Technical Report Analyzing Various Tracking Techniques

Dema Abu Adas, School of Computer Science, University of Guelph
Prepared for Dr. Stacey Scott, School of Computer Science, University of Guelph
April 28th, 2018

Abstract

The goal of this research was to find cost-efficient technology that can accurately monitor a large population of cattle in an outdoor environment. From the results of existing tracking systems, the use of computer vision was the best fit for the desired criteria as it would be ideal for an outdoor, large-spaced farm land. The use of thermal imaging analysis was also heavily researched as that was by far the most developed for tracking animals.

Introduction

As technology improves, the idea of smart farms become more desirable for farmers with large populations of cattle. Due to growing cattle farms, it becomes increasingly difficult for farmers to constantly monitor their stock to catch early on to early signs of sickness or abnormal behaviour. This report is specific for monitoring and tracking the behaviour of large-sized animals. The goal of this research was to find cost efficient, existing technology that can accurately monitor a large cattle size ideally in an outdoor environment. The four categories explored to find tracking and analyzing systems for were:

1. Physical exterior health of an animal
2. Interior health of an animal
3. Psychological health of an animal (anxiety, depression)

1. Different Types of Tracking Techniques

Table 1 Summary of various tracking techniques (Ibrahim, Venkat, Subramanian, Khader, & Wilde, 2016)

Type of Tracking	Summary	Pros	Cons
GPS	GPS is built into smartphones; can easily create an app that uses GPS technology	- Allows tracking of various types of devices and can give their actual positions continuously	- GPS devices would need to be attached to the animal; would not be cost efficient for the scalability of cattle
Bluetooth	Bluetooth technology is found in smartphones, laptops and Personal Digital Assistants (PDA). A test trial for tracking crowd density in a soccer championship public viewing in Belgium reported a 75% accuracy.	- Cost-effective - Low-power	- Short-range

Computer Vision	Tracking objects using computers is essential for areas where human monitoring would be difficult and tedious.	- Can be very efficient in tracking multiple objects at once since it would be heavily automated. (Ibrahim et al., 2016) - Cost efficient since the monitoring system would be external to the animal and stationed at every x area in a farm.	- Difficult because of the dynamic human motion (Ibrahim et al., 2016). - Hard because of clothing, partial occlusion and illumination conditions (Ibrahim et al., 2016). - Difficult to track multiple objects over time in scenes and to keep a constant identified of the targeted object (Ibrahim et al., 2016).
RFID	Small electronic device that has a small chip and antenna (Ibrahim et al., 2016).	- All you need is RFID tags and readers.	- less reliable. (Ibrahim et al., 2016) - processes less memory power despite being expensive (Ibrahim et al., 2016). - RFID tags are application specific.
IR Receiver and Transmitter	A researcher has implemented a system for crowd monitoring using an IR receiver and transmitter but is still in early stages (Ibrahim et al., 2016).	N/A	N/A

The use of tracking nodes (GPS or Bluetooth-based) on every animal is an extremely inefficient way of tracking since the cost of the tracking system is directly proportionate to the size of cattle. The research was leaned toward external tracking tools. Due to the large limitations of most external tracking tools, we then narrowed down the scope to computer vision tracking techniques as the area of expertise is emerging, and there are many different types of computer vision tracking techniques present currently and are being explored.

2. Computer Vision Based Tracking Techniques

2.1 Visible Light Cameras (RBG)

One research article applied an RGB camera to locate pigs and identified whether they were touching using concave points in video frames (Ju, Park, & Kim, 2017). This research targeted a larger population of pigs, specifically. The method by which livestock was determined to be touching was by taking the RGB shots, converting them into HSV colour space (Hue, Saturation, Value), and then binarizing the frame (Ju et al., 2017). Subsequently, the researchers applied a Gaussian Mixture Model and if the pig size is larger than average, the pigs would then be flagged as touching (Ju et al., 2017).

The application of the RGB camera to detect whether or not multiple pigs were touching was used in experimental trial with a camera placed four meters above the floor with specifications of 640 x 480 pixel and eight fps (frames per second) (Ju et al., 2017). The room size 4 by 3 m^2 with 22 pigs (Ju et al., 2017). The results of the research were that two pigs were caught touching. However, there were limitations in which the algorithm could not pick up whether more than two pigs were touching (Ju et al., 2017).

