

## 호모그래피 변환을 이용한 가시광 및 적외선 열화상 영상 정합

# Visible Light and Infrared Thermal Image Registration Method Using Homography Transformation

이상협<sup>1</sup>, 박장식<sup>2\*</sup>

Sang-Hyeop Lee<sup>1</sup>, Jang-Sik Park<sup>2\*</sup>

### 〈Abstract〉

Symptoms of foot-and-mouth disease include fever and drooling a lot around the hoof, blisters in the mouth, poor appetite, blisters around the hoof, and blisters around the hoof. Research is underway on smart barns that remotely manage these symptoms through cameras. Visible light cameras can measure the condition of livestock such as blisters, but cannot measure body temperature. On the other hand, infrared thermal imaging cameras can measure body temperature, but it is difficult to measure the condition of livestock. In this paper, we propose an object detection system using deep learning-based livestock detection using visible and infrared thermal imaging composite camera modules for preemptive response

*Keywords : Object Detection, Thermal Image, Image Matching, Homography Transform.*

---

<sup>1</sup> 경성대학교 전자공학과  
E-mail: soundslock05@naver.com

<sup>2\*</sup> 교신저자, 경성대학교 전자공학과, 교수  
E-mail: jsipark@ks.ac.kr

<sup>1</sup> Dept. of Electronic Eng., Kyung Sung University, Student

<sup>2\*</sup> Corresponding Author, Dept. of Electronic Eng., Kyung Sung University, Professor

## 1. 서론

객체 검출(Object detection)은 영상에서 가축, 사람 또는 차량 등의 특정 종류의 객체를 배경과 분리 검출하고, 보안감시, 교통서비스 등의 다양한 분야에 응용되고 있다[1-3]. 축산산업 분야에서는 사료의 양을 측정하여 공급하는 사료급이기 등의 축사 자동화와 가축의 움직임을 파악하고 가축의 상태를 파악하여 가축 관리 등 활발하게 적용되고 있다[4].

2014년 이후 매년 1만 마리 이상의 가축이 구제역으로 인하여 살 처분되는 상황으로 이로 인해 매년 3,000억 이상의 막대한 경제적 손실이 발생한다. 구제역의 증상은 가축의 발굽 주변에 발열과 침을 많이 흘리거나 입속에 물질 발생, 식욕 부진 등이며, 소는 발굽 주변에 물질이 발생하고, 돼지는 발굽 주변에 수포가 발생한다. 특히 가장 신속한 구제역 검진은 가축의 체온 변화를 통해 확인 가능하다[5]. 이러한 특징을 객체 검출을 통해 질병 개체에 대하여 선제적 대응하기 위해 기계학습(machine learning) 및 심층학습(Deep learning)기반의 질병 개체를 파악하기 위한 연구가 진행되고 있다.

기계학습 중에서 EM 알고리즘을 이용하여 정상 개체와 질병 개체의 생체 데이터 군집화(clustering)를 통해 질병 개체와 정상 개체를 구별하는 연구를 수행하였다[6]. Sujun 등은 가축의 둔부에 온도 감지 센서를 부착하여 온도에 따른 센서의 색상 변화를 CNN(Convolution neural network) 심층학습을 통해 가축의 온도를 측정하는 연구를 수행하였다[7].

기존의 가축 객체 검출의 경우 가시광 영상을 이용하였으나, 조명 환경의 영향이 많다. 특히 야간이나 날씨, 환경 등에 의한 조명 변화에 따른 객체 검출 성능에 영향을 받는다. 색상 정보가 많은 주간 영상의 경우 객체의 형태가 명확하여 객

체의 윤곽, 그림자 등 검출 되는 특징이 다양한 반면, 야간에는 낮은 조도로 인해 객체 주변 요소보다 높은 휘도의 배경으로 인하여 색상 정보를 획득하기 어렵기 때문에 객체의 특징 검출이 제한되어 객체의 검출이 어려움이 발생한다. 외부 조명환경의 변화에 성능이 저하되는 문제점을 해결하기 위해서 모델 구조를 변경하거나 적외선 열화상 영상을 이용한 객체 검출을 하는 연구가 이루어지고 있다. 또한 가축 관리를 위한 객체 검출도 중요하지만 구제역 등 가축 질병에 대한 선제적 대응을 위해서는 가축의 체온 측정도 필요하기 때문에 가시광 영상 및 적외선 열화상 영상을 동시에 활용하는 연구가 수행되었다.[8-9].

