Available at www.sciencedirect.com

INFORMATION PROCESSING IN AGRICULTURE 1 (2014) 73–81

journal homepage: www.elsevier.com/locate/inpa

Real-time recognition of sows in video: A supervised approach

Ehsan Khoramshahi [a,*], *Juha Hietaoja* [a], *Anna Valros* [b], *Jinhyeon Yun* [b], *Matti Pastell* [a]

[a] *Department of Agricultural Sciences, University of Helsinki, Finland*
[b] *Faculty of Veterinary Medicine, University of Helsinki, Finland*

ARTICLE INFO

Article history:
Received 19 May 2014
Received in revised form
8 July 2014
Accepted 18 July 2014
Available online 10 August 2014

Keywords:
Precision farming
Supervised classification
Real-time image-processing
Neural network

ABSTRACT

This paper proposes a supervised classification approach for the real-time pattern recognition of sows in an animal supervision system (asup). Our approach offers the possibility of the foreground subtraction in an asup's image processing module where there is lack of statistical information regarding the background. A set of 7 farrowing sessions of sows, during day and night, have been captured (approximately 7 days/sow), which is used for this study. The frames of these recordings have been grabbed with a time shift of 20 s. A collection of 215 frames of 7 different sows with the same lighting condition have been marked and used as the training set. Based on small neighborhoods around a point, a number of image local features are defined, and their separability and performance metrics are compared. For the classification task, a feed-forward neural network (NN) is studied and a realistic configuration in terms of an acceptable level of accuracy and computation time is chosen. The results show that the dense neighborhood feature (d.3×3) is the smallest local set of features with an acceptable level of separability, while it has no negative effect on the complexity of NN. The results also confirm that a significant amount of the desired pattern is accurately detected, even in situations where a portion of the body of a sow is covered by the crate's elements. The performance of the proposed feature set coupled with our chosen configuration reached the rate of 8.5 fps. The true positive rate (TPR) of the classifier is 84.6%, while the false negative rate (FNR) is only about 3%. A comparison between linear logistic regression and NN shows the highly non-linear nature of our proposed set of features.

© 2014 China Agricultural University. Production and hosting by Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, digital information in different forms, including digital images, has relentlessly covered many aspects of our daily lives. This rapid advancement is a natural result of Moore's law progression, as well as the establishment of standards in digital content [12].

Image processing (IP) has recently become an integral part of a variety of domains such as health, multimedia, agriculture, robotic, telecommunication, entertainment and many

* Corresponding author. Address: P.O. Box 28, Koetilantie 5, Department of Agricultural Sciences, University of Helsinki, Helsinki 00014, Finland. Tel.: +358 40 4444135.
E-mail address: ehsan.khoramshahi@helsinki.fi (E. Khoramshahi).
Peer review under the responsibility of China Agricultural University.

Production and hosting by Elsevier

http://dx.doi.org/10.1016/j.inpa.2014.07.002
2214-3173 © 2014 China Agricultural University. Production and hosting by Elsevier B.V. All rights reserved.

others [13]. Among them, modern agriculture has many challenges that could be appropriately addressed by IP techniques. A major challenge in this field is continuous monitoring of animals, which has the potential to improve animal welfare and a higher production level. The continuous supervision of animals by farm workers is expensive and expected to become even more costly in the near future [15]. On the other hand, digital videos include a considerable amount of data, while they are cheap sources of information, so efficient conversion of this data into information would serve many applications in precision livestock farming (PLF), including supervisory tasks.

In fact, a number of behavior analyzers (such as aggression detection [18] and sleeping detection [7]) are potentially eligible to be designed in a way to accept image-based features as extra inputs. IP techniques are indeed the core part to handle the complexities and produce the required information for decision support routines of a higher level.

A well-designed image/video analysis routine can be considered as an expert routine that extracts information about the current status and activity of animals (such as position and movements) in order to predict upcoming events. This means that the quality of an IP algorithm directly affects the quality of consecutive predictor routines.