2.2 Thermal Cameras

A study conducted in a zoological park in Australia monitored giraffes that were enclosed in the same living space as zebras (Dong et al., 2017). A single thermal camera (FLIR E60: 18mm FOL lens, night-vision support, and relative simple background segmentation) provided an automatic way to located giraffes in the enclosure. The average giraffe body temperatures were 38.5 ± 0.5°C (Dong et al., 2017). The use of thermal cameras was optimal considering the regions of temperature worked versus the atmosphere temperature. The researchers found that there were typically three ways to differentiate between animals: skin temperature, body size, or body shape (Dong et al., 2017). However, zebras have very similar body size and skin temperature, therefore the researchers were limited to only use body shape to differentiate between the two species.

Furthermore, the researchers decided to track the animal in two parts: the neck and body. They decided to use a machine learning algorithm that uses Support Vector Machine (SVM) (Dong et al., 2017). It was later decided that it was not feasible to process all detections simultaneously using a SVM as it exhaustively searches through every sub-window of each frame even though as seen in Table 2 the SVM detection algorithm had the highest recall, precision and accuracy results (Dong et al., 2017). Therefore, it was necessary for the film to be divided and processed one-frame-per-second (Dong et al., 2017). The researchers performed tests using different algorithms such as Viola-Jones and Template Matching to identify the differences between methods (Dong et al., 2017). The training and detection was implemented with the vision.CascadeObjectDetector in MATLAB Computer Vision System Toolbox (Dong et al., 2017).

Table 2 Accuracy of current detection algorithms (Dong et al., 2017)

Detection Algorithm	Recall	Precision	Accuracy
Template Matching	81.0%	51.9%	53.0%
SVM (Support Vector Machine)	83.5%	87.0%	78.6%
Viola-Jones	81.5%	83.4%	74.1%

The location of the giraffe was determined by the neck detection algorithm (Dong et al., 2017). One problem the researchers encountered was that the application of the object detection method onto every frame slowed the system down (Dong et al., 2017). In order to speed-up the algorithm, each frame was compared with the previous one using a similarity detector. The system then used all the information to calculate the distance between the giraffe and the camera (Dong et al., 2017).

Table 3 Results of tracking system created (Dong et al., 2017)

	Prediction		
		+	-
Truth	+	762	238
	-	168	133

One limitation in using thermal cameras is that if there were multiple animals within the same thermal temperature and body size/shape, it will be more difficult to identify the animal individually (Dong et al., 2017). In this specific research study, the overall algorithm was robust until the giraffe bent its neck, then it could not differentiate between a zebra or a giraffe (Dong et al., 2017). In Table 3, the results of the tracking system that was created for the giraffes shows the accuracy of the system. It was generally accurate, but there were still some limitations that does not make the system completely robust.

Another study that was aimed to help detect animals using thermography by UAVs was done using astronomical source detection software (Longmore et al., 2017). The researchers chose thermal cameras to identify the animals as there was a large contract between the warm-blooded animals and their background both during the day and at night that allowed for easy identification. (Longmore et al., 2017). The thermal camera they used was a FLIR, Tau 2 LWIR thermal imaging camera core with 7.5 mm lens and 640 x 512 pixels operating at 75 Hz (Longmore et al., 2017). The goal for this research was to make the project robust and cost-efficient, therefore the researches decided to use open source libraries to analyze the thermal images (Longmore et al., 2017).

Moreover, Python was their choice of programming language and Astropy was used as the open source module that astronomers also use for their thermal imagery analysis (Longmore et al., 2017). The researchers specifically used routines within photutils to identify and extract sources within the images. Using the *find_peaks* function, they were able to identify a sub-region in a 2D image (Longmore et al., 2017). The researchers also ended up using the OpenCV library, specifically the Histogram of Oriented Gradients (HOG) detector, which uses Support Vector Machines (SVM). Since the libraries relied on machine learning algorithms, the researchers had to train the system by showing it a set of images contain the object of interest and then not containing the object of interest (Longmore et al., 2017).