본 논문에서는 구제역 등의 가축 질병에 대하여 선제적으로 대응을 위해 가시광 및 적외선 열화상 복합 카메라 모듈을 활용하여 심층학습 기반 가축 객체 검출 및 호모그래피 변환을 적용하여 가시광 영상과 적외선 열화상 영상을 정합하는 방법을 제안한다. 가축 객체 검출은 가시광 영상과 적외선 열화상 영상 각각에 대해 심층 기반 객체 검출 모델인 YOLO v4(You only look ones)[10]을 사용하여 객체를 검출한다. 영상 정합(Image registration)은 가시광 영상과 적외선 열화상 영상의 공통된 특징을 추출하고, 검출된 특징을 특징 점으로 하여 상관관계를 계산하여 호모그래피 변환(Homography transformation)[11]을 통해 정합을 수행한다. 시물레이션을 통해 제안하는 영상 정합을 이용하는 방법이 단순 가시광 영상 또는 적외선 열화상 영상만을 활용한 학습에 비해서 성능 우수하고 가축 상태를 파악하는데 효과적임을 보인다.

## 2. 복합 영상 데이터셋 및 심층학습 모델

본 논문에서는 효율적인 축사 관리를 위하여

가시광 및 열화상 복합 카메라로부터 가시광 영상과 열화상 영상을 획득하고, YOLO 딥러닝 모델을 활용하여 객체를 검출하고 영상 정합하는 방법을 제안한다.

구제역 선제적 대응을 위한 호모그래피 변환을 이용한 복합 카메라 정합 영상 데이터셋에 대한 심층학습 기반의 객체 검출을 제안한다. 제안하는 알고리즘은 Fig. 1과 같다.

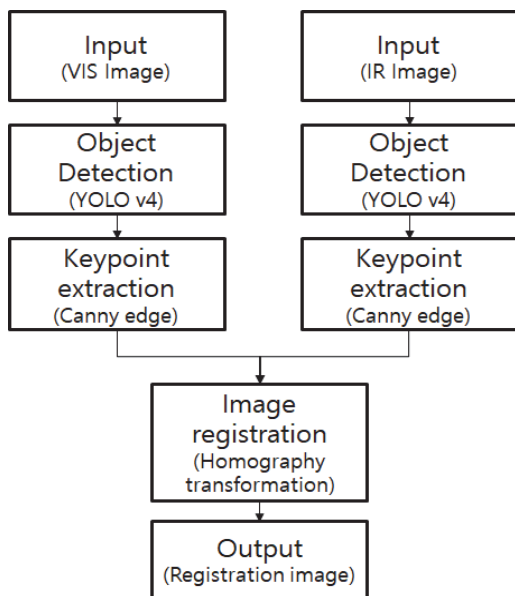


Fig. 1 Block diagram of the proposed algorithm

### 2.1 복합 카메라 영상 데이터셋

가축 객체 검출 학습과 시험을 위해 데이터셋은 가시광 및 열화상 복합 카메라를 사용하는 모듈을 이용하여 주간 우축사를 촬영한 영상 데이터셋을 구축하였다.

복합 카메라 영상 데이터셋은 각 카메라 영상에 대해 42건의 동영상에서 생성된 2,173장의 정지영상에서 가축 객체 5,000건으로 구성되어 있어, 총 82건의 영상, 4,346장의 정지영상, 가축

객체 10,000건으로 구성되어 있다. 가시광 영상은 색영상(Color image)이고, 해상도는 1920×1080 화소이다. 적외선 열화상 영상은 회색조 영상(Gray scale image)이고, 해상도는 640×512 화소이다.

Fig. 2은 가축 객체 검출 학습과 시험에 활용된 복합 카메라 데이터셋의 예이다. 오른쪽은 가시광 영상이며, 왼쪽은 열화상 영상이다.



Fig. 2 Examples of Visible light and infrared thermal multi camera dataset

### 2.2 심층학습 기반의 객체 검출 모델

객체 검출을 위해 사용된 심층학습 모델은 객체의 위치를 찾기 위한 영역제안(Region proposal)과 객체를 식별하기 위한 객체 분류(Classification)으로 구성된다.

