Combined Uncertainty Model for Best Wavelet Selection

Samer Arafat*, Marjorie Skubic*, and Kevin Keegan**
* Computer Engineering and Computer Science
** Department of Veterinary Medicine and Surgery
University of Missouri-Columbia

Abstract- This paper discusses the use of combined uncertainty methods in the computation of wavelets that best represent horse gait signals. Combined uncertainty computes a composite of two types of uncertainties, fuzzy and probabilistic. First, we introduce fuzzy uncertainty properties and classes. Next, the gait analysis problem is discussed in the context of correctly classifying wavelet-transformed sound gait from lame horse gait signals. Continuous wavelets are selected using generalized information theory-related concepts that are enhanced through the application of uncertainty management models. Our experimental results show that models developed by maximizing combined uncertainty produce better results, in terms of neural network correct classification percentage, compared to those computed using only fuzzy uncertainty.

Index Terms- maximum uncertainty, combined uncertainty, horse gait analysis, continuous wavelets, neural networks.

I. INTRODUCTION

Pal and Bezdek have reviewed several known measures of fuzzy uncertainty, and they have introduced two general classes, called multiplicative and additive, that satisfy five properties [1], [2]. Briefly speaking, using $\mathbf{H}^{\mu}(\mathrm{A})$, which is a measure of fuzzy uncertainty that is defined on a fuzzy set, A, and a membership function, μ, the five properties are (1) *Sharpness*, which requires zero uncertainty when an event is likely to occur (or not to occur), (2) *Maximality*, which requires maximum uncertainty values when the uncertainty in the occurrence and non-occurrence of an event is equally likely, i.e. when $\mu = 0.5$, as in Figure 1, (3) *Resolution*, which requires $\mathbf{H}^{\mu}(\mathrm{A}) \geq \mathbf{H}^{\mu}(\mathrm{A}^{*})$, such that $\mathbf{H}^{\mu}(\mathrm{A}^{*})$ is a sharpened (i.e. closer to either 0 or 1) version of $\mathbf{H}^{\mu}(\mathrm{A})$, (4) S*ymmetry*, that $\mathbf{H}^{\mu}(\mathrm{A}) = \mathbf{H}^{\mu}(1-\mathrm{A})$, and (5) *Valuation*, that $\mathbf{H}^{\mu}(\mathrm{A}\cup\mathrm{B}) + \mathbf{H}^{\mu}(\mathrm{A}\cap\mathrm{B}) = \mathbf{H}^{\mu}(\mathrm{A}) + \mathbf{H}^{\mu}(\mathrm{B})$. A plot for a measure that satisfies these properties is shown in Figure 1. The reader is referred to [1] for a detailed mathematical discussion of these properties.

The multiplicative class of fuzzy uncertainty is based on non-negative, monotone increasing concave functions. The additive class is broader and requires only non-negative concave functions. Several existing measures were shown to relate to these two classes [1], [2].

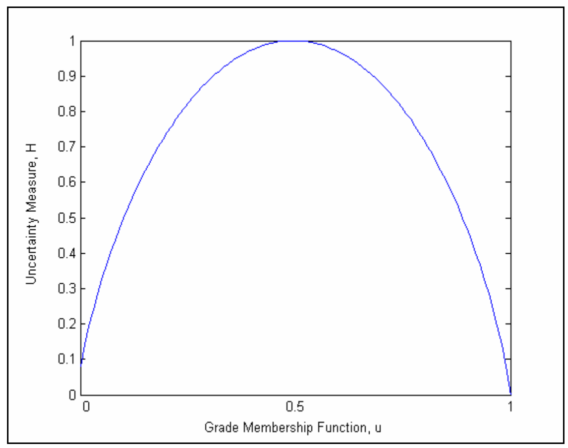


Fig. 1. A plot for a measure that satisfies all five fuzzy uncertainty properties.

