# Dairy Cattle Movement Detecting Technology Using Support Vector Machine

HaoEn Zhou*, Ling Yin, and CaiXing Liu

College of Informatics, South China Agricultural University,
510642 GuangZhou, China
Terren.chow@gmail.com, scauyin@163.com, liu@scau.edu.cn

**Abstract.** In this paper, the dairy cattle movement detecting technology based on 3-axis acceleration sensor information fusion is presented. For they show ideal performance in generalization and optimization, Support vector machines are used to build an information fusion model for dairy cattle's behavior classification. The data feature of the support vector machine fusion model is derived from 3-axis acceleration data. RBF function is used as the model's kernel function. The genetic algorithm is used to optimize the parameters of the kernel function. The training and testing results show that using genetic algorithm for kernel function parameter searching has good ability to optimize the fusion model.

**Keywords:** Dairy cattle movement detection, support vector machine, genetic algorithm.

## 1 Introduction

Automated animal behavior monitoring systems have become increasingly appealing for research and animal production management purposes [1]. Dairy cattle movement detecting plays a very important role in dairy cattle's physiology and health status monitoring.

There are already some approaches for dairy cattle movement detecting. Installing a pedometer on the cattle's leg is the most popular approach. By counting the number of the cattle's step in a specifically time, researcher can evaluate the active rate of that cattle [2-4]. This method has been implemented but its data model is not precise enough. Video-base Pattern recognition (PR) has been another hot point in the study of dairy cattle movement detecting. By applying PR algorithm on each video camera frame, Researcher can calculate the ratio of the cattle's active status and quiescent status based on the cattle's position changing rate. It has a precise data model but needs lots of computing resource and the deployment cost is too high. Object movement detecting using wireless sensor network and 3-axis acceleration sensor already received the attention in recent years.

---

* HaoEn Zhou is a graduate student in College of informatics, South China Agricultural University. He is interest in Wireless Sensor Network and Embedded System. This paper is supported by National 863 Fund under grant 2006AA10Z246.

P. Sénac, M. Ott, and A. Seneviratne (Eds.): ICWCA 2011, LNICST 72, pp. 23–32, 2012.

Support vector machines (SVMs) are known as a suitable learning method, which is based on statistical learning theory, is gaining applications in the areas of machine learning, computer vision and pattern recognition because of the high accuracy and good generalization capability [5]. Unlike artificial neural network (ANN), which uses traditional empirical risk minimization (ERM) to minimize the error on the training data, SVM uses structural risk minimization (SRM) principle to minimize an upper bound on the expected risk. Using SRM as the risk minimization principle improve the SVM's generalization performance compared to ANNs and make it suitable to deal with the object movement detecting problems [6-7]. Some paper presents[8-10] that the combination of SVM and WSN with acceleration sensor was applied to human movement detection and recognition and SVM performs relatively well across different scene. In [1], the author made the first trail at detecting dairy cattle movement using SVM and WSN. 3-axis acceleration is acquired and transmitted back to data processing station from a sensor network node mounted on the target cattle's neck. A multi-class SVMs is used to setup the movement type classification model and it can recognize up to 6 types of cattle's movement. But the parameters of the SVM's kernel function are selected empirically and they have not been optimized. In this paper, a dairy cattle movement detection technology based on 3-axis acceleration sensor and binary SVM is present. Parameters and the feature set of the SVM are optimized separately in order to improve the accuracy and the performance. The remainder of the paper is organized as follows: chapter II describes the binary SVM model for cattle's movement detection, chapter III discusses binary SVM's parameter optimization algorithm, chapter IV focuses on the details of applying both SVM and GA to cattle movement detection, and chapter V is the conclusion.

## 2 Binary Classification SVMs

SVMs are derived from statistical learning and VC-dimension theory. The k-class classification SVMs construct a decision function $f(x)=\operatorname{sgn}(g(x,\alpha))$, given a training data set $T=\{(x_1,y_2),\ldots,(x_i,y_i),(x_j,y_j)\}\in(X\times Y)^l$, where $x_i\in X=R^n, i=1,\ldots,l$ is a vector of length $d$ and $y_i\in\{1,\ldots,l\}$ is the class of the samples. After the training process, the decision function provides a specified class $y\in\{1,\ldots,k\}$ for an input $x\in R^n$.

