

# 고정 카메라 기반 비디오 모니터링 환경에서 GMM을 활용한 움직인 돼지 탐지

이세준, 유승현, 손승욱, 정용화, 박대희  
고려대학교 컴퓨터융합소프트웨어학과  
e-mail: llss22@korea.ac.kr

## GMM-based Moving Pigs Detection under Static Camera-based Video Monitoring

Sejun Lee, Seunghyun Yu, Seungwook Son, Yongwha Chung, Daihee Park  
Dept. of Computer Convergence Software, Korea University

### 요 약

고정 카메라 환경에서 움직이는 객체만을 탐지하는 것은 비디오 모니터링의 중요한 응용 분야이다. 본 논문에서는 비디오의 특성인 움직임 정보가 포함된 영상에서 GMM을 이용하여 움직인 돼지와 움직이지 않은 돼지의 위치를 대략적으로 구분하고, 추가적인 영상 처리 기법과 딥러닝 기반 객체 탐지기를 적용한 박스 단위 객체 탐지 결과를 활용하여 움직인 돼지의 외곽선을 보정한다. 돈사에서 촬영된 비디오 데이터로 실험한 결과, 제안 방법은 효과적으로 움직인 돼지를 탐지할 수 있음을 확인하였다.

### 1. 서론

비디오의 특성인 움직임 정보가 포함된 영상에서 GMM(Gaussian Mixture Model)[1]을 이용하여 움직인 객체를 탐지하는 문제는 매우 오래된 영상 처리 문제이지만, 아직도 그 정확도가 미흡하다는 한계가 있다. 본 연구에서는 최근 CNN(Convolutional Neural Network) 기술의 발전으로 정확도가 획기적으로 개선된 딥러닝 기반 객체 탐지기[2]의 결과를 활용하여 영상 처리 결과의 정확도를 개선코자 한다.

사실 객체 탐지 등 다양한 영상 처리 문제에 딥러닝 기법이 적용되고 있으나, 다중 객체 추적 등 실제 비디오 모니터링 응용에 적용하기 위해서는 정확도와 계산시간 사이의 트레이드오프를 고려해야 한다. 특히 스마트 축산과 관련하여 많은 관련 연구가 발표되고 있으나[3], 실질적인 비디오 분석을 위해서는 보다 정확하고 실시간 처리가 가능한 방법의 개발이 필요하다.

본 연구에서는 돈사내 설치된 카메라로부터 입력되는 비디오로부터 움직인 돼지를 탐지하는 효과적인 방법을 제안한다. 즉, 비디오의 특성인 움직임 정보가 포함된 영상에서 GMM을 적용하여 움직인 돼지의 대략적인 위치를 결정한 후, 딥러닝 기반 객체 탐지기를 적용한 박스 단위 객체 탐지 결과를 활용하여 움직인 돼지의 외곽선을 보정한다. 특히 현재 CNN 기반의 탐지기중 처리속도 대비 정확도가 가장 뛰어나다고 알려진 YOLOv4[4] 탐지 결과를

활용하여 질감 정보 추출용 파라미터를 보정함으로써 움직인 돼지를 보다 정확히 탐지한다. 제안 방법의 효과를 확인하기 위하여 실제 돈사에 설치된 고정 카메라에서 획득된 비디오에서 추출한 데이터로 실험한 결과, 제안 방법에 의한 움직인 돼지 탐지가 매우 효과적임을 확인하였다.

### 2. 제안 방법

본 연구에서는 움직인 객체를 탐지하기 위해 unsupervised data에서 효과적으로 작동하는 clustering 알고리즘 중 하나인 GMM(Gaussian Mixture Model) 알고리즘을 적용하여 동영상에서 움직임 정보를 추출한다. 하지만 GMM 결과에는 구조물의 미세한 움직임, 배경에 나타나는 광원의 변화, 이동하지 않는 객체의 작은 움직임 등의 불필요한 정보가 다량으로 검출되는 등 오탐지 문제와 움직였다고 판단된 객체의 일부가 나타나지 않는 미탐지 문제가 발생한다(그림 1 및 2 참조).