Complex information systems are expected to have several types of uncertainties, such as fuzzy, probabilistic, and non-specificity [1], [2]. *Fuzzy Uncertainty* deals with the imprecision or vagueness associated with the occurrence of an event. In contrast, *Probabilistic Uncertainty* models the uncertainty of an event belonging to a crisp set. A model that integrates these two uncertainty types may be called *Combined Uncertainty* (named Total Uncertainty in [1] and [2]). Although models that use combined uncertainty may be difficult to interpret [1], this paper shows that such models may offer a more complete model of system uncertainty and provide a better measure for wavelet selection. In this paper, we compare the use of a combined uncertainty model with that of either fuzzy or probabilistic uncertainty models for best wavelet selection in analyzing equine gait signals.

The paper is organized as follows. Section II briefly discusses horse gait analysis and its classification methodology. Section III provides details for our implementation of combined uncertainty models. Section IV explores experimental results, and finally, Section V concludes with final remarks and recommendations for future work.

II. HORSE GAIT ANALYSIS

Subjective visual observation is the commonly used form of gait lameness and neurological evaluation in humans and animals. It is based on the examiner's ability to form a comprehensive image of how a normal subject moves and then how this movement is changed with lameness. Visual assessment is subjective and, as such, is difficult because the examiner needs to consider several different and rapidly changing bodily movement patterns [3] [4] [5] [6] [7].

To address this limitation, we are studying objective computational methods for recognition and analysis of mild forelimb horse gait lameness. Analyzing mild lameness is a difficult problem because signal characteristics are relatively similar to those of sound gait. We have established a generalized approach to classify horse lameness into one of three classes: sound, right-sided forelimb and left-sided forelimb.

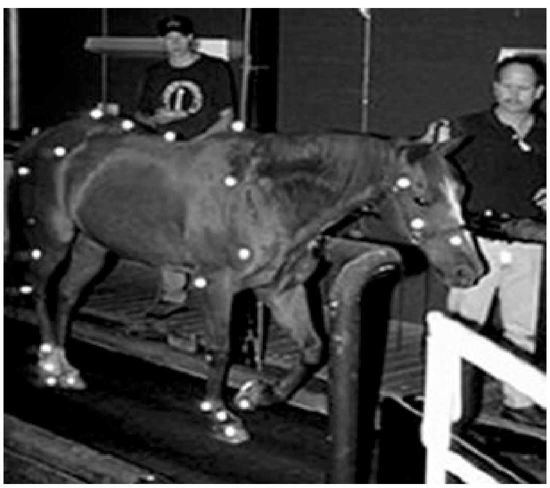


Fig. 2. A Horse trotting on the treadmill. Markers attached to the horse appear as white dots in the picture.

Twelve horses were used to collect a navicular data set. These horses were affected with navicular disease, a degenerative bone disease that causes forelimb foot lameness. Gait data is collected from markers that are fixed at selected body parts of a horse trotting on a high-speed treadmill (Figure 2). The information is captured from reflective markers using a 5-camera system that is connected to a computer. Kinematic data is extracted in the form of 3 dimensional coordinates collected at 120 Hz.

Next, the kinematic signals are transformed using continuous wavelets that best represent the corresponding signals. The best representation is the one that maximizes wavelet-transformed information content using measures of uncertainty. The wavelets are selected from a database consisting of 65 wavelets (from the Matlab Toolbox[1]) that represent a wide spectrum of known wavelet classes.

We define a grade membership function, $\mu_{kj} : C_k \rightarrow \{0,1\}$, which computes the possibility of occurrence of a wavelet-transformed signal's maximum movement such that,

$$\mu_{kj} = \frac{c_{kj}^2 - \min(\mathbf{C}_k^2)}{\max(\mathbf{C}_k^2) - \min(\mathbf{C}_k^2)} \qquad (1)$$

where,

$\mathbf{C}_k^2 = \left\{c_{k1}^2, c_{k2}^2, \ldots, c_{kN}^2\right\}$,

k is the k^{th} signal data sequence,

j = 1, …, N data points.

