

## 실시간 물체 인식 및 추적을 위한

### 다중 스케일 특징점 매칭\*

백운혁<sup>o</sup> 우운택

광주 과학기술원 U-VR 연구실

{wbaek, wwoo}@gist.ac.kr

### Multi-Scale Feature Matching

### for Real-time Object Recognition and Tracking

Woonhyuk Baek<sup>o</sup> Woontack Woo

GIST U-VR Lab

#### 1. 서론

영상의 특징점과 디스크립터는 컴퓨터 비전 분야에서 물체 인식 및 파노라마 영상제작, 로봇 위치인식과 같이 많은 응용분야에서 활용되고 있는 기본 기술이다. 이러한 특징점과 디스크립터에 대한 연구는 영상의 시점 변화 및 크기 변화에 강건하게 물체를 인식하거나 추적 할 수 있으며, 부분적인 가려짐에도 강건하게 물체를 인식 할 수 있다[1]. 이러한 특징점과 디스크립터는 시점 및 크기 변화에 따라 안정적으로 찾을 수 있는 특징점과 특징점 주변의 영상을 이용해 디스크립터를 계산해 내었다. 최근에는 그래픽 전용 하드웨어를 이용해 실시간으로 특징점과 디스크립터를 계산해 물체 인식 및 추적하는 연구[3]가 진행되었으며, 모바일 기기에서 물체 인식 및 추적을 하기 위한 연구[4]도 진행되었다.

본 논문에서는 스케일정보를 갖지 않는 특징점을 다중 크기 영상으로 학습시키는 방법을 활용하여 실시간으로 물체를 인식 및 추적하는 방법을 제안한다. 제안된 방법은 그래픽 전용 하드웨어와 같이 특별한 장치를 사용하지 않으며, 스케일 정보를 계산하는 과정을 사전 학습으로 해결함으로써 기존의 방법과 유사한 수준의 정확성을 유지하며 실시간으로 물체를 인식 및 추적할 수 있는 장점을 갖는다.

#### 2. 다중 스케일 특징점 매칭 방법

기존의 물체 인식에 사용된 특징점의 경우 특징점의 스케일 변화에 강건하게 하기 위해 입력영상을 다중 크기로 변화시켜 크기변화에 강건한 특징점을 찾아 디스크립터를 생성하거나 추출된 특징점의 스케일 정보를 계산하였다. 이와 같이 스케일 영역에서 지역적 최대값을 갖는 크기변화에 강건한 특징점을 찾는 것이나 특징점의 스케일을 계산하는 것은 어려울 뿐만 아니라 많은 계산시간이 필요하게 된다.

실시간으로 물체를 인식하고 추적하는 방법에는 D. Wagner가 제안한 방법[4]이 있다. 이 방법은 E. Rosten이 제안했던 FAST 코너 특징점[5]을 이용해 빠르게 특징점을 추출하여 사용하였으며, 추출된 특징점 주변 영상 패치를 이용해 SIFT[1] 디스크립터를 생성해 사용하였다. 이러한 방법의 경우에는 스케일 정보를 계산하지 않는 FAST코너 특징점을 사용한다. 이렇게 스케일 정보가 없는 특징점을 사용했기 때문에, 입력 영상의 물체에 크기 변화가 생겼을 경우 인식에 실패를 하게 된다.

본 논문에서 이러한 크기변화에 따른 물체 인식을 실패하는 문제를 없애기 위해 크기 정보가 없는 특징점의 주변 영상 패치를 다양한 크기로 변화를 주어 SURF[2] 디스크립터를 생성하며, 생성된 디스크립터를 군집화 한다. 이렇게 학습시간에 다양한 크기 변화를 갖는 데이터셋을 만들어 내어 스케일정보에 강건한 인식과 동시에 실행 시간에 스케일 정보를 계산하는 과정을 제거하여 빠른시간에 인식이 가능하도록 한다. 또한, 학습을 위한 영상은 단순한 다중 스케일 영상이 아닌 시점변화 포함한 다중 스케일 영상을 만들어 학습에 사용하여 시점 변화에 강건하게 물체를 인식 할 수 있다.

\* 본 연구는 문화체육관광부 및 한국문화콘텐츠진흥원의 문화콘텐츠기술연구소육성사업의 결과로 수행되었음.