Two test trials were sent out to capture both humans and cows up to a maximum altitude of 120 meters (Longmore et al., 2017). One limitation the researchers found is that the data has to be manually transferred to a computer to run the image analysis tools on (Longmore et al., 2017). The researchers found that the images taken were optimal when the drone was close to the ground until 80 meters up (Longmore et al., 2017). They came to a conclusion that the algorithm used was 70% accurate with ± 10% uncertainty depending on the background of the image (Longmore et al., 2017). The majority of algorithm failing was due to cows being huddle together and/or overlapping in the image.

Furthermore, a study that targeted automatic dairy cattle body conditions assessment using thermal imagery used a measurement called Body Condition Scoring (BCS), which is a score to estimate the mobilization of energy reserves of cattle or the degree of fatness/thinness using a 5-point scale (Halachmi et al., 2013). BCS influences productivity, reproduction, health and longevity (Halachmi et al., 2013). The technology used in this research was RFID tags placed on each cow for identification. Thermal cameras were also used to take videos and then selects the cow's best frame. The cow's area was subsequently cropped and the undesired objects were erased (Halachmi et al., 2013). The cow's BCS is calculated by parabola fitting to calculate the cows contour and then calculate the mean absolute error. Finally, the MAE is converted to 1-5 BCS scale and the BCS score is added to the unique cow ID (Halachmi et al., 2013).

The thermal camera that was used was an InfraCAM SD thermal camera made by FLIR systems equipped with a focal plane array detector with a resolution of 120 x 120 pixels and a spectral range of 7.5 – 13 μm (Halachmi et al., 2013). It was attached to the barn ceiling, 310 cm above a weigh station at the exit of the milking parlor (Halachmi et al., 2013). The image pre-processing and selection of frame from the video was done using Matlab software, as well as the object extraction process which involved three steps:

1. Converting RBG to BW format
2. Measure properties of image regions using the Regionprops Matlab funtions
3. Label connected components in 2-D binary image using bwlab Matlab function

(Halachmi et al., 2013)

The setup was tested with cows and the difference in parabolas, as shown in Figure 1, distinguished the relationship between parabola shape and cow size. The system was particularly oblivious when a cow is thin as the green line would be the outline of and the blue line would be the parabola fit made for the cow. A cow is said to be healthier when the parabola is wider. (Halachmi et al., 2013).

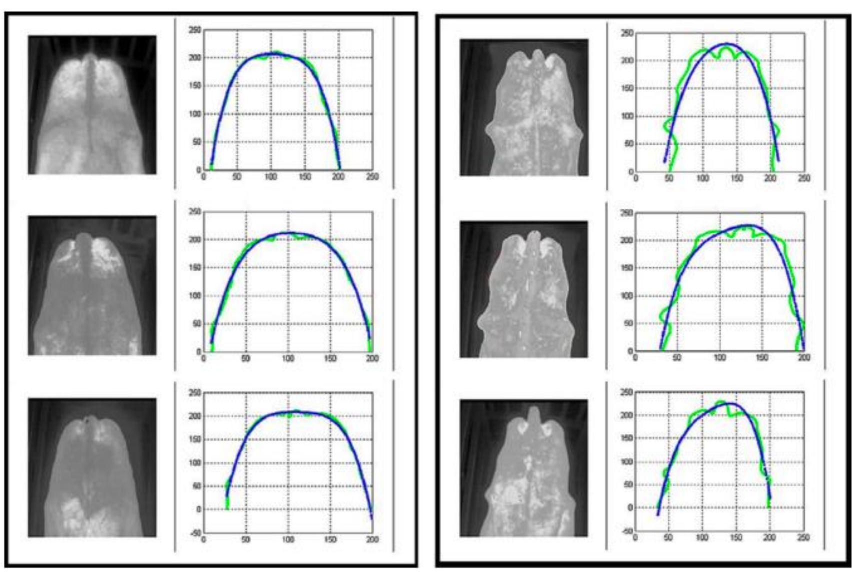

Figure 1: Analysis of thick versus thin cow (Halachmi et al., 2013)

2.3 Infrared Cameras

The advancements in wearable emergency-contact technology have provided the elderly who have difficulty walking with reassurance that they could reach help if an accident would occur. However, the biggest limitation is that an elderly person would have to manually press a button indicating that they have fallen in order to receive help. In order to overcome such limitations, an automatic system to detect falls was created (Planinc & Kampel, 2013). Advantages of using the Kinect was that the foreground/background was already preprocessed using both the official Microsoft SDK and OpenNI, as well as the coordinates of the where the person was located was easily identifiable and accessed (Planinc & Kampel, 2013).

