

중첩신경망을 리용한 시야불변인 시각대상인식의 한가지 방법

최진혁, 어경진

선행연구[1, 2]에서는 시야에 기초한 3차원대상인식이나 검색방법들이 많이 연구되었다. 다중시야에 기초한 3차원대상표현의 식별능력이 매우 높으므로 시야에 따르는 대상의 자세가 효율적으로 추정되어야 한다.

본문에서는 3차원대상인식을 위한 시야변화에 따르는 대상의 유형식별능력을 보다 높일수 있는 다중시야중첩신경망을 제안하였다.

1. 다중시야중첩신경망에 의한 3차원대상인식의 구성체계

2개 화상에 기초한 다중시야중첩신경망(CNN: Convolution Neural Network(PMV-CNN)) 구성은 학습단계와 검사단계로 이루어져있다. 매 대상에 대한 3개의 시야가 먼저 선택된 다음 자료층에서 2개의 시야로 갈라진다. 매 대상에 대하여 많은 시야들을 선택할수 있으나 망이 기억용량을 많이 차지하는것으로 하여 3개의 시야를 선택하는것이 합리적이다. 또한 3개의 시야를 선택하면 시야들사이의 보충관계를 유지할수 있다.

CNN 1은 다중시야로부터 병렬처리를 진행하고 같은 파라미터들을 공유하며 5개의 중첩층들로 구성된다. CNN 1을 통과한 후 시야-예비저장층에서 시야별에 따르는 요소별 최대값을 최대응답값특징사영으로 얻고 두 대상들에 대한 비교를 진행한다. 최대응답시야들은 2개의 내적층으로 구성된 CNN 2를 통과한다. 학습시 정의 정, 정의 부의 표본들을 포함한 쌍표본들을 입력으로 리용한다.

2개의 화상렬이 입력으로 되는 다중시야중첩신경망구성은 특징융합층에 의하여 결합된 다음 전결합층과 유연최대층이 검사표본들의 표식을 예측하는데 리용된다. 다중시야로부터 얻은 정보는 한 화상에 효율적으로 루적되어 복합특징표현을 이룬다. 다중시야중첩신경망쌍구성을 검사하기 위하여 각이한 종류들에서 선택된 적은 개수의 자료모임을 리용하여 대상인식성능을 평가한다.

PMV-CNN에 기초한 3차원대상인식의 구성도를 그림에 보여주었다.

학습에서는 많은 쌍표본들이 생성되므로 매 대상에 대하여 우연적으로 3개의 시야를 선택하며 검사에서는 무리짓기연산을 진행한다.

2. CNN에 기초한 심층망구성

제안된 심층망의 구성이 일반적인 중첩신경망과 다른 점은 다음과 같다.

① 시야-예비저장

시야별에 따르는 특징사영을 추출한 후 1개 대상에 대한 여러가지 시야들사이의 호상관계를 모형화하여야 한다. 시야예비저장층은 다중시야들사이의 국부적인 최량선별을 실현하는것을 목적으로 한다.

전통적인 중첩신경망구조에서 최대예비저장이나 평균예비저장은 매 사영층에서 리용되어 부분표본화된 화상들이 얻어진다. 그러나 시야예비저장에서는 매 시야별로 사영 특징화상들을 얻은 다음 각이한 시야들가운데서 매 사영화상들의 요소별최대값을 계산한다. 여러 시야에 해당하는 특징사영들이 한가지 시야에 대한 최대응답특징사영으로 전환되므로 시야선별과 시야들사이의 련관을 고려한다.