In the case of binary classification problem, the main idea of binary classification SVMs is, given a training data set, constructing a hyper plane to separate the data set into two classes $y\in\{1,-1\}$ so that the margin between the two classes is maximal. It is equals to the optimization problem:

$$\min_{W\in H,b\in R,\xi\in R^l}\frac{1}{2}\|W\|^2+C\sum_{i=1}^{l}\xi_i\ . \tag{1}$$

with constrains (2),

$$\begin{gathered} y_i\left(\left(W \bullet \Phi\left(x_i\right)+b\right)\right) \geq 1-\xi_i, i=1, \ldots, l \\ \xi_i \geq 0, i=1, \ldots, l \end{gathered} \quad . \tag{2}$$

where $W$ is the weight vector and $b$ is the bias term. $\Phi\left(x_i\right)$ is the non-liner mapping function which maps the input space $R^n$ to the high dimensional Hilbert space $H$ :

$$\begin{gathered} \Phi : X \subset R^n \rightarrow H \\ x \rightarrow \Phi(x) \end{gathered} \quad . \tag{3}$$

In order to solve this problem, a dual problem is constructed:

$$\min_{\alpha} \frac{1}{2} \sum_{i=1}^{l} \sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j K\left(x_i, x_j\right)-\sum_{j=1}^{l} \alpha_j \quad . \tag{4}$$

with constrains (5),

$$\begin{gathered} \sum_{i=1}^{l} y_i \alpha_i=0 \\ 0 \leq \alpha_i \leq C, i=1, \ldots, l \end{gathered} \quad . \tag{5}$$

where $K\left(x_i, x_j\right)$ is the kernel function. It is the inner product of two $\Phi(x)$ mappings. That is $K\left(x_i, x_j\right)=\Phi\left(x_i\right) \bullet \Phi\left(x_j\right)$. In this pager, the radial basis function (RBF) :

$$R(xi, xj)=\exp\left(-\|xi-xj\| / 2 \delta^2\right) \quad . \tag{6}$$

is selected as the kernel function. $\delta$ is a kernel function parameter called kernel width and it must be selected by user in advance.

Suppose that $\alpha^*=\left(\alpha_1^*, \ldots, \alpha_l^*\right)^T$ is the optimal solution to optimization problem (4). The decision function is constructed:

$$f(x)=\operatorname{sgn}\left(\sum_{i=1}^{l} a_i^* y_i K\left(x, x_j\right)+b^*\right) \quad . \tag{7}$$

where $b^*=y_j-\sum_{i=1}^{l} y_i \alpha_i K\left(x_i, x_j\right)$ .

## 3 Optimizing SVM Parameters Using Genetic Algorithm

The genetic algorithm simulates Darwinian theory of evolution using highly parallel, mathematical algorithm that, transform a set (population) of mathematical object (typically string of 1’s and 0’s) into a new population, using operators such as reproduction, mutation, crossover. Genetic algorithms by their very nature consider a wide range of the search space for a particular problem. Unlike other techniques, they search from a population of pointes not from a single point. So they are less likely to become trapped on false optimization peaks. Using simple operations, genetic algorithm are able to rapidly optimize design parameter after examining only a small fraction of the search space.

In this paper, the SVM punishment parameter $C$ and the kernel width of the RBF kernel function $\delta$ are optimized by genetic algorithm. The values of the two parameters are directly code in the chromosomes with real data. The encoding form of the two parameters is $\left[[C_i][\delta_i]\right], i=1,\ldots,m$, where $m$ is the number of sample the initial population. The proposed model used to select the parent chromosome is roulette-wheel method. Crossover and mutation method are used to modify the chromosome. The single best chromosome in each generation is survives to the next generation.

The function of adaptation value is used to advise the selection of the parent to produce the next generation. It also used to generate the probability of the crossover and the mutation operation during the reproduction process. The function of adaptation value is determined by the expectation output and the actual output of the SVM's decision function (7):

$$g(C,\delta)=\left(\sum_{i=1}^{m}|y_i|-\frac{1}{2}\sum_{i=1}^{m}|y_i-f(x_i)|\right)\Big/\sum_{i=1}^{m}|y_i| \quad . \tag{8}$$

Crossover is a process which exchange the gene of two matching individual at a specified probability $P_C$ . In this paper the $P_C$ is:

$$P_C=\max\left(g_i^t,g_j^t\right)\Big/g_{\max}^t \tag{9}$$

where $g_i^t, g_j^t$ is the adaptation value of the two individual parent and $g_{\max}^t$ is the maximum adaptation value in the parent's population.