The way the algorithm was designed to recognize a person was by looking at the orientation of the person's major axis and the height of their spine relative to the ground floor (Planinc & Kampel, 2013). The orientation of the major axis was based off of the position of the body's joints. The algorithm was also able to distinguish between a person falling on the ground versus a person lying on the ground using the height of the spine (Planinc & Kampel, 2013). It first mapped out the three-dimensional body-joint coordinates to the two-dimensional depth image and then calculated the features using image coordinates (Planinc & Kampel, 2013). Subsequently, it analyzed the features directly in the three-dimensional space using world coordinates (Planinc & Kampel, 2013). The difference between image coordinates and world coordinates is that image coordinates involves the orientation of the major axis being calculated using the coordinate of the head, shoulder, spine, hip, and knee (Planinc & Kampel, 2013). The least squares algorithm is then applied to fit a straight line, and the angle between the line and horizontal axis is calculated (Planinc & Kampel, 2013). On the other hand, the world coordinate uses the estimated three-dimensional ground floor and the calculated spine distance in relation to the floor (Planinc & Kampel, 2013). Thus, if the major orientation of the person is parallel and the height of the spine is near the ground, then the algorithm would detect that a fall has occurred (Planinc & Kampel, 2013).

Another group of researchers interested in black cattle body shape and temperature for the quality of the growth period of cattle conducted a research project using Kinect and thermal cameras to monitor progress. (Kawasue, Win, Yoshida, & Tokunaga, 2017) The researchers used 3 KINECT sensors to detect 3D data simultaneously. Therefore, the cattle did not need to be stopped while measurements were taken, and it was also adequate for animal sizing (Kawasue, Win, Yoshida, & Tokunaga, 2017). The body shape data was used to monitor the cattle's developmental progress (Kawasue, Win, Yoshida, & Tokunaga, 2017). The cattle are reconstructed on the computer and the measurements of weight, size, posture, body shape, and body temperature were recorded (Kawasue, Win, Yoshida, & Tokunaga, 2017).

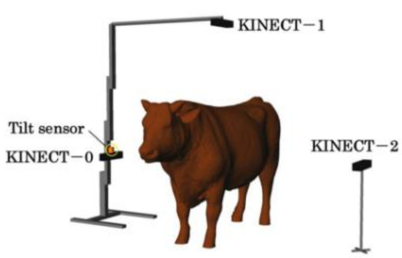

Figure 2 Set up for the system (Kawasue, Win, Yoshida, & Tokunaga, 2017)

Figure 2 illustrates the set-up of the system. The limitation would be that the system cannot be changed and moved around much and requires the cow to be in an ideal position that all cameras would be able to sense. One way to work around that is if the set-up was near an eating location where the cow can eat as the system is checking its vitals.

2.4 Combining Multiple Camera Types

The research study was performed as a part of a mountain rescue mission using an Unmanned Arial Vehicle (UAV) and sensors with real-time video and image data (Egglestone, Ansell, & Cook, 2013). The company, E-MIGS, provided the drone which had two types of cameras attached to it: the Panasonic Kumix 1080HD that had been reduced to Pal format for transmission and the Miracle Thermal Camera that transmitted data in real time to a UAV control station (Egglestone, Ansell, & Cook, 2013). The UAV was able to fly up to 400 feet and took images every 1.5 seconds (Egglestone, Ansell, & Cook, 2013). The drone set-up was tested in Patterdale over a large period of time to ensure its ability to properly function in multiple weather types (Egglestone, Ansell, & Cook, 2013).