그림 1. 영상처리 적용 전 원본 이미지



그림 2. 그림 1 영상에 대한 GMM 탐지 결과

이러한 문제들은 움직인 객체 탐지를 정상적으로 이루어지지 않게 하는 원인이 될 수 있으므로, 본 연구에서는 이를 보정하기 위해 픽셀의 질감 정보와 배경 이미지, 그리고 YOLOv4의 탐지 box 정보를 활용하여 GMM 알고리즘의 결과를 개선하는 방법을 제안한다. 그림 3은 본 논문에서 제안하는 방법의 전체적인 흐름을 보여준다.

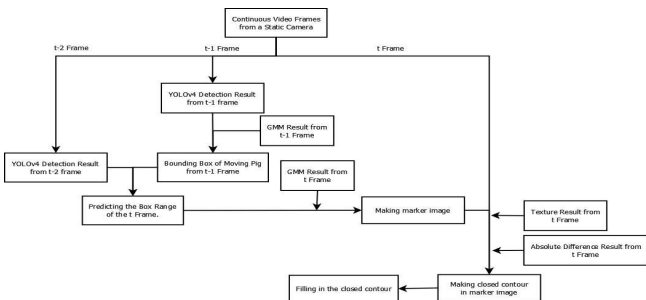


그림 3. 영상 처리 및 YOLOv4 탐지 결과를 활용한 움직인 돼지 탐지 방법

우선 GMM 결과에서 노이즈를 제거하기 위해 침식(erosion) 연산을 반복한다(본 논문에서는 7번). 실험 결과의 그림 8은 침식 연산의 반복 횟수에 따른 탐지 결과의 변화를 보여준다. 그림 8과 같이 침식 연산을 반복함에 따라 움직이지 않는 돼지와 이미지 중앙에 존재하는 노이즈가 사라지는 것을 확인할 수 있다. 남은 픽셀들은 확실한 전경을 나타낸다고 가정된 뒤, 이후 알고리즘에서 전경을 탐지하는 기준으로 사용한다. 기준점은 객체의 탐지를 나타내는 255의 픽셀 값으로 표현한다.

그리고 전경 픽셀의 영역을 예측하기 위해 YOLOv4의 탐지 결과인 box 정보를 이용한다. YOLOv4 탐지기의 정확도가 100%는 아니지만, 테스트 환경에서 mAP(IOUS) 기준 96% 이상의 정확도로 측정되었기에 본 연구에서는 YOLOv4 탐지 결과를 신뢰할 수 있는 정보라고 가정한다. 또한 YOLOv4의 box 내부에만 객체가 존재하고 box 외부에는 객체가 존재하지 않는다고 가정한다. Box의 범위는 t-2 번째 프레임과 t-1 번째 프레임의 box 정보를 비교하여 box의 이동 방향을 확인한 뒤, box의 이동 방향의 범위를 늘려 t 번째 프레임에서 객체가 존재할 것으로 예측되는 예측 box를 구성한다. 예측 box의 외부 픽셀은 확실히 배경임을 나타내는 검정색(본 연구에서는 0의 픽

셀 값)으로 설정하고, 예측 box의 내부 픽셀이면서 기준점이 아닌 픽셀은 확실하지 않다고 표시할 수 있는 임의의 정수(본 연구에서는 128의 픽셀 값)로 표시한다. 설명을 위하여 이와 같이 세 가지 영역으로 구분된 이미지를 ‘마커 이미지’로 명명한다.