Then the overlapping of two parent $\left[[C_i][\delta_i]\right]^t$ and $\left[[C_j][\delta_j]\right]^t$ is:

$$\begin{aligned}\left[[C_i][\delta_i]\right]^{t+1}&=P_C\left[[C_j][\delta_j]\right]^t+(1-P_C)\left[[C_i][\delta_i]\right]^t\\ \left[[C_j][\delta_j]\right]^{t+1}&=P_C\left[[C_i][\delta_i]\right]^t+(1-P_C)\left[[C_j][\delta_j]\right]^t\end{aligned} \tag{10}$$

Since the $P_C$ is determine by the high adaptive value of the two parent, the individual who has higher adaptive value can keep more information about itself and pass it to the next generation.

Mutation is a process which changes some genes of an individual randomly at a specified probability $P_M$ . In this paper, the $P_M$ is defined as:

$$P_M\begin{cases}0.8\times\left(g_a^t-g_i^t\right)\Big/g_a^t & g_i^t\le g_a^t\\ 0.2\times\left(g_{\max}^t-g_i^t\right)\Big/g_{\max}^t & others\end{cases} \tag{11}$$

where $g_a^t$ is the average adaptation value of the parent's population. The $P_M$ is higher when the individual has lower adaptive value. It makes the individual more stable than the individual who has a lower adaptive value. It also give a larger chance to make the individual who has a lower adaptive value become better, or in other hands, make them be eliminated more faster.

## 4 Dairy Cattle Movement Detecting Using SVM and GA

The movement of dairy cattle is diversity and complicated, that has many kinds of action states, such as includes walking, canter, sprint, and so on. The dairy cattle movement detecting system is detecting action state of the cattle at all times. Sensor must be installed on the target cattle for action recording. The actions of the target cattle are then converted to a series of sensor output. These sensor data are transmitted to the background station. A classifier running on background station classifies automatically these data in to different classes so the action states of that cattle can be recognized. The behavior of the cattle in a specific moment can be recorded as moving when the action state falls into either walking, canter or sprint class at that time.

### 4.1 Data Collection Model

The cattle's actions data collection model is built on the Wireless Sensor Network (WSN) technology. There are three kinds of network node in the data collection network: sensor node, repeat node and the sink node. A sensor node is installed on the right hand side of the cattle's neck (figure 1). It captures the outputs of the sensor mount on it and sends them to the sink node deployed near the cattle at a rate of 10Hz. The sink node transmits the sensor data to the background station where the classifier is running on though wire network such as Ethernet or RS485 field network. The repeat node acts as a range extender and is placed between the sensor node and the sink node. It receives the data from the sensor node and sends them to the sink node simultaneously.

The radio system of the network nodes is a combination of a Texas Instrument (TI)'s CC2500 single-chip wireless transceiver and a CC2591 wireless signal amplifier. The radio system is running on 2.4GHz wireless channel at 15dBm power level. The maximum communication range between sensor node and sink node can beup to 400 meters if a repeat node is placed between them. The wireless network protocol stack that the nodes are running on is a reduced version of TI's SimpliciTI WSN protocol. The maximum data rate of the network is 250Kbit/s and the maximum hops between the sensor node and sink node are 2 when a repeat node is deployed between them. Data from sensor node can only be transmitted to the repeat node or the sink node. Data hops between two sensor nodes of the same kind not allowed.

In this paper, the sensor used for cattle's action recording is Analog Deceive International (ADI)'s AXD330L 3-axis acceleration sensor. Its measurement range is ±3g and its sensitivity is 300mV/g. The sensor is mounted on the center of the sensor node' PCB. After the installation of sensor node, the positive direction of the sensor's X-axis is point to the tail of the cattle and the Z-axis is point to the neck of the cattle (figure 2).

