 새로운 콘텐츠(동영상, 가상 3D 물체 등)를 삽입하는 기술로 활용이 가능하다[1]. 또한, 물체 영상을 검색 키워드로 사용하는 내용기반 검색엔진[2]이나 로봇 비전 연구에서 영상 획득 장치로 특정한 물체를 인식하여 상호작용을 구현하는 연구[3] 등에서 다양하게 활용이 가능하다.

Affine 변환이 주로 발생하는 영상들 간의 매칭에는 국소적인 특징을 이용한 SIFT알고리즘이 널리 사용되고 있다[4]. SIFT알고리즘은 서로 다른 크기의 이미지에 가우시안 함수를 적용하여 크기 변화에 강건한 후보점을 추출하고 추출한 후보점을 정제한 후 각 후보점에 대하여 방위와 크기를 할당한다. 이와 같은 국소 특징을 이용한 SIFT 알고리즘은 이미지의 회전 변화, 크기변화에는 강건하게 적용되지만, 물체 자체의 3차원적 회전 변화에는 동일한 물체일지라도 인식에 어려움이 있다[5][6]. 이는 국소 후보점을 이용하는 SIFT Descriptor가 물체의 3차원적 회전변화, 즉 perspective transform을 수반하는 포괄적인 변화에는 강인하지 않는 특성을 지니고 있기 때문이다.

위 문제를 해결하기 위한 다양한 연구가 진행되고 있다.

Hamit Soyel은 SIFT 알고리즘 기반 얼굴 인식을 위하여 다른 시점의 카메라 뷰와 기존의 SIFT 매칭을 개선한 새로운 방법을 제안하였다[5]. Ran Zhou는 SIFT알고리즘을 이용하여 얼굴인식을 하면서 인식을 개선하기 위해 3차원 회전을 적용한 후 특징점 매칭을 하는 방법에 대해 제안하였다[6].

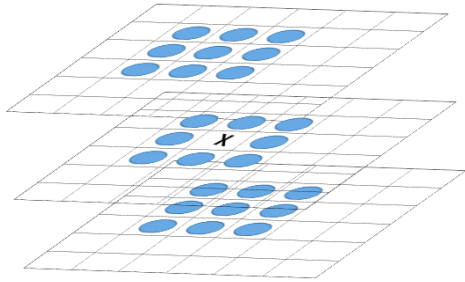
본 논문에서는 특정 회전축을 기준으로 물체의 회전에 대한 database을 구성하여 임의의 회전을 가한 물체의 회전을 추정함으로써 기존 물체인식에 어려움이 있던 물체의 3차원 회전 변화를 보완하는 Descriptor 정의 방법을 제안한다.

2. SIFT알고리즘 기반 3차원 물체 인식

2.1 SIFT알고리즘의 개요

SIFT(Scale-Invariant Feature Transform) 알고리즘은 물체인식 분야에서 이미지의 회전 및 크기 변화에 강건한 국소 특징점을 정의하기 위해서 자주 사용되는 알고리즘이다. 국소 특징점을 정의하기 위하여 후보점을 검출하는 과정에는 DoG(Difference of Gaussian)를 이용한다.

검출된 후보점 중에서 크기변화에 강건한 후보점을 정의하기 위하여 그림1과 같이 현재 Pixel에서 다른 크기의 인접한 Octave 이미지와의 주변 26개의 픽셀값을 비교하여 가장 큰 값을 갖는 후보점을 찾는다.



(그림 1) 크기변화에 강건한 후보점 찾기

검출한 후보점에 대하여 지역화하기 어려운 키포인트를 제거하는 작업을 한다. 먼저 낮은 대비를 갖는 후보점들을 제거한다. 이때는 DoG의 출력값을 검사하여 노이즈에 민감하게 반응하는 불안정한 점들을 제거한다. 그리고 에지 선상의 후보점들은 계산량을 줄이기 위하여 Hessian Matrix을 이용하여 제거한다. 그 후, 각 후보점들에 대하여 이미지 회전변화에 강건하도록 각각 방향성과 크기를 정의한다.

회전에 대하여 36개의 bin을 설정한 후 누적 히스토그램을 통하여 후보점의 대표적인 방향성을 계산할 수 있다. 이를 Keypoint Descriptor로 구성한 후 매칭 시 유클리디언 거리 계산을 통하여 가장 유사한 후보점을 찾는 데 사용할 수 있다.

2.2 SIFT알고리즘의 문제점 분석

SIFT알고리즘을 이용하여 3차원 회전을 적용한 물체에 대한 매칭을 하면, 동일한 물체일지라도 기존에 정의했던 SIFT 특징점들의 매칭에 어려움이 있음을 확인할 수 있다.

