

MEASURING PIG WELFARE BY AUTOMATIC MONITORING OF STRESS CALLS

Gerhard Manteuffel, Peter Christian Schön

Forschungsinstitut für die Biologie landwirtschaftlicher Nutztiere, Forschungsbereich Verhaltensphysiologie
Wilhelm-Stahl-Allee 2, D-18196 Dummerstorf, Germany E-Mail: manteuff@fhn-dummerstorf.de

Abstract: Vocalizations of animals are results of particular emotional states. Hence, stress screams of pigs may be indicators of disturbed welfare. Our objective was to develop a system that monitors and records the amount of stress calls and can be employed in environments of breeding, transportation and slaughtering. Using a combination of sound analysis by linear prediction coding and artificial neuronal networks we were able to detect stress vocalizations of pigs in noisy stables with only small recognition errors (< 5 %). The programmed system (STREMODU: stress monitor and documentation unit) running on a commercial laptop is insensitive to environmental noise, human speech and pig vocalizations other than screams. As a stand-alone device it can be routinely used for objective measurements of acute stress occurring in various farming environments and conditions. Since the system can be trained to various animal vocalizations an extended use for other purposes (c.f. oestrus detection) is also well within its scope.

Introduction

Animal vocalizations are voluntarily produced utterances which are able to indicate specific emotional states occurring spontaneously or induced by external events (Jürgens, 1979; Weary & Frazer, 1995; Schrader & Todt, 1998). In contrast to human language where meanings may change, animal vocalizations are usually produced according to fixed programs developed during phylogeny and achieved in ontogeny. Hence, in animals vocalizations are specifically and invariantly attributed to particular inner states, may be with the exception of apes, where a more creative use might be possible under certain circumstances.

Changing emotional states can be the cause or the effect of physiological and / or behavioural reactions. The latter can be measured and taken as a reference for the species-specific and individual meaning of a vocalization (Schrader & Rohn, 1997). The dependence of animal vocalization on psycho-physiological conditions makes sound analysis a well-suited tool for non-invasive judgments of welfare and stress. The procedure consists of two elements: a syntactic which describes the formal features and a semantic which derives the meaning from coincident behavioral and physiological reactions.

In farm animals housing conditions may produce stress and impairments of welfare if biological needs remain transiently or permanently unfulfilled. Pigs specifically can easily be stressed which leads to an activation of sensory and limbic brain centres and, eventually, to reactions of the sympathico-adrenomedullary and the hypothalamo-pituitary-adrenal axes. In parallel motor centres may be triggered controlling behavioural components of the stress reaction. One of these is vocalization performed by laryngeal and pharyngeal muscles. In pigs the result is a rather sustained scream containing high frequency bands with some dynamic components.

Up to now, methods of analysis do not exist that allow the recognition of arbitrary sounds because general solutions for extraction and classification are virtually absent. False or too few features may result in insufficient classification and too many features can overflow most computers, especially if real-time capability is demanded. Hence, differentiated well-adapted procedures are required for nearly every classification task (Schön et al., 1999).

Here, a procedure is presented that allows the recognition of stress screams of domestic pigs. It combines Linear Prediction Coding (LPC) with an artificial neural network (Schön et al.

2001) and is able to calculate and record in real time while neglecting other vocalizations and sounds (patent pending).

Methods

LPC uses signal changes instead of the signal itself. The continuous analog time signal is digitized and sampled. From the series of samples a probe $x(n)$ is taken together with the preceding one $x(n-1)$ and the linear prediction of $x(n)$ is calculated. Prediction error $e(n)$ is minimized using the coefficient a_1 such that

$$e(n) = x(n) - a_1x(n-1).$$

Taking an increased number of preceding samples and respective coefficients $a_1 \dots a_p$ leads to an improvement of the procedure:

$$e(n) = x(n) - a_1x(n-1) - a_2x(n-2) - \dots - a_px(n-p).$$

Z-transformation

$$E(z) = X(z) - \sum_{k=1}^p a_k z^{-k} * X(z)$$

delivers a formal equivalence to the source-filter-model of the vocal tract (Fant, 1970) where the LPC-coefficients correspond to the filter coefficients of the vocal tract. Hence the model is sensitive to variations of the resonance frequencies of the vocal tract and indirectly to the motor efforts which are required to obtain them. Polynomial development allows to calculate the LPC-spectrum from the LPC-coefficients. The resulting LPC-sonograms resemble sonograms based on Fast Fourier Transformation (FFT) but are based on distinctly fewer parameters (Fig. 1). In a typical stress vocalization of pigs frequency bands occur which may vary in time but usually display a more or less continuous course.

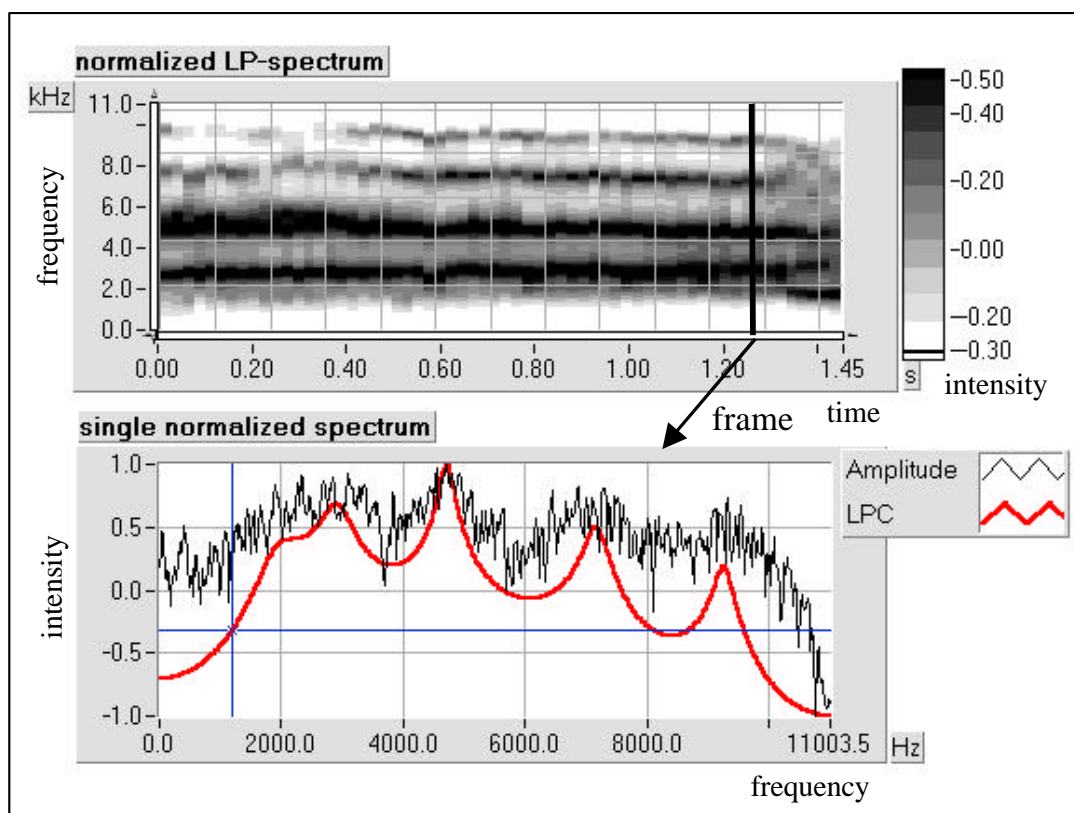


Fig. 1. LPC-analysis of a pig stress call. Top: Time dependent course of the LPC-spectra as a LPC-sonogram. Bottom: Amplitude- and LPC- spectrum (bold line) of a single time window as indicated.