YOLO는 영역제안과 객체 분류를 동시에 수행하는 단일 단계 객체 검출의 대표적인 모델이며, 실시간 처리가 필요한 곳에 사용된다. 본 논문에서 사용된 YOLOv4를 사용한다[12].

Fig. 3은 가시광 영상 및 적외선 영상에 대한 객체 검출 결과이다.



(a)



(b)

Fig. 3 Object detection results for livestock  
(a) Visible image, (b) Infrared thermal image

### 3. 영상 정합 및 호모그래피 변환

#### 3.1 영상 정합

영상 정합은 서로 다른 두 영상에 동일한 영역을 찾는 것으로 주로 컴퓨터 비전 분야에서 많이 다루어지고 있다. 두 영상 사이에 밝기, 명암, 색상 등 광학적 특성, 가로와 세로 위치, 크기, 기하학적 특성 등 차이가 날 수 있다. 이를 다양한 방법을 통해 동일한 영역을 찾는 기법이다[13].

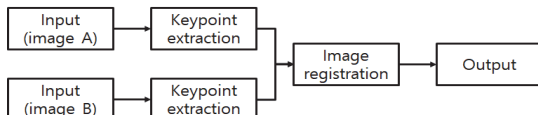


Fig. 4 Block diagram of keypoint extraction based image registration

Fig. 4은 특징 기반의 영상 정합을 나타낸다. 두 입력 영상에 대해 각각 특징점 추출(Keypoint extraction)을 수행하고 이를 이용해 영상 정합을 수행한다.

특징 기반 영상 정합은 선택된 점들에 대해 점들 사이의 대응관계를 통하여 영상 전체의 대응과 정합을 얻는 방식이다. 점들을 선택할 때는 균등한 간격으로 선택할 수 있다.

특징 점 주변의 화소 값 자체를 비교하는 것으로, 이 때 점 주변의 화소들을 모아서 얻은 작은 영상을 패치라고 한다. 서로 다른 두 영상에서 얻은 특징 점의 패치를 모두 비교해서 유사도가 높은 패치들이 많다면 두 영상의 정합을 달성할 수 있다. 두 패치의 유사도를 측정하는 방법에는 SSD(Sum of Squared Difference), SAD(Sum of Absolute Difference)가 있다. SSD는 두 패치에서 같은 위치에 있는 픽셀들끼리의 차이를 구하고 제곱합을 구하는 것이고 SAD는 이의 절댓값을 합하는 것이다. SSD와 SAD의 수식은 식 (1) 과 식 (2)와 같다.

$$SSD = \sum_{x,y} I_1(x,y) - I_2(x,y)^2 \quad (1)$$

$$SAD = \sum_{x,y} |I_1(x,y) - I_2(x,y)| \quad (2)$$

두 패치에 대해 구한 SSD나 SAD 값이 작을수록 두 패치는 비슷한 패치라고 볼 수 있다. 임계값을 지정하여 두 패치가 유사한지 아닌지 결정한다. 따라서 패치의 크기와 유사도의 임계값 두 가지가 패치 기반 영상 정합의 성능을 결정짓는 중요한 요소가 된다.

#### 3.2 호모그래피 변환

호모그래피 변환은 한 평면을 다른 평면에 투

영시켰을 때 투영된 대응점들 사이에 일정한 변환 관계이다.  $3 \times 3$  행렬로 표현되며 대응점들의 동차 좌표 표현에 대해 성립하는 변환관계이다. 수식으로 표현하며 다음 식 (3)과 같이 표현된다.

$$w \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

여기서  $(x, y)$ 는 이전 위치이고,  $(x', y')$ 는 변환된 위치이다. 한 평면 위의 점들이 각기 다른 평면 위의 점들로 투영되었다면 대응점들 사이에는 항상 행렬  $H$ 가 존재하고 또 유일하게 존재한다. 호모그래피 변환을 결정하기 위해서는 임의의 4개의 대응쌍을 이용하여 구한  $H$ 는 다른 모든 점들에 대해서도 동일하게 적용되어야 한다[14].