그림 2에서는 SIFT알고리즘을 이용하여 특징점을 찾아낸 결과를 보인다. 이 결과를 보면 왼쪽 이미지에서는 132개가 특징점들이 정의되었고, 오른쪽에서는 138개가 정의되었다. 그림 3은 그림 2에서 정의한 두 이미지의 SIFT 특징점을 이용하여 매칭을 수행한 결과를 보인다. 이 결과를 보면 서로 매칭된 SIFT 특징점들의 개수가 단지 2개임을 알 수 있다. 이는 복잡한 3차원 구조를 가진 물체가 회전에 의해 국소 특징이 사라진 영상으로 투영될 경우 단순한 SIFT 알고리즘만으로 매칭이 어렵다는 것을 보여준다.



(그림 2) 물체에 대한 SIFT 특징점 정의



(그림 3) SIFT 특징점을 이용한 매칭 결과

3. 특징점 Appearance Model 기반 3차원 물체 인식

3.1 Appearance Model의 개요

본 논문에서 정의하는 특징점 Appearance Model은 인식할 물체에 대하여 회전축을 고정하고 학습 영상을 구성한 것을 의미한다. 특정 물체를 Y축 회전을 기준으로 일정 각도마다 학습 영상을 구성하면 물체의 3차원 회전 정보를 2차원 이미지에 투영시켜 특징점을 획득할 수 있다.



(그림 4) Y축 회전을 기준으로 구성한 Appearance Model

3.2 특징점 Appearance Model 기반 3차원 물체 인식 알고리즘

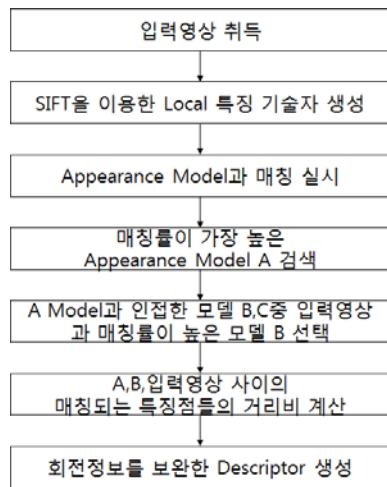
본 논문에서는 물체의 3차원 회전을 인식하기 위하여 물체의 Y축을 기준으로 16방향 Appearance Model을 구성한다. 그 후 Appearance Model과 입력영상의 매칭률이 가장 높은 모델 A를 선택한다. 모델 A를 찾기 위하여 다음과 같은 수식을 이용한다.

$$\text{매칭률} = \frac{\text{Model과 매칭된 특징점 개수}}{\text{입력영상의 SIFT 특징점 개수}} \times 100$$

가장 높은 매칭률을 보이는 Appearance Model과 인접한 Model중 높은 매칭률을 나타내는 Appearance Model를 이용하여 입력영상의 회전각을 추정할 수 있다. 위에서 찾은 두 개의 Model에서 회전각의 크기에 따라 각각의 모델 A,B라고 하면 회전각은 다음과 같이 구할 수 있다.

$$\text{회전각} = \angle A + \frac{m_b * \text{Bin.Size}}{m_a + m_b}$$

두 모델 A,B의 매칭률을 입력영상의 회전에 대한 거리 비로 비례식을 나타낸 것이다. 위에 표시한 m_a 는 Model A의 매칭률이고, m_b 는 Model B의 매칭률, Bin.Size 는 모델 사이의 각도 크기인 22.5이다. 위 식을 이용하면 임의의 입력영상에 대하여 회전각을 추정할 수 있다. 이렇게 추정된 각도를 SIFT Descriptor에 보완하여 입력한다. 그림 5는 알고리즘에 대한 전반적인 흐름도를 나타낸다.



(그림 5) 알고리즘 흐름도

4. 실험 결과

물체의 3차원 Y축 회전을 실험하기 위하여 Y축 회전이 용이한 음료수병을 물체로 하여 실험을 하였다. 실험 환경은 물체의 회전정보를 추정하는 것을 목표로 하기 때문에 세그멘테이션이 용이하도록 배경을 깨끗하게 설정하였다. 촬영한 이미지의 크기를 640×360으로 설정하였으며, 16방향으로 22.5°마다 촬영하여 Appearance Model을 구성한다.

우측 표는 Appearance Model과 임의의 Y축 회전을 가

한 입력 영상의 SIFT특징점 매칭률에 대한 결과이다. S1, S2는 각각 임의의 회전을 가한 입력영상들이다. 회색으로 음영표시 된 부분은 가장 높은 매칭률을 갖는 Appearance Model A와 A에 인접한 모델중 더 높은 매칭률을 보이는 Appearance Model B를 나타낸다. 이 두 모델을 이용하여 입력영상에 대한 회전각을 추정 할 수 있다.