이후 마커 이미지의 예측 영역에서 객체의 경계선을 생성한다. 여기에는 두 가지의 정보를 사용하는데, 그 첫 번째는 픽셀의 질감 정보이다. 각각의 픽셀 값만으로는 한 픽셀이 어떤 부분을 나타내는지 알 수 없지만, 주변 픽셀의 최댓값, 최솟값을 비교하면 객체나 배경의 중앙에 있는 픽셀과 경계면에 있는 픽셀을 구분할 수 있다. 본 연구에서는 픽셀 주변의 7 × 7 크기의 block을 만들어 값을 비교하고, 이를 통해 객체의 대략적인 경계선을 생성한다. 두 번째로 사전에 생성한 배경 이미지를 이용하여 원본 이미지와의 픽셀값 차이를 구한 후 원본 픽셀값과의 비율을 이용해 배경으로 추정되는 픽셀을 제거한다. 본 연구에서는 해당 원본 픽셀에서 배경 픽셀의 차를 원본 픽셀 값으로 나눈 값이 기준 수치(본 연구에서는 0.25)보다 낮을 경우 해당 픽셀이 배경이라고 가정한다. 이러한 두가지 정보 각각은 완전히 닫힌 경계선을 구할 수 없지만, 병행해서 사용하면 각각의 에러를 보완한 닫힌 경계선의 생성이 가능하다. 마지막으로, 기준으로 사용하는 픽셀에서부터 시작하여 생성된 객체의 경계 내부를 255의 픽셀 값으로 변경한다.

### 3. 실험 결과

본 실험은 경상남도 하동군에 위치한 바른양돈 돈사에서, 절반의 영역만 모니터링하는 카메라에서 획득된 비디오로 수행되었다. 카메라는 돈사의 중앙을 기준으로 2.1 m 높이의 기둥에 약 45도 각도로 설치되어있으며, 이를 통해 12시간 분량의 1920 × 1080 해상도의 돼지들 영상 데이터를 획득하였고, 약 9만장의 키 프레임을 추출하였다 [5]. 또한, YOLOv4의 처리 속도를 고려하여 영상의 해상도를 512 × 288 해상도로 변경하였다. 또한, tilted-down-view로 촬영된 영상으로 인한 앞과 뒤쪽 돼지들의 크기 차이를 균일하게 하기 위해 영상의 시점을 변환하여 영상의 뒷부분의 비교적 작은 객체는 확장하고 앞부분의 비교적 큰 객체는 축소하여 top-view 형태로 변경하였다. 마지막으로, 본 실험은 Intel(R) Core(TM) i7-9700 CPU @ 3.00GHz, 64GB RAM, GeForce RTX 2080 GPU, CUDA 10.0 환경에서 수행되었다.

실험 결과, GMM 결과(그림 2 참조)와 비교하여 제안 방법(그림 4~7 참조)이 매우 효과적으로 움직인 돼지를 탐지할 수 있음을 확인하였다. 또한, GMM 알고리즘의 runtime은 평균 10.08ms(99.15 FPS)로 측정되었으며, 제안 방법의 runtime은 평균 4.87ms(205.33FPS)로 측정되었다. 즉, GMM을 포함한 전체 runtime이 14.95ms(66.88FPS)로, 제안 방법이 실시간(30 FPS, 33ms 기준) 환경에서도 돈사내 움직인 돼지를 탐지할 수 있음을 확인하였다.