## 4. 시뮬레이션 결과 및 검토

성능 비교를 위하여 단일 영상만 이용한 심층 학습 기반의 객체 검출과 복합 카메라 데이터셋의 객체 검출의 미검출 성능과 온도 측정 유무를 비교 분석한다. 학습 및 추론 그리고 영상정합 후처리 시뮬레이션은 리눅스 Ubuntu 18.04, PyTorch, CUDA 10.2, CuDNN 7.2, Caffe2, Dual Quadro P5000 GPU 환경에서 수행하였다.

### 4.1 객체 검출

심층 학습 기반의 객체 검출 모델은 YOLO v4를 사용한다. 학습에 사용된 데이터는 동일한 환경에 대해 가시광과 적외선 열화상 복합 카메라 모듈로 촬영된 우축사 영상을 사용한다. 가시광 영상 및 적외선 열화상 영상 학습에 각 정지영상

1,717장과 가축객체 3,500건을 사용한다. 시험을 위해 575장의 정지영상과 가축 객체 1,500건을 사용한다.

### 4.2 특징점 추출 및 정합 위치 선정

특징점 추출은 윤곽선을 특징점으로 하여 추출한다. 윤곽선 검출은 Canny 윤곽선 검출 기법을 통해 윤곽선을 추출한다. 두 영상의 해상도가 일치하지 않기 때문에, 해상도가 높은 가시광 영상에 대해 적외선 열화상 영상을 슬라이딩 윈도우 (Sliding Window) 기법을 통해 1화소씩 움직여 모든 픽셀에 대해 상관관계를 계산하여 상관관계 값이 가장 높은 값을 가지는 위치를 선정한다.

Fig. 5은 가시광 영상과 적외선 열화상 영상의 윤곽선 검출 결과이다.



(a)



(b)

Fig. 5 Object detection results for livestock  
(a) Visible image, (b) Infrared thermal image

### 4.3 호모그래피 변환 기반 영상 정합

특징점 추출과 정합 위치 선정을 통해 선정된 위치에 대하여 호모그래피 변환을 수행한다. 호모그래피 변환은 적외선 열화상 영상의 각  $x, y, z$  축에 대하여  $-15^\circ$  부터  $15^\circ$  까지  $0.5^\circ$  씩 회전을 수행하여 상관관계를 계산하여 가장 높은 값을 가지는 회전에 대해 영상 정합을 수행한다. 이후 가시광 영상 및 적외선 열화상 영상의 객체 검출 결과에 대하여 영상 매핑(Mapping)을 수행한다. 영상 맵핑은 두 영상의 모든 객체에 대해 IoU (Intersection over Union)을 계산하여 0.75 이상의 값을 가지는 객체는 동일한 객체로 판단하여 Fig. 6과 같이 초록색으로 표시한다.



Fig. 6 Objects detected simultaneously in visible light and thermal image are displayed in green

심층학습 기반의 객체 검출 정확도는 기존의 단순 영상 정합과 제안하는 호모그래피 변환을 이용한 영상 정합 영상에 대해 FPPI (False Positive per Image) 0.1 기준에서의 미검출율(Miss rate)로 측정하였고, Table 1에 표기하였다.

적외선 열화상 영상이 미검출율 0.169로 가장 성능이 높고 온도도 측정이 가능하지만 색상 정보가 부족하여 주변 정보를 알 수 없는 단점을 보인다. 제안하는 정합 영상의 경우 미검출율 0.171로 적외선 열화상 영상과 0.002의 차이로 큰 차이를

나타내지 않는다. 하지만 제안하는 정합 영상은 색상 정보도 획득 가능하며 온도도 측정 가능하다.

Table 1. Object detection performance measurement according to deep learning model

Videos	Miss rate	Temp
VIS	0.186	-
IR	0.169	○
VIS + IR	0.171	○

### 5. 결론

본 논문에서는 구제역 선제적 대응을 위해 가시광 및 적외선 열화상 복합 카메라 모듈을 사용하여 심층학습 기반 객체 검출 및 호모그래피 변환을 적용하여 가시광 영상과 적외선 열화상 영상을 정합하는 방법을 제안한다.

객체 검출에는 2,173장의 정지영상과, 가축 객체 5,000건에 대하여 YOLO v4 모델을 사용한다. 적외선 열화상 영상만 사용하는 객체 검출에 비해 미검출율이 0.002 높지만, 색상 정보를 알 수 있다는 장점을 가지고 있다.