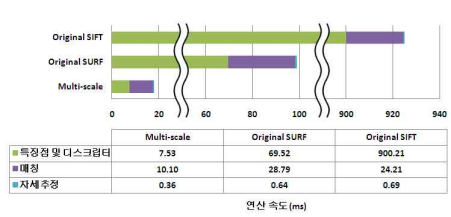
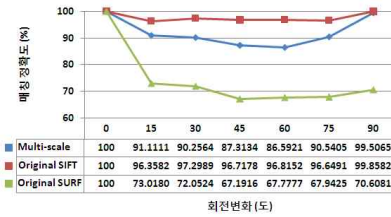
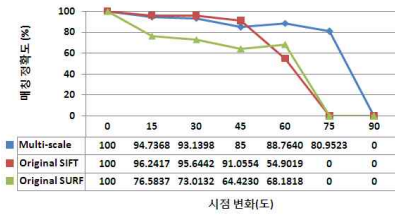


그림 1. 시점 변화에 따른 매칭 정확도    그림 2. 회전 변화에 따른 매칭 정확도    그림 3. 연산속도 비교 (millisecond)

제안된 다중 스케일 영상 매칭방법을 적용하여 실시간으로 물체를 인식 및 추적을 위해 연산 속도와 정확도를 기존의 방법과 비교하여 실험하였다. 모든 실험에는 640x480 크기의 영상과 2.9GHz의 CPU를 갖는 PC를 사용하였으며 허용 오차 범위를 2픽셀로 정의하고 추출된 특징점과 디스크립터 매칭의 정확도를 살펴 보았다.

그림 1은 시점변화에 따른 정확도를 보인 것이다. 45도 이하의 시점변화가 적은 영상의 경우 제안된 방법은 SURF보다 높은 정확도를 보이며, SIFT와 유사한 수준의 정확도를 보이고 있다. 45도 이상의 시점변화가 큰 영상의 경우에는 SIFT와 SURF보다 높은 정확도를 보이며, 시점 변화가 적은 영상과 유사한 수준의 정확도를 유지하고 있다. 특히 75도의 시점변화와 같이 극심한 시점변화를 갖는 경우에도 물체를 인식 할 수 있다. 이는 기존의 방법들과는 달리 학습 시간에 크기 변화 뿐 아니라 시점 변화에 따른 영상을 함께 학습하여 극심한 시점변화에서도 안정적으로 특징점을 매칭하고, 물체를 인식 할 수 있다. 그림 2의 경우에는 회전 변화에 따른 정확도를 보인 것이다. SIFT보다는 떨어지는 정확도를 보이고 있지만, 평균적으로 90%이상의 정확도를 유지하고 있다.

다중스케일 매칭 방법은 영상 1장의 매칭 시간은 18ms(millisecons)이하로 실시간으로 물체를 인식하고 추적할 수 있게 된다. 그림 3은 물체를 인식하기 위한 과정의 단계별로 필요한 시간을 측정된 것이다. 제안한 방법은 실행시간에 특징점의 크기정보를 계산하는 과정을 사전학습으로 해결하여 기존의 SURF에 비해 약 5배정도 빠를 뿐 아니라 SIFT에 비해 약 100배 이상 빠른 연산속도를 보여주고 있다.

### 3. 결론

본 논문에서는 크기 정보가 없는 특징점과 다중 스케일 특징점 매칭 방법에 대해 제안하였다. 이는 기존의 특징점 매칭 방법에서 특징점의 크기정보를 계산하는 시간을 줄여줌으로서 실시간으로 특징점을 찾고 디스크립터를 계산하여 물체를 인식하였다. 이렇게 특징점의 크기정보를 계산하지 않기 위해 특징점을 다중 스케일로 학습하여 사용하였다. 이로 인해 매칭되는 특징점의 개수가 적어져 텍스처가 적은 물체를 안정적으로 물체를 인식 및 추적하기 어려워 졌다. 향후에는 시점 변화 외의 텍스처가 적은 영상을 포함한 다양한 환경 변화에도 안정적으로 물체를 인식하고 추적할 수 있는 방법에 대한 연구 및 개발이 요구된다.

### 4. 참고문헌

- [1] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, pp. 99-110, 2004.
- [2] Herbert Bay, Tinne Tuytelaars and Luc V. Gool, "SURF: Speeded Up Robust Features," European Conference on Computer Vision, pp. 404-417, 2006.
- [3] Sudipta N. Sinha, J. Frahm and M. Pollefeys and Y. Genc, "GPU-based Video Feature Tracking and Matching," In Workshop on Edge Computing Using New Commodity Architecture, 2006.
- [4] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond and D. Schmalstieg, "Pose tracking from natural features on mobile phones," Mixed and Augmented Reality, pp. 125-134, 2008.
- [5] E. Rosten, and T. Drummond, "Machine learning for high-speed corner detection," In Proc. of ECCV, pp. 430-443, 2006