**Fig. 1.** Installation of the Sensor node

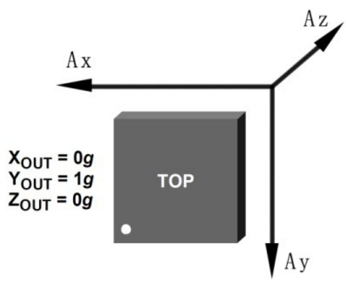


**Fig. 2.** The directions of the sensor's axes after installation

The central control unit of the network nodes is TI's ultra-low power MCU MSP430F169. The MCU keeps sampling the sensor's analog outputs of each axis and converts them digital results at intervals of 100ms. The mapping between the sensor's analog output and the MCU's digital result is:

$$N_{ADC} = 4095 \times \frac{V_{IN} - V_{R^-}}{V_{R^+} - V_{R^-}} \quad . \tag{12}$$

Where $N_{ADC}$ is the MCU's digital output, $V_{IN}$ is the sensor's analog output, $V_{R^-}$ and $V_{R^+}$ are the reference voltages of the MCU. In MSP430F169, $V_{R^-}$ is 0Volt and $V_{R^+}$ is 2.5Volt.

### 4.2 Cattle's Behavior Classification

The cattle's behavior classification is performed by a binary SVM classifier running on the back ground station. After training and testing, the binary SVM classifies the sensor data received from the sink node into two classes: quiescence and movement using the decision function (7). In order to prevent the over fitting problem in the training process, a cross-validation procedure is performed. In the cross-validation procedure, the training set is divided into $n$ subsets of equal size. Sequentially one subset is tested using the classifier trained on the remaining $n-1$ subset. The parameters of the SVM in each subset training process are searched by the GA evolutionary process. These parameters include the SVM punishment parameter $C$ and the kernel width of the RBF kernel function $\delta$ .

The data set used for the training process is the sensor data whose class labels are already know. The class labels of the data in the training set are marked up manually by examining the cattle's monitoring video. The features of the sensor data are $\left\{ax, ay, az, \sqrt{ax^2+ay^2+az^2}, \max, \min,\right\}$, where $ax, ay, az$ is the acceleration value for each of the three axes of the acceleration sensor. $\max$ is the maximum value of the three axes. $\min$ is the minimum value of the three axes.

The cross-validation procedure with the GA parameter searching process is achieved by VC++ program. The SVM training and classification process are achieved with the help of LIBSVM library version 3.0. The integrated process of the cattle's behavior classification is conducted as follows:

Step1: Divide the training set into $n$ subsets.
Step2: Select a subset of training data.
Step3: Generate an initial population of SVM parameter.
Step4: SVM is trained by the initial population of SVM parameter base on the remaining $n-1$ subsets.
Step5: Calculate the error and the adaptation value by (8).
Step6: $P_C$ and $P_M$ are calculated by (9) and (11). Inherit, crossover and mutate the parameters. The next generation population produces.
Step7: SVM is retrained by the new population of SVM parameter on the remaining subsets.
Step8: Repeat step5 to step7 until $m$ generations have been evolved.
Step9: SVM is test on the subset selected in step1. Calculate the classification error rate.
Step10: Step2 to step9 is repeated on another subset until all $n$ subset have been selected once.
Step11: Use the trained SVM of the lowest error rate test as the final SVM.
Step12: Use the final SVM to classify sensor data received from sink node.

The number of subsets $n$ is determined by the total number of the sensor data in the training set. Suppose the sample period of the sensor is $p$ and the number of the samples in a subset is $N$ . The total sampling time $T$ of the subset is:

$$T = p \times N \tag{13}$$

A larger value of $N$ will lead to better training accuracy but consume more time in the training process. The parameter $N$ should be chose carefully to keep the balance of the training time and the training accuracy.

### 4.3 Experiment

Experiments have been done on three individual dairy cattle. The number of data in the training set is 20480. They are divided into 10 subsets. The sampling period of the sensor node on each cattle is 100ms. In the training set, the movement class is label as 1 and the quiescence class is label as -1. The number of evolved generation is 500. Table 1 shows the training result of the three cattle.

**Table 1.** Training results of the three cattle

| No. | Movement samples | Quiescence samples | Best parameters $C, \delta$ | Best accuracy |
|---|---|---|---|---|
| 1 | 15663 | 4817 | 0.55,0.27 | 97.5% |
| 2 | 9561 | 10919 | 0.47,0.34 | 96.8% |
| 3 | 7440 | 13040 | 0.56,0.31 | 98.1% |

After training, the classification SVM for each cattle is tested. 6 testing sets of different size are applied on each cattle and the accuracy is collected. The class labels of each data in the testing set are marked up manually by examining the monitoring video, just the same way as generating the training set. The number of the data in the test sets is selected randomly. The test results are show in table 2 to table 4.