Next, fuzzy uncertainty is computed using $\mathbf{H}(A)$, or $\mathbf{H}^{\mu}(A)$. For example, consider Shannon's probabilistic measure, $\mathbf{H}_S^p(P) = -\sum_j p_{kj}\log_2 p_{kj} + (1-p_{kj})\log_2(1-p_{kj})$. A fuzzy version of this measure was introduced by Deluca and Termini [9], and is plotted in Figure 1,

$$\mathbf{H}_{DTE}^{\mu} = -\sum_j \mu_{kj}\log_2 \mu_{kj} + (1-\mu_{kj})\log_2(1-\mu_{kj}) \qquad (2)$$

Given K data sequences and M wavelets (here, M=65), we define a function, **B**, that computes the best representative wavelet, Ψ_k, for the k^{th} data sequence such that the self-information associated with this data source, i.e. its uncertainty, is maximized:

$$\Psi_k = B(\Psi_{ki}) = \max_i \{H(\Delta_{ki})\} \qquad (3)$$

where,

$\Delta_{ki} = \{\mu(c_{kj})\}$ or $\{p(c_{kj})\}$, is either the possibility or probability vectors of the k^{th} signal, depending on the measure that is being applied.

i = 1, …, M.

Basically, the best wavelet is the one that computes a transformation with maximum uncertainty. The higher the uncertainty is, the higher an information system's self-information is [8].

A wavelet that best represents all available signal data is the one that is selected *most often* for k data sequences.

This generalized methodology can be used on any of the measured kinematic signals. For the experiments described in this paper, two signals were used, the vertical movement

[1] Matlab 6.0, The MathWorks, Inc., Natick, MA.

of the poll (head) and the right front foot. After the best wavelets for the poll and front foot data sequences are selected, signals are transformed using the respective best wavelets at a set of pre-selected wavelet scale values. Examples of such transformations using best wavelets are in Figures 3 and 4.

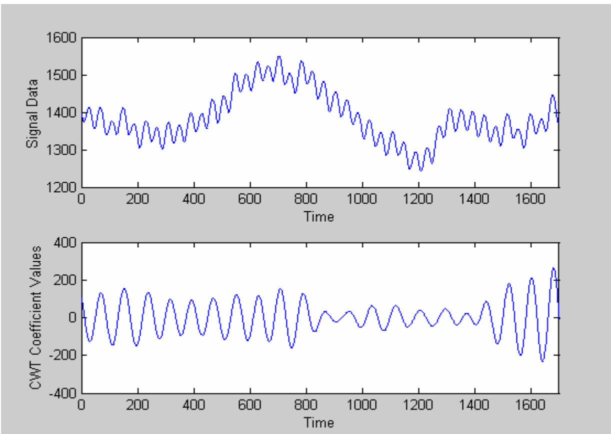


Fig. 3. (a) A typical raw signal of a horse's poll. This horse is lame on the left side. (b) the transformed signal at a scale value 64 and the Morlet wavelet.

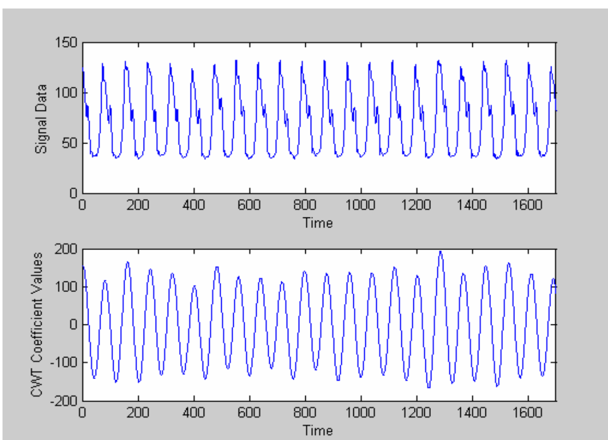


Fig. 4. (a) A typical raw signal of a horse's front foot. This horse is lame on the left side. (b) the transformed signal at a scale value 32 and the Mexican Hat wavelet.

The last stage in feature extraction is designed to capture temporal trends within each signal. A time-sequence (TS) process composition process is used to capture the temporal trends in a gait signal by building feature vectors such that 3, 5, or 7 adjacent transformed points are combined into each feature vector.