For practical purposes 12 LPC-coefficients, equivalent to the first 6 resonance frequencies, are calculated using time windows of 46.44 ms duration.

Artificial neuronal networks are proven tools for many classification tasks. They consist of one- to multi-dimensional groups of nodes ("neurons"), each calculating a scalar output value from input vectors of arbitrary dimensionality. The inputs excite the neurons by a mostly simple transfer function ("weight"). Each neuron sums its inputs and outputs the result by another transfer function. Since the input space of all input vectors usually is transferred to many or all neurons a single neuron does not represent a specific feature of the input signal but rather the sum of arbitrary components (sub-symbolic distributed signal processing). The output of the network is observed and can be adapted by changing the input weights as long as it has not reached the desired value ("supervised learning"). Alternatively other network-types may self-organize the input space according to specific rules ("unsupervised learning"). Basically both types of networks are suitable for classification and have their own advantages and disadvantages. We obtained good results using the supervised Perceptron type (Rosenblatt, 1962) or Self-Organizing Feature Maps (Kohonen, 1982) taking the LPC-coefficients of single time windows as 12-dimensional input vectors.

The procedures to obtain the vectors and for network programming were developed using the graphical language LabVIEW[®] with the supplementary software DataEngine V.i.[®]. A detailed description of the whole procedure can be found in Schön et al. (2001).

The system was trained with stress calls of pigs of various ages recorded in a noise-reduced chamber with a Sennheiser MKE 46 microphone and stored on a Sony DCT-790 DAT-Recorder (Schön et al., 1998). Stressors applied were immobilization of 10 piglets by holding them upright at the thorax and keeping them above the floor, of 10 growing pigs by forcing them on the back, and of 10 sows which were held with a nose snare. In addition, pig grunts and different types of noise occurring in pig plants were trained as examples of non-stress. The result is a labeled classification neural network with 193 neurons in 4 layers. Only the inputs which belong to the class "stress" are forwarded to the output and result in a registration on the display and a data record where the occurrence of stress screams is related to time. The whole system was termed STREMODO (stress monitoring and documentation, Fig. 2).

Results and performance

The system was tested in a first run using vocalizations and noise samples which were recorded together with the training sets but were not included in them (Schön et. al., 2001; Schön and Manteuffel, 2001). The vocalization and noise were played back via loudspeaker and classified in real time. The result is shown in table 1, demonstrating that very few misclassifications of stress calls (< 1 %) occurred and that also only few acoustical non-stress events were faultily recognized as stress calls.