기존의 가시광 영상의 객체 검출은 온도 정보를 알 수 없고, 적외선 열화상 영상은 가시광 영상에 비해 객체 검출 성능이 높고 온도 정보를 알 수 있으나 색상 정보가 없는 문제점이 있다. 이를 본 논문에서 제안하는 영상 정합을 통해 객체의 색상 정보 및 온도 정보를 활용하여 구제역에 대해 선제적 대응이 가능 할 것으로 판단된다.

### 감사의 글

본 논문은 부산광역시 및 (재)부산인재평생교육

진흥원의 BB21플러스 사업(2021년도)과 경상대학교 2020년도 교내특별연구 지원으로 수행되었음

## 참고문헌

- [1] Zaidi, S. S. A., Ansari, M. S., Aslam, A., Kanwal, N., Asghar, M., and Lee, B., "A Survey of Modern Deep Learning based Object Detection Models," *arXiv preprint arXiv:2104.11892*, (2021).
- [2] Park, S.-J., Han, S.-Y., Park, S.-B., and Kim, J.-H., "A Study on the Detection Method of Lane Based on Deep Learning for Autonomous Driving," *Journal of the Korean Society of Industry Convergence*, vol. 23, no. 6\_2, pp. 979-987, (2020).
- [3] J.-H. Lee and J.-S. Kim, "A Study on the Stability Control of Injection-molded Product Weight using Artificial Neural Network," *Journal of The Korean Society of Industry Convergence*, vol. 23, no. 5, pp. 773-787, Oct. 2020.
- [4] Qiao, Y., Su, D., Kong, H., Sukkarieh, S., Lomax, S., and Clark, C., "Data augmentation for deep learning based cattle segmentation in precision livestock farming," *In 2020 IEEE 16th International Conference on Automation Science and Engineering*, pp. 979-984. DOI: 10.1109/CASE48305.2020.9216758. (2020)
- [5] Jamal SM, Belsham GJ. "Foot-and-mouth disease: past, present and future," *Vet Res*, vol. 44, no. 1, pp. 116. DOI: 10.1186/1297-9716-44-116. PMID: 24308718; PMCID: PMC4028749. (2013)
- [6] W. Lee, S. Hwang and J. Kim, "Fast Detection of Disease in Livestock based on Machine Learning," *The 37th conference of Korea Institute of information and communication engineering*, vol. 19, no. 1, pp.294-297, (2015).
- [7] Chowdhury, S., Verma, B., Roberts, J., Corbet, N., and Swain, D., "Deep Learning Based Computer Vision Technique for Automatic Heat Detection in Cows," *2016 International Conference on Digital Image Computing: Techniques and Applications*, pp. 1-6, (2016).
- [8] Lee, Y., and Shin, J., "DNN Based Multi-spectrum Pedestrian Detection Method Using Color and Thermal Image," *Journal of Broadcast Engineering*, vol. 23, no. 3, pp. 361-368, (2018).
- [9] Xudong, Z., Xi, K., Ningning, F., and Gang, L., "Automatic recognition of dairy cow mastitis from thermal images by a deep learning detector," *Computers and Electronics in Agriculture*, vol. 178, (2020).
- [10] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M., "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, (2020).
- [11] Liu, J., Yang, S., Fang, Y., and Guo, Z., "Structure-guided image inpainting using homography transformation," *IEEE Transactions on Multimedia*, vol. 20, no. 12, pp. 3252-3265, (2018).
- [12] Wang, R. J., Li, X., and Ling, C. X., "Pelec: A real-time object detection system on mobile devices," *arXiv preprint arXiv:1804.06882*, (2018).
- [13] Borovec, J., Kybic, J., Arganda-Carreras, I., Sorokin, D. V., Bueno, G., Khvostikov, A. V., and et. al, "ANHIR: automatic non-rigid histological image registration challenge," *IEEE transactions on medical imaging*, vol. 39, no. 10, pp. 3042-3052, (2020).
- [14] Nguyen, T., Chen, S. W., Shivakumar, S. S., Taylor, C. J., and Kumar, V., "Unsupervised deep homography: A fast and robust homography estimation model," *IEEE Robotics and Automation Letters*, vol. 3, no. 3, pp. 2346-2353, (2018).