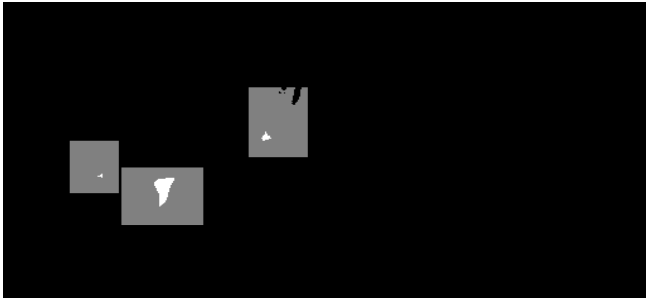


그림 4. 그림 1 영상에 대한 마커 이미지

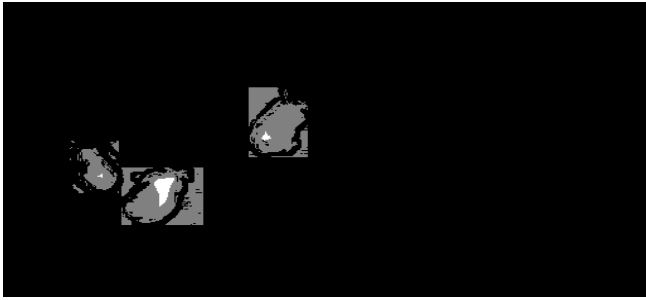


그림 5. 그림 1 영상에 대한 질감 정보를 이용한 객체 경계선 생성 결과

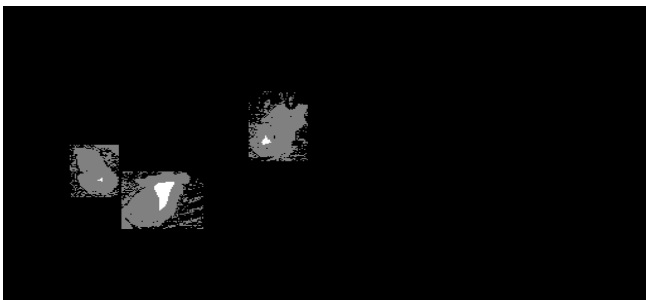


그림 6. 그림 1 영상에 대한 배경 정보를 이용한 객체 경계선 생성 결과



그림 7. 그림 1 영상에 대한 제안 방법 적용 결과



그림 8. 침식 연산 횟수에 따른 제안 방법의 적용 결과 변화 (위에서부터 침식연산 3회, 5회, 7회 반복)

#### 4. 결론

스마트 축산용 실시간 비디오 모니터링 응용을 위해서는 정확도-계산시간 간의 트레이드오프를 고려해야 한다. 본 연구에서는 실제 돈사에 설치된 고정 카메라에서 획득된 비디오 영상에서 움직인 돼지를 보다 정확히 탐지하기 위하여 GMM 기법과 CNN 기법을 동시에 활용하는 방법을 제안하였다. 사실 최근의 CNN 기반 딥러닝 기술의 발전으로 정지 영상 내 돼지에 대한 픽셀 단위 추출이 가능하지만, 이러한 픽셀 단위의 출력을 생성하기 위하여 많은 학습 데이터가 필요하며 처리 속도도 오래 걸리는 한계가 있다. 또한 동영상 내에서 움직인 돼지와 움직이지 않은 돼지를 구분하기 위해서는 더욱 많은 학습 데이터가 필요하고 계산량이 증가한다는 한계가 있다.

본 연구에서는 최근 정확도가 개선된 정지 영상용 CNN 기반 객체 탐지기[4]의 결과를 입력받아 질감 추출용 파라미터를 지속적으로 튜닝함으로써 GMM 기반 움직인 돼지 탐지의 정확도를 개선할 수 있는 방법을 제안하였다. 즉, 계산량 측면에서 비디오 데이터를 실시간으로 분석할 수 있는 딥러닝 기법이 아직 성숙되지 않은 상황에서, 제안 방법은 돈사 내 움직인 돼지들을 탐지할 수 있는 효과적인 방법이 될 수 있음을 확인하였다.

## 감사의 글

본 논문은 2020년도 과학기술정보통신부의 재원으로 연구개발특구진흥재단의 과학벨트성과확산지원사업 (1711123920)과 2021년도 정부(과학기술정보통신부)의 재원으로 한국연구재단-현장맞춤형 이공계 인재양성 지원사업의 지원을 받아 수행된 연구임 (No. 2019H1D8A1109907) 지원으로 수행된 연구결과임.

## 참고 문헌

- [1] C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," *Proc. of CVPR*, 1999.
- [2] L. Liu, et al., "Deep Learning for Generic Object Detection: A Survey," *International Journal of Computer Vision*, Vol. 128, pp. 261-318, 2020.
- [3] S. Matthews, et al., "Early Detection of Health and Welfare Compromises through Automated Detection of Behavioural Changes in Pigs," *The Veterinary Journal*, Vol. 217, pp. 43-51, 2016.
- [4] A. Bochkovskiy, C. Wang, and H. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [5] S. Yu, S. Son, H. Ahn, Y. Chung, and D. Park, "비디오에서 키프레임 추출을 위한 파라미터 설정" *Proc. of IPIU*, 2021.