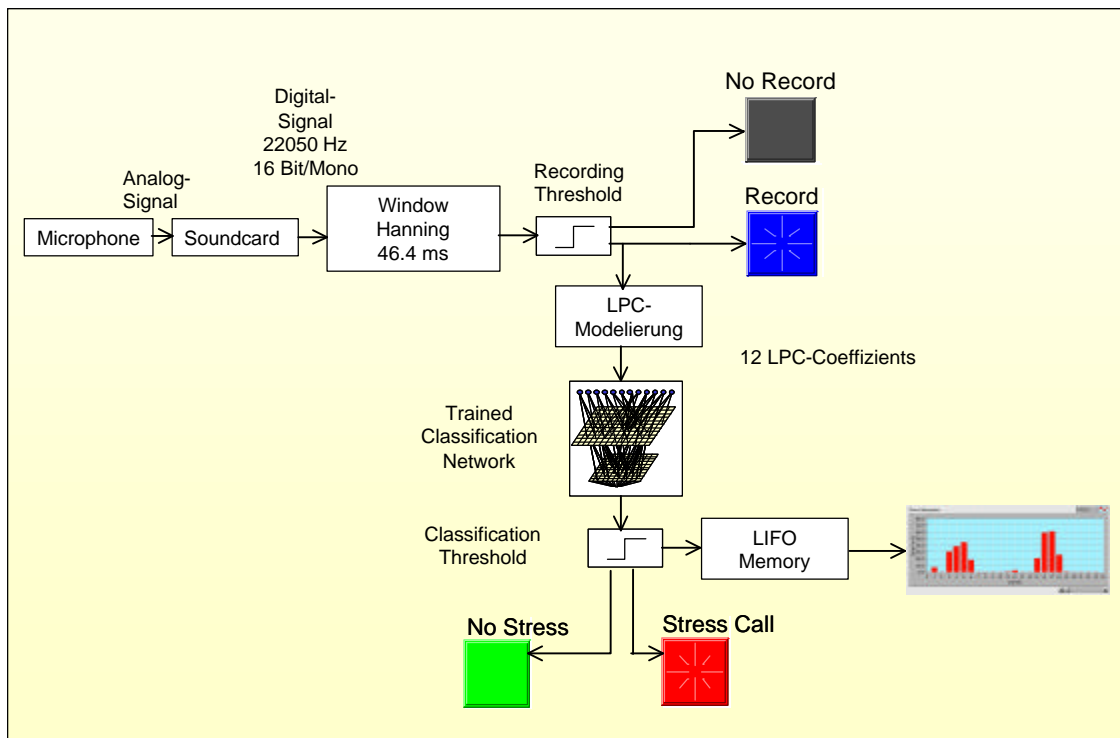


Fig. 2. General layout of STREMODO. The upper right signal indicates the activity of the system (record / no record). The bottom signals indicate recognition of a stress call being green (no stress) or red (stress). Occurrence of stress calls is written to a list and the rate is indicated in a separate window.

Table 1

Animals (Age)	n	Calls/Noise	Number of LPC-Vectors	Type	Misclassification (%)
Piglets (2 nd week)	10	Cries	1904	Stress	0.58
Piglets (5 th week)	7	Cries	2476	Stress	0.85
Piglets (2 nd week)	3	Grunts	171	no Stress	2.34
Piglets (5 th week)	3	Grunts	245	no Stress	2.04
Sows (1 st lactation)	5	Suckling grunts	60	no Stress	1.67
Noise	-	Stable-noise	1706	no Stress	1.23

Classification results of the STREMODO prototype

Subsequently the system was tested for its performance in a more realistic situation in a pig-plant in Mecklenburg-Vorpommern, Germany. We compared the occurrence of stress screams in two different feeding regimes. In one case the animal-to-feeding-place ratio was 6:1 with 24 fattening pigs (average weight 50 kg) in a 4.85 x 3.90 m² concrete / slatted floor pen (sensor-aided system where feed was added at a fixed time if the 1.50 m trough was emptied). In the second system feed was supplied once daily at an animal to feeding place ratio of 1:1 (4 m trough) in a 4.0 x 1.80 m² slatted floor pen with 11 animals (average weight 80 kg). For measurement the microphone of STREMODO was placed 2 m above floor level in the centre of the pen and the system was allowed to record for several hours. In parallel a video system

recorded the animals' behaviour and vocalization. At the resulting STREMODO records large increases in screaming during the first two feeding sessions were detected with the 6:1 feeding place ratio (Fig. 3a). The respective behavioural recordings displayed a high competition among the animals for the feed in these situations. In the third feeding episode the animals displayed less screaming and less competition, probably due to a more saturated state. In contrast, in the pen with the 1:1 feeding place ratio the animals showed no detectable increase in screaming at the feeding time as compared to other times when occasional fighting screams were detected (Fig. 3b). These results demonstrate that STREMODO is able to reveal the amount of stress vocalizations in noisy stable environments and can be used for continuous monitoring of porcine stress calls. The preliminary evaluation of the video registration of the animal behaviour and vocalization delivered a very good correspondence to the automatic STREMODO record and misclassifications of less than 5 %, as in the prototype laboratory tests.