**Table 2.** Test results of the cattle 1

| No. | Test data samples | Accuracy |
|---|---|---|
| 1 | 33458 | 94.3% |
| 2 | 21757 | 94.1% |
| 3 | 22378 | 90.5% |
| 4 | 30145 | 91.0% |
| 5 | 20732 | 89.7% |
| 6 | 17681 | 93.2% |

**Table 3.** Test results of the cattle 2

| No. | Test data samples | Accuracy |
|---|---|---|
| 1 | 11767 | 95.0% |
| 2 | 35233 | 83.1% |
| 3 | 14456 | 77.6% |
| 4 | 23788 | 92.3% |
| 5 | 33660 | 88.9% |
| 6 | 21100 | 97.7% |

**Table 4.** Test results of the cattle 3

| No. | Test data samples | Accuracy |
|---|---|---|
| 1 | 14876 | 96.1% |
| 2 | 34332 | 96.4% |
| 3 | 32543 | 97.3% |
| 4 | 23266 | 92.2% |
| 5 | 17557 | 92.7% |
| 6 | 18684 | 94.5% |

The test results show that the accuracy of the classification is acceptable in most of the time. But there are some cases that the accuracy is rather low compared to other tests. For example, the accuracy of third test for cattle 3 is only 77.6%. One of the main reason is the network transmit error. Data packet lost makes some data in those tests become unavailable. Since all of the features of the unavailable data are zero, the classification SVM's output is fixed to -1. The action state of the cattle is fixed to quiescence when the sensor data is unavailable. Classification error occurs when the actual state of the cattle is movement but the sensor data is lost. The Quality of Service (QoS) of the network is important. The network stack still needs improvements to make the data transmit more reliable.

## 5 Conclusion

For better performance and higher accuracy, dairy cattle movement detecting technology base on 3-axis acceleration sensor information fusion is studied. The sensor data is collected through wireless sensor network. Binary classification SVM is used to build up the information fusion model. The Genetic Algorithm is used for SVM parameter searching in the crossover validation procedure. Test results show that the system has satisfactory accuracy and the quality of the data is a key external factor to improve the accuracy of the system.

## References

1. Paula, M., Mikko, J.: Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. J. Applied Animal Behaviour Science 119, 32–38 (2009)
2. Yoshioka, H., Ito, M., Tanimoto, Y.: Effectiveness of a Real-time Radiotelemetric Pedometer for Estrus Detection and Insemination in Japanese Black Cows. J. Journal of Reproduction and Development 56, 351–355 (2010)
3. Ungar, E., Schoenbaum, I., Henkin, Z.: Inference of the Activity Timeline of Cattle Foraging on a Mediterranean Woodland Using GPS and Pedometry. J. Sensors 11, 362–383 (2011)
4. van Eerdenburg, F.J.C.M.: The pedometer, an automated aid in the detection of estrous. J. Veterinary Quarterly 30, 49–57 (2008)
5. Cristianini, N., Shawe-Taylor, J.: Methodology of support vector machine. Machine press, Beijing (2005)
6. Hairong, W., Dongmei, L., Yun, W., Weiguo, Y.: Fire Detecting Technology of Information Fusion using Support Vector Machines. In: 2010 International Conference on Artificial Intelligence and Computational Intelligence, pp. 194–198 (2010)
7. Liu, X.W., Wang, H.: Multiclass Support Vector Machines Theory and Its Data Fusion Application in Network Security Situation Awareness. In: 3rd International Conference on Wireless Communications, Networking and Mobile Computing, ShangHai, pp. 6349–6352 (2007)

8. Zhao, M., Shi, Z.: Human Activity Recognition with User-Free Acceleration in the Sensor Networks. In: 2005 International Conference on Neural Networks and Brain, BeiJing, pp. 1212–1217 (2005)
9. Guraliuc, A.R., Serra, A.A., Nepa, P., Manara, G., Potorti, F.: Detection and classification of human arm movements for physical rehabilitation. In: 2010 Antennas and Propagation Society International Symposium, pp. 1–4 (2010)
10. Joohyun, H., Namjin, K., Eunjong, C., Taesoo, L.: Classification Technique of Human Motion Context based on Wireless Sensor Network. In: 27th IEEE Engineering in Medicine and Biology Annual Conference, ShangHai, pp. 5201–5202 (2005)