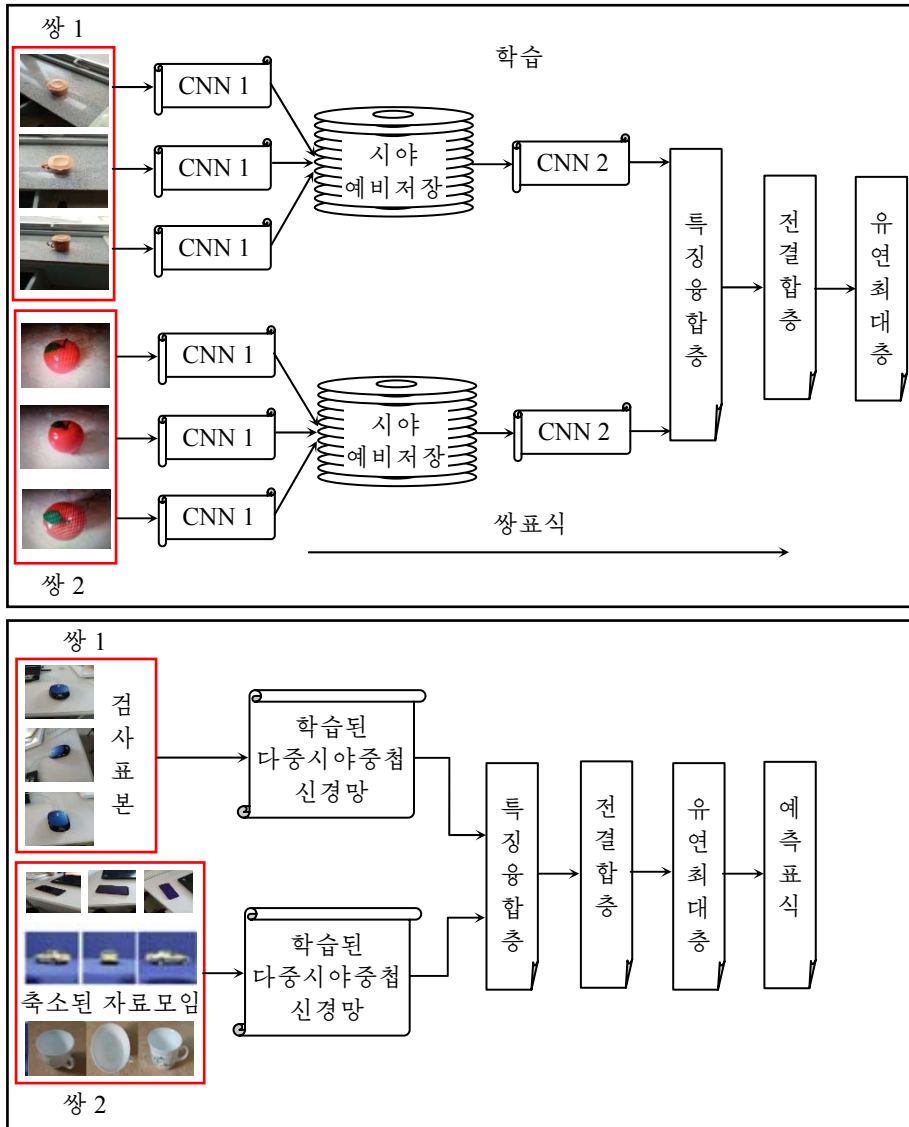


그림. PMV-CNN에 기초한 3차원대상인식의 구성도

② 특징융합층

각이한 대상들에 대한 다중시야표본들이 쌍을 이룬 중첩신경망의 입력으로 된다. 보다 안정되고 밀집된 특징표현을 얻기 위하여 특징을 학습시켜 인식성능을 높인다. 두 대상들이 하나의 유형에 속한다면 일정한 공간에서 어느 정도 학습된 특징표현이 가까이에 배치되며 서로 다른 유형에 속해있다면 멀리 떨어져있게 된다. 그러므로 특징융합층에서

이 두 특징표현을 연결한 다음 융합된 특징표현이 두 대상들이 같은 유형에 속하거나 속하지 않는가를 결정할수 있는 새로운 특징으로 리용된다.

③ 축소된 자료모임

PMV-CNN에서 쌍구조가 설계되었으므로 2개의 쌍표본들이 리용되어야 하지만 검사 과정에는 하나의 화상모임들만을 얻을수 있다. 학습자료모임에서 선택된 표본들로 축소된 자료모임을 검사과정에 리용한다. 구체적으로 유형마다 10개의 대상들을 우연적으로 선택한 다음 이 표본들로 축소된 자료모임을 구성한다. 대상별로 3개의 표본들을 각각 선택하여 쌍표본들을 구성한다. 이 두 대상들이 같은 유형에 속한다면 쌍표본들은 정의 표본으로 된다.

결국 학습에서는 원래 학습자료모임에서 구성된 두 쌍표본들을 리용한다. 검사에서 하나는 검사자료모임에서 1개 대상의 표본들이고 다른 하나는 축소된 자료모임에서 1개 대상의 표본들이다.

3. 다중시야중첩신경망을 리용한 3차원대상분류의 성능평가실험

다중시야중첩신경망구조를 리용하여 3차원대상의 분류방법의 정확성평가를 실험을 통하여 진행하였다. 매 대상에서 80%의 표본들은 학습에 리용한다. 남은 표본들 가운데서 유형마다 매 대상별시야의 50%에 해당하는 표본들은 축소된 자료모임으로, 50%에 해당하는 표본들은 검사자료모임으로 리용한다.

실험에서는 정의 표본들과 부의 표본들의 개수비율을 1:3으로 설정하였다. 검사의 안정성을 확증하기 위하여 매 대상의 나머지시야들을 3개의 부분무리들로 무리짓기하여 대상을 표현한다.

다중시야자료모임에서의 표본정보를 표 1에 보여주었다.

표 1. 다중시야자료모임에서의 표본정보

원래의 정보				쌍을 이룬 다중시야정보				
자료모임	유형/ 개	대상/ 개	대상별 시야/개	표본	시야별 쌍/개	다중시야 (한 대상) /개	쌍을 이룬 다중시야 (두 대상)/개	쌍을 이룬 다중시야(모든 대상) /개
자료모임 1	8	80	41	3 280	3	10 660	10 660×10 660	3 160×10 660×10 660
자료모임 2	47	549	60	32 940	3	34 220	34 220×34 220	150 426×34 220×34 220

표 1에서 보여준것처럼 제안된 다중시야중첩신경망은 보충관계를 고려하면서도 학습 표본의 개수를 충분히 늘이는것으로 하여 학습성능을 보다 높일수 있게 한다.

2개의 자료모임에 대한 인식방법들의 성능비교결과를 표 2에 보여주었다.

표 2. 2개의 자료모임에 대한 인식방법들의 성능비교

방법	자료모임 1/%	자료모임 2/%
선행방법[1]	12.4	47.5
선행방법[2]	93.8	37.2
제안된 방법	95.0	63.8

표 2에서 보여준것처럼 제안된 방법이 선행방법보다 유형분류정확도가 훨씬 더 높다는것을 알수 있다.

맺 는 말

일반적인 중첩신경망에 비해볼 때 촬영대상의 각이한 자세변화에 따르는 다중시야쌍의 중첩신경망을 리용하면 시각대상인식의 분류성능이 보다 높아진다.

참 고 문 헌

- [1] Nazim Ashraf et al.; Computer Vision and Imaging Understanding, 123, 1, 41, 2014.
- [2] Zhigang Tu et al.; Pattern Recognition, 79, 1, 32, 2018.

주체109(2020)년 11월 5일 원고접수

A Method of Visual View Invariant Object Recognition by Using Convolution Neural Network

Choe Jin Hyok, O Kyong Jin

In this paper, we proposed multi-view CNN architecture to recognize 3D objects and evaluated performance of object classification through experiments.

Keywords: view invariant, 3D object recognition, multi-view CNN