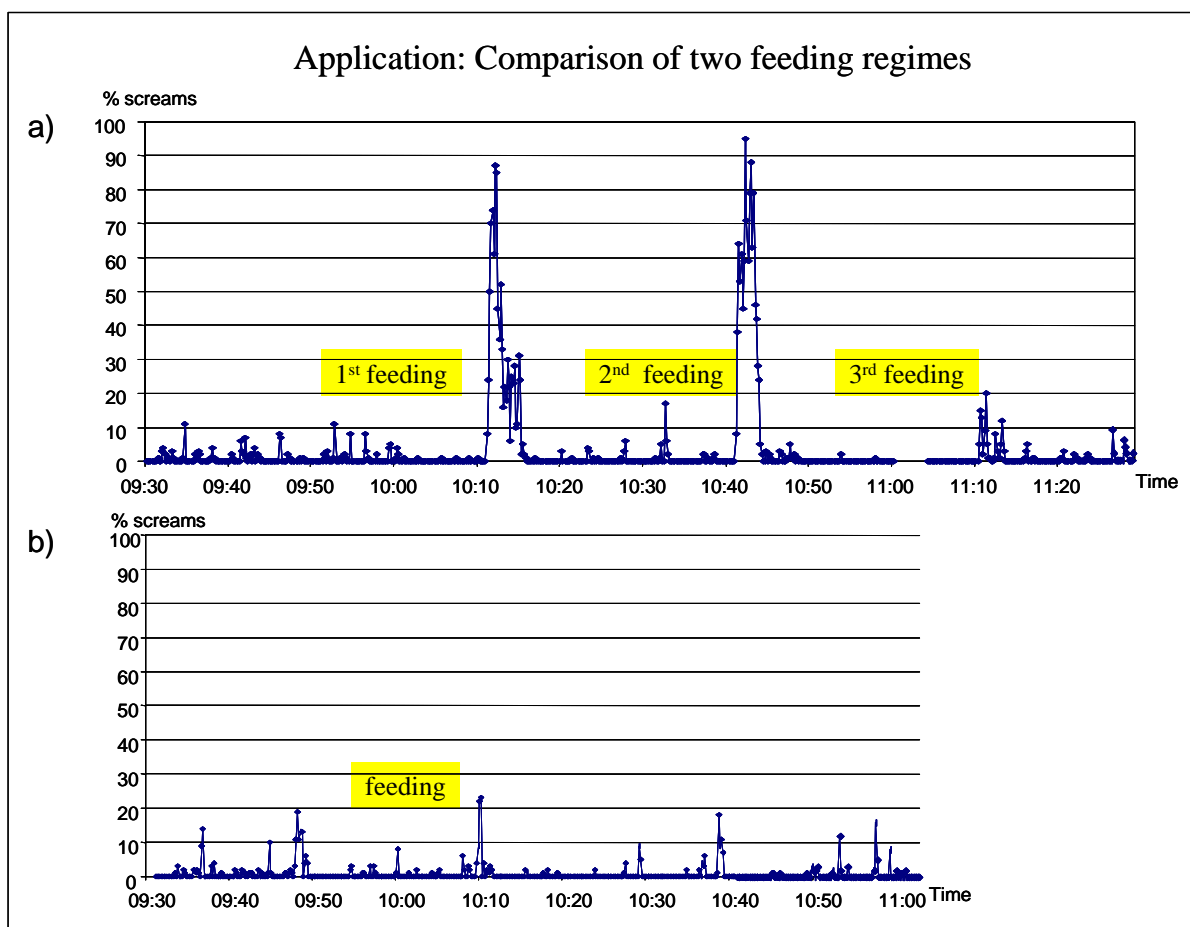


Fig. 3. The temporal percentage of the duration of stress vocalization as recorded by STREMODO.