The role of automatic image-based monitoring is especially emphasized for high-risk situations (e.g. farrowing), where the lack of human supervision could lead to animal injuries or even death. In such cases, early warning from an image-based system allows an efficient intervention to improve the production performance and animal welfare. Therefore, IP techniques should be integrated into such predictive routines to help farms by taking care of individuals or groups of animals automatically.

The very first step in the process of designing a video-based animal behavior classifier is finding the animal's patterns efficiently. This problem could generally be considered as a supervised machine learning (ML) task, because there is prior geometric and/or radiometric knowledge about the underlying pattern, but it can also be treated as a semi-supervised approach by combining on-line estimations.

Many IP methods during the last decade have been employed to locate animals in videos. For example, the combination of likelihood ratios and shading was applied by Hu and Xin [9] to segment pigs from the background. Chen et al. [3,4] implemented an averaging/thresholding routine inside an FPGA to detect animals. Haar of Oriented Gradients (HOG) was employed by Zhang et al. [19] to capture the shape and texture features of the animal's head. Viazzi et al. [16] used the frame difference to extract the pig image. Ahrendt et al. [1] developed an image-based real-time tracking algorithm for pigs. They employ support map segments to build a Gaussian model of the individual pigs. A combination of fuzzy-c means clustering, morphological operation and blob analysis has been studied for segmentation by Kaiyan et al. [11]. To classify aggression behavior among pigs, Viazzi et al. [17] defined a set of images features for Motion History Image (MHI). They recognized the positive cases by employing Linear Discriminant Analysis (LDA).

During recent years, a group of IP methods have been developed and successfully employed in many real-time foreground/background classification tasks (see e.g. [6,8,20,21]).

Background subtraction has been widely studied and used in real applications since the late 1990s. The first versions were mainly based on simple averaging. Those approaches estimated the background by an average over the time and then subtracted the current image from the average to derive an estimation of dynamic objects. Subsequently, statistical approaches have been introduced which have noticeably improved the classification quality. They were mainly based on a normality assumption about the distribution of gray levels. Single Gaussian and the mixture of two Gaussian (MOG) were the very first models. Single Gaussian has been proved to work well, especially for indoor situations with a small change in illumination, but it was unable to work for those situations where the color density of the background and foreground was close. MOG has since been developed, and is now able to approximate more complicated distributions. Many branches from the first MOG have been introduced to improve the defects of the original algorithm. For example Friedman and Russell [6] improved it by computing the effect of shadow by adding a third cluster. He used the expectation–maximization (EM) technique, such as the classic MOG approach, in order to fit his clustering model to the data. His approach has been proved to work well for the situations where there is enough information for different labels (shadow, pattern, and background).

Despite the considerable progress that has been made in recent background-subtraction techniques, they mainly works well in conditions where the target objects are uniformly visible to the camera, and active enough to capture the movement variance (e.g. movements of cars or people), whereas their performance can be significantly affected when the scene becomes static, or the target pattern is covered by some fixed elements (such as crate's bars). In order to overcome this problem, we here propose an algorithm for the special condition of an animal supervision (asup) system that is able in work in static conditions where the target patterns of an asup are covered by some elements and divided into small non-connected pieces.

In our paper we examine the above specific situation by: proposing a set of image-based features according to local neighborhoods, investigating their separability, comparing them in terms of accuracy and required computational time for classification with a feed-forward neural network (ffd-NN), and selecting an optimized configuration for ffd-NN in order to stay in a computationally acceptable timing frame.

The motivation for our research is to prepare algorithmic and structural software platforms for automatic predictions of a domestic animal's behaviors. To enable this research, a C++ image-processing framework has been developed with the aim of reducing the complexity of design by deploying a dynamic-programming scheme.

The first part of this article is a preliminary description of the feature sets. The supervised classification scheme is proposed next. Finally, the set of features are compared in detail, and the process of finding the optimized feature set is discussed.

The main objective of this work is to propose a supervised scheme for the real-time detection of sows in live video streams. The contributions of this paper include: (1) comparing a set of image-based features for the pattern recognition

Download English Version:

https://daneshyari.com/en/article/4493369

Download Persian Version:

https://daneshyari.com/article/4493369

Daneshyari.com