One of the limitations of the UAV was that it needed Wi-Fi to retrieve and display the images that we sent to the web server (Egglestone, Ansell, & Cook, 2013). Using UAVs especially the E-MIGS was difficult to constantly maintain as the battery life for the UAV was 10 minutes and had to be changed and charged frequently (Egglestone, Ansell, & Cook, 2013).

Another study that aimed to minimize cost and energy consumption when monitoring cattle used mobile device collecting data and a cloud platform (Lindgren, Zaitov, & Mitkov, 2016). The mechanism behind the experiment involved multiple devices that transmitted with low energy and over short ranges to exchange information between devices as well as a base station with long-range internet connection (Lindgren, Zaitov, & Mitkov, 2016). The researcher also used a cloud service analysis tool that helped collect data to alert farmers (Lindgren, Zaitov, & Mitkov, 2016).

The prototype was implemented and tested, and each animal had their own sensor (Texas Instruments sensor tag with compatible radio) that logged information (Lindgren, Zaitov, & Mitkov, 2016). If another animal's sensor was nearby, they would be an exchange of data. This system was tested with ten cows, each with their own communication device and had six nodes deployed around the farm land (Lindgren, Zaitov, & Mitkov, 2016).

Conclusion

There was a large variety of research studies conducted for tracking and monitoring animals. The results delivered were not very robust, as tracking objects with a camera seems to be very difficult in terms of identifying the correct image analysis with a short margin of error. The most result driven experiments were done using thermal cameras as it is very easy to use thermal cameras both in an indoor and outdoor environment and as well with warm-blooded animals that have a contracting temperature with the environment. As the future advances, thermal cameras have a great potential to be a efficient automated tracking tool for large animals.

References

Dong, R., Carter, M., Smith, W., Joukhadar, Z., Sherwen, S., & Smith, A. (2017). Supporting animal welfare with automatic tracking of giraffes with thermal cameras. *Proceedings of the 29th Australian Conference on Computer-Human Interaction - OZCHI '17*, (November), 386–391. https://doi.org/10.1145/3152771.3156142

Egglestone, P., Ansell, D., & Cook, C. (2013). Demo. *Proceedings of International Conference on Making Sense of Converging Media - AcademicMindTrek '13*, 286–287. https://doi.org/10.1145/2523429.2523431

Halachmi, I., Klopčič, M., Polak, P., Roberts, D. J., & Bewley, J. M. (2013). Automatic assessment of dairy cattle body condition score using thermal imaging. *Computers and Electronics in Agriculture*, *99*, 35–40. https://doi.org/10.1016/j.compag.2013.08.012

Ibrahim, A. M., Venkat, I., Subramanian, K. G., Khader, A. T., & Wilde, P. De. (2016). Intelligent Evacuation Management Systems. *ACM Transactions on Intelligent Systems and Technology*, *7*(3), 1–27. https://doi.org/10.1145/2842630

Ju, M., Park, D., & Kim, H. (2017). Touching-Pigs Segmentation using Concave Points in Continuous Video Frames, 14–18.

Kawasue, K., Win, K. D., Yoshida, K., & Tokunaga, T. (2017). Black cattle body shape and temperature measurement using thermography and KINECT sensor. *Artificial Life and Robotics*, *22*(4), 464–470. https://doi.org/10.1007/s10015-017-0373-2

Lindgren, A., Zaitov, A., & Mitkov, B. S. (2016). Opportunistic iot for monitoring of grazing cattle. *Proceedings of the Eleventh ACM Workshop on Challenged Networks - CHANTS '16*, 31–32. https://doi.org/10.1145/2979683.2979697

Longmore, S. N., Collins, R. P., Pfeifer, S., Fox, S. E., Mulero-Pázmány, M., Bezombes, F., … Wich, S. A. (2017). Adapting astronomical source detection software to help detect animals in thermal images obtained by unmanned aerial systems. *International Journal of Remote Sensing*, *38*(8–10), 2623–2638. https://doi.org/10.1080/01431161.2017.1280639

Planinc, R., & Kampel, M. (2013). Introducing the use of depth data for fall detection. *Personal and Ubiquitous Computing*, *17*(6), 1063–1072. https://doi.org/10.1007/s00779-012-0552-z