- a) Restrictive feeding using a 6:1 relation of animals / feeding-place. (trough-sensor-feeding)*
b) Trough-feeding with a 1:1 relation

Discussion

The STREMODO system for recognition of stress that is expressed by vocalization consists of an initial step of extraction of the elementary characteristic syntactic elements using LPC-analysis and a subsequent classification by a trained artificial neuronal network. Its performance is well within the range of benchmarks given in the respective literature (c.f. Dreyfus, 1992; Gramß and Strube, 1992). Hence, the system is suited to detect stress calls of

pigs with a well-satisfying accuracy, as was also shown by practical application in a farming environment. The automatic protocol on the time and amount of stress calls makes it suitable for objective welfare monitoring in commercial pig housing. In addition, it can be applied in behavioural research when during long term experiments stress vocalization may be a reasonable parameter.

For these applications two specific characteristics of STREMODO are particularly advantageous. The system is considerably insensitive to noise and non-stress vocalizations which are always occurring in pig housings. Further, being an automat it is always ready for monitoring due to its real-time performance and delivers a list of its results for later evaluation. At the present developmental stage STREMODO runs on a laptop (Maxdata Artist Stanford 13.3). For most practical purposes, however, a special low-cost stand-alone unit which is robust against dust and harmful gas frequently found in stables will be more favourable.

In part the efficiency of STREMODO is due to the relatively simple approach which is possible because stress screams of pigs are quite sustained and not very much modulated. If modulations occur at all, their temporal order is largely irrelevant. In the case of other applications a more extended network for analysis with greater emphasis on temporal parameters will be necessary, which is not a fundamental problem, however. Since in important mammalian farm animals most vocalizations which may indicate emotional states are nevertheless comparatively simply structured, rapid adaptation of the system to such species is possible. In particular, the rutting vocalization of cows seems to be well detectable.

References

- Dreyfus, G. (1992) Neural Networks for the automatic recognition of handwritten digits. In: Applications of Neural Networks (H.G. Schuster, Ed.), Verlag Chemie, Weinheim, pp 35-60
- Fant, G. (1970) Acoustic Theory of Speech Production. Mouton, The Hague.
- Gramß, T.; Strube, H.W. (1992) Word recognition with a fast learning neural net. In: Applications of Neural Networks (H.G. Schuster, Ed.), Verlag Chemie, Weinheim, pp 223-237
- Jürgens, U. (1979). Vocalization as an emotional indicator. A neuroethological study in the squirrel monkey. Behaviour 69, 88-117
- Kohonen, T. (1982) Self-organized formation of topologically correct feature maps. Biol. Cybern. 43, 59-69
- Rosenblatt, F. (1962) Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms. Spartan Books, Washington D.C.
- Schön, P.C.; Puppe, B.; Manteuffel, G. (1998) A sound analysis system based on LabVIEW[®] applied to the analysis of suckling grunts of domestic pigs *Sus scrofa*. Bioacoustics 9, 119-133
- Schön, P.C.; Puppe, B.; Manteuffel, G. (1999) Common features and individual differences in nurse grunting of domestic pigs (*Sus scrofa*): A multi-parametric analysis. Behaviour 136, 49-66
- Schön, P.C.; Puppe, B.; Manteuffel, G. (2001) LPC-analysis and Self Organizing Feature Map (SOFM) as tools to classify stress calls of domestic pigs (*Sus scrofa*), J. Acoust. Soc. Am. 110, 1425-1431
- Schön, P.C.; Manteuffel, G. (2001) Nichtinvasive Beurteilung emotionaler Stressbelastung von Nutztieren mittels Vokalisationserkennung durch ein künstliches neuronales Netzwerk. In: Proc. 15. IGN-Tagung: Tierschutz und Nutztierhaltung (D. Schäfer and E. v. Borell, eds.) pp 104-109, Halle, Germany

- Schrader, L.; Rohn, C. (1997) Lautäußerungen von Hausschweinen als Indikator für Stressreaktionen. *Landbauforsch. Völkenrode* 47, 89-95
- Schrader, L.; Todt, D. (1998) Vocal quality is correlated with levels of stress hormones in domestic pigs. *Ethology* 104, 859-876
- Weary, D. M.; Fraser, D. (1995) Calling by domestic piglets: reliable signals of need? *Anim. Behav.* 50, 1047-1055.