Finally, a neural network is trained using back propagation. The input of the neural network consists of feature vectors for the transformed, TS composed, poll and front foot signals. The output of the neural network consists of three units that represent 3 classes, right forelimb lameness, left forelimb lameness, and sound movement. The neural network captures the temporal correlation between the two signals, i.e., the poll and front foot. The neural network's performance is computed using the percentage of correctly classified horse data. We used 6 horse data sets per class to train the neural network and 2 horse data sets per class for testing. For one full round of experiments, we rotated this selection of training and test horse data, until each horse data has been used at least twice for training/testing. Therefore, the experiments were repeated 8 times within each round. Seven rounds of experiments have been computed and a median of correct classification percentage (CCP) of testing feature vectors has been recorded. Experimental results have shown that wavelet selection using probabilistic or fuzzy uncertainty produces a better CCP compared to a standard method that uses visual inspection to select a wavelet that "looks" most similar to signal characteristics. A plot of this and other section's experiments is in Section III, below.

However, this intermediate conclusion poses interesting questions: can the wavelet selection process be improved if both kinds of uncertainties, fuzzy and probabilistic, are combined? If so, what are possible models for such fusion of information? These and other conclusions [1] motivate experimentation with combined uncertainty models.

III. Combined Uncertainty Model

Pal and Bezdek addressed this question and concluded that the formulation of such integration is difficult, but important [1]. Also, they commented that previous works that either added or multiplied together these uncertainties are hard to interpret. Finally, they observed that much of the work they reviewed did not incorporate a third type of uncertainty, *resolutional uncertainty*, or *non-specificity*, which reflects the ambiguity in specifying the exact solution. This paper experiments with combined uncertainty that is computed by multiplying probabilistic uncertainty with fuzzy uncertainty measures.

Given K sets of gait data sequences (or signals), consider the k^{th} wavelet-transformed data sequence, $\{c_{kj}\}$, which occurs at a corresponding possibility vector, $\{\mu(c_{kj})\}$. A measure of fuzzy uncertainty, $\mathbf{H}^{\mu}(A)$, may be computed using either one of the following three fuzzy measures that Pal and Bezdek reviewed from the literature [1], [2]:

1. The first to measure fuzzy uncertainty without reference to probabilities were Deluca and Termini [9], and is listed in Equation 2, above.

2. Another measure was developed by Pal and Pal [10]:

$$\mathbf{H}^{\mu}_{\text{PPE}}(A) = K\sum_{j}\mu_{kj}e^{1-\mu_{kj}} + (1-\mu_{kj})e^{\mu_{kj}} \tag{4}$$

3. A more recent measure was developed by Pal and Bezdek such that it satisfies their definition for additive fuzzy measures [1], [2]:

$$\mathbf{H}^{\mu}_{\alpha\text{QE}}(\alpha, A) = K\sum_{j}\mu_{kj}{}^{\alpha}(1-\mu_{kj})^{\alpha} \tag{5}$$

where,

$\alpha \in (0,1)$,

K is a normalization constant.

Equation 5 may be written as:

$$\mathbf{H}^{\mu}_{\alpha\text{QE}}(\alpha, A) = K\sum_{j}\mathbf{S}^{\mu}_{\alpha\text{QE}},$$

where, $\mathbf{S}^{\mu}_{\alpha\text{QE}} = \mu_{kj}{}^{\alpha}(1-\mu_{kj})^{\alpha}$.

Therefore, α is used as a "sensitivity to μ_{kj}" tune controller. If α is close to 0, $\mathbf{S}_{\alpha\text{QE}}$ is insensitive to changes in μ_{kj}. If α is close to 1, $\mathbf{S}_{\alpha\text{QE}}$ becomes sensitive to changes in μ_{kj}. An $\alpha = 0.5$ balances out these two trends [1]. We used an $\alpha = 0.5$ in this study's experiments.

Next, we define p_{kj}, a probabilistic uncertainty measure that computes the probability of occurrence for c_{kj} such that,

$$p_{kj} = \frac{c_{kj}{}^{2}}{\|\mathbf{C}_k\|^{2}} \tag{6}$$

Finally, we define the combined uncertainty, $\mathbf{H}^{\text{com}} = \mathbf{H}^{\mathbf{com}}(\mathbf{A}, \mathbf{P})$, associated with the occurrence of c_{kj} such that,

$$\mathbf{H}^{\text{com}} = \mathbf{H}^{p} \cdot \mathbf{H}^{\mu}\text{, or} \tag{7}$$

$$\mathbf{H}^{\mathbf{com}}(\mathbf{A}, \mathbf{P}) = \mathbf{H}^{\mathbf{p}}(\mathbf{P}) \cdot \mathbf{H}^{\mu}(\mathbf{A})$$

where,

$\mathbf{H}^{p}$ is a probabilistic measure.