<표 1> 샘플과 Appearance Model 매칭 결과

단위(%)

model \ 샘플	물체1		물체2		물체3		물체4	
	S1	S2	S1	S2	S1	S2	S1	S2
0도	25	10	36	26	6	3	29	33
22.5도	29	17	27	34	3	2	40	46
45도	18	31	19	41	4	4	42	29
67.5도	14	36	12	17	8	4	22	18
90도	12	14	29	22	7	4	17	18
112.5도	7	7	34	26	6	4	18	17
135도	14	9	32	28	5	6	26	17
157.5도	9	10	20	32	5	4	19	14
180도	7	28	15	15	4	3	15	11
202.5도	7	10	20	19	5	4	10	8
225도	6	10	21	20	11	9	10	11
247.5도	6	7	17	18	14	17	9	14
270도	4	5	16	17	13	29	13	13
292.5도	4	2	19	20	28	19	11	13
315도	9	3	12	22	20	8	21	17
337.5도	16	5	36	25	7	6	27	25

그림6은 Y축 회전 이외의 임의의 회전을 적용한 입력영상과 기존의 Appearance model 사이의 매칭률을 나타낸다. 매칭률이 0~2% 사이로 물체 인식을 위한 SIFT 특징점 매칭이 잘 되지 않으므로 인식에 어려움이 크다는 것을 알 수 있다.



(그림 6) 임의의 3차원 회전영상 매칭

5. 결론

본 논문에서는 기존 SIFT 알고리즘을 이용하여 물체 인식이 불가능한 non-Affine transform 문제를 해결하기 위해 특징점 Appearance Model을 이용한 3차원 물체 인식 기술에 대해 제안하였다. Appearance Model을 생성하여 매칭을 하면 기존의 지역 특징점만으로 인식이 어려웠던 영상이 유사한 Appearance Model과 매칭이 되서 물체를 인식할 수 있다.

Appearance Model을 8방향으로 구성하여 매칭할 경우 물체의 Affine Transform 문제를 추정하기에는 모델사이 이미지의 변화가 크다. 임의의 입력 영상과 Appearance Model간에 인식이 쉽지 않으므로 물체 인식을 위한 모델은 16방향 이상의 모델로 설정해야 한다.

Appearance Model의 개수를 증가시킬수록 중간 이미지에 대한 정보를 보유하고 있으므로 회전축이 고정된 Affine Transform에 대해서 인식률을 높일 수 있었다.

현재는 회전축을 Y축으로 고정하여 물체의 3차원 Y축 회전에 대한 인식이 가능함을 알 수 있었지만 임의의 다른 축을 기준으로 회전을 가하게 되면 지역 특징점인 SIFT 특징점의 매칭이 어려웠다.

향후에는 물체의 다양한 회전에 도 인식이 가능하도록 현재 정의된 Descriptor에 글로벌한 정보를 추가하여 지역 특징점들의 매칭이 좀 더 용이하도록 Descriptor을 보완하거나 현재 추출한 지역정보들을 종합하여 3차원 모델을 구축하는 방법에 대하여 연구를 진행할 것이다.

Acknowledgment

“본 연구는 교육과학기술부와 한국연구재단의 지역혁신인력양성사업으로 수행된 연구결과임”

“본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학IT연구센터육성 지원사업의 연구결과로 수행되었음”(NIPA-2013-H0301-13-3005)

참고문헌

- [1] Mu Li, Shantanu Rane, and Petros Boufounos, “Quantized embeddings of scale-invariant image features for mobile augmented reality”, IEEE International Workshop on Multimedia Signal Processing (MMSP 2012), September 2012. pp. 1-6.
- [2] Xiaoou Tang, “IntentSearch: Capturing User Intention for One-Click Internet Image Search”, IEEE Transactions on pattern analysis and machine intelligence, Vol.34 No 7, July 2012. pp.1342-1353.
- [3] Junji Satake, Masaya Chiba, and Jun Miura, “A SIFT-Based Person Identification using a Distance-Dependent Appearance Model for a Person

Following Robot”, Robotics and Biomimetics(ROBIO), 2012 IEEE International Conference on, pp.962-967

[4] D. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision, Vol.60 Issue 2, November 2004. pp.91-110.

[5] Hamit Soyel and Hasan Demirel, “Improved SIFT Matching for Pose Robust Facial Expression Recognition”, Automatic Face& Gesture Recognition and Workshops, 2011 IEEE International Conference on, pp.585-590

[6] Ran Zhou, Jie Wu, Qing He, Chao Hu and Zhuliang Ye, “Approach of Human Face Recognition Based on SIFT Feature Extraction and 3D Rotation Model”, Information and Automation(ICIA), 2011 IEEE International Conference on, pp.476-479