One probabilistic measure is the one defined by Equation 6, and therefore we use

$$\mathbf{H}^{p} = p_{kj} \tag{8}$$

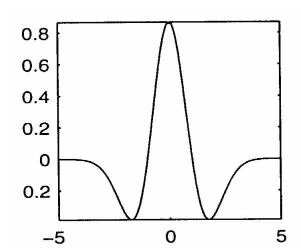
We observe that in the case where $\mathbf{H}^{\mu} = \mathbf{H}^{\mu}_{\text{DTE}}$, i.e., $\mathbf{H}^{\text{com}} = p_{kj} \cdot \mathbf{H}^{\mu}_{\text{DTE}}$, then this is the same as $\mathbf{H}^{\text{com}}_{\text{DTE}}$, which was developed in [9] and reviewed in [1].

We use this new model to select wavelets that best transform horse gait signals, in a maximum uncertainty sense. The discussion in the next section compares experimental results using fuzzy uncertainty versus combined uncertainty models.

IV. Experiments And Results

Experiments were computed using either our best wavelet selection algorithm with an uncertainty measure, or using a standard method that selects wavelets by inspecting the wavelets, visually, and deciding which wavelets are most similar to the raw vertical poll and front foot signals. The Matlab graphical display was used to select wavelets that appear to be similar to raw signal characteristics. The wavelets selected were the Biorthogonal 1.3 wavelet for the poll, at a scale value of 64, and the Symlet 7 wavelet for the front foot, at a scale value of 32.

Using the navicular data set and the best representative wavelet selection (maximum uncertainty selection) method described in Section II, wavelets were selected using either fuzzy uncertainty, probabilistic uncertainty, or combined uncertainty, by applying Equations 1-8, above. We note that either one of the 3 fuzzy measures, $\mathbf{H}^{\mu}_{\text{DTE}}$, $\mathbf{H}^{\mu}_{\text{PPE}}$, or $\mathbf{H}^{\mu}_{\alpha\text{QE}}$ produced the same results. $\mathbf{H}^{p}_{s}$ was used to compute probabilistic uncertainty.



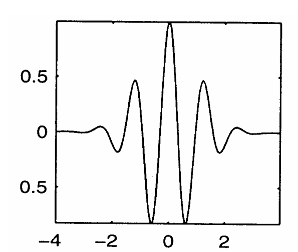


Fig. 5 (a) The Mexican Hat Wavelet (b) The Morlet Wavelet

When using either fuzzy uncertainty or probabilistic uncertainty, models produced the same results, and the two wavelets selected were the Morlet wavelet, selected for the poll signal at scale 64, and the Gaussian 2 wavelet, selected for the front foot signal, selected at scale 32. Finally, using the combined uncertainty model (CUM), the two wavelets selected were the Morlet wavelet, selected for the poll signal at scale 64, and the Mexican Hat wavelet at scale 32 (these two wavelets are shown in Figure 5).

A neural network was trained using the transformed data and CCP values of horse data were computed. Using the cross validation method, described above, in Section III, experimental results were computed and are plotted in Figure 6. An 83% CCP was computed using wavelets selected with combined uncertainty, compared to 72% CCP using wavelets selected with only one measure of uncertainty.

Results shown in Figure 6 show that there is a noticeable improvement in performance when wavelets are selected using combined uncertainty models compared to those selected using only probabilistic or fuzzy uncertainties.

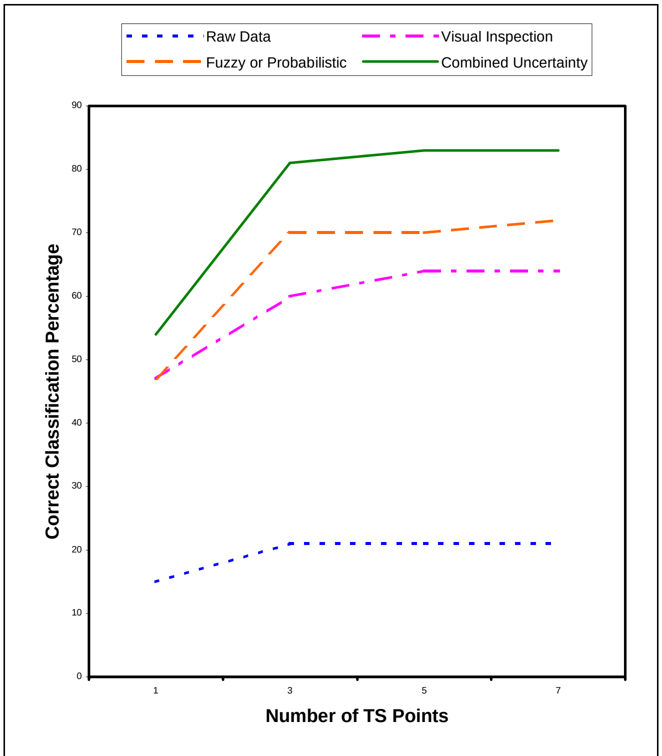


Fig. 6 Correct Classification Percentages for 5 sets of preprocessing methods. The top curve shows results for the BWS using Combined Uncertainty. The lowest curve shows results for training with untransformed signals.

This suggests that, on one hand, there are at least two types of uncertainties involved in the computation of horse gait signals. On the other hand, the results suggest that a model integrating both probabilistic and fuzzy uncertainties show promise for the formulation of a more complete uncertainty model for a complex information system.

V. Conclusion

We reviewed three measures of fuzzy uncertainty that are available in the literature and showed that a multiplicative combination that incorporates probabilistic and fuzzy uncertainties produced higher performance when applied to the horse gait classification problem. These results show promise for maximizing combined uncertainties, which better model self-information for a complex information system.

A third type of uncertainty includes non-specificity, or resolutional uncertainty, which was not addressed by this paper. Integrating the three types of uncertainty would compute Total Uncertainty, which would be an interesting topic for future research.

References

[1] N. Pal and J. Bezdek, “Measuring Fuzzy Entropy,” *IEEE Transactions on Fuzzy Systems*, vol. 2, no. 2, May 1994.

[2] N. Pal and J. Bezdek, “Quantifying Different Facets of Fuzzy Uncertainty,” *Fundamental of Fuzzy Sets*, Chapter 9, 2000.

[3] S. Holzreiter and M. Kohle, “Assessment of Gait Patterns using Neural Networks,” *Journal of Bioinformatics*, vol. 26, no. 6, pp. 645-651, 1993.

[4] H. M. Lakany, “A Generic Kinematic Pattern for Human Walking,” *Neurocomputing*, vol. 35, pp. 27-54, 2000.

[5] C. N. Kobluk, D. Schnurr, and F. D. Horney, “Use of high speed cinematography and computer generated gait diagrams for the study of equine hind limb kinematics,” *Equine Vet Journal*, vol. 21, no. 1, pp. 185-188, 1989.

[6] Davis, R. B., “Reflections on Clinical Gait Analysis,” *Journal of Electromygraphy and Kinesiology*, vol. 7, No. 4, pp. 251-257, 1997.

[7] M. Saleh and G. Murdoch, “In defense of gait analysis,” *Journal of Bone and Joint Surgery*,” vol. 67, no. 2, pp. 237-241, 1985.

[8] C. E. Shannon and W. Weaver, The *Mathematical Theory of Communication*, Urbana, The University of Illinois Press, 1949.

[9] A. Deluca and S. Termini, “A definition of non-probabilistic entropy in the setting of fuzzy set theory,” *Information and Control*, vol. 20, pp. 301-312, 1972.

[10] N. R. Pal and S. K. Pal, “Object-Background Segmentation using New Definitions of Entropy,” *IEE Proceeding*, pp. 284-295, 1989.